Eidgenössische Technische Hochschule Zürich

# Analysis I und II

Vorlesungsskript 2017/2018

Manfred Einsiedler und
Andreas Wieser

1. Juni 2018

## Vorwort

Willkommen an der ETH Zürich (oder zumindest zu Ihrem Studium dieses Vorlesungsskripts der ETH Zürich). Dieses Skript wurde erstmals im Studienjahr 2016/2017 für den Vorlesungszyklus Analysis I und II der Studiengänge Interdisziplinäre Naturwissenschaften Bachelor, Physik Bachelor und Mathematik Bachelor erstellt. Im zweiten Jahr der Entwicklung dieses Skripts werden wir vermehrt interaktive Inhalte in der eSkript-Version dieses Textes einsetzen.

Die Analysis I/II und Lineare Algebra I/II Vorlesungen bilden gemeinsam das Fundament der Mathematik, wie sie an der ETH und auch an anderen Universitäten unterrichtet wird. Die allermeisten weiteren Mathematik- und Physik-Vorlesungen werden auf diesen vier Vorlesungen aufbauen und deren Themengebiete erweitern – ohne aber die Zeit zu haben, diese grosszügig zu wiederholen. Damit dieser „theoretische Turm“, in dem eine Vorlesung auf die nächste aufbaut, stabil bleibt, muss das Fundament umso stabiler sein. Für die Analysis I/II-Vorlesung und für dieses Skript bedeutet das unter anderem, dass mathematische Exaktheit gross geschrieben wird.

Wie wir sehen werden, enthalten die Analysis I/II-Vorlesungen viele Themen rund um die und insbesondere auch die Differential- und Integralrechnung. Manche dieser Themen kennen Sie wahrscheinlich schon aus dem Gymnasium; wir werden trotzdem kein grosses Vorwissen voraussetzen (abgesehen von einem intuitiven Verständnis von Variablen und einer gewissen Übung für algebraische Umformungen). Auch werden Sie einen wichtigen Unterschied bemerken: obwohl wir versuchen werden, so viele Themen wie möglich mit Beispielen zu belegen, werden diese viel weniger im Mittelpunkt stehen, als Sie dies wahrscheinlich gewohnt sind. Ebenso ist das Einüben der erlernten Algorithmen und Rechenmethoden nicht der Fokus unserer Vorlesung und wird Ihnen überlassen. Vielmehr werden die Beispiele auch für Sie unübliche Formen annehmen, wenn mitunter das Beispiel darin besteht, die Existenz der Exponentialfunktion zu zeigen oder die Stirling-Formel für das asymptotische Verhalten von $n!$ zu bestimmen. Im Mittelpunkt steht vielmehr der Aufbau der mathematischen Theorie, die notwendig ist, um Themen wie Differentialgleichungen und mehrdimensionale Integralsätze zu erreichen und zu verstehen.

Wie bereits erwähnt, hat die eSkript-Version dieses Skripts weitere Funktionalitäten, die von einem PDF-File nicht unterstützt werden können.

- (Applets) Viele der Themen werden anschaulicher wenn die Objekte nicht nur in einem Bild dargestellt werden, sondern wenn sich dieses Bild auch verändern lässt. Wir setzten hierfür GeoGebra ein um in mehreren Applets die Theorie anschaulicher darzustellen.

  Hier war es uns ein Anliegen, dass die Bedienung sowohl auf einem PC als auch auf mobilen Geräten uneingeschränkt möglich ist. Leider haben wir es noch nicht geschafft, dass die Verwendung dieser Applets innerhalb des eSkripts auf allen Geräten ohne Probleme abläuft. Wenn Sie beim Laden eines Applets innerhalb des eSkripts zum Beispiel auf einem mobilen Gerät auf Probleme stossen, dann können sie stattdessen das Wort

„Applet“ anklicken, womit das Applet auf einer eigenen Webseite geladen wird und besser funktionieren sollte. Des Weiteren ist das PDF an der gleichen Stelle auch mit dem Applet verlinkt und der Link erlaubt Ihnen, dass Sie mit dem GeoGebra-Applet weiter experimentieren oder dieses verändern können.

Wir wollen aber auch erwähnen, dass diese Hilfen zwar den Einstieg erleichtern sollen, aber das Ziel sein sollte, dass Sie anschliessend auch weitere Theorien ohne derartige Applets verstehen und sich vorstellen können. Das beste Verständnis eines Themas erhält man, wenn man sich die Beweise aneignet und sich ein eigenes Bild davon macht, was in dem Beweis eigentlich passiert.

- (Annotation) Mit Hilfe des Servers hypothes.is können alle Studenten der Analysis-Vorlesung Kommentare am Rand des eSkripts anbringen. Diese Funktionalität versteckt sich meist rechts oben. Wir empfehlen Ihnen gleich nach Erhalt des Zugangscodes für die hypothesis-groups einen Account anzulegen, damit Sie Zugang zu allen Diskussionen haben. Es kann sein, dass Sie einige Zeit nach Erstellen des Accounts Probleme haben, sich bei Ihrem Account einzuloggen (zum Beispiel auf Safari). Grund dafür kann sein, dass in den Privacy-Settings Cookies nur für besuchte Webseiten erlaubt sind. In diesem Fall entweder die Privacy-Settings ändern oder die Seite hypothes.is kurz besuchen, anschliessend klappt der login auch innerhalb des eSkripts wieder für einige Zeit. Es gibt mehrere Arten von Kommentaren:

  - (public) Kommentare, die für jeden sichtbar sind. Für fachliche Fragen an andere Studenten, Assistenten und die Professoren. hypothes.is unterstützt ein etwas eingeschränktes „latex“, das heisst, dass Sie kleinere Formeln ohne grosse Probleme in den Diskussionen am Rand eingeben können.
  - (privat) Ihre eigenen Kommentare, die für den Rest der Welt unsichtbar sind.
  - (spoiler) Wenn Sie Lösungen zu einer Übung präsentieren wollen, oder Hinweise weiter ausbauen wollen, so sollten Sie dies hier machen. Denn viele Studenten wollen vielleicht selber zuerst über ein Problem nachdenken (was zu empfehlen ist) bevor sie eine Lösung sehen wollen.
  - (tippfehler) Für Rückmeldungen zu Tippfehler (von fachlichen Fehler bis zu Rechtschreibfehler) im Skript.
  - (technik) Für Meldungen, wenn Formeln falsch oder gar nicht angezeigt werden, Applets auf gewissen Platformen nicht laufen, oder andere technische Probleme, die weder Tippfehlern sind noch mit dem Inhalt des Skripts zu tun haben. Zum Beispiel wollten wir, dass man in einem Applet nur die relevanten Punkte bewegen kann und nicht jedes einzelne Objekt. Wenn Sie etwas finden, dass man bewegen kann obwohl dies inhaltlich keinen Sinn macht oder sogar bis zum Neustart der Webseite das Applet zerstört, bitte hier melden.

  Allgemein gilt, dass Sie Besitzer Ihrer Kommentare sind und wir diese nicht löschen können. Technisch/rechtlich die einzige Möglichkeit für uns ist ein kompletter Neustart

der Webseite unter einer anderen Addresse, was alle Kommentare unabhängig vom Inhalt löscht. Da dies während des akademischen Jahres nicht erwünschenswert ist, bitten wir Sie ihre Kommentare höflich zu formulieren und auf das Fach zu beschränken.

- (Aufgaben und Multiple Choice) Im Skript werden auch einige Übungsaufgaben und Fragen gestellt, die Sie zur Beschäftigung mit den besprochenen Themen anregen sollen. Manchmal haben die Übungsaufgaben auch Lösungen, meist nur Hinweise, aber die Fragen vom Typ Multiple Choice haben immer die richtige Antwort und eine kurze Erklärung. Da der Lernerfolg sich aber nur dann einstellen kann, wenn Sie sich vor dem Lesen der Lösung, des Hinweises oder der Antwort mit der Aufgabe beschäftigen und versuchen diese eigenständig zu lösen, finden Sie diese Hilfen nur im eSkript nach einem zusätzlichen Freiklicken des entsprechenden Texts.


## Danksagung

Wir möchten uns bei Menny Akka für viele Vorschläge von Beispielen für die ersten Kapitel des Skripts sowie bei Manuel Lüthi, Roland Prohaska und Philipp Wirth für zahlreiche Korrektur- und Verbesserungsvorschläge bedanken. Des Weiteren wollen wir uns bei Carina Heiss, Anh Huy Truong und Anian Altherr für die Erstellung der verschiedenen geogebra-Applets bedanken. Unser Dank gilt auch den Kommentaren der folgenden Studentinnen und Studenten, Hilfsassistentinnen und Hilfsassistenten der Analysis-Vorlesung: Donior Akhmedov, Samet Armagan, Theodor Batel, Joel Beimler, Vladislav Bunkin, Chiara Caduff, Leon Carl, Maximilian Daschner, Janek Denzler, Till Dieminger, Marc Dollmann, Charlotte Dombrowsky, Alessandro Fasse, Rainer Feichtinger, Atsuhiro Funatsu, Maxim Gerspach, James Golub, Marcel Graetz, Oliver Gräub, Sebastián Guerrero Soriano, Anna Gütl, Jenny Held, Johannes Hruza, Ivan Ilak, Marco Isele, Erik Jahn, Richard Karl, Oriel Kiss, Can Knaut, Benedikt König, Thea Kosche, Raphael Leu, Clemens Macho, Nadine Merk, Marco Moldenhauer, Andreas Mono, Chrishon Nilanthan, Helena Obrist, Mak Planincic, Oliver Rietmann, Quentin Roubaty, Luca Rüegg, Katharina Sauder, Pascal Schmuki, Janik Schüttler, Tolga Sevim, Nicole Sieber, Alexander Smirnow, Felix Sefzig, Taro Spirig, Vinzenz Stampf, Lidia Stocker, Martin Stoller, Stephan Tornier, Moritz Wittmer, Benjamin Zayton und David Zollikofer.

# Inhaltsverzeichnis

**1 Einführung*** **1**

1.1 Quadratur der Parabel . . . 1
1.2 Einige Tipps . . . 6
1.3 Logische Begriffe . . . 8
1.3.1 Aussagenlogik und die Boolesche Algebra . . . 8
1.3.2 Prädikatenlogik . . . 11
1.4 Mengenlehre und Abbildungen . . . 16
1.4.1 Naive Mengenlehre . . . 16
1.4.2 Abbildungen . . . 21
1.4.3 Relationen . . . 33
1.4.4 Mächtigkeit . . . 40
1.5 Zahlenmengen . . . 48
1.5.1 Konstruktion der ganzen Zahlen aus den natürlichen Zahlen . . . 49
1.5.2 Teilbarkeit und Kongruenzen . . . 51
1.6 Beweise . . . 53
1.6.1 Widerspruchsbeweise . . . 53
1.6.2 Kontraposition . . . 54
1.6.3 Induktionsbeweise . . . 55
1.6.4 Das Schubfachprinzip . . . 56
1.6.5 Weitere Methoden . . . 57
1.6.6 Beweise finden . . . 58
1.6.7 Beweise aufschreiben . . . 58
1.6.8 Beweise lesen . . . 59
1.6.9 Prädikatenlogik vs Umgangssprache . . . 59
1.7 Weitere Lernmaterialien . . . 61
1.7.1 Verwendung des Kapitels . . . 61
1.7.2 Flächeninhalt . . . 62
1.7.3 Logik . . . 62
1.7.4 Funktionen und Relationen . . . 63
1.7.5 Beweismethoden . . . 64
1.7.6 Geometrische Probleme . . . 65
1.7.7 Übungen zu Primzahlen . . . 65

1.7.8 Lernkarten . . . 67

**2 Die reellen Zahlen** **68**

2.1 Die Axiome der reellen Zahlen . . . 68
2.1.1 Körperaxiome . . . 68
2.1.2 Angeordnete Körper . . . 74
2.1.3 Das Vollständigkeitsaxiom . . . 76
2.1.4 Eine erste Anwendung der Vollständigkeit . . . 77
2.1.5 Verwendung der reellen Zahlen und der Axiome . . . 78
2.2 Die natürlichen Zahlen . . . 80
2.2.1 Definition der natürlichen Zahlen und vollständige Induktion . . . 80
2.2.2 Die ganzen Zahlen . . . 85
2.2.3 Die rationalen Zahlen . . . 86
2.2.4 Division mit Rest und Anfänge der Zahlentheorie* . . . 87
2.2.5 Verwendung der ganzen Zahlen und deren Eigenschaften . . . 88
2.3 Die komplexen Zahlen . . . 89
2.3.1 Verwendung der komplexen Zahlen . . . 94
2.4 Intervalle und der Absolutbetrag . . . 95
2.4.1 Intervalle . . . 95
2.4.2 Der Absolutbetrag auf den reellen Zahlen . . . 96
2.4.3 Der Absolutbetrag auf den komplexen Zahlen . . . 98
2.5 Maximum und Supremum . . . 102
2.5.1 Maximum und Minimum . . . 102
2.5.2 Supremum und Infimum . . . 102
2.5.3 Uneigentliche Werte, Suprema und Infima . . . 105
2.5.4 Verwendung des Supremums und des Infimums . . . 106
2.6 Konsequenzen der Vollständigkeit . . . 108
2.6.1 Das Archimedische Prinzip . . . 108
2.6.2 Häufungspunkte einer Menge . . . 110
2.6.3 Intervallschachtelungsprinzip . . . 111
2.6.4 Überabzählbarkeit . . . 113
2.6.5 Die Cantor-Menge* . . . 114
2.7 Modelle und Eindeutigkeit der Menge der reellen Zahlen* . . . 118
2.7.1 Ebene Geometrie und die Zahlengerade . . . 118
2.7.2 Dezimalbrüche . . . 118
2.7.3 Dedekind-Schnitte . . . 118
2.7.4 Vervollständigung der rationalen Zahlen . . . 119
2.7.5 Definition mittels Steigungen . . . 119
2.7.6 Eindeutigkeit . . . 120
2.8 Weitere Lernmaterialien . . . 124
2.8.1 Verwendung des Kapitels . . . 124
2.8.2 Weitere Übungsaufgaben . . . 125

2.8.3 Multiple-Choice Fragen . . . 128
2.8.4 Lernkarten . . . 128

**3 Funktionen und die reellen Zahlen 129**
3.1 Summen und Produkte . . . 129
3.1.1 Rechenregeln für die Summe . . . 131
3.1.2 Rechenregeln für das Produkt . . . 133
3.1.3 Die geometrische Summe . . . 134
3.2 Polynome . . . 135
3.2.1 Polynomdivision . . . 138
3.2.2 Nullstellen und Interpolation . . . 139
3.2.3 Algebraische und transzendente Zahlen . . . 141
3.3 Die Fakultät und der Binomialsatz . . . 142
3.3.1 Fakultät . . . 142
3.3.2 Binomialkoeffizienten . . . 143
3.3.3 Der binomische Lehrsatz . . . 145
3.3.4 Eine Summe von Binomialkoeffizienten* . . . 146
3.4 Reellwertige Funktionen . . . 151
3.4.1 Beschränktheit . . . 151
3.4.2 Monotonie . . . 152
3.5 Stetigkeit . . . 155
3.5.1 Komplex-wertige Funktionen . . . 160
3.6 Der Zwischenwertsatz . . . 161
3.7 Der Satz über die Umkehrabbildung . . . 164
3.7.1 Wurzeln aus natürlichen Zahlen . . . 166
3.8 Stetige Funktionen auf kompakten Intervallen . . . 168
3.8.1 Beschränktheit . . . 168
3.8.2 Maximum und Minimum . . . 169
3.8.3 Gleichmässige Stetigkeit . . . 170
3.9 Weitere Lernmaterialien . . . 174
3.9.1 Verwendung des Kapitels . . . 174
3.9.2 Weitere Übungsaufgaben . . . 174
3.9.3 Lernkarten . . . 177

**4 Das Riemann-Integral 178**
4.1 Treppenfunktionen und deren Integral . . . 178
4.1.1 Zerlegungen . . . 179
4.1.2 Das Integral einer Treppenfunktion . . . 180
4.2 Definition des Riemann-Integrals . . . 184
4.3 Erste Integrationsgesetze . . . 189
4.3.1 Linearität . . . 189
4.3.2 Monotonie . . . 191

4.3.3 Teilintervalle . . . 193
4.4 Anwendungen . . . 196
4.4.1 Intervallfunktionen . . . 196
4.4.2 Flächeninhalt . . . 197
4.4.3 Masse, Momente und Schwerpunkt . . . 198
4.4.4 Geleistete Arbeit . . . 198
4.4.5 Vorteil des Integralbegriffs . . . 199
4.5 Integrierbarkeit monotoner Funktionen . . . 200
4.6 Integration von Polynomen . . . 203
4.7 Integrierbarkeit stetiger Funktionen . . . 206
4.8 Weitere Lernmaterialien . . . 208
4.8.1 Verwendung des Kapitels . . . 208
4.8.2 Weitere Übungsaufgaben . . . 208
4.8.3 Multiple-Choice Fragen . . . 209
4.8.4 Lernkarten . . . 209
4.9 Einschub: Mehrdimensionale Integrale* . . . 210
4.9.1 Definition mittels Treppenfunktionen . . . 210
4.9.2 Iterierte Integrale . . . 211
4.9.3 Schwerpunkt eines Körpers . . . 213
4.9.4 Polarkoordinaten . . . 215
4.9.5 Kugelkoordinaten . . . 218
4.9.6 Zusammenfassung . . . 221

**5 Folgen und Grenzwerte 223**
5.1 Konvergenz von Folgen . . . 223
5.1.1 Zusammenhang zur Stetigkeit . . . 228
5.1.2 Teilfolgen . . . 229
5.1.3 Cauchy-Folgen . . . 231
5.1.4 Reduktion auf reelle Folgen . . . 231
5.2 Reelle Folgen . . . 233
5.2.1 Monotone Folgen . . . 234
5.2.2 Limes superior und Limes inferior . . . 235
5.2.3 Konvergente Teilfolgen . . . 238
5.2.4 Reelle Cauchy-Folgen . . . 241
5.2.5 Uneigentliche Grenzwerte . . . 242
5.2.6 Ein Diagramm für die Zusammenhänge der Begriffe und Sätze . . . 243
5.3 Die Exponentialfunktion . . . 244
5.3.1 Eine Interpretation . . . 244
5.3.2 Konvergenz der Folge . . . 244
5.3.3 Inversionsformel . . . 247
5.3.4 Additionsformel . . . 247
5.3.5 Stetigkeit . . . 248

5.3.6 Strenge Monotonie . . . 248
5.3.7 Surjektivität . . . 248
5.3.8 Der Logarithmus und Potenzen . . . 249
5.4 Grenzwerte von Funktionen . . . 252
5.4.1 Grenzwerte und punktierte Umgebungen . . . 252
5.4.2 Links- und rechtsseitige Grenzwerte . . . 254
5.4.3 Einseitige Stetigkeit und Sprungstellen . . . 255
5.4.4 Die Bewegung nach Unendlich . . . 256
5.4.5 Umgebungsfilter und Konvergenz entlang eines Filters* . . . 256
5.4.6 Einige Rechenbeispiele . . . 259
5.5 Riemann-Summen . . . 261
5.6 Landau Notation . . . 265
5.7 Normen und Konvergenz auf Vektorräumen . . . 268
5.7.1 Die euklidsche Norm auf $\mathbb{C}^d$ . . . 269
5.7.2 Normäquivalenz . . . 271
5.7.3 Folgenkonvergenz . . . 272
5.7.4 Konvergenz in endlich dimensionalen Vektorräumen . . . 272
5.7.5 Stetigkeit . . . 275
5.7.6 Vektorwertige Integrale . . . 275
5.8 Weitere Lernmaterialien . . . 277
5.8.1 Verwendung des Kapitels . . . 277
5.8.2 Übungen . . . 278
5.8.3 Lernkarten . . . 280

**6 Reihen, Funktionenfolgen und Potenzreihen 281**
6.1 Reihen . . . 281
6.1.1 Reihen mit nicht-negativen Gliedern . . . 285
6.1.2 Bedingte Konvergenz . . . 288
6.1.3 Alternierende Reihen . . . 291
6.1.4 Das Cauchy-Kriterium . . . 292
6.2 Absolute Konvergenz . . . 294
6.2.1 Hinreichende Kriterien für absolute Konvergenz . . . 294
6.2.2 Umordnen von Reihen . . . 296
6.2.3 Produkte . . . 297
6.3 Konvergenz von Funktionenfolgen . . . 301
6.3.1 Punktweise Konvergenz . . . 301
6.3.2 Gleichmässige Konvergenz . . . 303
6.4 Potenzreihen . . . 307
6.4.1 Konvergenzradius . . . 307
6.4.2 Addition und Multiplikation . . . 309
6.4.3 Stetigkeit bei Randpunkten . . . 310
6.5 Die (komplexe) Exponentialabbildung . . . 312

6.5.1 Darstellung durch die Potenzreihe . . . 312
6.5.2 Die komplexe Exponentialreihe . . . 313
6.5.3 Die Additionsformel . . . 314
6.5.4 Der Absolutbetrag der Exponentialabbildung . . . 315
6.6 Trigonometrische Funktionen . . . 316
6.6.1 Additionsformeln . . . 317
6.6.2 Die Kreiszahl . . . 318
6.6.3 Tangens und Cotangens . . . 320
6.6.4 Polarkoordinaten und Multiplikation auf den komplexen Zahlen . . . 321
6.6.5 Zwei Logarithmen auf der komplexen Zahlenebene . . . 324
6.7 Integration von Potenzreihen . . . 327
6.7.1 Die hyperbolischen Funktionen . . . 328
6.8 Ziffernentwicklungen und fraktale Konstruktionen* . . . 330
6.8.1 Die Cantor-Menge . . . 331
6.8.2 Cantors Teufelstreppe . . . 332
6.8.3 Peanos raumfüllende Kurve . . . 333
6.9 Weitere Lernmaterialien . . . 338
6.9.1 Verwendung des Kapitels . . . 338
6.9.2 Übungen . . . 339
6.9.3 Lernkarten . . . 341

**7 Differentialrechnung 342**
7.1 Die Ableitung . . . 342
7.1.1 Definition und geometrische Interpretation . . . 342
7.1.2 Beispiele und Ableitungsregeln . . . 344
7.1.3 Extremwerte . . . 352
7.1.4 Stetige Differenzierbarkeit . . . 353
7.1.5 Ableitungen höherer Ordnung . . . 354
7.2 Zentrale Sätze der Differentialrechnung . . . 358
7.2.1 Der Mittelwertsatz . . . 358
7.2.2 Korollare des Mittelwertsatzes und Kurvendiskussion . . . 360
7.2.3 Konvexität . . . 362
7.2.4 Mittelwertsatz nach Cauchy . . . 366
7.2.5 Regel von de l'Hôpital . . . 367
7.3 Trigonometrische Funktionen . . . 371
7.3.1 Sinus und Arkussinus . . . 371
7.3.2 Kosinus und Arkuskosinus . . . 372
7.3.3 Tangens und Arkustangens . . . 373
7.3.4 Kotangens und Arkuskotangens . . . 375
7.3.5 Ein physikalisches Beispiel* . . . 375
7.3.6 Verwendung der trigonometrischen Funktionen . . . 377
7.4 Hyperbolische Funktionen . . . 378

7.4.1 Der Areasinus Hyperbolicus . . . 378
7.4.2 Der Areakosinus Hyperbolicus . . . 378
7.4.3 Der Areatangens Hyperbolicus . . . 379
7.5 Erste Differentialgleichungen . . . 380
7.5.1 Differenzengleichungen . . . 381
7.5.2 Stammfunktionen . . . 383
7.5.3 Lineare Differentialgleichungen erster Ordnung . . . 384
7.5.4 Zweite Ordnung . . . 387
7.6 Weitere Lernmaterialien . . . 397
7.6.1 Verwendung des Kapitels . . . 397
7.6.2 Übungen . . . 398
7.6.3 Lernkarten . . . 402

**8 Die Ableitung und das Riemann-Integral 403**
8.1 Der Fundamentalsatz der Integral- und Differentialrechnung . . 403
8.1.1 Differentiation von Potenzreihen . . . 406
8.1.2 Die alternierende harmonische Reihe . . . 408
8.1.3 Die Leibniz-Reihe . . . 409
8.2 Integrationsmethoden . . . 410
8.2.1 Partielle Integration . . . 411
8.2.2 Substitution . . . 412
8.2.3 Integration rationaler Funktionen . . . 413
8.2.4 Trigonometrische Substitution . . . 417
8.2.5 Weitere Integrationsmethoden . . . 419
8.2.6 Neue Funktionen . . . 420
8.2.7 Das bestimmte Integral . . . 421
8.2.8 Leibniz-Notation . . . 422
8.3 Anwendungen . . . 424
8.3.1 Flächeninhalte . . . 424
8.3.2 Bogenlänge . . . 425
8.3.3 Wegintegrale von Vektorfeldern . . . 428
8.3.4 Volumen von Rotationskörpern . . . 429
8.3.5 Oberflächen von Rotationskörpern . . . 430
8.4 Das uneigentliche Integral . . . 433
8.4.1 Uneigentliche Integrationsgrenzen . . . 433
8.4.2 Das Integral über unbeschränkte Funktionen . . . 436
8.4.3 Die Gamma-Funktion* . . . 438
8.5 Taylor Approximation . . . 441
8.5.1 Analytische Funktionen . . . 445
8.5.2 Konvergenzgeschwindigkeit des Newton-Verfahrens* . . . 446
8.6 Numerische Integration . . . 449
8.6.1 Landau-Notation II . . . 452

8.7 Drei asymptotische Formeln . . . 454
8.7.1 Das Wallissche Produkt . . . 454
8.7.2 Stirling-Formel . . . 455
8.7.3 Asymptotik der harmonischen Reihe* . . . 457
8.8 Weitere Lernmaterialien . . . 459
8.8.1 Verwendung des Kapitels . . . 459
8.8.2 Übungen . . . 459
8.8.3 Lernkarten . . . 462

**9 Metrische Räume 463**
9.1 Konvergenz in metrischen Räumen . . . 463
9.1.1 Definition und erste Beispiele . . . 463
9.1.2 Konvergenz von Folgen . . . 467
9.1.3 Der Raum der stetigen Funktionen . . . 468
9.1.4 Ein kurzer Überblick . . . 469
9.1.5 Euklidische Norm . . . 470
9.1.6 Wie sehen metrische Räume aus? . . . 471
9.2 Topologische Grundbegriffe . . . 472
9.2.1 Offene und abgeschlossene Teilmengen . . . 472
9.2.2 Häufungspunkte und Dichtheit . . . 476
9.2.3 Zusammenhang . . . 477
9.3 Stetigkeit . . . 480
9.3.1 Definition und Charakterisierungen . . . 480
9.3.2 Zwei stärkere Stetigkeitsbegriffe . . . 482
9.3.3 Konstruktion von stetigen Funktionen . . . 482
9.3.4 Eine Bemerkung zur Stetigkeit . . . 484
9.3.5 Stetigkeit und Zusammenhang . . . 485
9.4 Vollständigkeit . . . 489
9.4.1 Der Banachsche Fixpunktsatz . . . 490
9.4.2 Wahrscheinlichkeitsmatrizen* . . . 491
9.5 Kompaktheit . . . 495
9.5.1 Definitionen . . . 495
9.5.2 Äquivalenzen . . . 497
9.5.3 Endlich-dimensionale Vektorräume . . . 501
9.5.4 Bilder von kompakten Mengen . . . 502
9.5.5 Operatornorm von Matrizen . . . 502
9.5.6 Gleichmässige Stetigkeit und Oszillation . . . 503
9.6 Fundamentalsatz der Algebra . . . 506
9.7 Der Raum der stetigen Funktionen . . . 508
9.8 Konstruktion der reellen Zahlen* . . . 510
9.9 Weitere Lernmaterialien . . . 515
9.9.1 Verwendung des Kapitels . . . 515

9.9.2 Weitere Normen und Metriken . . . 515
9.9.3 Weitere Übungen . . . 517
9.9.4 Lernkarten . . . 522

**10 Mehrdimensionale Differentialrechnung 523**
10.1 Die Ableitung . . . 524
10.1.1 Der Definitionsbereich . . . 524
10.1.2 Lineare Abbildungen . . . 525
10.1.3 Definitionen . . . 525
10.1.4 Reduktion der Dimension . . . 529
10.2 Die Kettenregel und der Mittelwertsatz . . . 532
10.2.1 Verknüpfungen differenzierbarer Funktionen . . . 532
10.2.2 Geometrische Interpretation der mehrdimensionalen Kettenregel . . . 534
10.2.3 Der Mittelwertsatz . . . 535
10.3 Höhere Ableitungen und Taylor-Approximation . . . 538
10.3.1 Definition und Eigenschaften der höheren partiellen Ableitungen . . . 538
10.3.2 Mehrdimensionale Taylor-Approximation . . . 540
10.4 Extremwerte . . . 545
10.4.1 Beweis des Kriteriums für Definitheit* . . . 549
10.5 Parameterintegrale . . . 551
10.5.1 Die Bessel-Differentialgleichung* . . . 555
10.6 Wegintegrale . . . 558
10.6.1 Skalare Wegintegrale . . . 558
10.6.2 Wegintegrale von Vektorfeldern . . . 559
10.7 Konservative Vektorfelder . . . 562
10.7.1 Integrabilitätsbedingungen . . . 563
10.8 Weitere Lernmaterialien . . . 567
10.8.1 Verwendung des Kapitels . . . 567
10.8.2 Übungen . . . 567
10.8.3 Lernkarten . . . 569

**11 Anfänge der Differentialgeometrie 570**
11.1 Sätze zur impliziten Funktion und zur inversen Abbildung . . . 570
11.1.1 Satz zur impliziten Funktion . . . 572
11.1.2 Satz zur inversen Abbildung . . . 578
11.1.3 Polar- und Zylinderkoordinaten . . . 581
11.1.4 Kugelkoordinaten . . . 582
11.2 Teilmannigfaltigkeiten des Euklidschen Raumes . . . 583
11.2.1 Definition und Beispiele . . . 583
11.2.2 Niveaumengen als Teilmannigfaltigkeiten . . . 587
11.2.3 Tangentialraum und Tangentialbündel . . . 590
11.3 Extremwertprobleme . . . 594

11.3.1 Extrema auf kompakten Teilmengen . . . 594
11.3.2 Extrema mit Nebenbedingungen und Lagrange-Multiplikatoren . . . 595
11.3.3 Diagonalisierbarkeit symmetrischer Matrizen* . . . 598
11.3.4 Eine hinreichende Bedingung für Lagrange-Multiplikatoren* . . . 600
11.4 Weitere Lernmaterialien . . . 603
11.4.1 Verwendung des Kapitels . . . 603
11.4.2 Übungen . . . 603
11.4.3 Lernkarten . . . 604

**12 Mehrdimensionale Integralrechnung 605**
12.1 Das Riemann-Integral für Quader . . . 605
12.1.1 Definition . . . 605
12.1.2 Erste Eigenschaften . . . 609
12.2 Riemann-Integrierbarkeit und Stetigkeit . . . 612
12.2.1 Stetige Funktionen sind Riemann-integrierbar . . . 612
12.2.2 Sandwich-Charakterisierung von Riemann-Integrierbarkeit . . . 613
12.2.3 Integrierbarkeit über Riemann-Summen . . . 616
12.3 Das Lebesgue-Kriterium . . . 619
12.3.1 Nullmengen . . . 619
12.3.2 Charakterisierung mittels Stetigkeit . . . 622
12.3.3 Oszillation einer Funktion . . . 622
12.3.4 Das Lebesgue-Kriterium ist notwendig . . . 623
12.3.5 Das Lebesgue-Kriterium ist hinreichend . . . 624
12.4 Das Riemann-Integral über Jordan-messbare Mengen . . . 626
12.5 Der Satz von Fubini . . . 633
12.6 Mehrdimensionale Substitutionsregel . . . 641
12.6.1 Lineare Substitution . . . 641
12.6.2 Substitution mit Diffeomorphismen . . . 645
12.6.3 Beweis der Substitutionsregel mit kompakten Träger . . . 648
12.6.4 Beispiele . . . 657
12.7 Uneigentliche Mehrfachintegrale . . . 660
12.7.1 Ausschöpfungen und uneigentliche Integrale . . . 661
12.7.2 Substitution für uneigentliche Riemann-Integrale . . . 664
12.7.3 Eine Anwendung: das Basler Problem* . . . 666
12.8 Weitere Lernmaterialien . . . 669
12.8.1 Verwendung des Kapitels . . . 669
12.8.2 Übungen . . . 669
12.8.3 Lernkarten . . . 671

**13 Mehrdimensionale Integralsätze 672**
13.1 Divergenzsatz und der Satz von Green . . . 672
13.1.1 Der Divergenzsatz für Rechtecke und Bereiche unter einem Graphen . . 673

13.1.2 Glatt berandete Bereiche . . . 681
13.1.3 Glatt berandete Bereiche in der Ebene . . . 684
13.1.4 Divergenzsatz für glatt berandete Bereiche in der Ebene . . . 690
13.1.5 Rotation und der Satz von Green . . . 692
13.1.6 Eine Anwendung: der Jordansche Kurvensatz für glatte Kurven* . . . 693
13.2 Oberflächenintegrale . . . 696
13.2.1 Flächen und deren Karten . . . 696
13.2.2 Skalares Oberflächenintegral . . . 697
13.2.3 Flussintegrale entlang Oberflächen . . . 699
13.3 Der Divergenzsatz im dreidimensionalen Raum . . . 706
13.3.1 Der Divergenzsatz für Bereiche unter Graphen . . . 706
13.3.2 Der Divergenzsatz auf glatt berandeten Bereichen . . . 709
13.4 Der Laplace-Operator* . . . 712
13.4.1 Schwingungen der Trommel . . . 713
13.4.2 Harmonische Funktionen . . . 714
13.5 Der Satz von Stokes im dreidimensionalen Raum . . . 717
13.5.1 Die Wirbelstärke . . . 717
13.5.2 Glatt berandete orientierbare Flächen . . . 718
13.5.3 Der Satz von Stokes . . . 718
13.6 Zwei Werkzeuge der weiteren Analysis* . . . 724
13.6.1 Glättung durch Faltung . . . 724
13.6.2 Glatte Partitionen der Eins . . . 725
13.7 Konservative Vektorfelder* . . . 728
13.7.1 Homotopie . . . 728
13.7.2 Integrabilität . . . 729
13.7.3 Invarianz unter Homotopien . . . 730
13.7.4 Beweis der Konservativität . . . 732
13.8 Integrale und alternierende Tensorprodukte* . . . 733
13.8.1 Das skalare Integral über eine Teilmannigfaltigkeit . . . 733
13.8.2 Das alternierende Tensorprodukt . . . 734
13.8.3 Differentialformen . . . 738
13.9 Weitere Lernmaterialien . . . 740
13.9.1 Verwendung des Kapitels . . . 740
13.9.2 Übungen . . . 740
13.9.3 Lernkarten . . . 742

**14 Gewöhnliche Differentialgleichungssysteme 743**
14.1 Mehrdeutigkeit der Lösung . . . 743
14.1.1 Separierbare Differentialgleichungen: ein Leibniz Kochrezept . . . 746
14.2 Lineare Differentialgleichungen mit konstanten Koeffizienten . . . 747
14.2.1 Die Lösungen für die homogene Gleichung . . . 747
14.2.2 Eine Lösung der inhomogenen Differentialgleichung . . . 752

14.2.3 Lösen des Anfangswertproblems . . . 753
14.3 Differentialgleichungsysteme . . . 755
14.3.1 Umwandlung einer Differentialgleichung höherer Ordnung . . . 755
14.3.2 Lineare autonome Differentialgleichungssysteme . . . 756
14.3.3 Beispiele linearer autonomer Differentialgleichungssysteme . . . 760
14.3.4 Ein Beispiel eines nicht-linearen Systems . . . 763
14.4 Existenz- und Eindeutigkeitssatz von Picard-Lindelöf . . . 767
14.4.1 Mögliches Verhalten in der Nähe eines Fixpunktes* . . . 773
14.4.2 Stetige Abhängigkeit* . . . 774
14.4.3 Ein weiteres Beispiel: der rotierende Tropfen* . . . 775
14.5 Lineare Differentialgleichungssysteme* . . . 780
14.6 Weitere Lernmaterialien . . . 782
14.6.1 Verwendung des Kapitels . . . 782
14.6.2 Übungen . . . 782
14.6.3 Lernkarten . . . 782

**A Changelog** **783**

**B Christmas-Special*** **785**

**C Summer-Special 2018*** **799**
C.1 Vektorfelder und Flüsse auf einer Teilmannigfaltigkeit . . . 799
C.2 Geodätischer Fluss auf dem Tangentenbündel . . . 800
C.3 Physikalisch relevante Modifikationen . . . 802
C.4 Dynamische Systeme . . . 803
C.4.1 Existenz minimaler Teilmengen . . . 805
C.5 Der geodätische Fluss auf (Quotienten) der Ebene . . . 808
C.6 Ein Quotient einer nicht-euklidschen Ebene . . . 810

**D Kapitelübergreifende Fragen** **814**
D.1 Wie kann man Konvergenz einer Folge zeigen? . . . 814
D.2 Warum sind kompakte konvexe Mengen Jordan-messbar? . . . 816
D.3 Eine Anekdote zur Jordan-Messbarkeit . . . 817

**E Prüfungsrelevant** **819**
E.1 Rechnungen und Anwendungen der Theorie . . . 819
E.1.1 Erstes Semester, Übungen . . . 820
E.1.2 Zweites Semester, Übungen . . . 820
E.2 Vorlesung und Theorieteil . . . 821
E.2.1 Probeprüfung und Katalog . . . 823

**F Summer-Special 2017*** **824**
F.1 Der Satz des gekämmten Igels . . . . . 824
F.2 Der Brouwersche Fixpunktsatz . . . . . 825
F.3 Brouwerscher Fixpunktsatz via Igelsatz . . . . . 826
F.4 Brouwerscher Fixpunktsatz via Divergenzsätze . . . . . 829
F.5 Beweis des Igelsatzes . . . . . 831
F.6 Erweiterungen und Anwendungen . . . . . 834

**Liste der Symbole** **837**

**Literatur** **838**

# Kapitel 1

# Einführung*

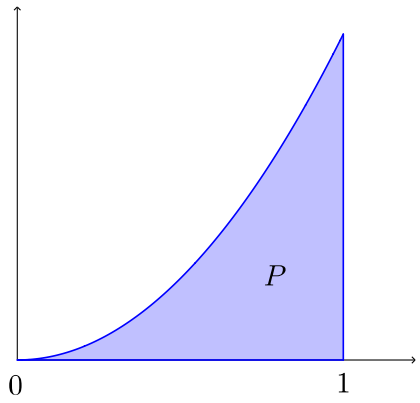
**Dieses Kapitel ist für den Aufbau notwendig, da es die im Weiteren benötigte Sprache entwickelt, doch werden die entsprechenden Themen nicht in der Analysis-Jahresprüfung 08/18 und 02/19 abgefragt.*

## 1.1 Quadratur der Parabel

Als Beispiel, wie wir hier denken und vorgehen wollen, aber auch als Einleitung in die Integralrechnung, werden wir uns in diesem Abschnitt mit dem Bereich

$$P = \left\{(x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq 1,\ 0 \leq y \leq x^2\right\} \tag{1.1}$$

unter der Parabel zwischen 0 und 1 beschäftigen und dessen Flächeninhalt berechnen. Dieser Flächeninhalt wurde als erster krummlinig begrenzter Bereich schon von Archimedes (ca. 287–ca. 212 v.Chr.) im 3. Jahrhundert v.Chr. bestimmt. (Historisch Interessierten empfehlen wir auch den Podcast der BBC über Archimedes, wobei man bei der zwanzigsten Minute einsteigen kann, wenn man wenig Zeit hat.) Wir wollen für die Flächenberechnung davon ausgehen, dass wir wissen, was die Symbole in der Definition in Gleichung (1.1) bedeuten und dass $P$ gerade den Bereich in folgender Figur (Figur 1.1) beschreibt.[1]




Figur 1.1: Der Bereich $P$.

[1] Insbesondere nehmen wir vorläufig an, dass wir die Menge der reellen Zahlen $\mathbb{R}$ bereits kennen.

Natürlich ist die Berechnung des Flächeninhalts von $P$ keine Herausforderung und innerhalb von Sekunden möglich, wenn wir das bestimmte (Riemann-) Integral und die dazugehörigen Rechenregeln verwenden. Wir wollen dies jedoch nicht als bekannt voraussetzen, da wir das Integral erst in etwa einem Monat einführen und verstehen werden.

Genau genommen müssen wir uns vor der Berechnung folgende fundamentale Frage stellen:

*Was ist eigentlich ein Flächeninhalt?*

Wenn wir diese Frage nicht genau beantworten können, dann können wir eigentlich nicht wissen, was es bedeutet, den Flächeninhalt von $P$ zu berechnen. Deswegen relativieren wir unser Ziel in folgender Weise – eine Proposition ist ein mathematischer Satz, also eine mathematische Aussage, mittlerer Bedeutung:

**Proposition 1.1** (Flächeninhalt unter der Parabel)**.** *Angenommen es gibt einen Begriff eines Flächeninhalts für Bereiche in $\mathbb{R}^2$, der folgende Eigenschaften erfüllt:*

- *Der Flächeninhalt des sogenannten abgeschlossenen Rechtecks*

$$[a,b] \times [c,d] = \big\{(x,y) \in \mathbb{R}^2 \mid a \leq x \leq b,\ c \leq y \leq d\big\}$$

  *und des sogenannten offenen Rechtecks*

$$(a,b) \times (c,d) = \big\{(x,y) \in \mathbb{R}^2 \mid a < x < b,\ c < y < d\big\}$$

  *ist gleich $(b-a)(d-c)$, wobei $a,b,c,d$ reelle Zahlen sind mit $a \leq b, c \leq d$.*

- *Falls $G$ ein Bereich in $\mathbb{R}^2$ ist und $F$ ein in $G$ enthaltener Bereich ist, dann ist der Flächeninhalt von $F$ kleiner oder gleich dem Flächeninhalt von $G$.*

- *Für Bereiche $F, G$ in $\mathbb{R}^2$ ohne gemeinsame Punkte ist der Flächeninhalt des vereinigten Bereiches $F \cup G$ die Summe der Flächeninhalte von $F$ und $G$.*

*Dann ist der Flächeninhalt von $P$ wie in Gleichung* (1.1) *(falls überhaupt definiert) gleich $\frac{1}{3}$.*

In anderen Worten: wir haben die Frage, ob es einen Begriff des Flächeninhalts gibt und für welche Bereiche dieser definiert ist, offengelassen, wollen aber zeigen, dass $\frac{1}{3}$ der einzige "vernünftige" Wert für den Flächeninhalt von $P$ darstellt. Die Idee des Beweises wird im eSkript anhand eines Applets visualisiert.

**Applet 1.2** (Abschätzung eines Flächeninhaltes)**.** *Wir verwenden jeweils bis zu 1000 Rechtecke um den Flächeninhalt von unten und von oben abzuschätzen. Im Beweis unten werden wir aber unbegrenzt viele Rechtecke verwenden und können damit den Flächeninhalt ohne jegliche Unschärfe genau bestimmen.*

Für den Beweis von Proposition 1.1 benötigen wir ein Lemma (auch Hilfssatz genannt):

**Lemma 1.3** (Summenformel mittels Induktion)**.** *Sei $n \geq 1$ eine natürliche Zahl. Dann gilt*

$$1^2 + 2^2 + \cdots + (n-1)^2 + n^2 = \frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6}. \tag{1.2}$$

*Beweis (mittels vollständiger Induktion).* Für $n = 1$ ist die linke Seite von Gleichung (1.2) gleich 1 und die rechte Seite gleich $\frac{1}{3} + \frac{1}{2} + \frac{1}{6} = 1$. Also stimmt Gleichung (1.2) für $n = 1$. *Dieser Beweisschritt wird Induktionsanfang genannt.*

Angenommen wir wissen bereits, dass Gleichung (1.2) für die natürliche Zahl $n$ gilt. Wir wollen nun zeigen, dass daraus folgt, dass Gleichung (1.2) auch für $n + 1$ gilt. Die linke Seite von Gleichung (1.2) ist gegeben durch

$$\begin{aligned} 1^2 + 2^2 + \cdots + n^2 + (n+1)^2 &= (1^2 + 2^2 + \cdots + n^2) + (n+1)^2 \\ &= \frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6} + (n+1)^2 \\ &= \frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6} + n^2 + 2n + 1 \\ &= \frac{n^3}{3} + \frac{3n^2}{2} + \frac{13n}{6} + 1 \end{aligned}$$

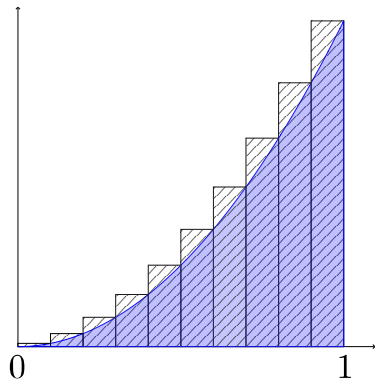
wobei wir Gleichung (1.2) für die Zahl $n$ verwendet haben. Die rechte Seite von Gleichung (1.2) ist gegeben durch

$$\begin{aligned} \frac{(n+1)^3}{3} + \frac{(n+1)^2}{2} + \frac{n+1}{6} &= \frac{n^3 + 3n^2 + 3n + 1}{3} + \frac{n^2 + 2n + 1}{2} + \frac{n+1}{6} \\ &= \frac{n^3}{3} + \frac{3n^2}{2} + \frac{13n}{6} + 1 \end{aligned}$$

Damit ist gezeigt, dass die linke und die rechte Seite von Gleichung (1.2) auch für $n + 1$ übereinstimmen. *Dieser Beweisschritt wird Induktionsschritt genannt.*

Es folgt, dass Gleichung (1.2) wegen dem Induktionsanfang für $n = 1$ stimmt und daher auch für $n = 2$ wegen dem Induktionsschritt und weiter für $n = 3$ wieder wegen dem Induktionsschritt. Fährt man so weiter, erhält man (1.2) für jede natürliche Zahl. *Wir sagen, dass Gleichung* (1.2) *mittels vollständiger Induktion für alle natürlichen Zahlen* $n \geq 1$ *folgt. Des Weiteren deuten wir das Ende des Beweises mit einem kleinen Quadrat an.* □

*Beweis von Proposition 1.1.* Wir nehmen an, dass es einen Begriff des Flächeninhalts mit den Eigenschaften in der Proposition gibt und dieser für $P$ definiert ist. Angenommen $I$ ist der Flächeninhalt von $P$. Wir überdecken $P$ für eine gegebene natürliche Zahl $n \geq 1$ mit Rechtecken wie in folgender Figur (Figur 1.2).

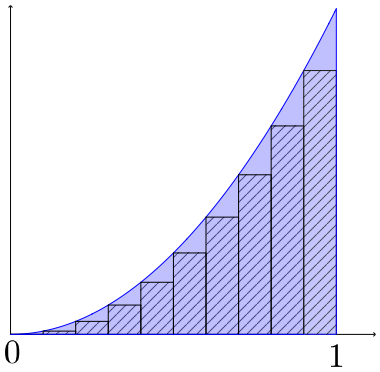


Figur 1.2: Die Überdeckung von $P$ mit Rechtecken für $n = 10$.

Wir erhalten aus den angenommenen Eigenschaften des Flächeninhalts und Lemma 1.3, dass

$$\begin{aligned} I &\leq \frac{1}{n}\frac{1}{n^2} + \frac{1}{n}\frac{2^2}{n^2} + \cdots + \frac{1}{n}\frac{n^2}{n^2} = \frac{1}{n^3}(1^2 + 2^2 + \cdots + n^2) \\ &= \frac{1}{n^3}\left(\frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6}\right) = \frac{1}{3} + \frac{1}{2n} + \frac{1}{6n^2} \leq \frac{1}{3} + \frac{1}{n} \end{aligned}$$

Wir bemerken, dass die Geradenstücke, bei denen sich die Rechtecke berühren, Flächeninhalt 0 haben und wir sie also einfach ignorieren dürfen. (Wieso? – Hier die Details zu ergänzen ist eine gute Übung.). Verwenden wir hingegen Rechtecke wie in Figur 1.3 erhalten wir ebenso




Figur 1.3: Von $P$ überdeckte Kollektion von Rechtecken für $n = 10$.

$$\begin{aligned} I &\geq \frac{1}{n}\frac{0}{n^2} + \frac{1}{n}\frac{1^2}{n^2} + \cdots + \frac{1}{n}\frac{(n-1)^2}{n^2} = \frac{1}{n^3}(1^2 + \cdots + (n-1)^2) \\ &= \frac{1}{n^3}(1^2 + \cdots + (n-1)^2 + n^2 - n^2) = \frac{1}{n^3}\left(\frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6} - n^2\right) \\ &\geq \frac{1}{n^3}\left(\frac{n^3}{3} - n^2\right) = \frac{1}{3} - \frac{1}{n} \end{aligned}$$

Zusammenfassend gilt also

$$-\frac{1}{n} \leq I - \frac{1}{3} \leq \frac{1}{n}$$

für alle natürlichen Zahlen $n \geq 1$. Die einzige Zahl, die kleiner als $\frac{1}{n}$ und grösser als $-\frac{1}{n}$ ist für alle natürlichen Zahlen $n \geq 1$, ist die 0. Dies ist anschaulich relativ klar (siehe unten) und wird später aus dem *Archimedischen Prinzip* folgen, welches wir in zwei Wochen ausführlich besprechen werden. Daher gilt $I = \frac{1}{3}$ und die Proposition folgt. □

Wir haben in obigem Beweis folgenden Satz benötigt:

**Satz** (Eine Version des Archimedischen Prinzips)**.** *Wenn $x \in \mathbb{R}$ die Ungleichung $-\frac{1}{n} \leq x \leq \frac{1}{n}$ für alle natürlichen Zahlen $n$ erfüllt, dann ist $x = 0$.*

Warum ist dies „anschaulich klar“? Stellen Sie sich $x \geq 0$ in der Dezimaldarstellung vor. Verwenden wir die Annahme für $n = 11$, so sehen wir, dass $x$ von der Form $x = 0.0a_2a_3\ldots$ sein muss für vorerst unbekannte Ziffern $a_2, a_3, a_4, \ldots$. Verwenden wir nun $n = 10^2 + 1$, dann sehen wir, dass $x$ nach dem Komma mindestens 2 Nullen haben muss (das heisst, $a_2 = 0$). Da aber die Annahme ebenso für $n = 10^k + 1$ für eine beliebige natürliche Zahl $k$ gilt, sehen wir, dass $x$ unendlich viele Nullen nach dem Komma haben muss. Also ist $x$ Null. Falls $x \leq 0$, so erfüllt $-x$ die Ungleichung $0 \leq -x \leq \frac{1}{n}$ und insbesondere ist $-x$ und damit $x$ gleich Null nach vorherigem Argument. Dies ist wohlgemerkt kein Beweis – wir werden später einen vollständigen Beweis des Archimedischen Prinzips erbringen.

*Bemerkung.* Wie schon erwähnt, haben wir die Frage, ob es einen Flächeninhalt für Bereiche im $\mathbb{R}^2$ gibt, nicht beantwortet. Wir haben auch nicht genau beschrieben, was denn eigentlich Bereiche im $\mathbb{R}^2$ sind; wir sind aber implizit davon ausgegangen, dass Bereiche jene Teilmengen des $\mathbb{R}^2$ sind, denen wir einen Flächeninhalt zuordnen können. Diese grundlegenden Fragen werden zum Teil in Analysis I und II mit den Begriffen des Riemann-Integrals und der Jordanmessbaren Mengen beantwortet. Des Weiteren werden diese Fragen in grösserer Allgemeinheit in der Vorlesung der Mass- und Integrationstheorie im vierten Semester und anderen weiterführenden Vorlesungen im dritten oder vierten Jahr des Mathematikstudiums besprochen.

## 1.2 Einige Tipps

*„Aller Anfang ist schwer.“*

Wie schon erwähnt, werden Sie einen grossen Unterschied zwischen der Schulmathematik und der Mathematik an Universitäten bemerken. Letztere verwendet auch ihre eigene Sprache, die Sie erst erlernen müssen. Je schneller Sie diese Aufgabe auf sich nehmen, desto mehr werden Sie von den Vorlesungen mitnehmen. Wir besprechen in den nächsten Abschnitten die wichtigsten Grundbegriffe, dennoch empfehlen wir Ihnen auch das Buch [SS12] von Schichl und Steinbauer (vor allem Kapitel 1-4), das eine ausführliche Einführung in das mathematische Arbeiten darstellt. Dies bringt uns gleich zum nächsten Tipp.

*„Es gibt keinen Königsweg zur Mathematik“*
(überliefertes Zitat von Euklid zu dem ägyptischen König Ptolemäus I)

Man kann Mathematik nicht durch Zusehen erlernen; genauso wenig wie Sie Tennis oder Skifahren erlernen können, in dem Sie sich alle verfügbaren Turniere oder Weltmeisterschaften im Fernsehen anschauen. Vielmehr sollten Sie Mathematik wie eine Sprache erlernen und eine Sprache bringt man sich bei, in dem man sie benützt. Besprechen Sie mit Kollegen die Themen der Vorlesungen. Erklären Sie sich gegenseitig die Beweise aus der Vorlesung oder die Lösung der Übungsbeispiele. Vor allem aber sollten Sie so viele Übungen wie nur möglich lösen; nur so können Sie sich sicher sein, dass Sie die Themen gemeistert haben.

Es ist in Ordnung, wenn Sie in kleinen Gruppen an den Übungen arbeiten. Dies hat sogar den Vorteil, dass durch die Diskussionen in der Gruppe die Objekte der Vorlesungen lebendiger werden. Sie sollten jedoch sicherstellen, dass Sie die Lösungen komplett verstehen, erklären können, und ähnliche Probleme anschliessend selbstständig lösen können.

*“Wer fragt, ist ein Narr für eine Minute.*
*Wer nicht fragt, ist ein Narr sein Leben lang.”*
(Konfuzius)

Stellen Sie so viele Fragen, wie nur möglich und stellen Sie sie dann, wenn sie auftauchen. Wahrscheinlich haben viele Ihrer Kollegen die gleiche Frage oder haben das Problem noch gar nicht bemerkt. Dem Dozierenden oder dem Hilfsassistenten wird so ermöglicht, gleichzeitig bei vielen ein Problem zu beheben und Probleme bei den Studierenden zu erkennen, wo sie oder er dachte, dass keine vorhanden seien. Des Weiteren will das gute Formulieren von Fragen geübt sein; das erste Studienjahr ist der ideale Zeitpunkt dies zu tun.

Zuletzt möchten wir erwähnen, dass zahlreiche Foren und Blogs sehr hilfreiche Ratschläge liefern, wie man richtig ins Studium einzusteigen hat und wie man Mathematik zu lernen hat. Ein Beispiel eines Blogs mit guten Tipps ist der Blog von Terence Tao. (Terence Tao ist Fields-Medaillenträger und einer der bedeutendsten lebenden Mathematiker.) Eine Frage, die Terry Tao behandelt, ist

*„Does one have to be a genius to do maths?“*

Genauso wie in anderen Gebieten (zum Beispiel im Sport – siehe zum Beispiel das Buch „Was heisst schon Talent?“ von M. Syed) ist die klare Antwort *Nein*. Im Zuge seiner Erklärung verweist Terry Tao auch auf diesen Artikel. Die Lektüre dieses Artikels möchten wir Ihnen sehr empfehlen – zumindest bis zum ersten Erscheinen des Wortes „muscle“.

VO: Analysis 20.09.17

## 1.3 Logische Begriffe

Die logischen Begriffe und die dazugehörende Theorie der Logik haben sich seit der Antike stark geändert, doch seit der Begriffsschrift von Frege (1848-1925) im Jahre 1879 nicht mehr so stark. Da ein Verständnis dieser „logischen Sprache“ für die Mathematik notwendig ist, wollen wir diese hier besprechen. Für historisch Interessierte verweisen wir auch auf den Podcast der BBC, den man sich anhören kann, wenn man schon zu müde ist, um Übungen zu lösen oder man Probleme hat einzuschlafen.

Mathematische Aussagen sind, wenn ein klar definierter Rahmen gegeben ist und der Wert für etwaige Variablen bekannt ist, entweder wahr oder falsch. Diese strenge Sichtweise ist notwendig, damit wir von Bekanntem ausgehend wahre Aussagen (Lemmata und Propositionen zum Beispiel) folgern können, auf die wiederum Verlass ist (das heisst, ohne dass man Spezialfälle ausschliessen muss). Sie haben vielleicht schon einmal Sudoku gespielt und wissen, wie schwerwiegend ein kleiner Denkfehler werden kann.

### 1.3.1 Aussagenlogik und die Boolesche Algebra

Da eine mathematische Aussage nur zwei „Werte“ annehmen kann, wahr (w) und falsch (f), können wir gewissermassen „Rechenoperationen“ für solche Aussagen sehr einfach und schnell definieren. Dies wurde erstmals von George Boole im Jahre 1847 formal ausgeführt (siehe [Boo47]) und stellt sowohl für die Grundlagen der Mathematik als auch für die Informatik und Elektrotechnik einen historisch bedeutsamen Schritt dar. Im Folgenden werden mathematische Aussagen mit Grossbuchstaben A, B, ... dargestellt.

Die einfachste Operation ist die **Negation einer Aussage**: Sei A eine Aussage. Die Negation von A ist die Aussage „A gilt nicht“, die wir kurz mit „Nicht A“ bezeichnen und als „¬A“ schreiben. Die Aussage „¬A“ hat den Wert f, falls A den Wert w hat und den Wert w, falls A den Wert f hat. Wir stellen dies in einer sogenannten **Wahrheitstabelle** wie folgt dar:

| $A$ | $\neg A$ |
|---|---|
| $w$ | $f$ |
| $f$ | $w$ |

Wir verwenden das Symbol „=“ für die Gleichheit und können es gebrauchen, um Aussagen wie „$x = y$“ zu bilden. Des Weiteren schreiben wir „$x \neq y$“ für die Negation „$\neg(x = y)$“ der Gleichheit.

Sei nun B eine weitere Aussage. Die **Und-Operation** angewendet auf A und B ist die Aussage „sowohl A als auch B gelten“, kurz „A und B“ und symbolisch geschrieben „A∧B“.

Die zugehörige Wahrheitstabelle ist

| $A$ | $B$ | $A \wedge B$ |
|---|---|---|
| $w$ | $w$ | $w$ |
| $w$ | $f$ | $f$ |
| $f$ | $w$ | $f$ |
| $f$ | $f$ | $f$ |

Die **Oder-Operation** angewendet auf die beiden Aussagen A,B ergibt die Aussage „A gilt oder B gilt“, kurz „A oder B“ und symbolisch geschrieben „A∨B“. Die zugehörige Wahrheitstabelle ist:

| $A$ | $B$ | $A \vee B$ |
|---|---|---|
| $w$ | $w$ | $w$ |
| $w$ | $f$ | $w$ |
| $f$ | $w$ | $w$ |
| $f$ | $f$ | $f$ |

Man sollte an dieser Stelle bemerken, dass, wie oben ersichtlich ist, das logische „Oder“ (und daher auch das mathematische „Oder“) das einschliessende „Oder“ ist, welches auch zutrifft, falls beide Aussagen zutreffen. Dies unterscheidet sich teilweise vom umgangssprachlichen Gebrauch des Wortes „oder“, das oft auch exklusiv gemeint ist (als „entweder ... oder“). Die Aussage „entweder A oder B“ kann über die drei Grundoperationen Negation, Und und Oder definiert werden (siehe Übung 1.4).

Die Algebraisierung der Logik hat den entscheidenden Vorteil, dass wir mittels Klammern klar darstellen können, wie Verknüpfungen von logischen Aussagen zu verstehen sind. In der Umgangssprache gibt es keine Klammern und Aussagen der Form

„*Sokrates hat einen Hund und*
*ein Hund ist das beste Haustier oder*
*eine Katze ist das beste Haustier.*“

sind zweideutig. (Inkludiert obiges die Behauptung „Sokrates hat einen Hund“ oder kann es auch sein, dass die Aussage „Sokrates hat keinen Hund und eine Katze ist das beste Haustier“ gilt?)

Wir können Verknüpfungen (wenn nötig mit Klammern) verwenden, um neue logische Operationen zu definieren. Beispielsweise ist die logische **Implikation** „A $\Longrightarrow$ B“ als die Aussage „(¬A)∨B“ definiert. „A $\Longrightarrow$ B“ wird als „A impliziert B“ ausgesprochen und auch in diesem Sinne verwendet. Denn falls A wahr ist und „$A \Longrightarrow B$“ wahr ist, dann muss auch B wahr sein. Auffallend sind folgende Eigenheiten:

- Jegliche Kausalität oder auch der konkrete Zusammenhang zwischen den beiden Aussagen wird ignoriert.

- Die Implikation ist immer richtig, falls die Annahme A falsch ist. Also ist zum Beispiel die Aussage

$$„(0 = 1) \implies \text{(die Welt ist eckig)}“ \tag{1.3}$$

dadurch wahr.

Die **logische Äquivalenz** der Aussagen A und B wird als „A $\iff$ B“ geschrieben, als „A genau dann wenn B“ oder als „A ist äquivalent zu B“ ausgesprochen und durch die Aussage „(A $\implies$ B)$\wedge$(B $\implies$ A)“ definiert.

Es ist hilfreich, sich die logischen Operationen auch als elektrische Schaltungen vorzustellen (siehe zum Beispiel Sektion 3.1 in [SS12]). Mit dieser Analogie im Hinterkopf sind die Eigenheiten der obigen Diskussion bzgl. Kausalität oder das Beispiel (1.3) leichter verständlich, denn einer Schaltung ist egal, wer (Frau, Mann, Kind, Hund oder Maschine) den Schalter betätigt oder warum (absichtlich oder unabsichtlich) dieser betätigt wurde. Auch unsere Beweise werden manchmal keine echte Kausalität herstellen, sondern nur aufzeigen, dass es keine andere Möglichkeit gibt.

Sei A eine Aussage. Die Aussage „A$\vee$($\neg$A)“ ist eine sogenannte **Tautologie**, denn sie trifft unabhängig vom Wahrheitswert von A stets zu. In der Tat, wenn A wahr ist, dann ist „A$\vee$($\neg$A)“ wahr und wenn A falsch ist, dann ist „$\neg$A“ wahr und „A$\vee$($\neg$A)“ auch wahr.

**Übung 1.4** (Logische Implikationen und Tautologien)**.** *Seien A,B zwei Aussagen.*

*(i) Bestimmen Sie die Wahrheitstabellen zu „$A \implies B$“ und zu „$A \iff B$“.*

*(ii) Definieren Sie die logische Operation des* ***auschliessenden Oder*** *„A XOR B“ mittels der obigen Operationen. Die Aussage „A XOR B“ soll zutreffen, wenn A oder B zutrifft, aber nicht beide zutreffen.*

*(iii) Überprüfen Sie, dass die Aussagen „$A \iff (\neg(\neg A))$“ und „$(A \wedge B) \iff (\neg(\neg A \vee \neg B))$“ Tautologien sind.*

Wir werden anstelle der Symbole „$\neg, \wedge, \vee, \implies, \iff$“ meist die Worte „nicht, und, oder, impliziert, ist äquivalent zu“ verwenden, wobei „oder“ immer einschliessend zu verstehen ist. Da wir keine Prioritätsregel (wie die bekannte Punkt- vor Strichrechnung) festgelegt haben, sollten wir stets Klammern verwenden um Zweideutigkeiten zu vermeiden. Eine übliche Prioritätsregel ist „$\wedge$“ vor „$\vee$“, womit die Aussage „$A \wedge B \vee C$“ als „$(A \wedge B) \vee C$“ zu verstehen wäre. Nichtsdestotrotz werden wir letztere Schreibweise bevorzugen und nur die Prioritätsregel verwenden, dass die Operation $\neg$ immer Vorrang hat. Also kann zum Beispiel „($\neg$A)$\vee$B“ auch als „$\neg$A$\vee$B“ geschrieben werden.

Seien A, B zwei Aussagen. Eine sehr nützliche Tautologie ist die Aussage

$$„(A \implies B) \iff (\neg B \implies \neg A)“ \tag{1.4}$$

Um zu überprüfen, dass es sich bei (1.4) tatsächlich um eine Tautologie handelt, verwenden wir zuerst die Definition der logischen Implikation und sehen, dass die linke Seite „$\neg$A$\vee$B“

ist. Die rechte Seite wiederum ist „$\neg(\neg B)\vee\neg A$“, was zu „$B\vee\neg A$“ und damit auch zu „$\neg A\vee B$“ äquivalent ist.

Die Tautologie (1.4) wird oft in Beweisen verwendet: Angenommen wir wollen zeigen, dass die Aussage A die Aussage B impliziert. Dann ist es manchmal einfacher und wegen (1.4) stets ausreichend zu zeigen, dass falls B nicht zutrifft auch A nicht zutrifft (dies wird Kontraposition genannt, siehe auch Abschnitt 1.6.2).

An dieser Stelle sollte man davor warnen, die Aussagen „$A \Longrightarrow B$“ und „$\neg A \Longrightarrow \neg B$“ zu verwechseln; sie sind grundverschieden (auch wenn der Unterschied in politischen Diskussionen manchmal gern ignoriert wird). Wir empfehlen Ihnen, Beispiele mathematischer und nicht-mathematischer Art zu finden, die den Unterschied zwischen „$A \Longrightarrow B$“ und „$\neg A \Longrightarrow \neg B$“ gut belegen.

**Übung 1.5** (Ritter und Knappen - aus [Smu78])**.** *Auf der Insel der Ritter und Knappen ist jeder Einwohner entweder ein Ritter oder ein Knappe (und jeder weiss den Status von allen anderen Einwohnern). Es ist wichtig zu wissen, dass*

- *Ritter immer die Wahrheit sagen.*
- *Knappen immer lügen.*

*Sie werden auf Einwohner der Insel treffen und Ihre Aufgabe ist bei jedem zu entscheiden, ob er ein Ritter oder ein Knappe ist.*

*(i) Sie treffen Johannes und Willhelm auf der Insel. Johannes sagt „Willhelm und ich sind Ritter.“ Willhelm sagt „Das ist eine Lüge, Johannes ist ein Knappe!“ Was sind sie?*

*(ii) Sie treffen Gildas, Ergard und Telones auf der Insel. Gildas sagt „Seien Sie vorsichtig, wir sind nicht alle drei Ritter.“ Ergard sagt: „Wir sind auch nicht alle drei Knappen.“ Telones sagt „Hören Sie nicht auf sie, ich bin der einzige Ritter.“ Was sind sie?*

*(iii) Sie treffen Heinrich und Arthur auf der Insel. Heinrich murmelt irgend etwas Unverständliches. Arthur sagt „Er sagte, er sei ein Knappe. Das ist er sicher - vertrauen Sie ihm nicht!“ Was sind sie?*

VO: Lineare Algebra 20.09.17

### 1.3.2 Prädikatenlogik

Die logischen Begriffe, die wir oben vorgestellt haben, sind zwar grundlegend für alles weitere, aber nicht genügend komplex um interessante Aussagen zu bilden. Oft werden Aussagen formuliert, die für Elemente einer bestimmten Menge (an Zahlen, Vektoren, Äpfeln, ...; mehr zu Mengen später) wahr oder falsch sein können. Zum Beispiel könnte $x$ für eine natürliche (oder rationale, reelle, ...; mehr zu Zahlen später) Zahl stehen und „$x = x^2$“ ist dann eine Aussage über $x$, die wahr oder falsch sein kann. Für derartige Aussagen „$A(x)$“ über Elemente $x$ einer Menge $X$ gibt es nun zwei weitere fundamentale Operationen, sogenannte Quantoren, um weitere Aussagen zu bilden.

Der **Allquantor**, geschrieben $\forall$, wird verwendet um eine Aussage über alle Elemente von $X$ zu treffen. Genauer steht die Aussage „$\forall x \in X : A(x)$“ für die Aussage „Für alle $x$ in $X$ gilt die Aussage $A(x)$“. Diese Aussage kann wahr oder falsch sein. Zum Beispiel ist die Aussage

$$„\forall n \in \mathbb{N} : n = n^2“$$

falsch, aber die Aussage

$$„\forall n \in \mathbb{N} : (n = n^2 \implies n = 1)“$$

ist richtig (denn wir werden $\mathbb{N}$ als $\{1, 2, 3, \ldots\}$ definieren).

Der **Existenzquantor**, geschrieben $\exists$, wird verwendet um auszudrücken, dass es ein $x \in X$ mit einer gewissen Eigenschaft gibt. Genauer steht „$\exists x \in X : A(x)$“ für „Es gibt ein $x \in X$, für das die Aussage $A(x)$ gilt“. Zum Beispiel ist die Aussage

$$„\exists n \in \mathbb{N} : n = n^2“$$

richtig, da $n = 1$ eine natürliche Zahl ist, die $1^2 = 1$ erfüllt. Weiters ist auch die Aussage

$$„\exists n \in \mathbb{N} : (n = n^2 \implies n = 1)“$$

richtig, was man ähnlich wie zuvor sieht oder alternativ wie folgt: $n = 5$ erfüllt dass $5 \neq 25 = 5^2$, womit die Implikation „$5 = 5^2 \implies 5 = 1$“ wahr ist.

*Bemerkung* (Alternative Notation). Anstelle der Symbole $\forall$ respektive $\exists$ werden teilweise auch die weniger gebräuchlichen Symbole $\bigwedge$ respektive $\bigvee$ als Synonyme verwendet. Letztere haben den Vorteil, dass sie andeuten, dass $\bigwedge$ ein erwachsen gewordenes $\wedge$ und $\bigvee$ ein erwachsen gewordenes $\vee$ darstellt. Um diesen Kommentar besser zu verstehen, nehmen wir an, dass $X = \{x_1, x_2, x_3, \ldots, x_n\}$ eine endliche Menge ist, die genau die Elemente $x_1, \ldots, x_n$ enthält. Dann gilt

$$„\left(\bigwedge x \in X : A(x)\right) \iff (A(x_1) \wedge A(x_2) \wedge \ldots \wedge A(x_n))“$$

und

$$„\left(\bigvee x \in X : A(x)\right) \iff (A(x_1) \vee A(x_2) \vee \ldots \vee A(x_n))“$$

Später werden wir sehen, dass die Symbole $\cap, \bigcap$ für den Durchschnitt von Mengen den Symbolen $\wedge, \forall$ und die Symbole $\cup, \bigcup$ für die Vereinigung von Mengen den Symbolen $\vee, \exists$ entsprechen.

Kombinieren wir $\forall$ und $\exists$ so treten sehr schnell für den allgemeinen Sprachgebrauch vielleicht subtile aber für die Logik und die Mathematik fundamentale Eigenheiten der Quantoren zu Tage: Seien $X, Y$ Mengen und für jedes $x \in X$ und $y \in Y$ sei $A(x, y)$ eine Aussage. Dann haben die Aussagen

$$„\forall x \in X \ \exists y \in Y : A(x, y)“$$

und

$$„\exists y \in Y\ \forall x \in X : A(x,y)“$$

sehr verschiedene Bedeutungen.

**Beispiel 1.6** (Nicht-vertauschbare Quantoren)**.** *Angenommen $X$ steht für die Menge der Studienanfänger an der ETH, $Y$ für die Menge der Vorlesungen für Studienanfänger. Zu $x \in X$ und $y \in Y$ sei $A(x,y)$ die Aussage „Student $x$ interessiert sich für die Vorlesung $y$“. Dann ist „ $\forall x \in X\ \exists y \in Y : A(x,y)$“ hoffentlich wahr (oder es gibt Studenten, die sich für das völlig falsche Studium entschieden haben). Die Aussage „ $\exists y \in Y\ \forall x \in X : A(x,y)$“ ist jedoch falsch, denn es gibt sicher zwei Studienanfänger, die keine gemeinsame Vorlesung besuchen und auch komplett verschiedene Interessen haben.*

Ein mathematisches Beispiel für das gleiche Phänomen:

**Beispiel 1.7** (Nicht-vertauschbare Quantoren)**.** *Sei $Y = X$ eine beliebige Menge. Die Aussage „ $\forall x \in X\ \exists y \in Y : x = y$“ ist dann sicher richtig, da wir für jedes $x \in X$ einfach $y = x$ wählen können. Die Aussage „ $\exists y \in X\ \forall x \in X : x = y$“ hingegen ist nur für sehr spezielle Mengen richtig, welche?*

Diese Beispiele sind möglicherweise zu einfach; wir werden jedoch der Problematik, dass man einen Existenz- mit einem Allquantor nicht vertauschen darf, in komplizierten Situationen wieder begegnen. Für zwei Quantoren der gleichen Sorte ist die Situation einfacher:

**Wichtige Übung 1.8** (Vertauschbarkeit von Quantoren)**.** *Seien $X, Y$ zwei Mengen und für jedes $x \in X$ und $y \in Y$ sei „$A(x,y)$“ eine Aussage. Überzeugen Sie sich oder noch besser eine Mitstudentin/einen Mitstudenten davon, dass*

$$„(\forall x \in X\ \forall y \in Y : A(x,y)) \iff (\forall y \in Y\ \forall x \in X : A(x,y))“$$
$$„(\exists x \in X\ \exists y \in Y : A(x,y)) \iff (\exists y \in Y\ \exists x \in X : A(x,y))“$$

*wahre Aussagen sind.*

Auf Grund der Aussage in Übung 1.8 werden wir, gegeben eine Aussage „$A(x,y)$“ über zwei Elemente einer Menge $X$, anstelle von der Aussage „$\forall x \in X\ \forall y \in X : A(x,y)$“ auch „$\forall x, y \in X : A(x,y)$“ und anstelle von „$\exists x \in X\ \exists y \in X : A(x,y)$“ auch „$\exists x, y \in X : A(x,y)$“ schreiben.

Wir bemerken noch eine sonderbare **Eigenheit des Allquantors**. Sei $X$ die Menge, die keine Elemente enthält (die sogenannte „leere Menge“ - siehe nächsten Abschnitt). Dann ist „$\forall x \in X : A(x)$“ immer wahr, unabhängig davon, welche Aussage $A(x)$ ist.

Man kann die Quantoren und die logischen Operationen auf sehr viele Arten kombinieren und sehr schnell sehr komplizierte Aussagen bilden. Später werden wir viele Beispiele sehen. Als nächstes möchten wir hier aber noch die Negation von Quantoren besprechen. Intuitiv ist die Negation des Allquantors ein Existenzquantor und die Negation des Existenzquantors

ein Allquantor. Besser und genauer ausgedrückt gelten für eine beliebige Aussage „$A(x)$“ über Elemente einer Menge $X$ die folgenden Aussagen:

$$\begin{aligned} &„\neg(\forall x \in X : A(x)) \iff \exists x \in X : \neg A(x)“ \\ &„\neg(\exists x \in X : A(x)) \iff \forall x \in X : \neg A(x)“ \end{aligned} \tag{1.5}$$

Beispielsweise ist die Negation von „Auf jedem Planeten herrscht Gravitation“ die Aussage „Es gibt einen Planeten, auf dem keine Gravitation herrscht“. Bei mehreren Quantoren verhält sich die Negation ähnlich wie ein Minus-Symbol vor einem geklammerten Ausdruck und kann „nach Innen“ geschoben werden. Wir empfehlen dies in folgender Übung zu überdenken.

**Wichtige Übung 1.9** (Negation und verknüpfte Quantoren)**.** *Sei $X$ eine Teilmenge der reellen Zahlen $\mathbb{R}$. Bilden Sie die Negation folgender Aussagen und versuchen Sie die Negation so weit wie möglich nach „rechts“ zu verschieben (obwohl die Aussagen geometrische Bedeutung haben, muss man diese weder kennen noch verstehen, um die Aufgabe zu lösen):*

- „$\forall y \in \mathbb{R}\ \forall \varepsilon > 0\ \exists x \in X : |x - y| < \varepsilon$“
- „$\forall y \in X\ \exists \varepsilon > 0\ \forall x \in X : |x - y| < \varepsilon \implies x = y$“
- „$\forall \varepsilon > 0\ \forall x \in X\ \exists y \in X : (y \neq x) \wedge |x - y| < \varepsilon$“

*Hier ist „$\forall \varepsilon > 0 : A(\varepsilon)$“ eine gebräuchliche Kurzform für „$\forall \varepsilon \in \mathbb{R} : \varepsilon > 0 \implies A(\varepsilon)$“ oder anders ausgedrückt für „$\forall \varepsilon \in \{x \in \mathbb{R} \mid x > 0\} : A(\varepsilon)$“. Die Aussage „$\exists \varepsilon > 0 : A(\varepsilon)$“ steht für „$\exists \varepsilon \in \mathbb{R} : \varepsilon > 0 \wedge A(\varepsilon)$“ oder äquivalenterweise für „$\exists \varepsilon \in \{x \in \mathbb{R} \mid x > 0\} : A(\varepsilon)$“. Sie können für die Negation von „$|x - y| < \varepsilon$“ auch einfach „$\neg(|x - y| < \varepsilon)$“ schreiben, doch ist dies auf Grund üblicher Definitionen und Eigenschaften von $\mathbb{R}$ äquivalent zu „$|x - y| \geq \varepsilon$“.*

Ein dritter Quantor, der häufig verwendet wird, ist der **Quantor der eindeutigen Existenz**, geschrieben $\exists!$ . Dieser lässt sich mit Hilfe der beiden Quantoren $\forall$ und $\exists$ wie folgt definieren: Die Aussage „$\exists! x \in X : A(x)$“ wird durch

$$„(\exists x \in X : A(x)) \wedge (\forall x, y \in X : (A(x) \wedge A(y) \implies x = y))“$$

definiert. Sie bedeutet, dass es ein $x$ in $X$ gibt, das die Aussage $A(x)$ erfüllt und dass dieses $x$ durch die Aussage eindeutig bestimmt ist. Es gibt also kein weiteres Element $y \in X$, welches nicht gleich $x$ ist und $A(y)$ erfüllt.

**Übung 1.10.** *Formulieren Sie ein konkretes Beispiel zu obigem. In anderen Worten: finden Sie eine Aussage über Elemente einer Menge, die von genau einem Element erfüllt wird.*

**Eine Warnung:** Sei $X$ eine Menge und $A(x)$ eine Aussage über Elemente von $X$. Nach dem Satz „Es gibt ein eindeutig bestimmtes $x_0$ in $X$, das $A(x_0)$ erfüllt.“ wird oft dieses $x_0$ in weiteren Argumenten verwendet. Da $x_0$ eindeutig festgelegt ist, ist dies akzeptabel. Manchmal wird aber auch nach dem Satz „Es gibt ein $x$ in $X$, das $A(x)$ erfüllt.“ so ein $x$ in der Argumentation benötigt. Hier sollte man erwähnen, dass so ein $x$ gewählt wird. Des

Weiteren sollte man, wenn nötig, sicher stellen, dass das darauf Folgende nicht (oder nur auf unwesentliche Weise) von der Wahl von $x$ abhängt.

**Übung 1.11.** *Zum Spass: Betrachten Sie den Satz „Everybody loves my babe, but my babe loves no one but me“ (leicht adaptiert nach Louis Armstrong) vom streng logischen Standpunkt. Handelt er von einem Liebespaar?*

VO: Analysis 21.09.17

## 1.4 Mengenlehre und Abbildungen

### 1.4.1 Naive Mengenlehre

Georg Cantor definierte 1895 in [Can95] den Begriff einer Menge wie folgt:

> „*Unter einer Menge verstehen wir jede Zusammenfassung $M$ von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denken (welche Elemente von $M$ genannt werden) zu einem Ganzen.*“

Also sind die zentralen Annahmen der naiven Mengenlehre (oder Cantorschen Mengenlehre)

(1) Eine **Menge** besteht aus beliebigen unterscheidbaren Elementen.

(2) Eine Menge ist unverwechselbar durch ihre **Elemente** bestimmt.

(3) Jede Eigenschaft $A(x)$ definiert die **Menge der Objekte** $\{x \mid A(x)\}$, die diese Eigenschaft besitzen (das heisst, die Aussage $A(x)$ erfüllen).

Manchmal schreiben wir eine Menge durch eine konkrete Auflistung ihrer Elemente ($M = \{x_1, x_2, \ldots, x_n\}$ zum Beispiel). Die **leere Menge**, geschrieben als $\{\}$ oder auch $\emptyset$, ist die Menge, die keine Elemente enthält. Mengen werden teilweise auch **Familien** (oder Kollektionen, Ansammlungen) genannt. Wir schreiben „$x \in X$“, falls $x$ ein Element der Menge $X$ ist. Je nach Zusammenhang nennen wir die Elemente mitunter auch Punkte, Zahlen oder Vektoren. Falls $x$ kein Element der Menge $X$ ist (also $\neg(x \in X)$), so schreiben wir auch $x \notin X$.

**Definition 1.12** (Inklusionen)**.** Wir sagen, dass eine Menge $P$ **Teilmenge** einer Menge $Q$ ist und schreiben $P \subseteq Q$, falls für alle $x \in P$ auch $x \in Q$ gilt (in Prädikatenlogik $\forall x \in P : x \in Q$). Wir sagen, dass $P$ eine **echte Teilmenge** von $Q$ ist und schreiben $P \subsetneq Q$, falls $P$ eine Teilmenge von $Q$ ist, aber nicht gleich $Q$ ist. Wir schreiben $P \not\subseteq Q$, falls $P$ keine Teilmenge von $Q$ ist (also $\neg(P \subseteq Q)$ gilt).

Äquivalente Formulierungen für „$P$ ist eine Teilmenge von $Q$“ sind „$P$ ist in $Q$ **enthalten**“ und „$Q$ ist eine **Obermenge** von $P$“, was wir auch als „$Q \supseteq P$“ schreiben. Die Bedeutung der Aussage „$Q$ ist eine **echte Obermenge** von $P$“, geschrieben $Q \supsetneq P$, ergibt sich nun implizit.

**Übung 1.13.** *Seien $P, Q$ Mengen. Formuliere die Aussagen „$P$ ist eine echte Teilmenge von $Q$“ und „$P$ ist keine Teilmenge von $Q$“ in Prädikatenlogik.*

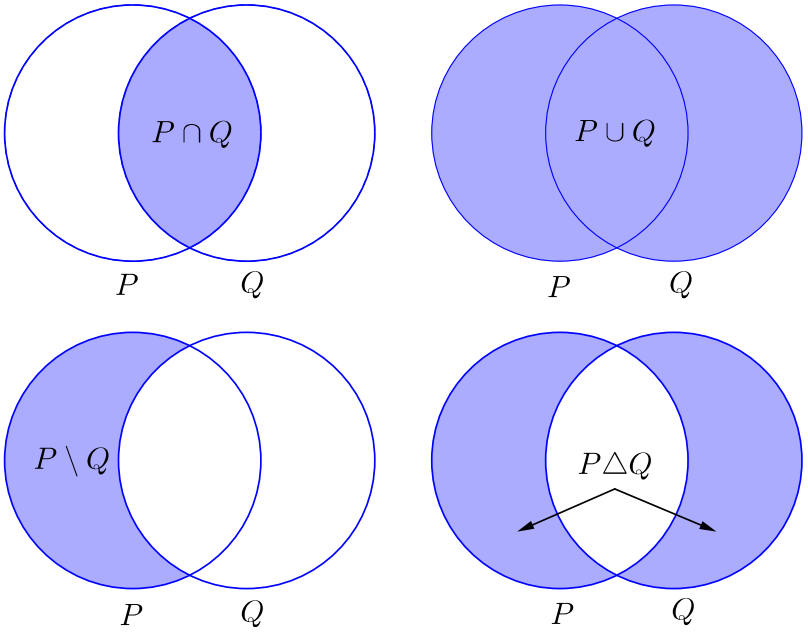
Wegen der zweiten Annahme der naiven Mengenlehre sind zwei Mengen $P, Q$ genau dann gleich, wenn $P \subseteq Q$ und $Q \subseteq P$ gelten. Insbesondere könnte $\{x, y\} = \{z\}$ gelten, falls $x = y = z$ ist (es gibt für Elemente einer Menge keine Vielfachheiten). Im Folgenden führen wir geläufige Konstruktionen von Mengen aus gegebenen Mengen ein.

**Definition 1.14** (Mengenoperationen)**.** Seien $P, Q$ zwei Mengen. Der **Durchschnitt** $P \cap Q$, die **Vereinigung** $P \cup Q$, das **relative Komplement** $P \setminus Q$ und die **symmetrische Differenz**

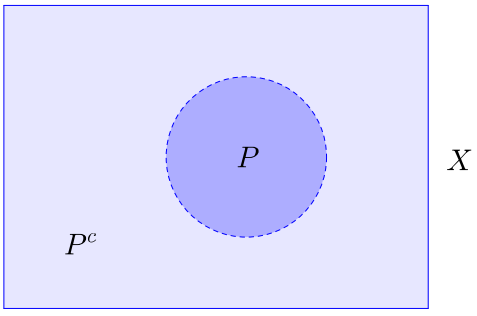
$P \triangle Q$ sind durch

$$\begin{aligned} P \cap Q &= \{x \mid x \in P \wedge x \in Q\} \\ P \cup Q &= \{x \mid x \in P \vee x \in Q\} \\ P \setminus Q &= \{x \mid x \in P \wedge x \notin Q\} \\ P \triangle Q &= P \setminus Q \cup Q \setminus P = \{x \mid x \in P \text{ XOR } x \in Q\} \end{aligned}$$

definiert. Wenn aus dem Zusammenhang klar ist, dass alle betrachteten Mengen Teilmengen einer gegebenen (Ober-) Menge $X$ sind, dann ist das **Komplement** $P^c$ von $P$ (in $X$) definiert durch $P^c = X \setminus P$. Dies ist alles in den Bildern 1.4 und 1.5 veranschaulicht.




Figur 1.4: Schnitt, Vereinigung, Komplement beziehungsweise symmetrische Differenz von $P, Q$. Skizzen dieser Art nennen sich auch Venndiagramme.




Figur 1.5: Das Komplement $P^c = X \setminus P$ von $P$ in $X$.

**Übung 1.15** (Distributivgesetze)**.** *Zeigen Sie die folgenden Distributivgesetze:*

*(i) Seien $A, B, C$ Aussagen. Zeigen Sie*

$$(A \vee B) \wedge C \iff (A \wedge C) \vee (B \wedge C)$$
$$(A \wedge B) \vee C \iff (A \vee C) \wedge (B \vee C)$$

*(ii) Seien $P, Q, R$ Mengen. Zeigen Sie*

$$(P \cup Q) \cap R = (P \cap R) \cup (Q \cap R)$$
$$(P \cap Q) \cup R = (P \cup R) \cap (Q \cup R)$$

*Können Sie auch Kommutativ- und Assoziativgesetze formulieren?*

**Übung 1.16** (Gesetze von De Morgan)**.** *Seien $P, Q$ Teilmengen einer Menge $X$. Zeigen Sie den folgenden Spezialfall der De Morganschen Gesetze:*

$$(P \cup Q)^c = P^c \cap Q^c,$$
$$(P \cap Q)^c = P^c \cup Q^c.$$

*Erläutern Sie Ihr Vorgehen an einem Bild im Stile von Figuren 1.4 und 1.5.*

Wir verallgemeinern nun zwei Begriffe aus Definition 1.14:

**Definition 1.17** (Beliebige Vereinigungen und Schnitte)**.** Sei $\mathcal{A}$ eine Kollektion von Mengen. Dann definieren wir die **Vereinigung** resp. den **Schnitt** der Mengen in $\mathcal{A}$ als

$$\bigcup_{A \in \mathcal{A}} A = \{x \mid \exists A \in \mathcal{A} : x \in A\}$$
$$\bigcap_{A \in \mathcal{A}} A = \{x \mid \forall A \in \mathcal{A} : x \in A\}$$

Vergewissern Sie sich an dieser Stelle, dass diese beiden Begriffe zu Definition 1.14 kompatibel sind. Falls wir die Vereinigung nicht über alle Mengen in $\mathcal{A}$ nehmen wollen, sondern nur über solche, die eine gewisse Eigenschaft $E(A)$ erfüllen, dann schreiben wir auch

$$\bigcup_{A \in \mathcal{A} : E(A)} A = \bigcup_{A \in \{B \in \mathcal{A} \mid E(B)\}} A$$

und analog für den Durchschnitt. Falls $\mathcal{A} = \{A_1, A_2, \ldots\}$, dann schreiben wir auch

$$\bigcup_{n \geq 1} A_n = \bigcup_{n=1}^{\infty} A_n = \bigcup_{A \in \mathcal{A}} A = \{x \mid \exists n \in \mathbb{N} : x \in A_n\}$$

und

$$\bigcap_{n \geq 1} A_n = \bigcap_{n=1}^{\infty} A_n = \bigcap_{A \in \mathcal{A}} A = \{x \mid \forall n \in \mathbb{N} : x \in A_n\}.$$

**Wichtige Übung 1.18** (De Morgansche Gesetze)**.** *Sei $\mathcal{A}$ eine Kollektion von Mengen. Zeigen Sie die allgemeine Form der De Morganschen Gesetze:*

$$\left(\bigcup_{A\in\mathcal{A}} A\right)^c = \bigcap_{A\in\mathcal{A}} A^c$$
$$\left(\bigcap_{A\in\mathcal{A}} A\right)^c = \bigcup_{A\in\mathcal{A}} A^c.$$

**Übung 1.19.** *Was ist die Menge*

$$\bigcap_{n=1}^{\infty}\left\{x\in\mathbb{R}\;\middle|\; -\frac{1}{n}\le x\le\frac{1}{n}\right\},$$

*wobei $n$ wie in Definition 1.17 über alle natürlichen Zahlen läuft, und wie nennen wir dies?*

Für unsere Zwecke wird folgende Definition ebenfalls nützlich sein:

**Definition 1.20** (Disjunktheit)**.** Zwei Mengen $A, B$ heissen **disjunkt**, falls $A\cap B=\{\}$. In diesem Fall wird ihre Vereinigung $A\cup B$ **disjunkte Vereinigung** genannt und auch als $A\sqcup B$ geschrieben. Für eine Kollektion $\mathcal{A}$ von Mengen, sagen wir, dass die Mengen in $\mathcal{A}$ **paarweise disjunkt** sind, falls für alle $A_1, A_2\in\mathcal{A}$ mit $A_1\neq A_2$ gilt $A_1\cap A_2=\{\}$. Die Vereinigung der Mengen in $\mathcal{A}$ wird dann auch **disjunkte Vereinigung** genannt und als $\bigsqcup_{A\in\mathcal{A}} A$ geschrieben.

Wir wenden uns nun einer weiteren, intuitiv vielleicht bekannten, Mengenkonstruktion zu:

**Definition 1.21** (Kartesisches Produkt)**.** Für zwei Mengen $X$ und $Y$ ist das **kartesische Produkt** $X\times Y$ die Menge aller geordneten Paare $(x,y)$ wobei $x\in X$ und $y\in Y$. In Symbolen,

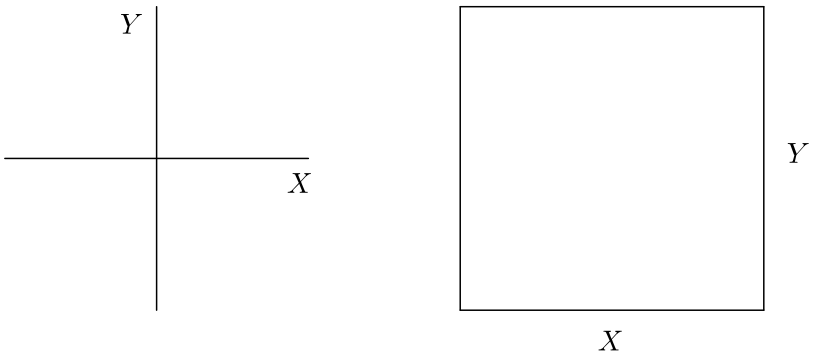
$$X\times Y=\{(x,y)\mid x\in X \text{ und } y\in Y\}.$$

Allgemeiner ist das kartesische Produkt von $n$-Mengen $X_1, X_2, \ldots, X_n$ definiert als

$$X_1\times X_2\times\cdots\times X_n=\{(x_1,x_2,\ldots,x_n)\mid x_1\in X_1, x_2\in X_2,\ldots,x_n\in X_n\}.$$

Weiters definieren wir $X^n$, für eine natürliche Zahl $n$, als das $n$-fache kartesische Produkt von $X$ mit sich selbst. Das heisst, $X^2=X\times X$, $X^3=X\times X\times X$ und so weiter.

Graphisch (meist schematisch) wird das kartesische Produkt $X\times Y$ zweier Mengen $X$, $Y$ wie in folgendem Bild dargestellt.




Figur 1.6: Zwei Darstellungen von $X \times Y$.

Wir wollen kurz ein paar Beispiele dieser Konstruktion erwähnen.

**Beispiel 1.22** (Einige kartesische Produkte)**.** *Falls $X = \{0, 1\}$ ist, dann ist*

$$X^2 = \{0, 1\}^2 = \{(0, 1), (1, 0), (0, 0), (1, 1)\}.$$

*Falls $X = \{a, b, c, \ldots, z\}$ die Menge der Buchstaben im Alphabet ist, so kann man $X^n$ mit der Menge aller möglichen (potenziell sinnfreien) Wörter der Länge $n$ identifizieren. Die Menge der deutschen Wörter der Länge $n$ bildet eine echte Teilmenge von $X^n$ unter dieser Identifikation.*

**Übung 1.23** (Digitale Uhr)**.** *Die Menge der von einer digitalen Uhr angezeigten Uhrzeiten lässt sich als kartesisches Produkt zweier Mengen auffassen. Wie?*

**Übung 1.24** (Schnitte von Rechtecken)**.** *Seien $X, Y$ Mengen und $A, A'$ Teilmengen von $X$, $B, B'$ Teilmengen von $Y$. Zeigen Sie die Formel*

$$(A \times B) \cap (A' \times B') = (A \cap A') \times (B \cap B').$$

*Überzeugen Sie sich auch davon, dass es keine ähnliche Formel für die Vereinigung gibt.*

Unsere vorerst letzte Konstruktion von Mengen ist in der nächsten Definition gegeben:

**Definition 1.25** (Potenzmenge)**.** Für eine Menge $X$ ist die **Potenzmenge** $\mathcal{P}(X)$ durch die Menge all ihrer Teilmengen gegeben, das heisst

$$\mathcal{P}(X) = \{Q \mid Q \text{ ist eine Menge und } Q \subseteq X\}.$$

**Übung 1.26.** *Bestimmen Sie $\mathcal{P}(\{\})$ und $\mathcal{P}(\mathcal{P}(\{\}))$.*

Der Grund, warum obiges die naive Mengenlehre genannt wird, ist, dass sie Anfang des zwanzigsten Jahrhunderts durch die **axiomatische Mengenlehre** (ZF- oder ZFC-Mengenlehre abhängig von einem weiteren Axiom - ZF steht für Zermelo-Fraenkel) ersetzt wurde. Denn die naive Mengenlehre kann sehr schnell zu Widersprüchen geführt werden:

**Beispiel 1.27** (Russell 1903, [Rus03]). *Sei $R$ die Menge aller Mengen, die sich nicht selbst enthalten, also*

$$R = \{X \mid X \text{ ist eine Menge und } X \notin X\}.$$

*Stellt man nun die Frage, ob $R$ selbst zu $R$ gehören kann, so erhält man „$R \in R \iff R \notin R$“. Dies kann aber nicht gelten und ist also ein Widerspruch der naiven Mengenlehre.*

Dieses Beispiel wird **Russell-Paradox** genannt (siehe auch diesen Link) und ist eine Version des sogenannten Lügner-Paradoxons („Dieser Satz ist falsch.“ oder „Ich lüge immer.“). Es entstand, da wir ohne vernünftige Einschränkungen das Bilden beliebiger neuer Mengen erlaubten. In der axiomatischen Mengenlehre werden nur gewisse Konstruktionen (wie zum Beispiel das kartesische Produkt oder die Potenzmenge) erlaubt, was dazu führt, dass die Menge $R$ aus dem Russell-Paradox keine Menge mehr ist. Dies löst den Widerspruch auf: $R$ kann sich somit nicht selbst enthalten, da $R$ ja nur Mengen enthält. $R$ wird in der axiomatischen Mengenlehre eine **Klasse** genannt. Weiters dürfen Klassen nicht als Elemente von Mengen oder Klassen auftreten.

Im folgenden bedienen wir uns der naiven Mengenlehre, da wir die Zeit nicht aufbringen können, die axiomatische Mengenlehre ausreichend zu besprechen und da wir auf jeden Fall nur “vernünftige” Mengenkonstruktionen verwenden werden. (Dies macht vom streng logischen Standpunkt wenig Sinn, ist aber ein notwendiger Kompromiss.)

**Übung 1.28** (Der Barber aus Sevilla). *Bart, der Barber in dem Dorf Sevilla, rasiert alle Männer von Sevilla, die sich nicht selbst rasieren. Sonst rasiert Bart aber niemanden. Rasiert Bart sich selbst?*

VO: Analysis, Lineare Algebra 24.09.17

### 1.4.2 Abbildungen

Der Begriff der Funktion ist für die Analysis, die Lineare Algebra, die Mathematik allgemein und ihre Anwendungen unentbehrlich.

**Definition 1.29** (Funktionen und erste dazugehörige Begriffe). In Worten ausgedrückt ist eine **Funktion** $f$ von der Menge $X$ nach der Menge $Y$ eine Zuordnung, die jedem $x \in X$ ein eindeutig bestimmtes $y = f(x) \in Y$ zuweist. Wir schreiben $f : X \to Y$ für eine Funktion von $X$ nach $Y$ und sprechen manchmal auch von einer **Abbildung** oder einer **Transformation**. Zu einer Funktion $f : X \to Y$ wird $X$ der **Definitionsbereich** von $f$ und $Y$ der **Wertebereich** (oder **Zielbereich**) von $f$ genannt. Im Zusammenhang mit der Funktion $f$ wird ein Element $x$ des Definitionsbereichs auch **Argument** und ein von der Funktion angenommenes Element $y = f(x) \in Y$ für ein $x \in X$ auch **Wert** der Funktion genannt.

Zwei Funktionen $f_1 : X_1 \to Y_1$ und $f_2 : X_2 \to Y_2$ gelten als gleich, falls $X_1 = X_2$, $Y_1 = Y_2$ und $f_1(x) = f_2(x)$ für alle $x \in X_1$ gilt. Das **Bild einer Funktion** $f : X \to Y$ ist die Menge

$$f(X) = \{y \in Y \mid \exists x \in X : y = f(x)\}.$$

Falls $f : X \to Y$ eine Funktion und $A$ eine Teilmenge von $X$ ist, dann definieren wir die **Einschränkung** $f|_A : A \to Y$ durch $f|_A(x) = f(x)$ für alle $x \in A$. Umgekehrt wird die Abbildung $f$ auch als eine **Fortsetzung** der Funktion $f|_A$ bezeichnet. Des Weiteren definieren wir das **Bild der Teilmenge** $A$ unter $f$ als

$$f(A) = f|_A(A) = \{y \in Y \mid \exists x \in A : y = f(x)\}\,.$$

**Beispiel 1.30** (Erste Funktionen)**.**

*(a) Sei $X$ eine Menge. Die **Identitätsabbildung** $\mathrm{id}_X : X \to X$ ist die Funktion definiert durch $\mathrm{id}_X(x) = x$ für alle $x \in X$.*

*(b) Seien $X, Y$ Mengen und sei $y_0 \in Y$. Dann ist $f : X \to Y$ definiert durch $f(x) = y_0$ für alle $x \in X$ eine Funktion. Solche Funktionen werden **konstante Funktionen** genannt.*

*(c) Sei $A \subseteq X$ eine Teilmenge und $Y$ eine Menge, die $\{0, 1\}$ enthält. Die Funktion*

$$\begin{aligned} \mathbb{1}_A &: X \to Y \\ \mathbb{1}_A(x) &= \begin{cases} 0 & \text{für } x \notin A \\ 1 & \text{für } x \in A \end{cases} \end{aligned}$$

*wird die **charakteristische Funktion** von $A$ genannt. Das Bild von $\mathbb{1}_A$ ist genau dann ganz $\{0, 1\}$, wenn die Menge $A$ weder leer noch ganz $X$ ist (wieso?).*

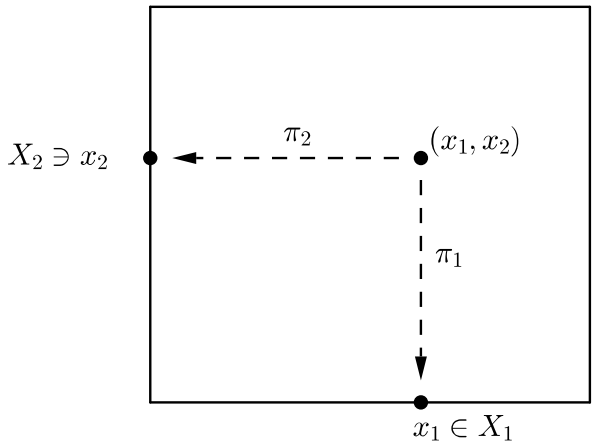
*(d) Für zwei Mengen $X_1, X_2$ definieren wir die **Projektionen***

$$\begin{aligned} \pi_1 &: X_1 \times X_2 \to X_1 \\ \pi_1\big((x_1, x_2)\big) &= x_1 \end{aligned}$$

*und*

$$\begin{aligned} \pi_2 &: X_1 \times X_2 \to X_2 \\ \pi_2\big((x_1, x_2)\big) &= x_2. \end{aligned}$$

*Das Bild von $\pi_1$ ist $X_1$, falls $X_2$ nicht leer ist, und ebenso ist $\pi_2(X_1 \times X_2) = X_2$, falls $X_1 \neq \{\}$. Wir schreiben manchmal auch $\pi_{X_1}$ anstelle von $\pi_1$ und $\pi_{X_2}$ anstelle von $\pi_2$.*




Oft, aber bei weitem nicht immer, wird eine Funktion durch eine Formel wie zum Beispiel $y = x^2$ definiert, doch ohne Angabe des Definitionsbereichs und des Wertebereichs wäre die Definition der Funktion unvollständig (siehe auch Beispiel 1.33). Wir diskutieren hier kurz geläufige Art und Weisen, diesen Mangel zu beheben. Ist $f : X \to Y$ eine Funktion, so schreibt man auch

$$f : X \to Y,\ x \mapsto f(x)$$

oder

$$\begin{aligned} f : X &\to Y \\ x &\mapsto f(x), \end{aligned}$$

wobei $f(x)$ eine konkrete Formel sein könnte. Beispielsweise wäre $f : \mathbb{R} \to \mathbb{R},\ x \mapsto x^2$ eine vollständig definierte Funktion in dieser Notation. Alternativ schreibt man auch

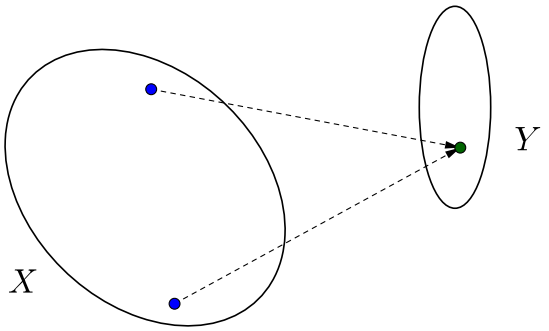
$$x \in X \mapsto f(x) \in Y,$$

also zum Beispiel $x \in \mathbb{R} \mapsto x^2 \in \mathbb{R}$, und hat auch hier Definitions- und Wertebereich angegeben. Wir sprechen „$\mapsto$“ unter anderem als „wird abgebildet auf‘ aus.

**Definition 1.31** (Drei Eigenschaften von Funktionen)**.** Sei $f : X \to Y$ eine Funktion.

- Die Funktion $f$ heisst **injektiv** oder eine **Injektion** falls für alle $x_1, x_2 \in X$ gilt, dass $f(x_1) = f(x_2) \implies x_1 = x_2$.
- Die Funktion $f$ ist **surjektiv**, eine **Surjektion** oder eine Funktion von $X$ **auf** $Y$, falls $f(X) = Y$.
- Die Funktion $f$ heisst **bijektiv**, eine **Bijektion** oder eine **eineindeutige Abbildung**, falls $f$ surjektiv und injektiv ist.

Eine Funktion $f : X \to Y$ ist also nicht injektiv, falls zwei verschiedene Elemente $x_1, x_2 \in X$ existieren mit $f(x_1) = f(x_2)$. Im Bild:




**Applet 1.32** (Funktionen auf Mengen mit höchstens drei Elementen)**.** *In diesem Applet kann man verschiedene Funktionen $f$ von einer Menge $X$ mit höchstens drei Elementen nach einer Menge $Y$ mit höchstens drei Elementen definieren. Hierbei kommt es zu verschiedenen Eigenschaften der Funktion.*

**Beispiel 1.33** (Ist Quadrieren bijektiv?)**.** *Ist $x \mapsto x^2$ injektiv, surjektiv oder bijektiv? Diese Frage macht wie oben erklärt (noch) keinen Sinn, da weder Definitionsbereich noch Wertebereich festgelegt wurden und wir diese zur Beantwortung der Frage kennen müssen. Folgende Beobachtungen bestätigen, dass die Antwort zur Frage in der Tat von der Wahl des Definitionsbereichs und des Wertebereichs abhängt:*

*(a) Die Funktion $f : \mathbb{R}_{\geq 0} \to \mathbb{R},\ x \mapsto x^2$ ist injektiv, denn falls $x, y$ zwei nicht-negative Zahlen mit $x < y$ sind, so gilt auch $x^2 < y^2$. Genauso folgt aus $x > y$, dass $x^2 > y^2$ ist. Sie ist jedoch nicht surjektiv: $-1$ ist nicht Element des Bildes von $f$, da Quadrate von reellen Zahlen niemals negativ sind.*

*(b) Die Funktion $f : \mathbb{R} \to \mathbb{R}_{\geq 0},\ x \mapsto x^2$ ist surjektiv (es existiert eine Wurzel), sie ist aber nicht injektiv, denn $f(1) = 1 = (-1)^2 = f(-1)$.*

*(c) Die Funktion $f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0},\ x \mapsto x^2$ ist bijektiv.*

*Wir werden diese Aussagen später noch ausführlicher beweisen.*

Sei $f : X \to Y$ eine bijektive Funktion. Dann existiert zu jedem $y \in Y$ ein eindeutig bestimmtes $x \in X$, so dass $y = f(x)$: Da $f$ surjektiv ist, existiert ein solches $x$ und da $f$ injektiv ist, kann es höchstens ein solches $x$ geben. Daher rührt auch die Bezeichnung „eineindeutige Abbildung“ für eine Bijektion. Insbesondere kann man wie folgt eine Abbildung definieren.

**Definition 1.34** (Umkehrabbildung)**.** Sei $f : X \to Y$ eine Bijektion. Die **Umkehrabbildung** (**Inverse Abbildung** oder auch kurz ***f* invers**) $f^{-1} : Y \to X$ von $f$ ist die Abbildung, welche jedem $y \in Y$ das eindeutig bestimmte Element $x \in X$ mit $y = f(x)$ zuweist.

An dieser Stelle sollte man noch erwähnen, dass $f^{-1}$ nicht durch eine Formel gegeben sein muss, selbst wenn $f$ durch eine Formel definiert war. Des Weiteren wollen wir anmerken, dass

sich allgemeiner für eine injektive Funktion $f : X \to Y$ eine Umkehrabbildung mit eingeschränktem Definitionsbereich $f^{-1} : f(X) \to X$ definieren lässt. In der Tat ist die Funktion $x \in X \mapsto f(x) \in f(X)$ (also „$f$ mit eingeschränktem Wertebereich“) bijektiv.

**Definition 1.35** (Verknüpfung)**.** Angenommen $f : X \to Y$ und $g : Y \to Z$ sind Funktionen. Dann ist die **Verknüpfung** $g \circ f : X \to Z$ (gesprochen $g$ nach $f$ oder $g$ Ring $f$) für alle $x \in X$ durch $g \circ f(x) = g(f(x))$ definiert.

Wir bemerken, dass für eine bijektive Abbildung $f : X \to Y$ und ihre Umkehrabbildung $f^{-1} : Y \to X$ die Identitäten

$$f \circ f^{-1} = \mathrm{id}_Y, \quad f^{-1} \circ f = \mathrm{id}_X$$

erfüllt sind – eben per Konstruktion von $f^{-1}$ (siehe auch Übung 1.44).

**Beispiel 1.36** (Assoziativgesetz für Verknüpfungen)**.** *Seien $f : X_1 \to X_2$, $g : X_2 \to X_3$, $h : X_3 \to X_4$ Funktionen. Dann können wir sowohl die Funktion $h \circ (g \circ f) : X_1 \to X_4$ als auch die Funktion $(h \circ g) \circ f : X_1 \to X_4$ betrachten. Glücklicherweise sind die Klammern irrelevant, denn es gilt*

$$h \circ (g \circ f)(x) = h(g \circ f(x)) = h(g(f(x)) = h \circ g(f(x)) = (h \circ g) \circ f(x)$$

*für alle $x \in X_1$. Deswegen schreiben wir einfach $h \circ g \circ f : X_1 \to X_4$.*

**Beispiel 1.37** (Verknüpfung ist nicht kommutativ)**.** *Für zwei Funktionen $f : X \to Y$ und $g : Y \to Z$ macht zwischen $g \circ f$ und $f \circ g$ im Allgemeinen nur $g \circ f$ Sinn, da $f$ nicht auf $Z$ definiert sein muss. Dies kann man sich mit einer Analogie aus der Informatik illustrieren. Seien zwei Programme $P_1, P_2$ gegeben, so dass $P_1$ eine Zahl als Input und einen Zeichenkette als Output hat und $P_2$ eine Zeichenkette als Input und eine Zeichenkette als Output hat. In diesem Fall kann man $P_2$ den Output von $P_1$ als Input übergeben; umgekehrt jedoch kann man $P_1$ den Output von $P_2$ nicht übergeben, da $P_1$ nur mit einer Zahl etwas anzufangen weiss.*

*Nimmt man nun an, dass $X = Y = Z$, so dass sowohl $g \circ f$ als auch $f \circ g$ Sinn machen und beide Abbildungen auf $X$ definiert sind, dann gilt die Gleichheit $g \circ f = f \circ g$ im Allgemeinen trotzdem nicht. Wir nennen ein geometrisches Beispiel. Seien $X = Y = Z = \mathbb{R}^2$, $f$ die Spiegelung um die $y$-Achse und $g$ die Verschiebung (Translation) um $1$ in positiver Richtung entlang der $x$-Achse. Der Ursprung $(0, 0)$ wird unter $g \circ f$ auf $(1, 0)$ abgebildet und unter $f \circ g$ auf $(-1, 0)$ abgebildet, also gilt $g \circ f \neq f \circ g$.*

Eine Funktion, für die Wertebereich gleich dem Definitionsbereich ist, lässt sich auch mit sich selbst verknüpfen.

**Definition 1.38** (Iterationen einer Funktion)**.** Sei $f : X \to X$ eine Funktion von einer Menge $X$ auf sich selbst. Wir definieren die **Iterationen der Funktion** $f$ durch

$$f^{\circ 1} = f : X \to X, \quad f^{\circ 2} = f \circ f : X \to X, \quad f^{\circ 3} = f \circ f \circ f : X \to X$$

und rekursiv auch

$$f^{\circ(n+1)} = f \circ f^{\circ n} : X \to X$$

für alle natürlichen Zahlen $n \in \mathbb{N}$. Wir sprechen $f^{\circ n}$ auch als „$f$ hoch $n$“ aus und setzen des Weiteren $f^{\circ 0} = \mathrm{id}_X$.

Wir bemerken, dass man die Iteration einer Funktion $f : X \to X$ oft einfach auch durch $f^n = f^{\circ n}$ bezeichnet.[2]

**Übung 1.39** (Beispiele und eine Eigenschaft von Iterationen)**.**

*(i) Sei $X = \{a, b, c\}$ eine Menge mit 3 verschiedenen Elementen und $f : X \to X$ die Funktion definiert durch $f(a) = b$, $f(b) = c$ und $f(c) = a$. Zeigen Sie, dass $f$ bijektiv ist. Beschreiben Sie die Umkehrabbildung und alle Iterationen von $f$ explizit.*

*(ii) Iterieren Sie die Funktion $g : x \in \mathbb{R} \to 3x \in \mathbb{R}$ und beschreiben Sie intuitiv, wie sich $f^{\circ n}(x)$ für ein $x \in \mathbb{R}$ und verschiedene $n \in \mathbb{N}$ verhält.*

*(iii) Iterieren Sie die Funktion $g : x \in \mathbb{R} \to x^2 \in \mathbb{R}$ und beschreiben Sie intuitiv, wie sich $f^{\circ n}(x)$ für ein $x \in \mathbb{R}$ und verschiedene $n \in \mathbb{N}$ verhält.*

*(iv) Sei $X$ eine Menge und $f : X \to X$ eine Transformation auf $X$. Zeigen Sie, dass für alle $m, n \in \mathbb{N}_0$ gilt $f^{\circ m} \circ f^{\circ n} = f^{\circ(m+n)}$.*

Für eine Abbildung $f : X \to X$ auf einer endlichen Menge $X$ sind Injektivität und Surjektivität äquivalent (siehe auch Abschnitt 1.4.4). Für unendliche Mengen ergeben sich jedoch weitere Möglichkeiten, die auf den ersten Blick ungewöhnlich erscheinen könnten.

**Beispiel 1.40** (Hilbert–Hotel)**.** *Die Abbildung*

$$f : n \in \mathbb{N} \mapsto n + 1 \in \mathbb{N}$$

*ist zwar injektiv aber nicht surjektiv, denn $1 \notin f(\mathbb{N})$. Dies kann man sich zur Veranschaulichung wie folgt vorstellen: Angenommen ein Hotel (das sogenannte **Hilbert Hotel**) ist voll belegt und hat für jede natürliche Zahl $n$ ein Hotelzimmer mit der Türnummer $n$ (entlang eines wirklich sehr langen Ganges der Reihe nach nummeriert). Dann kann man trotzdem noch einen Gast unterbringen. Dazu muss die Rezeption bloss dem Gast im ersten Zimmer einen Brief mit folgender Botschaft aushändigen:*

> *„Wir bedauern, Ihnen mitzuteilen, dass wir Ihnen ein neues Zimmer zuweisen müssen. Bitte übergeben Sie diesen Brief dem Gast im nächsten Zimmer und übernehmen Sie dieses Zimmer, sobald es frei ist.“*

*Daraufhin wird das erste Zimmer frei sein, wo dann der neue Gast untergebracht werden kann. Ebenso bekommen alle alten Gäste jeweils ein neues Zimmer.*

[2] Wir wollen die Notation $f^{\circ n}$ verwenden, um Verwirrungen mit der $n$-ten Potenz einer reellwertigen Funktion zu vermeiden.

*Genauso gibt es auch surjektive Abbildungen $\mathbb{N} \to \mathbb{N}$, die nicht injektiv sind, zum Beispiel*

$$g : \mathbb{N} \to \mathbb{N},\ n \mapsto \begin{cases} 1 & \text{falls } n = 1 \\ n-1 & \text{falls } n \neq 1 \end{cases}.$$

*Das Gedankenspiel von Hilberts Hotel lässt sich durchaus erweiteren. Wir verweisen dazu auf die nächste Übung und diesen Kurzfilm, der folgende Übung auflöst und erweitert.*

**Übung 1.41** (Hilberts Bus)**.** *Wir setzen das Gedankenspiel von Hilberts Hotel fort: Stellen Sie sich vor, dass ein sehr langer, voll besetzter Bus mit Gästen vor Hilberts vollem Hotel auffährt, der für jede natürliche Zahl einen entsprechend nummerierten Sitzplatz aufweist. Wie können Sie die neuen Gäste alle im Hotel unterbringen? Welche Abbildung $h : \mathbb{N} \to \mathbb{N}$ können Sie verwenden, um im Hotel Platz zu beschaffen?*

Sowohl die Eigenschaften in Definition 1.31 als auch die Verknüpfung in Definition 1.35 sind so grundlegend, dass wir nicht umhin kommen, einige fundamentale Tatsachen zu beweisen:

**Lemma 1.42** (Eigenschaften von Verknüpfungen)**.** *Seien $f : X \to Y$ und $g : Y \to Z$ Funktionen.*

*(i) Falls $f$ und $g$ injektiv sind, dann ist auch $g \circ f$ injektiv.*

*(ii) Falls $f$ und $g$ surjektiv sind, dann ist auch $g \circ f$ surjektiv.*

*(iii) Falls $f$ und $g$ bijektiv sind, dann ist auch $g \circ f$ bijektiv und es gilt $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.*

*Beweis.* (i): Angenommen $f$ und $g$ sind injektiv und $g \circ f(x_1) = g \circ f(x_2)$ für zwei Elemente $x_1, x_2 \in X$. Wegen $g \circ f(x_1) = g(f(x_1))$ und $g \circ f(x_2) = g(f(x_2))$ muss $f(x_1) = f(x_2)$ gelten, da $g$ injektiv ist. Da $f$ injektiv ist, gilt $x_1 = x_2$. Dies zeigt, dass $g \circ f$ injektiv ist.

(ii): Angenommen $f$ und $g$ sind surjektiv und $z \in Z$ ist ein beliebiges Element. Da $g$ surjektiv ist, können wir ein $y \in Y$ mit $g(y) = z$ wählen. Da auch $f$ surjektiv ist, gibt es ein $x \in X$ mit $f(x) = y$ und damit $g \circ f(x) = g(f(x)) = g(y) = z$. Also existiert für jedes Element $z \in Z$ ein Element $x \in X$ mit $g \circ f(x) = z$ und daher ist $g \circ f$ surjektiv.

(iii): Angenommen $f$ und $g$ sind bijektiv. Der erste Teil von (iii) folgt direkt, denn nach (i) ist $g \circ f$ injektiv und nach (ii) surjektiv. Wir verifizieren nun die behauptete Formel. Sei $z \in Z$. Wir müssen zeigen, dass $f^{-1}(g^{-1}(z))$ ein Punkt (und damit der Punkt) ist, der unter $g \circ f$ auf $z$ abgebildet wird. Tatsächlich gilt wegen $f(f^{-1}(y)) = y$ für alle $y \in Y$ und $g(g^{-1}(z)) = z$ für alle $z \in Z$ auch

$$g \circ f(f^{-1}(g^{-1}(z)) = g(f(f^{-1}(g^{-1}(z)))) = g(g^{-1}(z)) = z$$

und (iii) ist bewiesen. □

In der folgenden Übung kombinieren wir die Begriffe aus den Definitionen 1.31 und 1.35 ein weiteres Mal:

**Wichtige Übung 1.43** (Weitere Eigenschaften von Verknüpfungen)**.** *Seien $f : X \to Y$ und $g : Y \to Z$ Funktionen.*

(i) *Zeigen Sie, dass $g$ surjektiv ist, falls $g \circ f$ surjektiv ist. Überzeugen Sie sich davon, dass in diesem Fall $f$ nicht unbedingt surjektiv sein muss.*

(ii) *Zeigen Sie, dass $f$ surjektiv ist, falls $g \circ f$ surjektiv und $g$ injektiv ist.*

(iii) *Zeigen Sie, dass $f$ injektiv ist, falls $g \circ f$ injektiv ist. Überzeugen Sie sich davon, dass in diesem Fall $g$ nicht unbedingt injektiv sein muss.*

(iv) *Zeigen Sie, dass $g$ injektiv ist, falls $g \circ f$ injektiv ist und $f$ surjektiv ist.*

**Wichtige Übung 1.44** (Bijektivität und Existenz einer Umkehrabbildung)**.** *Sei $f : X \to Y$ eine Funktion. Wenn $f$ bijektiv ist, dann gelten $f^{-1} \circ f = \mathrm{id}_X$ und $f \circ f^{-1} = \mathrm{id}_Y$, wie schon bemerkt wurde. Wir zeigen hier die „Umkehrung“: Angenommen es existiert eine Funktion $g : Y \to X$, so dass $g \circ f = \mathrm{id}_X$ und $f \circ g = \mathrm{id}_Y$. Zeigen Sie unter Verwendung der letzten Übung, dass $f$ bijektiv ist und dass $g = f^{-1}$ gilt. Des Weiteren, überprüfen Sie, dass $f^{-1} : Y \to X$ ebenfalls bijektiv ist. Was ist die Umkehrabbildung von $f^{-1}$?*

Folgende Definition ist auch für nicht bijektive Funktionen nützlich:

**Definition 1.45** (Urbilder einer Funktion)**.** Für eine Funktion $f : X \to Y$ und eine Teilmenge $B \subseteq Y$ definieren wir das **Urbild** $f^{-1}(B)$ von $B$ unter $f$ als

$$f^{-1}(B) = \{x \in X \mid f(x) \in B\}.$$

Beispielsweise gilt für eine Funktion $f : X \to Y$, dass $f^{-1}(\{\}) = \{\}$ und $f^{-1}(Y) = X$. In der nächsten Übung prüfen wir einige Eigenschaften der Definition des Urbilds.

**Wichtige Übung 1.46** (Verhalten von Bildern und Urbildern unter Mengenoperationen)**.** *Gegeben sei eine Funktion $f : X \to Y$ und Teilmengen $A, A' \subseteq X$ und $B, B' \subseteq Y$.*

(i) *Zeigen Sie, dass $f(f^{-1}(B)) \subseteq B$ gilt. Unter welcher Bedingung an $f$ gilt auf jeden Fall Gleichheit?*

(ii) *Zeigen Sie, dass $f^{-1}(f(A)) \supseteq A$ gilt. Unter welcher Bedingung an $f$ gilt auf jeden Fall Gleichheit?*

(iii) *Zeigen Sie die Gleichungen*

$$f(A \cup A') = f(A) \cup f(A'), \quad f^{-1}(B \cup B') = f^{-1}(B) \cup f^{-1}(B').$$

(iv) *Zeigen Sie, dass $f(A \cap A') \subseteq f(A) \cap f(A')$ und dass Gleichheit gilt, wenn $f$ injektiv ist. Verifzieren Sie, dass in diesem Fall auch $f(A \setminus A') = f(A) \setminus f(A')$ gilt.*

(v) *Zeigen Sie, dass $f^{-1}(B \cap B') = f^{-1}(B) \cap f^{-1}(B')$ und $f^{-1}(Y \setminus B) = X \setminus f^{-1}(B)$ gelten.*

(vi) *Verallgemeinern Sie die Regeln $f^{-1}(B \cup B') = f^{-1}(B) \cup f^{-1}(B')$ und $f^{-1}(B \cap B') = f^{-1}(B) \cap f^{-1}(B')$ für beliebige Vereinigungen resp. Schnitte.*

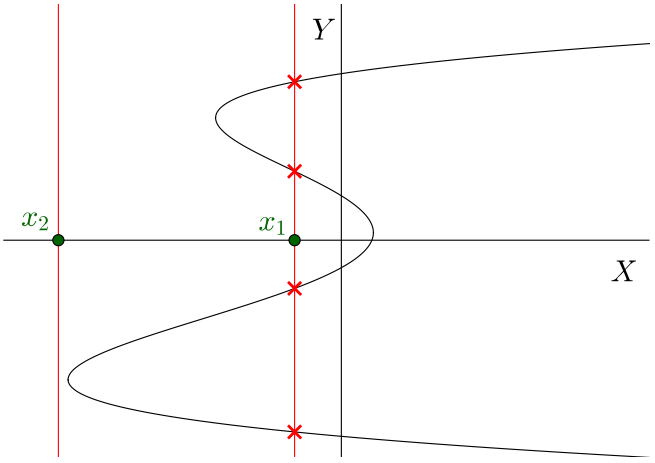
*Zusammenfassend sollten Sie sich merken, dass die Urbildoperation mit allen von uns besprochenen mengentheoretischen Operationen (unter anderem Vereinigung, Durchschnitt und Komplement) verträglich ist, während dies die Bildoperation nur für die Vereinigung oder unter gewissen Bedingungen erfüllt.*

Mit Hilfe des folgenden Begriffs lässt sich der Funktionsbegriff, der oben in Worten eingeführt wurde, formalisieren:

**Definition 1.47** (Graph einer Funktion)**.** Sei $f : X \to Y$ eine Funktion. Der **Graph** von $f$ ist

$$\operatorname{graph}(f) = \{(x, y) \in X \times Y \mid y = f(x)\} \subseteq X \times Y.$$

Der Graph einer Funktion enthält viel Information über die Funktion selbst. Unter anderem lässt sich aus dem Graphen bereits erkennen, ob die Funktion wohldefiniert ist (das heisst in der Tat eine Funktion ist) oder nicht. Beispielsweise kann folgende Teilmenge von $X \times Y$ nicht der Graph einer Funktion von $X$ nach $Y$ sein:




Figur 1.7: Die hier dargestellte Teilmenge von $X \times Y$ ist nicht der Graph einer Funktion von $X$ nach $Y$. Denn man müsste dem grün markierten Punkt $x_1$ in $X$ vier verschiedene Punkte in $Y$ zuweisen, was der Definition einer Funktion widerspricht. Genauso problematisch ist der Punkt $x_2 \in X$, dem gar kein Punkt zugewiesen wird. Standardmässig nimmt man bei der Darstellung des Graphen einer Funktion den Definitionbereich als die Horizontale und den Wertebereich als die Vertikale.

Welche Teilmengen von $X \times Y$ Graphen einer Funktion sein können und welche nicht, wollen wir im Folgenden erklären. Wir beginnen damit, das in Figur 1.7 aufgetauchte Problem formal zu interpretieren:

**Übung 1.48** (Graphen und Projektion auf den Definitionsbereich)**.**

*(i) Sei $f : X \to Y$ eine Funktion. Wie in Beispiel 1.30 bezeichnen wir mit $\pi_1 : X \times Y \to X$ die Projektion auf $X$, die durch $(x, y) \mapsto x$ definiert ist und mit $\pi_2 : X \times Y \to Y$ die Projektion auf $Y$. Zeigen Sie, dass $\pi_1|_{\operatorname{graph}(f)} : \operatorname{graph}(f) \to X$ bijektiv ist und finden Sie*

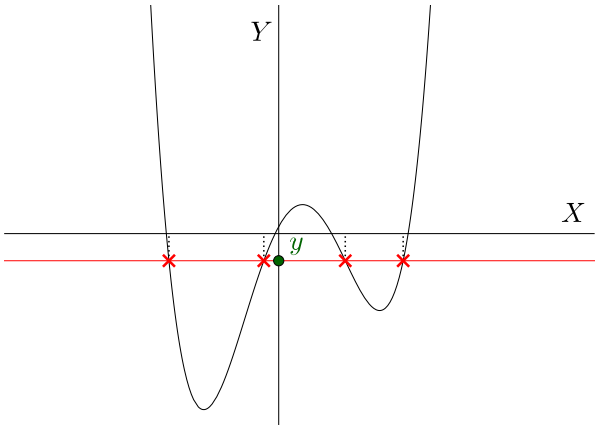
*die Umkehrabbildung. Beweisen Sie des Weiteren die Identität*

$$\pi_2 \circ (\pi_1|_{\mathrm{graph}(f)})^{-1} = f,$$

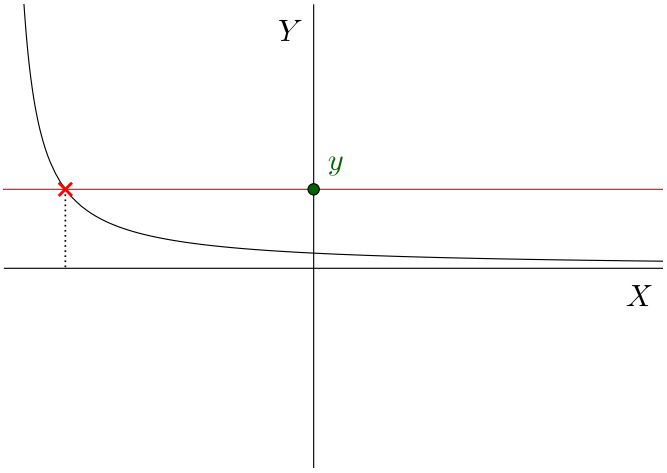
*welche erklärt, wie man aus dem Graphen einer Funktion, die Funktion zurückerhalten kann.*

*(ii) Zeigen Sie, dass eine Teilmenge $\Gamma \subseteq X \times Y$ genau dann ein Graph einer Funktion $f : X \to Y$ ist, wenn $\pi_1|_\Gamma : \Gamma \to X$ bijektiv ist. Verifizieren Sie auch, dass in diesem Fall $f$ eindeutig bestimmt ist.*

Auch die Eigenschaften aus Definition 1.31 lassen sich im Graphen erkennen. Unter anderem ist ersichtlich, ob eine Funktion injektiv ist oder nicht:

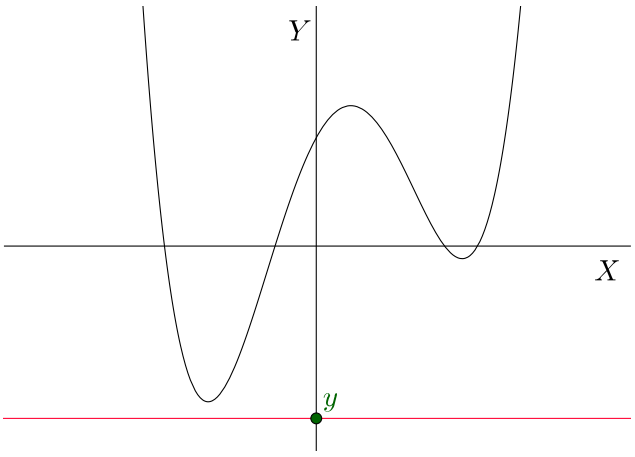



Figur 1.8: Dies ist der Graph einer nicht injektiven Funktion. Der grün markierte Punkt $y$ in $Y$ wird unter der Funktion von vier verschiedenen Elementen in $X$ angenommen.

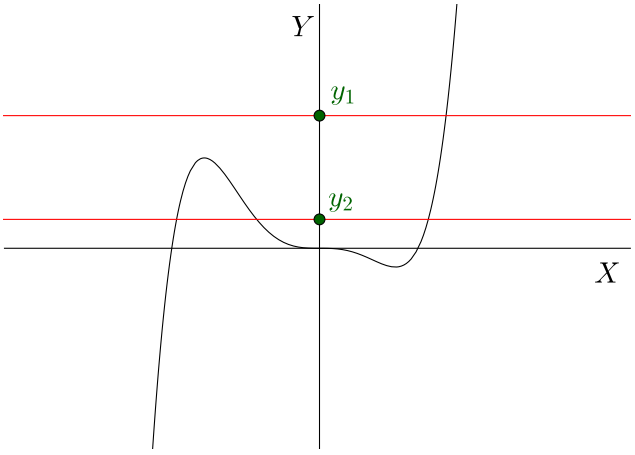



Figur 1.9: Dies ist der Graph einer injektiven Funktion. Der grün markierte Punkt $y \in Y$ ist das Bild von nur einem Punkt in $X$. Dies ist immer noch der Fall, wenn $y$ verschoben wird (bis auf die Tatsache, dass $y$ dann vielleicht gar nicht mehr im Bild liegt). In anderen Worten: jede horizontale Linie schneidet den Graphen in höchstens einem Punkt.

Genauso lässt sich beurteilen, ob eine Funktion surjektiv ist oder nicht - siehe Figuren 1.10 und 1.11. Wir formalisieren dies in Übung 1.50.




Figur 1.10: Dies ist der Graph einer nicht-surjektiven Funktion. Der grün markierte Punkt $y$ in $Y$ wird unter der Funktion von keinem Element in $X$ angenommen.




Figur 1.11: Dies ist der Graph einer surjektiven Funktion. Jede horizontale Linie (wie zum Beispiel jene durch $y_1$ oder $y_2$) schneidet den Graphen in mindestens einem Punkt.

**Applet 1.49** (Einschränkungen einer Funktion)**.** *Wir betrachten eine Funktion $f : I \to \mathbb{R}$, die auf einem Intervall $I$ in $\mathbb{R}$ definiert ist. Durch Einschränken der Funktion auf Teilintervalle $[a, b] \subseteq I$ des Definitionsbereichs oder durch Einschränken des Wertebereichs auf ein Intervall $[c, d] \subseteq \mathbb{R}$ können wir mitunter Injektivität, Surjektivität, oder auch Bijektivität der neuen Funktion erreichen.*

**Wichtige Übung 1.50** (Eigenschaften des Graphen unter Projektion auf den Wertebereich)**.** *Sei $f : X \to Y$ eine Funktion.*

- *Zeigen Sie, dass $f$ injektiv ist genau dann, wenn $\pi_2|_{\mathrm{graph}(f)}$ injektiv ist.*
- *Zeigen Sie, dass $f$ surjektiv ist genau dann, wenn $\pi_2|_{\mathrm{graph}(f)}$ surjektiv ist. Intuitiv bedeutet letzteres, dass die Projektion des Graphen $\mathrm{graph}(f)$ auf $Y$ jedes Element erreicht.*
- *Angenommen $f$ ist bijektiv. Wie findet man den Graphen von $f^{-1}$?*

Sei $f : X \to Y$ eine Funktion. Wir verwenden manchmal auch die Notation $\{f(x) \mid x \in X\}$ für die Bildmenge $f(X)$. Allgemeiner schreiben wir auch

$$\{f(x) \mid x \in X \wedge A(x)\} = \big\{y \mid y \in Y \wedge \exists x \in X : \big(f(x) = y \wedge A(x)\big)\big\}\,,$$

wobei $A(x)$ irgend eine Aussage über Elemente $x$ von $X$ ist. In dieser Gleichung haben wir rechts genau nach dem Schema in der dritten Annahme an die Mengenlehre (siehe Abschnitt 1.4.1) eine Menge definiert und verwenden dies, um den kürzeren Ausdruck links zu definieren.

**Applet 1.51** (Bilder und Urbilder)**.** *Wir betrachten eine Funktion $f : I \to \mathbb{R}$, die auf einem Intervall $I$ in $\mathbb{R}$ definiert ist. Je nach Einstellung des Applets wird entweder das Bild $f(A)$ eines Teilintervalls $A = [x_1, x_2] \subseteq I$ oder das Urbild $f^{-1}([y_1, y_2])$ eines Intervalls $[y_1, y_2] \subseteq \mathbb{R}$ dargestellt.*

Gewisse Objekte bezeichnet man als **kanonisch** oder **natürlich**, wenn sie unabhängig von jeglicher Wahl sind. Dieser Begriff drängt sich insbesondere in der Linearen Algebra auf. Wir möchten dies an einem Beispiel erläutern.

**Beispiel 1.52** (Kanonische Abbildungen)**.** *Sei $X$ eine Menge und $A \subseteq X$ eine Teilmenge. Wir betrachten injektive Abbildungen $A \to X$ und versuchen darunter eine kanonische auszumachen. Da $A \subseteq X$ eine Teilmenge ist, ist die injektive Abbildung $A \to X$, die man am schnellsten hinschreiben kann, die **Inklusionsabbildung***

$$\iota : A \to X,\ x \mapsto x$$

*am natürlichsten. Denn wir mussten dazu weder spezifische Elemente von $A$ noch von $X$ wählen. Es gibt nun aber ganz viele andere injektive Abbildungen von $A$ nach $X$. Diese involvieren jedoch alle eine (recht zufällige) Regel.*

*Wir untersuchen das konkrete Beispiel $A = \{a\}$ und $X = \{a, b, c\}$. Dann gibt es 3 verschiedene injektive Abbildungen $A \to X$. Nebst der Inklusionsabbildung (die $a$ auf $a$ abbildet) sind dies*

- *die Funktion $A \to X$ mit $a \mapsto b$ und*
- *die Funktion $A \to X$ mit $a \mapsto c$.*

*Für diese beiden Abbildungen mussten wir jedoch die Menge $X$ kennen und spezifische Elemente von $X$ (das wären $b, c$) auswählen, was bei der Inklusionsabbildung nicht der Fall war. In diesem Sinne ist die Inklusionsabbildung unter den injektiven Abbildungen $A \to X$ die natürliche, die kanonische.*

*Ein weiteres interessanteres Beispiel bilden die in Beispiel 1.30 definierten **kanonischen Projektionen** $\pi_1 : X_1 \times X_2 \to X_1$ und $\pi_2 : X_1 \times X_2 \to X_2$ für zwei Mengen $X_1, X_2$.*

Für zwei Mengen $X, Y$ wird die Menge aller Abbildungen von $X$ nach $Y$ als $Y^X$ geschrieben, formaler

$$Y^X = \{f \mid f : X \to Y \text{ ist eine Funktion}\}\,.$$

Grund für die Notation ist unter anderem die folgende Behauptung, die man mittels Induktion beweisen kann: Falls $X$ und $Y$ endliche Mengen sind mit $m$ resp. $n$ Elementen für zwei natürliche Zahlen $m, n \in \mathbb{N}$, so hat $Y^X$ genau $n^m$ Elemente.

### 1.4.3 Relationen

**Definition 1.53** (Relationen)**.** Seien $X$ und $Y$ Mengen. Eine **Relation** auf $X \times Y$ ist eine Teilmenge $\mathcal{R} \subseteq X \times Y$. Wir schreiben auch $x\mathcal{R}y$ falls $(x, y) \in \mathcal{R}$ und verwenden oft Symbole wie $<, \ll, \leq, \cong, \equiv, \sim$ für Relationen. Falls $X = Y$ ist, dann sprechen wir auch von einer Relation auf $X$. Wenn $\sim$ (resp. $\cong$, ...) eine Relation ist, dann schreiben wir auch „$x \not\sim y$“ (resp. „$x \ncong y$“, ...) für „$\neg(x \sim y)$“ (resp. „$\neg(x \cong y)$“, ...).

**Übung 1.54** (Eine bekannte Relation)**.** *In einem gewissen Sinne haben wir schon Relationen betrachtet. Seien $X, Y$ Mengen und sei $\mathcal{G}$ eine Relation auf $X \times Y$, die die folgende Eigenschaft erfüllt:*

$$\forall x \in X \ \exists! y \in Y : x\mathcal{G}y$$

*Wie nennen wir eine solche Relation gemeinsam mit $X$ und $Y$?*

Wir werden noch weitere Beispiele von Relationen sehen, doch beschreibt die folgende Definition eine wichtige Klasse von Relationen:

**Definition 1.55** (Äquivalenzrelationen)**.** Eine Relation $\sim$ auf $X$ ist eine **Äquivalenzrelation**, falls folgende drei Eigenschaften erfüllt sind:

- Reflexivität: $\forall x \in X : x \sim x$.
- Symmetrie: $\forall x, y \in X : x \sim y \implies y \sim x$.
- Transitivität: $\forall x, y, z \in X : ((x \sim y) \wedge (y \sim z)) \implies x \sim z$.

Äquivalenzrelationen sind oft durch eine Gleichheit in gewissen Aspekten definiert und sollten als eine Form von einer Gleichheit angesehen werden. In der Tat bezieht sich der Ausdruck „das Gleiche“ in der deutschen Sprache auf eine Art „Äquivalenzrelation“ (die je nach Zusammenhang verschieden sein kann), wohingegen der Ausdruck „dasselbe“ nur für „ein und dasselbe“ Objekt verwendet werden sollte.

**Beispiel 1.56** (Beispiele von Äquivalenzrelationen)**.**

*(i) Das einfachste Beispiel einer Äquivalenzrelation auf einer beliebigen Menge $X$ ist die Gleichheit, also die Relation $\mathcal{R} = \{(x, y) \in X^2 \mid x = y\}$.*

*(ii) Ein weiteres allgemeines Beispiel ist die sogenannte triviale Relation $\mathcal{R} = X^2$, bezüglich der je zwei Elemente in $X$ äquivalent sind.*

*(iii) Wir betrachten ein Beispiel in der ebenen (euklidschen) Geometrie. Sei $X$ die Menge der Geraden in der Ebene. Zu zwei Geraden $G_1, G_2$ schreiben wir*

$$G_1 \sim G_2 \iff G_1 \text{ und } G_2 \text{ sind parallel.}$$

*Dies definiert eine Äquivalenzrelation auf $X$ (wieso?).*

*(iv) Äquivalenzrelationen werden in anderen Wissenschaften oder auch im Alltag häufig verwendet. Beispielsweise kann man auf der Menge aller Lebewesen eine Äquivalenzrelation definieren, in dem man zwei Lebewesen für äquivalent erklärt, wenn sie zur selben Art (oder Gattung oder Familie) gehören.*

**Übung 1.57** (Zwei weitere Beispiele)**.** *In dieser Übungen möchten wir weitere, allgemeine Beispiele von Äquivalenzrelationen besprechen. Sei $X$ eine Menge.*

*(i) (Zerquetschen einer Teilmenge) Sei $A \subseteq X$ eine Teilmenge. Zeigen Sie, dass die Relation gegeben durch*

$$x \sim_A y :\iff (x, y \in A) \vee (x = y)$$

*für $x, y \in X$ eine Äquivalenzrelation auf $X$ definiert.*

*(ii) (Äquivalenz über eine Abbildung) Sei $f : X \to Y$ eine Abbildung in eine weitere Menge $Y$. Wir definieren eine Relation auf $X$ durch*

$$x_1 \sim_f x_2 :\iff f(x_1) = f(x_2)$$

*für alle $x_1, x_2 \in X$. Zeigen Sie, dass $\sim$ eine Äquivalenzrelation auf $X$ definiert.*

**Übung 1.58** (Ein falscher Beweis)**.** *In dieser Aufgabe behaupten wir fälschlicherweise, dass jede symmetrische und transitive Relation $\sim$ auf einer Menge $X$ auch reflexiv ist (d.h. eine Äquivalenzrelation ist). Finden Sie den Fehler in folgendem „Beweis“:*

> *Sei $x \in X$ ein Element. Sei $y \in X$, so dass $x \sim y$. Wegen Symmetrie der Relation gilt also auch $y \sim x$. Folglich gilt unter Verwendung der Transitivität der Relation $(x \sim y) \wedge (y \sim x) \implies x \sim x$, was zu zeigen war.*

*Finden Sie ein Beispiel einer Relation, die symmetrisch und transitiv, aber nicht reflexiv ist.*

**Übung 1.59** (Beispiele allgemeiner Relationen)**.** *Finden Sie eine Relation auf den natürlichen Zahlen $\mathbb{N}$, die von den Eigenschaften einer Äquivalenzrelation*

- *nur die Symmetrie,*
- *nur die Transitivität,*
- *die Reflexivität und die Transitivität, aber nicht die Symmetrie*

*erfüllt.*

Wie schon erwähnt ist eine Äquivalenzrelation gewissermassen eine Form von Gleichheit. Dies lässt sich auch formalisieren:

**Definition 1.60** (Äquivalenzklassen und die Quotientenmenge)**.** Sei $\sim$ eine Äquivalenzrelation auf einer Menge $X$. Dann wird für $x \in X$ die Menge

$$[x]_\sim = \{y \in X \mid y \sim x\}$$

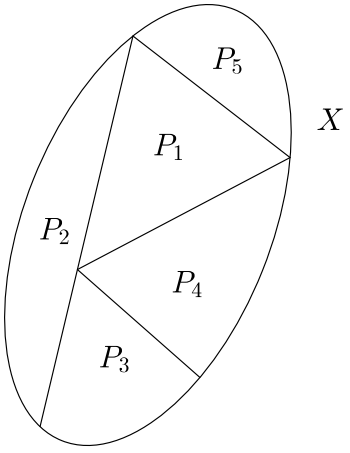
die **Äquivalenzklasse** von $x$ genannt. Weiters ist

$$X/_\sim = \{[x]_\sim \mid x \in X\}$$

der **Quotient** (oder die **Quotientenmenge**) von $X$ modulo $\sim$. Ein Element $x \in X$ wird auch **Repräsentant** seiner Äquivalenzklasse $[x]_\sim$ genannt.

Anschaulich gesprochen geben wir äquivalente Elemente von $X$ in ein und denselben Topf und nicht äquivalente Elemente in verschiedene Töpfe. In diesem Bild besteht die Äquivalenzklasse eines Elements aus allen Elementen, die im gleichen Topf sind. Der Quotient modulo $\sim$ wiederum ist die Menge der Töpfe. Ein alternativer Begriff ist in folgender Definition enthalten.

**Definition 1.61** (Partition)**.** Sei $X$ eine Menge und $\mathcal{P}$ eine Familie von nicht-leeren, paarweise disjunkten Teilmengen von $X$, so dass $X = \bigsqcup_{P \in \mathcal{P}} P$. Dann wird $\mathcal{P}$ eine **Partition** von $X$ genannt.




Figur 1.12: Schemenhafte Darstellung einer Partition $\mathcal{P} = \{P_1, \ldots, P_5\}$ einer Menge $X$.

**Proposition 1.62** (Korrespondenz zwischen Äquivalenzrelationen und Partitionen)**.** *Sei $X$ eine Menge. Dann entsprechen Äquivalenzrelationen auf $X$ und Partitionen von $X$ einander im folgenden Sinne: Für eine gegebene Äquivalenzrelation $\sim$ auf $X$ ist die Menge*

$$\mathcal{P}_\sim = \{[x]_\sim \mid x \in X\}$$

*eine Partition von $X$. Umgekehrt definiert für eine Partition $\mathcal{P}$ von $X$*

$$x \sim_{\mathcal{P}} y \iff \exists P \in \mathcal{P} : x \in P \wedge y \in P$$

*für $x, y \in X$ eine Äquivalenzrelation auf $X$. Des Weiteren sind die Konstruktion der Partition aus der Äquivalenzrelation und die Konstruktion der Äquivalenzrelation aus der Partition zueinander invers: Für jede Partition $\mathcal{P}$ von $X$ gilt $\mathcal{P}_{\sim_{\mathcal{P}}} = \mathcal{P}$ und für jede Äquivalenzrelation $\sim$ auf $X$ gilt $\sim_{\mathcal{P}_\sim} = \sim$.*

Sei $X$ eine Menge, $\sim$ eine Äquivalenzrelation auf $X$ und $\mathcal{P}_\sim$ wie in Proposition 1.62 definiert. Selbstverständlich gilt nach den Definitionen $\mathcal{P}_\sim = X/\sim$. Wir möchten jedoch zwei verschiedene Symbole mitführen, da wir $\mathcal{P}_\sim$ und $X/\sim$ jeweils verschieden interpretieren möchten. $\mathcal{P}_\sim$ werden wir stets als eine Kollektion von Teilmengen von $X$ auffassen; $X/\sim$ hingegen werden wir als einen neuen Raum erachten, wo die Punkte durch Identifikation („Aneinanderkleben“) von gewissen Punkten in $X$ entstanden sind. (Die Punkte in $X/\sim$ sind die Teilmengen von $X$, die in $\mathcal{P}_\sim$ enthalten sind).

**Applet 1.63** (Eine Äquivalenzrelation)**.** *Links wird eine Menge $X$ partitioniert, was einer Äquivalenzrelation $\sim$ auf $X$ entspricht. Rechts betrachten wir die Menge der Äquivalenzklassen, also den Quotienten von $X$ bzüglich $\sim$. Die Menge links könnte eine abstrakte Menge darstellen oder den Einheitskreis. Im letzteren Fall muss klar definiert sein, zu welcher Menge die Punkte der Kanten, bei denen der Kreis unterteilt wird, gehören. Wir haben dies im Beispiel mit Farben angedeutet.*

**Übung 1.64** (Zerquetschen einer Teilmenge)**.** *Charakterisieren Sie die Äquivalenzklassen der in Übung 1.57 (ii) definierten Äquivalenzrelation $\sim$. Zeigen Sie, dass $\mathcal{P}_\sim$ (wie in obiger Proposition definiert) eine Partition ist (ohne auf die noch nicht-bewiesene Proposition zurückzugreifen) und beschreiben Sie $X/\sim$ intuitiv.*

Für den Beweis der Proposition 1.62 ziehen wir folgende Behauptung vor:

**Behauptung.** *Sei $\sim$ eine Äquivalenzrelation auf $X$. Dann sind folgende drei Aussagen für alle $x, y \in X$ (paarweise) äquivalent:*

*(i)* $x \sim y$

*(ii)* $[x]_\sim = [y]_\sim$

*(iii)* $[x]_\sim \cap [y]_\sim \neq \{\}$

*Beweis der Behauptung.* Wir beweisen die Implikationen $(i) \implies (ii) \implies (iii) \implies (i)$, woraus folgt, dass alle drei Aussagen äquivalent sind. Seien $x, y \in X$.

Wir nehmen also zuerst an, dass $x \sim y$ gilt wie in $(i)$. Sei $z \in [x]_\sim$. Dann ist $z \sim x \sim y$ und also $z \sim y$ und $z \in [y]_\sim$ wegen der Transitivität. Die andere Inklusion folgt analog (durch Vertauschen von $x$ und $y$) und es gilt $[x]_\sim = [y]_\sim$ wie in $(ii)$.

Gilt $[x]_\sim = [y]_\sim$ wie in $(ii)$, so folgt $[x]_\sim \cap [y]_\sim \neq \{\}$ wie in $(iii)$ wegen Reflexivität.

Gilt $[x]_\sim \cap [y]_\sim \neq \{\}$ wie in $(iii)$, so existiert ein $z \in X$ mit $z \sim x$ und $z \sim y$. Aus Symmetrie und Transitivität folgt daher $x \sim y$ wie in $(i)$. □

*Beweis von Proposition 1.62.* Sei $\sim$ eine Äquivalenzrelation auf $X$. Dann gilt $\bigcup_{P \in \mathcal{P}_\sim} P = \bigcup_{x \in X} [x]_\sim = X$, da $x \in [x]_\sim$ für jedes $x \in X$. Paarweise Disjunktheit der Elemente von $\mathcal{P}_\sim$ gilt dank der Behauptung und es folgt, dass $\mathcal{P}_\sim$ eine Partition ist.

Sei nun $\mathcal{P}$ eine Partition von $X$ und sei die Relation $\sim_{\mathcal{P}}$ wie in der Proposition definiert. Es ist für jedes $x \in X$ auch $x \sim_{\mathcal{P}} x$, da es wegen $\bigcup_{P \in \mathcal{P}} P = X$ ein $P \in \mathcal{P}$ gibt mit $x \in P$; dies zeigt Reflexivität. Falls $x \sim_{\mathcal{P}} y$ für $x, y \in X$, dann folgt $y \sim_{\mathcal{P}} x$ direkt aus der Definition, das heisst $\sim_{\mathcal{P}}$ ist symmetrisch. Angenommen es gilt $x \sim_{\mathcal{P}} y$ und $y \sim_{\mathcal{P}} z$. Dann gibt es ein Partitionselement $P_1 \in \mathcal{P}$ mit $x, y \in P_1$ und $P_2 \in \mathcal{P}$ mit $y, z \in P_2$. Insbesondere ist $P_1 \cap P_2 \neq \{\}$ und daher $P_1 = P_2$ nach den Eigenschaften der Partition. Es folgt $x, z \in P_1$, $x \sim_{\mathcal{P}} z$ und die Transitivität der Relation $\sim_{\mathcal{P}}$. Daher ist $\sim_{\mathcal{P}}$ eine Äquivalenzrelation.

Für $x \in X$ ist die Äquivalenzklasse bezüglich $\sim_{\mathcal{P}}$ gegeben durch

$$[x]_{\sim_{\mathcal{P}}} = \{y \in X \mid y \sim_{\mathcal{P}} x\} = \bigcup_{P \in \mathcal{P} \wedge x \in P} P$$

Da aber die Elemente von $\mathcal{P}$ paarweise disjunkt sind, kann $x$ nur in einem Element enthalten sein. Insbesondere folgt, dass $[x]_{\sim_{\mathcal{P}}} \in \mathcal{P}$ das eindeutig bestimmte Element der Partition $\mathcal{P}$ ist, das $x$ enthält, und $\mathcal{P}_{\sim_{\mathcal{P}}} \subseteq \mathcal{P}$. Ist $P \in \mathcal{P}$ und $x \in P$, so gilt $[x]_{\sim_{\mathcal{P}}} = P$, also $\mathcal{P}_{\sim_{\mathcal{P}}} = \mathcal{P}$.

Ist umgekehrt $\sim$ eine Äquivalenzrelation auf $X$ und $\mathcal{P}_{\sim}$ die entsprechende Partition, dann gilt für alle $x, y \in X$

$$x \sim_{\mathcal{P}_{\sim}} y \iff [x]_{\sim} = [y]_{\sim} \iff x \sim y$$

nach der Definition von $\sim_{\mathcal{P}_{\sim}}$ und der Behauptung. □

Man verwendet Quotienten modulo Äquivalenzrelationen in der Mathematik oft für die Konstruktion von gewissen Räumen und auch von neuen Zahlenmengen. Wir betrachten zu letzterem ein einfaches und grundlegendes Beispiel: Wir konstruieren die rationalen Zahlen aus den ganzen Zahlen.

**Beispiel 1.65** (Konstruktion der rationalen Zahlen)**.** *Wir nehmen an, dass wir bereits die ganzen Zahlen $\mathbb{Z}$ und die Addition und Multiplikation auf $\mathbb{Z}$ mit allen üblichen Eigenschaften kennen und wollen damit die rationalen Zahlen $\mathbb{Q}$ definieren. Zu diesem Zwecke betrachten wir die Relation $\sim$ auf $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ definiert durch*

$$(m_1, m_2) \sim (n_1, n_2) \iff m_1 n_2 = n_1 m_2$$

*für $(m_1, m_2), (n_1, n_2) \in \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$. Diese Definition rührt daher, dass wir rationale Zahlen als Brüche von ganzen Zahlen auffassen möchten. Dabei müssen wir allerdings solche identifizieren, die „nach Kürzen“ gleich sind; zum Beispiel sollte gelten $\frac{10}{6} = \frac{5}{3}$. Allerdings wollen wir hier davon ausgehen, dass wir die rationalen Zahlen noch nicht kennen, weswegen wir anstatt Gleichungen der Form $\frac{m_1}{m_2} = \frac{n_1}{n_2}$ Gleichungen der Form $m_1 n_2 = n_1 m_2$ mit Ausdrücken innerhalb von $\mathbb{Z}$ betrachten (im Beispiel $10 \cdot 3 = 5 \cdot 6$).*

*Nun verifzieren wir, dass obige Relation tatsächlich eine Äquivalenzrelation ist. Seien dazu $(m_1, m_2), (n_1, n_2), (q_1, q_2) \in \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$.*

- *Reflexivität: $(m_1, m_2) \sim (m_1, m_2)$, denn $m_1 m_2 = m_1 m_2$.*

- *Symmetrie: Angenommen es gilt* $(m_1, m_2) \sim (n_1, n_2)$. *Dann ist per Definition also* $m_1 n_2 = n_1 m_2$, *was* $n_1 m_2 = m_1 n_2$ *und daher auch* $(n_1, n_2) \sim (m_1, m_2)$ *impliziert.*

- *Transitivität:* $(m_1, m_2) \sim (n_1, n_2)$ *und* $(n_1, n_2) \sim (q_1, q_2)$ *ergibt* $m_1 n_2 = n_1 m_2$ *und* $n_1 q_2 = q_1 n_2$. *Durch Multiplikation der ersten Gleichungen mit* $q_2$ *und der zweiten Gleichung mit* $m_2$ *erhalten wir*

$$m_1 n_2 q_2 = n_1 m_2 q_2 = q_1 n_2 m_2.$$

  *Da* $n_2$ *nicht Null ist, können wir in obiger Gleichung* $n_2$ *Wegstreichen (was eine der Eigenschaften von* $\mathbb{Z}$ *ist und* $\mathbb{Q}$ *nicht verwendet), womit sich* $m_1 q_2 = q_1 m_2$ *und damit* $(m_1, m_2) \sim (q_1, q_2)$ *ergibt.*

*Der Quotient* $(\mathbb{Z} \times (\mathbb{Z} \setminus \{0\}))/\sim$ *kann nun als Definition der rationalen Zahlen* $\mathbb{Q}$ *angesehen werden. Für eine Äquivalenzklasse* $[(m_1, m_2)]_\sim \in \mathbb{Q}$ *schreibt man wie üblich* $\frac{m_1}{m_2}$. *Damit gilt nun die Gleichung*

$$\frac{qm_1}{qm_2} = \frac{m_1}{m_2}$$

*für* $m_1 \in \mathbb{Z}$ *und* $q, m_2 \in \mathbb{Z} \setminus \{0\}$. *In der Tat bezeichnen nach Definition beide Seiten Äquivalenzklassen und die Gleichung gilt genau wenn*

$$(qm_1, qm_2) \sim (m_1, m_2)$$

*oder äquivalenterweise* $qm_1 m_2 = m_1 q m_2$ *erfüllt ist. Da letzteres gilt, erfüllen damit die so definierten rationalen Zahlen die üblichen Erweiterungs- und Kürzungsregeln.*

*Des Weiteren lässt sich* $\mathbb{Z}$ *als Teilmenge von* $\mathbb{Q}$ *auffassen. In der Tat ist die Abbildung*

$$\mathbb{Z} \to \mathbb{Q},\ m \mapsto \frac{m}{1}$$

*injektiv, denn die Gleichung* $\frac{m}{1} = \frac{n}{1}$ *ist für* $m, n \in \mathbb{Z}$ *per Definition genau dann erfüllt, wenn* $m = n$ *ist. Wir identifizieren* $\mathbb{Z}$ *mit dem Bild obiger Abbildung und schreiben insbesondere* $\frac{m}{1} = m$ *für* $m \in \mathbb{Z}$.

Wir wollen kurz zu allgemeinen Quotienten zurückkehren, also sei $X$ eine Menge und $\sim$ eine Äquivalenzrelation auf $X$. Häufig will man eine Funktion auf $X/\sim$ unter Verwendung der Elemente von $X$ (also der Repräsentanten der Äquivalenzklassen) definieren. Zum Beispiel möchten wir in obigem Beispiel in der Lage sein, zusätzliche Abbildungen auf $\mathbb{Q}$ zu definieren (unter anderem die Addition und die Multiplikation).

Konkreter, wenn $Y$ eine weitere Menge ist und $f : X \to Y$ eine Funktion ist, wollen wir möglicherweise durch

$$\bar{f} : X/_\sim \to Y,\ [x]_\sim \mapsto f(x)$$

eine Funktion definieren. Dies ist aber nur dann möglich, wenn $x_1 \sim x_2$ für $x_1, x_2 \in X$ (also $[x_1]_\sim = [x_2]_\sim$) auch $f(x_1) = f(x_2)$ impliziert. In diesem Fall ist $\bar{f}([x]_\sim)$ unabhängig von

der Wahl des Repräsentanten $x$ der Äquivalenzklasse $[x]_\sim$. Also definiert dies in der Tat eine Funktion $\bar{f}$, die jedem Element $[x]_\sim$ des Definitionsbereichs $X/\sim$ ein eindeutig bestimmtes Element $\bar{f}([x]_\sim)$ zuordnet. Zur Betonung dieser (für Funktionen notwendiger) Eigenschaft sagen wir in diesem Fall, dass $\bar{f}$ **wohldefiniert** ist.

**Übung 1.66** (Addition und Multiplikation auf den rationalen Zahlen)**.** *Wir definieren nun zusätzliche Strukturen auf $\mathbb{Q}$. Zeigen Sie, dass die Abbildungen*

$$+ : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}, \ \left(\frac{m}{n}, \frac{p}{q}\right) \mapsto \frac{m}{n} + \frac{p}{q} = \frac{mq + np}{nq}$$

*und*

$$\cdot : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}, \ \left(\frac{m}{n}, \frac{p}{q}\right) \mapsto \frac{m}{n} \cdot \frac{p}{q} = \frac{mp}{nq}$$

*wohldefiniert sind. Verifizieren Sie des Weiteren die Rechenregeln*

$$\frac{m}{n} \cdot \frac{n}{m} = 1, \quad \frac{a}{b} + \frac{-a}{b} = 0$$

*für $m, n, b \in \mathbb{Z} \setminus \{0\}$ und $a \in \mathbb{Z}$. Sie dürfen in dieser Aufgabe zwar alle üblichen Rechenregeln und Eigenschaften von $\mathbb{Z}$ verwenden, aber nicht jene von $\mathbb{Q}$ (da wir letztere ja definieren wollen).*

Folgendes Lemma beschreibt, unter welchen Bedingungen man eine Funktion aus Funktionen auf Teilmengen „zusammensetzen“ kann:

**Lemma 1.67** (Funktionen durch Fallunterscheidungen)**.** *Seien $X$ und $Y$ Mengen und sei $\mathcal{P}$ eine Partition von $X$. Angenommen es ist für jedes $P \in \mathcal{P}$ eine Funktion $f_P : P \to Y$ gegeben. Dann existiert eine eindeutige Funktion $f : X \to Y$, so dass $f|_P = f_P$ für jedes $P \in \mathcal{P}$ gilt.*

*Beweis.* Wir definieren die gewünschte Funktion $f : X \to Y$ wie folgt: Sei $x \in X$. Dann existiert genau ein $P \in \mathcal{P}$ mit $x \in P$, da $\mathcal{P}$ eine Partition ist. Wir setzen nun $f(x) = f_P(x)$ und haben also eine Funktion $f : X \to Y$ definiert. Diese erfüllt $f|_P = f_P$ nach Konstruktion. Zur Eindeutigkeit von $f$: Sei $g : X \to Y$ eine weitere Funktion mit $g|_P = f_P$. Sei $x \in X$ ein beliebiges Element und $P \in \mathcal{P}$ das Partitionselement mit $x \in P$. Dann gilt

$$g(x) = g|_P(x) = f_P(x) = f(x).$$

□

**Übung 1.68** (Funktionen durch Fallunterscheidungen)**.** *Zeigen Sie, dass die Bedingung in Lemma 1.67, dass $\mathcal{P}$ eine Partition ist, notwendig ist, durch Auffinden geeigneter Beispiele:*

- *Finden Sie Mengen $X, Y$, eine Kollektion $\mathcal{P}$ nicht paarweise disjunkter Teilmengen mit $X = \bigcup_{P \in \mathcal{P}} P$ und Funktionen $f_P : P \to Y$ für jedes $P \in \mathcal{P}$, so dass keine Funktion $f : X \to Y$ mit $f|_P = f_P$ für jedes $P \in \mathcal{P}$ existiert.*

- *Seien $X$ und $Y$ Mengen, $\mathcal{P}$ eine Kollektion paarweise disjunkter Teilmengen mit $X \neq \bigcup_{P \in \mathcal{P}} P$ und Funktionen $f_P : P \to Y$ für jedes $P \in \mathcal{P}$ gegeben. Zeigen Sie, dass mehrere Funktionen $f : X \to Y$ mit $f|_P = f_P$ für jedes $P \in \mathcal{P}$ existieren, falls $Y$ mehr als ein Element enthält.*

### 1.4.4 Mächtigkeit

Wir besprechen nun einen fundamentalen Begriff der Mengenlehre, den man sich intuitiv auch als Grössenvergleich vorstellen sollte:

**Definition 1.69** (Gleichmächtigkeit)**.** Zwei Mengen $X, Y$ sind **gleichmächtig**, geschrieben $X \sim Y$, falls es eine Bijektion $f : X \to Y$ gibt.

Der Begriff (die Relation) der Gleichmächtigkeit erfüllt auf der Klasse aller Mengen

- die Reflexivität, denn für jede Menge $X$ ist die Identitätsabbildung $\mathrm{id}_X : X \to X,\ x \mapsto x$ eine Bijektion,
- die Symmetrie, da für eine Bijektion $f : X \to Y$ die Umkehrabbildung $f^{-1}$ ebenso eine Bijektion ist (siehe auch Übung 1.44) und
- die Transitivität, da für zwei Bijektionen $f : X \to Y$ und $g : Y \to Z$ auch die Verknüpfung $g \circ f : X \to Z$ eine Bijektion ist (nach Lemma 1.42).

Also definiert die Gleichmächtigkeit eine Äquivalenzrelation auf der Klasse aller Mengen.

Diese entspricht eigentlich unserer alltäglichen Definition der Zahlen (genauer Kardinalzahlen im Gegensatz zu Ordinalzahlen). Beispielsweise gibt es in der Entwicklung eines Kindes einen Zeitpunkt, wo es erkennt, dass „zwei Brote“, „zwei Autos“ und „zwei Bälle“ etwas gemeinsam haben, was bei „drei Erwachsenen“ anders ist. Für grössere Zahlen verwenden wir anstatt dem Abzählen oft auch eine andere Methode um festzustellen, ob gleich viele Objekte einer Sorte wie Objekte einer anderen Sorte vorhanden sind. Zum Beispiel ist es einfach herauszufinden, ob man die richtige Anzahl Stühle bei einer Party bereitgestellt hat, in dem man die Gäste bittet, sich kurz alle gleichzeitig hinzusetzen. Die Bijektion in Definition 1.69 entspricht dem gleichen Zweck wie diese Bitte.

Für endliche Mengen $A = \{x_1, \ldots, x_m\}$ und $B = \{y_1, \ldots, y_n\}$ bestehend aus $m$ beziehungsweise $n$ verschiedenen Elementen (also soll $x_i \neq x_j$ für alle $1 \leq i < j \leq m$ und $y_k \neq y_\ell$ für alle $1 \leq k < \ell \leq n$ gelten) ist $A \sim B$ genau dann, wenn $m = n$ gilt. Diese Tatsache werden wir im Folgenden noch genauer besprechen.

**Definition 1.70** (Endliche Mengen)**.** Wir sagen, dass die Kardinalität der leeren Menge $\emptyset$ Null ist und schreiben $|\emptyset| = \#\emptyset = 0$. Sei nun $X$ eine beliebige Menge und $n \geq 1$ eine natürliche Zahl. Die Menge $X$ hat **Kardinalität** $n$, falls $X$ gleichmächtig zu $\{1, \ldots, n\}$ ist. In diesem Fall schreiben wir $|X| = \#X = n$. Des Weiteren ist $X$ eine **endliche Menge**, falls $|X| = n$ für eine nicht-negative ganze Zahl $n \in \mathbb{N}_0 = \{0, 1, 2, \ldots\}$. In diesem Fall sagen wir auch, dass $X$ **endliche Kardinalität** hat und wir schreiben $|X| < \infty$. Ist $X$ eine nicht-endliche Menge, so nennen wir $X$ eine **unendliche Menge** (oder sagen, dass $X$ **unendliche Kardinalität** hat) und schreiben $|X| = \infty$.

**Proposition 1.71** (Schubfachprinzip)**.** *Seien $m > n \geq 1$ natürliche Zahlen. Dann gibt es keine injektive Abbildung von $\{1, \ldots, m\}$ nach $\{1, \ldots, n\}$. Insbesondere sind zwei endliche Mengen $X$ und $Y$ genau dann gleichmächtig, wenn $|X| = |Y|$ ist.*

Wir bemerken, dass die Aussage des Schubfachprinzips sehr anschaulich ist. Angenommen man hat $n \geq 1$ Körbe und $m > n$ Äpfel, welche man in die Körbe legen will. Dann besagt das Schubfachprinzip, dass es immer einen Korb geben wird, in dem sich mindestens 2 Äpfel befinden werden und dies unabhängig davon, mit welcher Methode man die Äpfel in die Körbe verteilt. Ein Beispiel, wo man das Schubfachprinzips als Beweismethode verwenden kann, werden wir in Abschnitt 1.6.4 behandeln. Auch wenn die Aussage von Proposition 1.71 intuitiv klar ist, werden wir für einen rigorosen Beweis auf das Induktionsprinzip zurückgreifen.

*Beweis.* Wir beweisen zuerst mittels Induktion nach $n$, dass es keine injektive Abbildung von $\{1, \ldots, n+1\}$ nach $\{1, \ldots, n\}$ geben kann. Für $n = 1$ gibt es nur die (konstante) Abbildung $f : \{1, 2\} \to \{1\}$ gegeben durch $f(1) = 1$ und $f(2) = 1$, welche nicht injektiv ist. Dies beweist den Induktionsanfang.

Angenommen wir wissen bereits, dass es keine injektive Abbildung von $\{1, \ldots, n+1\}$ nach $\{1, \ldots, n\}$ gibt. Sei nun $f : \{1, \ldots, n+2\} \to \{1, \ldots, n+1\}$ eine beliebige Abbildung und sei $k = f(n+2)$. Falls es ein $j \in \{1, \ldots, n+1\}$ mit $f(j) = k = f(n+2)$ gibt, so ist $f$ nicht injektiv. Wir nehmen nun an, dass es ein derartiges $j$ nicht gibt.

Falls $k = n+1$ ist, so betrachten wir die eingeschränkte Funktion

$$g : \{1, \ldots, n+1\} \to \{1, \ldots, n\}\,,\ j \mapsto f(j).$$

Nach obiger Annahme, dass es kein $j \in \{1, \ldots, n+1\}$ mit $f(j) = n+1$ gibt, ist die Abbildung $g$ wohldefiniert. Nun zeigt aber die Induktionsvoraussetzung, dass $g$ nicht injektiv sein kann, also existieren $i \neq j$ in $\{1, \ldots, n+1\}$ mit $f(i) = g(i) = g(j) = f(j)$, wodurch auch $f$ nicht injektiv ist.

Falls $k < n+1$ ist, so verknüpfen wir $f$ mit der Vertauschung

$$\sigma : \{1, \ldots, n+1\} \to \{1, \ldots, n+1\}\,,\ j \mapsto \begin{cases} n+1 & \text{falls } j = k, \\ k & \text{falls } j = n+1, \\ j & \text{falls } j \neq k \text{ und } j \neq n+1. \end{cases}$$

und erhalten eine neue Funktion $F = \sigma \circ f$, welche nun $F(n+2) = \sigma(k) = n+1$ erfüllt. Auf Grund des oben behandelten Falles ist $F$ nicht injektiv. Da aber $\sigma$ bijektiv und $f = \sigma \circ F$ ist, ist auch $f$ nicht injektiv.

Dies schliesst den Induktionsschritt und damit den Induktionsbeweis ab. Wir haben also für eine natürliche Zahl $n \geq 1$ bewiesen, dass es keine injektive Abbildung von $\{1, \ldots, n+1\}$ nach $\{1, \ldots, n\}$ geben kann. Für $m > n \geq 1$ erhalten wir daraus, dass es keine injektive Abbildung von $\{1, \ldots, m\}$ nach $\{1, \ldots, n\}$ geben kann. (Wieso?)

Seien nun $X$ und $Y$ zwei endliche Mengen. Falls es eine bijektive Abbildung von $X$ nach $Y$ gibt, so existiert per Definition der Kardinalität endlicher Mengen auch eine bijektive Abbildung $f : \{1, \ldots, |X|\} \to \{1, \ldots, |Y|\}$. Auf Grund der Injektivität von $f$ gilt $|X| \leq |Y|$ und

auf Grund der Injektivität von $f^{-1}$ gilt $|Y| \leq |X|$. Also ist $|X| = |Y|$. Nimmt man umgekehrt an, dass $|X| = |Y|$ gilt, so folgt Gleichmächtigkeit von $X$ und $Y$ aus der Transitivität und der Definition der Kardinalität. □

**Übung 1.72** (Schubfachprinzip mittels surjektiven Funktionen)**.** *Seien $m > n \geq 1$ natürliche Zahlen. Zeigen Sie, dass es keine surjektive Abbildung von $\{1, \ldots, n\}$ nach $\{1, \ldots, m\}$ geben kann.*

**Übung 1.73.** *Sei $X$ eine endliche Menge und $f : X \to X$ eine Abbildung.*

*(i) Beweisen Sie, dass Injektivität, Surjektivität und Bijektivität von $f$ äquivalent sind.*

*(ii) Sei nun $f$ bijektiv. Zeigen Sie, dass es eine natürliche Zahl $n \geq 1$ gibt mit $f^{\circ n} = \mathrm{id}_X$.*

Der Begriff der Mächtigkeit ist aber vor allem für Mengen mit unendlich vielen Elementen interessant. Grund dafür ist, dass es „verschieden grosse“ unendliche Mengen gibt, wie das folgende Theorem von Cantor bestätigt. Ein Theorem ist eine mathematische Aussage grosser Wichtigkeit.

**Theorem 1.74** (Cantors Diagonalargument)**.** *Sei $X$ eine Menge. Dann ist $X$ nicht zu seiner Potenzmenge $\mathcal{P}(X)$ gleichmächtig. Insbesondere ist die Menge $\mathbb{N} = \{1, 2, 3, \ldots\}$ der natürlichen Zahlen nicht gleichmächtig zu $\mathcal{P}(\mathbb{N})$.*

**Übung 1.75** (Potenzmengen endlicher Mengen)**.** *Sei $n$ eine natürliche Zahl und sei $X$ eine Menge mit genau $n$ verschiedenen Elementen. Zeigen Sie, dass $|\mathcal{P}(X)| = 2^n$ gilt. Verwenden Sie dazu vollständige Induktion und schliessen Sie daraus Theorem 1.74 für den Fall, dass $X$ eine endliche Menge ist.*

**Übung 1.76.** *Sei $X$ eine Menge und sei $Y$ die Menge der Funktionen $X \to \{0, 1\}$. Zeigen Sie, dass $Y$ und $\mathcal{P}(X)$ gleichmächtig sind.*

*Beweis von Theorem 1.74 (per Widerspruch).* Wir nehmen an, dass es doch eine Bijektion $f : X \to \mathcal{P}(X)$ gibt und führen dies zu einem Widerspruch. Um diesen Widerspruch zu finden, definieren wir die Menge

$$A = \{x \in X \mid x \notin f(x)\}$$

aller Elemente $x$ in $X$, für die $x$ kein Element der Teilmenge $f(x) \subseteq X$ ist. Da nach Annahme $f : X \to \mathcal{P}(X)$ eine Bijektion ist und per Definition $A \in \mathcal{P}(X)$ ist, muss es ein $a \in X$ geben, so dass $A = f(a)$. Daraus folgt gemeinsam mit der Definition von $A$, dass

$$a \in A \iff a \notin f(a) \iff a \notin A$$

Dies ist aber absurd und daher existiert kein $a \in X$ mit $f(a) = A$. Dies ist ein Widerspruch zur Surjektivität von $f$; also kann es keine Bijektion $f : X \to \mathcal{P}(X)$ geben. □

Um zu sehen, inwiefern Theorem 1.74 ein Diagonalargument beinhaltet, laden wir Sie ein, folgende Übung zu lösen.

**Übung 1.77** (Die Diagonale in Cantors Diagonalargument)**.** *Betrachten Sie den Spezialfall $X = \mathbb{N}$ in Theorem 1.74 und nehmen Sie an, dass $\mathcal{P}(\mathbb{N}) \sim \mathbb{N}$, also $\mathcal{P}(\mathbb{N}) = \{A_1, A_2, A_3, \ldots\}$ für Teilmengen $A_n \subseteq \mathbb{N}$ für jede natürliche Zahl $n$. Für jedes solche $n$ lässt sich die Menge $A_n$ als die Folge in $0, 1$ auffassen, für welche die $m$-te Zahl genau dann $1$ ist, wenn $m \in A_n$ (vergleiche auch Übung 1.76 und Beispiel 1.30). Wir schreiben nun die Folgen in eine Tabelle der Art*

$$\begin{array}{cccccccc} A_1 & \leftrightarrow & \color{red}{0} & 1 & 0 & 1 & \ldots \\ A_2 & \leftrightarrow & 1 & \color{red}{0} & 0 & 0 & \ldots \\ A_3 & \leftrightarrow & 0 & 1 & \color{red}{1} & 1 & \ldots \\ A_4 & \leftrightarrow & 1 & 0 & 1 & \color{red}{1} & \ldots \\ & & \vdots & \vdots & \vdots & \vdots & \end{array}$$

*wobei die $n$-te Zeile der Menge $A_n$ entspricht. Dann lässt sich mit Hilfe der (in rot markierten) Diagonalen eine Folge konstruieren, deren zugehörige Menge nicht in der Liste $A_1, A_2, A_3, \ldots$ ist. Wie? Was hat das mit dem Beweis von Theorem 1.74 zu tun?*

**Übung 1.78** (Verschärfung von Cantors Diagonalargument)**.** *Formulieren Sie den Beweis von Theorem 1.74 so um, dass keine indirekte Annahme notwendig ist und zeigen Sie, dass eine Abbildung $f : X \to \mathcal{P}(X)$ nicht surjektiv sein kann.*

Da es aber für jede Menge $X$ eine injektive Abbildung $f : X \to \mathcal{P}(X)$ wie zum Beispiel $f : X \to \mathcal{P}(X),\ x \mapsto \{x\}$ gibt, drängt sich der Eindruck auf, dass $X$ gewissermassen „kleiner“ als $\mathcal{P}(X)$ ist.

**Definition 1.79** (Schmächtiger)**.** Seien $X$ und $Y$ zwei Mengen. Dann sagen wir, dass $X$ **schmächtiger** als (oder genau formuliert **höchstens so mächtig** wie) $Y$ ist und schreiben $X \lesssim Y$, falls es eine Injektion $f : X \to Y$ gibt. Wir sagen, dass $X$ **echt schmächtiger** (oder **weniger mächtig**) als $Y$ ist, falls $X$ schmächtiger als $Y$ ist und $Y$ nicht schmächtiger als $X$ ist ($X \lesssim Y$ und $Y \not\lesssim X$).

Eine Menge $X$ ist schmächtiger als eine Menge $Y$ genau dann, wenn $X$ gleichmächtig zu einer Teilmenge von $Y$ ist. Dies begründet die Terminologie.

**Übung 1.80** (Schmächtiger als Relation)**.** *Untersuchen Sie die Relation $\lesssim$ auf der Klasse aller Mengen auf ihre Eigenschaften (siehe auch Definition 1.55).*

Der oben eingeführte Begriff hat einen interessanten Zusammenhang zur Gleichmächtigkeit:

**Theorem 1.81** (Cantor, Schröder, Bernstein)**.** *Seien $X$ und $Y$ Mengen, so dass $X \lesssim Y$ und $Y \lesssim X$. Dann gilt $X \sim Y$.*

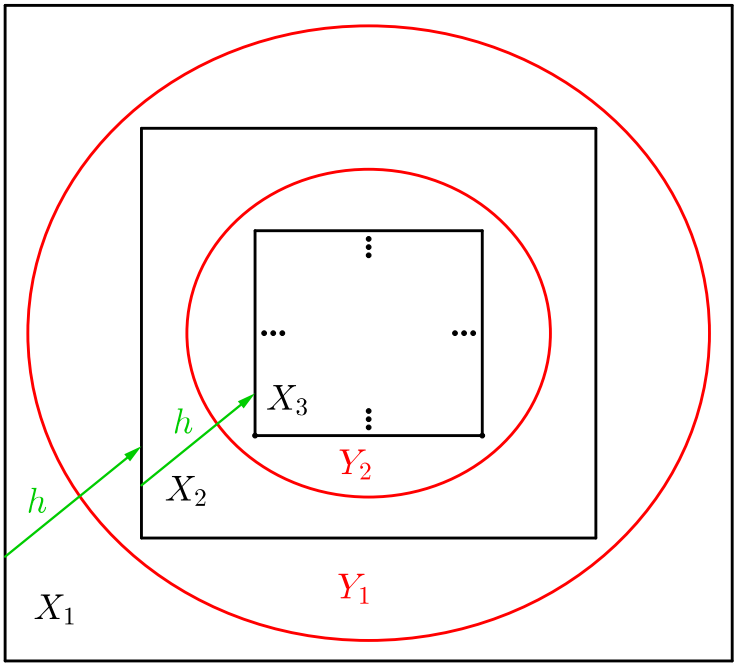
Das Theorem wurde 1887 von Cantor formuliert und von Dedekind unter Verwendung eines zusätzlichen Axioms im selben Jahr bewiesen. Dem 19-jährigen Studenten Bernstein gelang es schliesslich einen Beweis der Aussage zu liefern ohne jenes Axiom (das Auswahlaxiom) zu benutzen.

*Beweis von Theorem 1.81.* Seien $X, Y$ zwei Mengen mit $X \lesssim Y$ und $Y \lesssim X$ wie in Theorem 1.81. Dann gibt es injektive Funktionen $f : X \to Y$ und $g : Y \to X$. Wir definieren

$$X_1 = X, \quad Y_1 = g(Y), \quad X_2 = g \circ f(X).$$

Dann gelten die Inklusionen $X_1 \supseteq Y_1 \supseteq X_2$ und die Funktion $h : x \in X_1 \mapsto g \circ f(x) \in X_2$ ist eine Bijektion (mit Definitionsbereich $X_1$ und Wertebereich $X_2$, was für $g \circ f$ nicht zutrifft). Wir behaupten nun, dass es eine Bijektion $F : X_1 \to Y_1$ gibt. Für unsere Mengen $X, Y$ impliziert dies das Theorem, denn es gilt dann $X = X_1 \sim Y_1 = g(Y) \sim Y$, wobei die zweite Gleichmächtigkeit wegen der Bijektivität der Funktion $g(y) \in g(Y) \mapsto y \in Y$ gilt.

Zum Beweis der Behauptung nehmen wir also an, dass $X_1, Y_1, X_2$ drei Mengen sind mit $X_1 \supseteq Y_1 \supseteq X_2$ und dass es eine Bijektion $h : X_1 \to X_2$ gibt. Wir wollen daraus schliessen, dass $X_1 \simeq Y_1$ gilt.




Figur 1.13: Obwohl logisch gesehen unnötig, ist es hilfreich, sich den Beweis von Theorem 1.81 mit Hilfe dieses Bildes zu veranschaulichen und Parallelen zu Hilberts Hotel zu suchen.

Wir definieren $Y_2 = h(Y_1)$, $X_3 = h(X_2) = h^{\circ 2}(X_1)$ und allgemeiner für eine natürliche Zahl $n$

$$Y_{n+1} = h^{\circ n}(Y_1), \quad X_{n+1} = h^{\circ n}(X_1).$$

Wir zeigen nun mittels vollständiger Induktion, dass

$$X_1 \supseteq Y_1 \supseteq X_2 \supseteq Y_2 \supseteq \dots \tag{1.6}$$

gilt. Für $X_1 \supseteq Y_1 \supseteq X_2$ ist dies Voraussetzung; der Induktionsanfang ist also erledigt. Für den Induktionsschritt nehmen wir an, dass $X_1 \supseteq Y_1 \supseteq \ldots \supseteq Y_{n-1} \supseteq X_n$ für $n \geq 2$ schon bewiesen ist. Wenden wir auf $X_{n-1} \supseteq Y_{n-1} \supseteq X_n$ die Abbildung $h$ an, so erhalten wir

$$h(X_{n-1}) = X_n \supseteq h(Y_{n-1}) = Y_n \supseteq h(X_n) = X_{n+1}$$

womit Gleichung (1.6) mittels vollständiger Induktion gezeigt ist.

Wir definieren des Weiteren die Mengen

$$B_n = X_n \setminus Y_n, \quad K_n = Y_n \setminus X_{n+1}$$

für jede natürliche Zahl $n$, wobei „$B$“ für „bewegt“ und „$K$“ für „konstant“ steht. Für zwei natürliche Zahlen $m, n$ mit $1 \leq m < n$ gilt wegen (1.6), dass

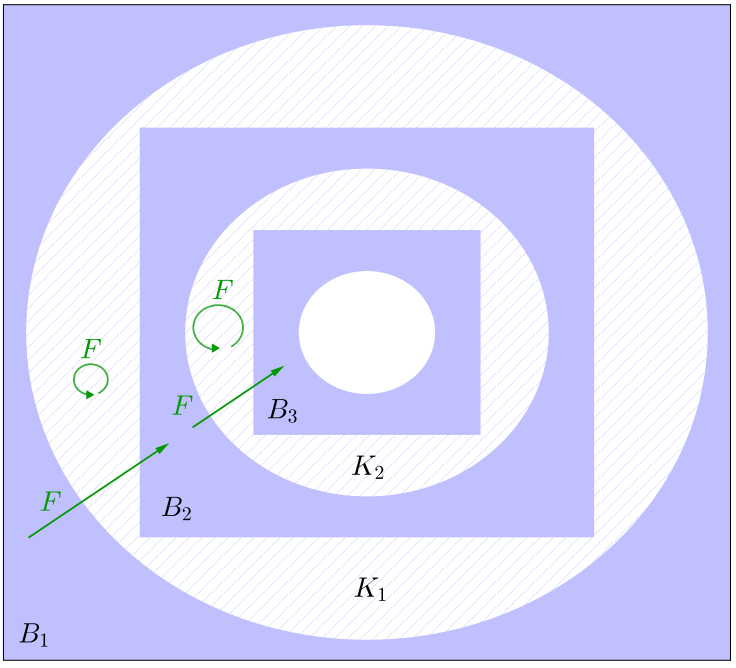
$$X_m \supseteq Y_m \supseteq X_{m+1} \supseteq X_n \supseteq Y_n.$$

Daher ist $B_m = X_m \setminus Y_m$ disjunkt zu $X_n$ und insbesondere auch disjunkt zu $B_n = X_n \setminus Y_n$ und $K_n = Y_n \setminus X_{n+1}$. Genauso zeigt man, dass $K_m = Y_m \setminus X_{m+1}$ disjunkt zu $B_n$ und $K_n$ ist.

Da $h$ injektiv ist, gilt

$$h(B_n) = h(X_n \setminus Y_n) = h(X_n) \setminus h(Y_n) = X_{n+1} \setminus Y_{n+1} = B_{n+1} \tag{1.7}$$

für alle natürlichen Zahlen $n$ nach Übung 1.46. (Stellen Sie sicher, dass Sie verstehen, wo und wie die Injektivität von $h$ hier verwendet wurde.)




Figur 1.14: Illustration der Abbildung $F$ und der Mengen $B_k$ für $k \in \mathbb{N}$.

Schlussendlich definieren wir

$$K_\infty = X_1 \setminus \bigcup_{n=1}^{\infty} (B_n \cup K_n),$$

womit $B_1, B_2, \ldots$ und $K_1, K_2, \ldots$ gemeinsam mit $K_\infty$ eine Partition $\mathcal{P}$ von $X_1$ definieren. Mittels Fallunterscheidung definieren wir nun eine Funktion

$$F : X_1 = K_\infty \sqcup \bigsqcup_{n=1}^{\infty} K_n \sqcup \bigsqcup_{n=1}^{\infty} B_n \to Y_1 = K_\infty \sqcup \bigsqcup_{n=1}^{\infty} K_n \sqcup \bigsqcup_{n=2}^{\infty} B_n$$

durch

$$F(x) = \begin{cases} h(x) & \text{falls } x \in B_n \text{ für ein } n \in \mathbb{N}, \\ x & \text{falls } x \in K_n \text{ für ein } n \in \mathbb{N} \text{ oder } x \in K_\infty \end{cases}$$

für alle $x \in X_1$. Es bleibt noch zu zeigen, dass $F$ bijektiv ist.

Zur Surjektivität: Per Definition ist $K_\infty = F(K_\infty) \subseteq F(X_1)$ und $K_n = F(K_n) \subseteq F(X_1)$ für alle $n$. Weiter gilt nach Gleichung (1.7) auch $B_{n+1} = h(B_n) = F(B_n) \subseteq F(X_1)$ für alle $n$ und damit ist ebenso $Y_1 = K_1 \cup K_2 \cup \ldots \cup K_\infty \cup B_2 \cup B_3 \cup B_4 \cup \ldots \subseteq F(X_1)$. Wir sehen also, dass $F$ surjektiv ist.

Zur Injektivität: Seien $x, y \in X_1$ mit $F(x) = F(y)$. Da in der Liste $F(K_1)$, $F(K_2)$, $\ldots$, $F(K_\infty)$, $F(B_1)$, $F(B_2)$, $\ldots$ die Mengen paarweise disjunkt sind, müssen $x$ und $y$ in der gleichen Menge der Partition $\mathcal{P}$ liegen. Falls jene Menge $K_\infty$ oder $K_n$ für ein $n$ ist, so gilt $x = y$. Ansonsten ist $h(x) = F(x) = F(y) = h(y)$ und die Gleichheit $x = y$ folgt aus der Injektivität von $h$. □

**Übung 1.82** (Verbindung zu Hilberts Hotel)**.** *Erklären Sie intuitiv, wo die Parallelen in obigem Beweis von Theorem 1.81 zu Hilberts Hotel sind. Fassen Sie dazu den Beweis in einigen wenigen Sätzen anhand des Bildes 1.14 oder dem folgenden Applet zusammen.*

**Applet 1.83** (Beweis des Satzes von Cantor-Schröder-Bernstein)**.** *Im Beweis von Theorem 1.81 wird die Menge $X$ in zwei Typen von Teilmengen zerlegt. Die konstruierte Funktion wird angedeutet, indem Bild (und Urbilder) von einem bewegbaren Punkt $x \in X$ eingezeichnet werden.*

Wie schon erwähnt wurde, ist der Begriff der Gleichmächtigkeit eine Äquivalenzrelation auf der Klasse der Mengen.

**Definition 1.84** (Kardinalität einer Menge)**.** Die zu einer Menge $X$ gehörenden Äquivalenzklasse bezüglich Gleichmächtigkeit wird die **Kardinalität** der Menge $X$ genannt und als $|X|$ geschrieben. Falls $X$ endlich ist und genau $n$ verschiedene Elemente hat, dann schreiben wir $|X| = n$. Eine Menge heisst **abzählbar unendlich**, falls sie die Kardinalität $|\mathbb{N}|$ hat oder in anderen Worten gleichmächtig zu $\mathbb{N}$ ist. Die Kardinalität von $\mathbb{N}$ wird auch $\aleph_0$, gesprochen **Aleph-**0, genannt.

**Übung 1.85** (Gitterpunkte in der Ebene)**.** *Zeigen Sie, dass* $\mathbb{N} \times \mathbb{N}$ *abzählbar unendlich ist. Sie können hierzu Theorem 1.81 verwenden: Eine Injektion* $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$ *finden Sie unter Verwendung der Tatsache, dass unendlich viele Primzahlen existieren. Alternativ lässt sich mit Hilfe eines Bildes tatsächlich eine explizite Bijektion* $\mathbb{N} \to \mathbb{N} \times \mathbb{N}$ *finden, wie? Zeigen Sie auch, dass das Produkt* $\mathbb{Z} \times \mathbb{Z}$ *abzählbar unendlich ist.*

Wie bereits erwähnt, kann man mit Hilfe des Begriffes der Schmächtigkeit Grössenvergleiche zwischen Kardinalitäten anstellen. Für uns ist aber abgesehen vom endlichen und abzählbar unendlichen Fall nur folgender Spezialfall interessant:

**Definition 1.86** (Das Kontinuum)**.** Eine Menge $X$ heisst **überabzählbar**, falls $\mathbb{N}$ schmächtiger ist als $X$, aber $\mathbb{N}$ nicht gleichmächtig zu $X$ ist. Nach Theorem 1.74 ist $\mathcal{P}(\mathbb{N})$ überabzählbar. Die Kardinalität von $\mathcal{P}(\mathbb{N})$ wird auch mit $\mathfrak{c}$ bezeichnet und das **Kontinuum** genannt.

**Übung 1.87** (Trichotomie der Kardinalitäten)**.**

*(i) Zeigen Sie, dass jede Menge entweder endlich, abzählbar unendlich oder überabzählbar ist. Gehen Sie dazu davon aus, dass* $X$ *eine nicht-endliche Menge ist und finden Sie eine injektive Abbildung* $\mathbb{N} \to X$*.*

*(ii) Zeigen Sie, dass es unendlich viele überabzählbare Kardinalitäten gibt.*

*Bemerkung.* Im Beweis von Übung 1.87 werden Sie wahrscheinlich abzählbar oft eine Wahl treffen müssen. Formal verwendet dies das sogenannte Auswahlaxiom der Mengenlehre (Axiom of Choice), welches wir implizit erlauben werden.

VO: 28.09.2017

## 1.5 Zahlenmengen

*„Die natürlichen Zahlen hat der liebe Gott gemacht,*
*alles andere ist Menschenwerk.“*
leicht adaptiert nach Kronecker (1823-1891)

In diesem Abschnitt diskutieren wir wahrscheinlich schon bekannte Zahlenmengen. Wir gehen zuerst auf die natürlichen Zahlen ein. Formal sind die natürlichen Zahlen durch folgendes (wirklich minimales) Axiomensystem definiert.

**Peano Axiome.** *Die* ***natürlichen Zahlen*** $\mathbb{N}$ *sind (in einem gewissen Sinne eindeutig) durch die folgenden Eigenschaften charakterisiert:*

*(i) Es existiert ein ausgezeichnetes Element* $1 \in \mathbb{N}$ *und eine injektive Abbildung* $\nu : \mathbb{N} \to \mathbb{N}$*, auch Nachfolgerfunktion genannt, so dass* $1 \notin \nu(\mathbb{N})$*.*

*(ii)* $\mathbb{N}$ *erfüllt das Induktionsaxiom: Ist* $A$ *eine Teilmenge von* $\mathbb{N}$*, die* $1$ *enthält und für alle* $n \in \mathbb{N}$ *die Eigenschaft „* $n \in A \implies \nu(n) \in A$*“ erfüllt, dann gilt* $A = \mathbb{N}$*.*

Die Nachfolgerfunktion $\nu$ sollte man sich als die Abbildung $n \in \mathbb{N} \mapsto n + 1 \in \mathbb{N}$ vorstellen, nur kennt man die Addition auf $\mathbb{N}$ noch nicht. Neben der Addition zweier natürlichen Zahlen kann man sich bei allen anderen Eigenschaften der natürlichen Zahlen, die man intuitiv kennt, fragen, wie sie zu definieren oder zu beweisen sind. Für eine detailliertere Diskussion dieser Frage verweisen wir auf [AE06]; wir möchten jedoch an einem elementaren Beispiel demonstrieren, wie man das Induktionsaxiom verwenden kann.

**Behauptung** (Nachfolgerzahlen)**.** *Sei* $n \in \mathbb{N}$ *eine natürliche Zahl verschieden von* $1$*. Dann ist* $n \in \nu(\mathbb{N})$*.*

*Beweis.* Wir betrachten die Menge $A = \nu(\mathbb{N}) \cup \{1\}$. Dann gilt $1 \in A$ per Definition und für $n \in A \subseteq \mathbb{N}$ gilt $\nu(n) \in A$ wieder per Definition von $A$. Nach dem Induktionsaxiom ist also $A = \mathbb{N}$ und die Behauptung folgt. □

Man kann ausgehend von den Peano-Axiomen sowohl Addition und Multiplikation mittels vollständiger Induktion (oder äquivalenterweise mittels Rekursion) definieren. Wir werden im Folgenden wieder von diesem Abstraktionslevel Abstand nehmen und davon ausgehen, dass wir die üblichen Eigenschaften der natürlichen Zahlen inklusive Addition und Multiplikation kennen und dass $\mathbb{N} = \{1, 2, 3, 4, \ldots\}$ (unter Verwendung der Axiome definiert man 2 als $\nu(1)$, 3 als $\nu(2)$ und so weiter). Dass die natürlichen Zahlen wirklich „natürlich“ sind, kann man sicherlich auf viele Arten belegen. Wie bereits besprochen, repräsentieren sie genau die Kardinalitäten aller endlichen nicht-leeren Mengen, was für die Menschheit schon lange von Bedeutung war.

Manchmal wollen wir vielleicht Abzählungen bei Null beginnen lassen und definieren deswegen (vergleiche Beispiel 1.88) unter Verwendung eines zusätzlichen Elements 0 die Menge der **nicht-negativen ganzen Zahlen** als

$$\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\} = \{0\} \sqcup \mathbb{N}.$$

Die Zahl 0 soll natürlich die üblichen Eigenschaften bezüglich Addition und Multiplikation erfüllen.[3] Die **ganzen Zahlen** sind durch

$$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\} = \mathbb{N} \sqcup \{0\} \sqcup \{-n \mid n \in \mathbb{N}\}$$

definiert. Die **rationalen Zahlen** sind gegeben durch

$$\mathbb{Q} = \left\{ \frac{m}{n} \mid m \in \mathbb{Z},\ n \in \mathbb{N} \right\}$$

und die reellen Zahlen werden mit $\mathbb{R}$ bezeichnet und enthalten $\mathbb{Q}$ und viele weitere Zahlen wie zum Beispiel $\sqrt{2}$, $\pi$, e, ....[4] Häufig stellt man sich $\mathbb{R}$ als die Zahlengerade vor. Wir werden in Kapitel 2 die Menge der reellen Zahlen $\mathbb{R}$ axiomatisch einführen und dann die Zahlenmengen $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ als Teilmengen von $\mathbb{R}$ nochmals ausführlich definieren.

*Bemerkung.* Eigentlich lautet das Zitat (siehe Seite 15 in [Web92]) von Kronecker zu Beginn dieses Abschnitts

*„Die ganzen Zahlen hat der liebe Gott gemacht, alles andere ist Menschenwerk.“*

Es kann jedoch sein, dass Kronecker eigentlich die natürlichen Zahlen gemeint hat. Auf jeden Fall wollen wir hier den Mönch und Historiker William of Malmesbury (ca. 1095-1143), der die ganze Zahl 0 als „dangerous Saracen magic“ bezeichnet hat, als grössere Autorität in dieser Frage ansehen. Scherz beiseite, die Menschheit hat in der Tat sehr lange gebraucht, um mit der Null und den negativen Zahlen zurecht zu kommen. Für einen geschichtlichen Exkurs zur „Zahl 0“ verweisen wir auf diesen Podcast der BBC, und zum Thema „Negative Zahlen“ auf die ersten 10-20 Minuten eines weiteren Podcasts der BBC.

### 1.5.1 Konstruktion der ganzen Zahlen aus den natürlichen Zahlen

In diesem Abschnitt möchten wir die ganzen Zahlen formaler mittels einer Äquivalenzrelation auf Paaren von natürlichen Zahlen konstruieren (in Analogie zu Beispiel 1.65).

**Beispiel 1.88** (Konstruktion der Menge der ganzen Zahlen)**.** *Wir nehmen an, dass wir bereits die Menge der nicht-negativen ganzen Zahlen* $\mathbb{N}_0 = \{0, 1, 2, 3, 4, \ldots\}$ *und die Addition auf* $\mathbb{N}_0$ *mit allen üblichen Eigenschaften kennen und wollen daraus die ganzen Zahlen definieren. Dazu betrachten wir eine Äquivalenzrelation auf* $\mathbb{N}_0^2$*: für* $(m_1, m_2), (n_1, n_2) \in \mathbb{N}_0^2$ *definieren wir*

$$(m_1, m_2) \sim (n_1, n_2) \iff m_1 + n_2 = n_1 + m_2 \tag{1.8}$$

*Motiviert ist das ganze dadurch, dass man eine ganze Zahl als Differenz von zwei nicht-negativen ganzen Zahlen auffassen kann. Diese sind aber nicht eindeutig gegeben; deswegen führt man auf Tupeln von natürlichen Zahlen obige Äquivalenzrelation ein. In der Tat hat man bei Definition* (1.8) *eigentlich* $(m_1, m_2) \sim (n_1, n_2) \iff m_1 - m_2 = n_1 - n_2$ *im Hinterkopf, darf dies aber formal nicht verwenden, da die Subtraktion (noch) nicht erlaubt ist. Wenn*

[3] Dies stellt eine Definition dar und muss nicht bewiesen werden.
[4] Wir werden im Laufe des ersten Semesters all diese Zahlen definieren.

*wir über* $[(m_1, m_2)]_\sim$ *sprechen, werden wir leicht schizophren an* $m_1 - m_2$ *denken, dies aber nicht verwenden.*

*Informell können wir uns das Tupel* $(m_1, m_2)$ *als einen Vektor mit Anfangspunkt* $m_1 \in \mathbb{N}_0$ *und Endpunkt* $m_2 \in \mathbb{N}_0$ *vorstellen, womit die Äquivalenzklassen allen Vektoren mit gleicher Länge und Richtung entspricht. Alternativ könnte* $m_1$ *der Kontostand vor und* $m_2$ *der Kontostand nach einer Transaktion darstellen und die Äquivalenzklasse entspricht dann dem Nettogewinn/-verlust.*

*Wir verifizieren nun, dass es sich bei* $\sim$ *effektiv um eine Äquivalenzrelation handelt. Es gilt für alle* $(m_1, m_2), (n_1, n_2), (q_1, q_2) \in \mathbb{N}_0^2$

- *Reflexivität:* $(m_1, m_2) \sim (m_1, m_2)$*, denn* $m_1 + m_2 = m_1 + m_2$*.*
- *Symmetrie:* $(m_1, m_2) \sim (n_1, n_2)$ *denn* $m_1 + n_2 = n_1 + m_2$ *impliziert* $n_1 + m_2 = m_1 + n_2$ *und daher auch* $(n_1, n_2) \sim (m_1, m_2)$*.*
- *Transitivität:* $(m_1, m_2) \sim (n_1, n_2)$ *und* $(n_1, n_2) \sim (q_1, q_2)$ *impliziert* $m_1 + n_2 = n_1 + m_2$ *und* $n_1 + q_2 = q_1 + n_2$*. Durch Summieren dieser Gleichungen ergibt sich*

  $$m_1 + n_2 + n_1 + q_2 = n_1 + m_2 + q_1 + n_2$$

  *und durch Wegstreichen von* $n_1 + n_2$ *(was eine der Eigenschaften von* $\mathbb{N}_0$ *ist und nicht die Subtraktion verwendet) erhalten wir* $m_1 + q_2 = q_1 + m_2$*, also* $(m_1, m_2) \sim (q_1, q_2)$

*Der Quotient* $\mathbb{N}_0^2/\sim$ *kann als Definition von* $\mathbb{Z}$ *angesehen werden, wobei wir die Äquivalenzklasse auch als* $[(m_1, m_2)]_\sim = m_2 - m_1$ *schreiben. Insbesondere identifizieren wir* $n \in \mathbb{N}_0$ *mit* $[(n, 0)]_\sim$ *und schreiben* $[(0, n)]_\sim$ *auch als* $0 - n = -n$ *für* $n \in \mathbb{N}$*. Dies macht nach folgender Übung Sinn.*

Genau wie in Beispiel 1.65 und der nachfolgenden Übung 1.66 lässt sich auf $\mathbb{Z}$ mit Hilfe der gegebenen Eigenschaften von $\mathbb{N}_0$ weitere Strukturen definieren.

**Übung 1.89** (Eine Partition der ganzen Zahlen)**.**

*(i) Zeigen Sie, dass die Abbildungen*

$$\iota_+ : n \in \mathbb{N}_0 \mapsto [(n, 0)]_\sim \in \mathbb{Z}$$
$$\iota_- : n \in \mathbb{N} \mapsto -n = [(0, n)]_\sim \in \mathbb{Z}$$

*injektiv sind und disjunkte Bilder* $\iota_+(\mathbb{N}_0), \iota_-(\mathbb{N})$ *haben, welche wir mit* $\mathbb{N}_0 = \iota_+(\mathbb{N}_0)$ *und* $-\mathbb{N} = \iota_-(\mathbb{N})$ *bezeichnen werden.*

*(ii) Zeigen Sie, dass* $\mathbb{Z} = \mathbb{N}_0^2/\sim = \iota_+(\mathbb{N}_0) \sqcup \iota_-(\mathbb{N})$ *gilt.*

**Übung 1.90** (Addition von ganzen Zahlen)**.**

*(i) Zeigen Sie, dass die Funktion* $f_{neg} : \mathbb{Z} = \mathbb{N}_0^2/\sim \to \mathbb{Z}$ *definiert durch*

$$f_{neg}([(m_1, m_2)]_\sim) = [(m_2, m_1)]_\sim$$

*für* $(m_1, m_2) \in \mathbb{N}_0^2$ *wohldefiniert und bijektiv ist. Was ist das Bild von* $\iota_+(\mathbb{N}_0)$ *und von* $\iota_-(\mathbb{N})$ *unter dieser Abbildung?*

*(ii) Zeigen Sie, dass die Funktion*

$$f_{add} : \mathbb{Z}^2 \to \mathbb{Z}, \quad ([(m_1, m_2)]_\sim, [(n_1, n_2)]_\sim) \mapsto [(m_1 + n_1, m_2 + n_2)]_\sim$$

*wohldefiniert ist. Verifizieren Sie, dass für jedes* $z \in \mathbb{Z}$ *gilt* $f_{add}(z, f_{neg}(z)) = [(0, 0)]_\sim$.

*(iii) Finden und definieren Sie weitere natürliche Abbildungen* $\mathbb{Z} \to \mathbb{Z}$ *oder* $\mathbb{Z}^2 \to \mathbb{Z}$.

Wir haben also die Menge der ganzen Zahlen $\mathbb{Z}$ aus den natürlichen Zahlen $\mathbb{N}$ und die Menge der rationalen Zahlen $\mathbb{Q}$ aus den ganzen Zahlen konstruiert, auf der wiederum die üblichen Rechenregeln gelten. Des Weiteren kann man aus $\mathbb{Q}$ die reellen Zahlen $\mathbb{R}$ konstruieren. Wir wollen dies jedoch erst im zweiten Semester durchführen und stattdessen die reellen Zahlen mittels ihren Axiomen im nächsten Kapitel einführen.

### 1.5.2 Teilbarkeit und Kongruenzen

Nachdem wir oben die Zahlenmengen $\mathbb{N}, \mathbb{N}_0, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ eingeführt haben, möchten wir in diesem Abschnitt die etwas verschiedene, „endlichen Zahlenmengen“ $\mathbb{Z}/q\mathbb{Z}$ zu $q \in \mathbb{N}$ einführen.

**Definition 1.91** (Teilbarkeit und Kongruenz)**.** Seien $a, b \in \mathbb{Z}$. Wir sagen, dass $b$ durch $a$ **teilbar** ist, und schreiben $a|b$, falls ein $k \in \mathbb{Z}$ existiert mit $ak = b$. Des Weiteren sind $a$ und $b$ **kongruent** modulo $q \in \mathbb{N}$, falls $b - a$ durch $q$ teilbar ist. In diesem Fall schreiben wir

$$a \equiv b \mod q.$$

Tatsächlich definiert Kongruenz modulo $q$ eine Äquivalenzrelation $\equiv$ (siehe Übung 1.93). Den Quotienten $\mathbb{Z}/\equiv$ bezeichnet man meist als $\mathbb{Z}/q\mathbb{Z}$ und die Äquivalenzklasse $[a]_\sim \in \mathbb{Z}/q\mathbb{Z}$ ist für $a \in \mathbb{Z}$ durch $a + q\mathbb{Z}$ gegeben. Genau wie die Zahlenmengen, die wir bereits kennen, verfügt die Menge $\mathbb{Z}/q\mathbb{Z}$ über zusätzliche Struktur wie Addition und Multiplikation.

**Applet 1.92** (Darstellung des Quotienten modulo Kongruenz)**.** *Wir stellen in diesem Applet den Quotienten* $\mathbb{Z}/q\mathbb{Z}$ *(für verschiedene Werte von* $q$*) dar. Es macht Sinn sich die Punkte* $\mathbb{Z}/q\mathbb{Z}$ *entlang eines Kreises vorzustellen, doch hat dies formal (vorerst) keine Bedeutung.*

**Übung 1.93** (Eigenschaften der Kongruenz)**.**

*(i) Zeigen Sie, dass Kongruenz modulo* $q$ *eine Äquivalenzrelation* $\equiv$ *auf* $\mathbb{Z}$ *darstellt.*

*(ii) Zeigen Sie, dass die Abbildungen*

$$(a + q\mathbb{Z}, b + q\mathbb{Z}) \in \left(\mathbb{Z}/q\mathbb{Z}\right)^2 \mapsto (a + b) + q\mathbb{Z} \in \mathbb{Z}/q\mathbb{Z}$$
$$(a + q\mathbb{Z}, b + q\mathbb{Z}) \in \left(\mathbb{Z}/q\mathbb{Z}\right)^2 \mapsto (a \cdot b) + q\mathbb{Z} \in \mathbb{Z}/q\mathbb{Z}$$

*wohldefiniert sind (man vergleiche auch mit Übung 1.90).*

Wie wir später sehen werden, ist $\mathbb{Z}/q\mathbb{Z}$ tatsächlich eine endliche Menge (mit Kardinalität $q$), auf der man ebenso gut rechnen kann, wie beispielsweise auf $\mathbb{Z}$. Grund dafür ist die Division mit Rest, die wir in Kapitel 2 im Detail besprechen werden. Unter anderem wegen dieser Endlichkeit ist $\mathbb{Z}/q\mathbb{Z}$ in vielen Gebieten der Mathematik, aber auch beispielsweise in der Informatik von grosser Nützlichkeit.

## 1.6 Beweise

VO: Abschnitte 1.6 und 1.7 sind zum Lesen gedacht.

Die Mathematik hat als eine von wenigen Wissenschaften die Eigenschaft, dass alle Resultate der Vergangenheit (abgesehen von menschlichen Fehlern und gewissen Perioden, in denen wieder nach der korrekten Methode gesucht wurde) nach wie vor richtig sind (und nicht bloss in erster Näherung). Das Erfolgsrezept dafür liegt im Aufbau der Mathematik. Das Fundament bilden die Axiome (beispielsweise jene der Mengenlehre oder die Peano-Axiome), die als bekannt und richtig angenommen werden. Mittels Definitionen werden neue Begriffe eingeführt und mittels Lemmata und Propositionen untersucht und in Verbindung zueinander gebracht, bis aus diesen Sätze und Theoreme bewiesen werden können. Selbstverständlich kann man bezweifeln, ob die so entstandenen Theorien interessant oder relevant sind. Über Interessen lässt sich natürlich streiten. Dennoch ist die Relevanz der heutigen mathematischen Theorien in vielen anderen Bereichen schwer zu bestreiten und dies obwohl wir in der Mathematik beispielsweise die 0 und die Zahl $10^{-10^{80}}$ (mit so vielen Nullen nach dem Komma wie nach derzeitigen Schätzungen Atome im Weltall) strikt unterscheiden. Zu diesem Thema möchten wir dieses Video, basierend auf dem lesenswerten Artikel „The unreasonable effectiveness of mathematics in the natural sciences“ von Eugene Wigner, empfehlen. In eine ähnliche Richtung geht das für ein allgemeines Publikum gedachte Video, das die (teilweise überraschende) Nützlichkeit der Mathematik in anderen Naturwissenschaften besonders hervorhebt.

Wir möchten an dieser Stelle einige Bemerkungen zu Beweisen machen. Ein Beweis, wie der Name schon sagt, soll die behauptete Aussage unter Verwendung von einfachen logischen Schritten aus vorher Bekanntem (zum Beispiel Axiomen) ableiten. Es gibt dafür einige Methoden, die wir zum Teil schon angewendet haben und die unter anderem auch verknüpft werden können.

Manche von Ihnen werden sich fragen, warum wir in dieser Vorlesung immer „diese Beweise“ besprechen müssen. Diese Frage lässt sich vielleicht mit der Frage an einen Physik-Professor vergleichen, wieso denn immer so viele Experimente durchgeführt werden müssen. Beweise sind im Aufbau der Mathematik unersetzlich. Selbst wenn Sie die hier entwickelte Theorie später möglicherweise nicht in der hier verwendeten Exaktheit und Genauigkeit benötigen, werden Sie bei aktiver Mitarbeit abstraktes Denken erlernen, welches für die Mathematik aber auch für andere Wissenschaften und in der Praxis äusserst nützlich sein wird.

### 1.6.1 Widerspruchsbeweise

Angenommen wir wollen eine Aussage $A$ beweisen. Dann ist es manchmal einfacher zu zeigen, dass $\neg A$ nicht sein kann, also zu einem Widerspruch führt, als direkt $A$ zu zeigen. Wir haben dafür schon ein Beispiel im Beweis von Theorem 1.74 gesehen, wo wir die Nicht-Existenz eines gewissen Objekts (dort einer Abbildung) behauptet haben. Wir nennen $\neg A$ die indirekte Annahme und den Beweis einen Widerspruchsbeweis. Man spricht auch vom „Satz vom ausgeschlossenen Dritten“, da entweder $A$ oder $\neg A$ gelten muss und es keine dritte

Möglichkeit gibt.[5]

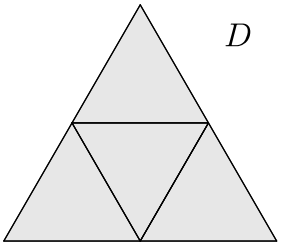
### 1.6.2 Kontraposition

Sind $A, B$ zwei Aussagen, dann sind die Aussagen $A \implies B$ und $\neg B \implies \neg A$ äquivalent. Die Aussage $\neg B \implies \neg A$ wird die Kontraposition der Aussage $A \implies B$ genannt. Wie man sich dies in einem Beweis zunutze machen kann, wollen wir in einem Beispiel illustrieren.

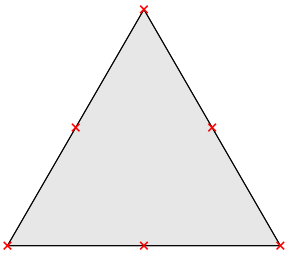
**Beispiel 1.94** (Überdeckungen mit Dreiecken – aus [Bla03], Kapitel 1, Seite 6)**.** *Sei $D$ das gleichseitige Dreieck mit Kantenlänge 2 und sei $a$ eine Zahl kleiner als 2. Wir betrachten folgende Aussagen:*

> *$A(a)$: $D$ lässt sich mit 4 gleichseitigen Dreiecken der Kantenlänge $a$ überdecken.*
>
> *$B(a)$: $D$ lässt sich mit 5 gleichseitigen Dreiecken der Kantenlänge $a$ überdecken.*

*Bei einer Überdeckung durch Dreiecke sind auch Überlappungen erlaubt. Wir behaupten nun, dass $\forall a > 0 : A(a) \iff B(a)$ gilt. Auf Grund deren Definitionen impliziert die Aussage $A(a)$ die Aussage $B(a)$; in Symbolen $A(a) \implies B(a)$. Wir fixieren $a > 0$ und beweisen nun die Implikation $B(a) \implies A(a)$, in dem wir die Kontraposition $\neg A(a) \implies \neg B(a)$ beweisen. Wir nehmen $\neg A(a)$ an, also dass sich $D$ nicht mit 4 gleichseitigen Dreiecken überdecken lässt. Dann muss $a < 1$ gelten, denn sonst könnte man $D$ wie folgt mit Dreiecken der Kantenlänge 1 (und also auch der Kantenlänge $a$) überdecken:*




*Gegeben eine Überdeckung von $D$ durch Dreiecke der Seitenlänge $a$, dann müssen die in obiger Grafik erhaltenen 6 Punkte (unten in rot markiert)*



*in jeweils verschiedenen gleichseitigen Dreiecken der Kantenlänge $a$ enthalten sein, da wir bereits erkannt haben, dass die Kantenlänge $a$ kleiner als 1 ist. Insbesondere sind mindestens 6 solche Dreiecke notwendig, um $D$ zu überdecken; also gilt $\neg B(a)$.*

[5] Man sollte dabei darauf achten, dass die Aussage $A$ nicht von einer freien Variablen $x$ abhängt, denn wenn die Aussage eigentlich $\forall x \in X : A(x)$ ist, dann ist die Negation $\exists x \in X : \neg A(x)$ und nicht $\forall x \in X : \neg A(x)$. Führt man also einen Widerspruchsbeweis, so sollte klar sein, was genau die getroffenen Annahmen sind.

### 1.6.3 Induktionsbeweise

Wir haben die Beweismethode der vollständigen Induktion schon im Beweis von Lemma 1.3 gesehen (und wissen jetzt, dass diese Methode genau einem definierenden Axiom von $\mathbb{N}$ entspricht). Diese wird verwendet, um eine Aussage $A(n)$ für alle natürlichen Zahlen $n$ zu zeigen. Der **Induktionsbeweis** hat zwei wichtige Teilschritte:

- **Induktionsanfang** (oder **Induktionsverankerung**): Man zeigt die Aussage $A(1)$.
- **Induktionsschritt**: Man zeigt, dass $A(n) \implies A(n+1)$ für alle natürlichen Zahlen $n$ gilt.

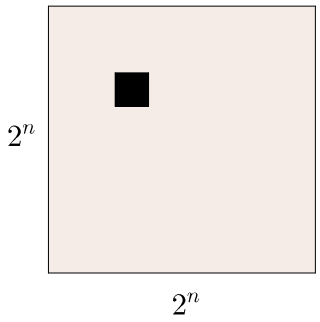
Man kann einen Beweis mit vollständiger Induktion mit einem rekursiven Algorithmus vergleichen, der die Aussage für alle natürlichen Zahlen beweist: Wenn man wissen will, warum die Aussage für $n = 10$ stimmt, dann zeigt der Induktionsschritt, dass die Aussage stimmt, weil sie schon für $n = 9$ richtig ist. Die Aussage für $n = 9$ stimmt wiederum, weil sie für $n = 8$ stimmt und so weiter bis man bei $n = 1$ angelangt ist. Der Fall $n = 1$ verankert (daher „Induktionsverankerung“) die Rekursion und den Beweis, da wir diesen Fall direkt und ohne Annahme einer anderen, noch zu verifizierenden Aussage überprüfen. Im nächsten Kapitel werden wir weitere Varianten des Induktionsbeweises kennenlernen.

**Übung 1.95** (Gauss'sche Summationsformel)**.** *Zeigen Sie mittels vollständiger Induktion, dass für jede natürliche Zahl $n \geq 1$ gilt*

$$1 + 2 + \cdots + (n-1) + n = \frac{n(n+1)}{2}$$

Wir möchten auch etwas „geometrischere“ Probleme behandeln.

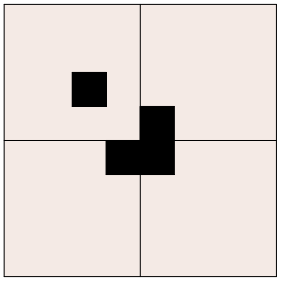
**Beispiel 1.96** (Eine geometrische Induktion)**.** *Sei $n$ eine natürliche Zahl. Wir betrachten das Quadrat mit Kantenlänge $2^n$, aus dem ein kleines Quadrat der Kantenlänge 1 entfernt wurde.*




*Bei diesen beiden Quadraten und auch bei allen noch zu erscheinenden Konstrukten möchten wir nur Ecken an ganzzahligen Stellen (d.h. in $\mathbb{Z}^2$) zulassen. Wir behaupten nun, dass sich das obige „Quadrat mit Loch“ durch Objekte der Art (rotieren ist erlaubt)*

*abdecken lässt und beweisen dies per Induktion über $n$. Zum Induktionsanfang $n = 1$: Das grössere Quadrat hat Kantenlänge 2, also ist das Bild*

*und die Aussage stimmt. Angenommen wir wissen, dass die Behauptung für $n$ korrekt ist. Wir zerlegen das Quadrat mit Kantenlänge $2^{n+1}$ in 4 Quadrate der Kantenlänge $2^n$ und entfernen aus einem dieser Quadrate ein kleines Quadrat der Kantenlänge 1. Zusätzlich entfernen wir aus den anderen Quadraten je ein kleines Quadrat wie in folgendem Bild*



*Die 4 so entstandenen Objekte (Quadrate mit Loch) lassen sich aber jeweils abdecken nach dem Induktionsschritt, also folgt die Behauptung.*

Vollständige Induktion ist ein unentbehrliches Hilfsmittel für Beweise in der Mathematik, genauso wie die Rekursion als Programmiermethode in der Informatik. Manchmal hat man jedoch das Gefühl, dass Induktion zwar den Zweck erfüllt (also das Lemma, die Proposition oder den Satz beweist), aber trotzdem nicht erklärt, warum eine Aussage richtig sein soll. Ein Beispiel dazu liefert Lemma 1.3, da wir aus dem Beweis beispielsweise nicht erkennen können, wie wir das Lemma für $1^4 + 2^4 + 3^4 + \cdots + n^4$ verallgemeinern könnten. Vielleicht würden Sie vermuten, dass diese Summe gleich einem Ausdruck der Form $\frac{n^5}{5} + an^4 + bn^3 + c^2 + dn + e$ für gewisse rationale Zahlen $a, b, c, d, e$ ist. Es fragt sich jedoch, wie wir diese Konstanten finden könnten. Die Induktionsmethode ist dabei nicht sehr hilfreich, da sie verwendet werden kann um die Wahrheit zu bestätigen, wenn man sie woanders bereits gefunden hat. Wir werden einen zweiten allgemeineren Beweis von Lemma 1.3 im Abschnitt 3.3 besprechen.

### 1.6.4 Das Schubfachprinzip

Mancher Existenzbeweis kann auf das Schubfachprinzip (Proposition 1.71) zurückgeführt werden. Ein Beispiel aus dem Alltag haben wir gleich nach Proposition 1.71 erklärt. Hier möchten wir das Schubfachprinzip an einem mathematischen Beispiel illustrieren.

**Beispiel 1.97.** *Wir betrachten die Abfolge von Zahlen*

$$1,\ 11,\ 111,\ 1111,\ 11111,\ \ldots$$

*und behaupten, dass es eine davon geben muss, welche durch 17 teilbar ist. Dazu definieren wir die Menge $R = \{0, \ldots, 16\}$ (die Reste mod 17) und die Schubfächer*

$$S_r = \{n \in \mathbb{N} \mid n - r \text{ ist durch } 17 \text{ teilbar}\}$$

*für $r \in R$. Nach Division mit Rest muss jede der Zahlen* $1,\ 11,\ 111,\ 1111,\ 11111,\ \ldots$ *in einem der Schubfächer $S_r$ enthalten sein und es gibt jeweils genau ein solches Schubfach. Da wir aber genau* 17 *Schubfächer haben und unendlich viele Zahlen darin verstauen wollen, müssen sicherlich zwei dieser Zahlen im gleichen Schubfach liegen. Formaler betrachtet man die Abbildung*

$$\{1,\ 11,\ 111,\ 1111,\ 11111,\ \ldots\} \to R, \quad x \mapsto r_x$$

*wobei $r_x \in R$ die eindeutig bestimmte Zahl mit $x \in S_{r_x}$ ist. Wie erwähnt kann diese Abbildung aber nicht injektiv sein. Es gibt also einen Rest $r \in R$ und zwei Zahlen $x, y \in \mathbb{N}$ von der Form* $111\ldots1$ *mit $x, y \in S_r$ und $y > x$. Die Differenz von $x$ und $y$ ist von der Form*

$$y - x = \underbrace{11\ldots1}_{=a}\underbrace{00\ldots0}_{n\ Nullen} = a \cdot 10^n,$$

*wobei $a$ in unserer Liste vorkommt und $n \in \mathbb{N}$. Des Weiteren ist $y - x$ wegen*

$$y - x = (y - r) - (x - r)$$

*durch* 17 *teilbar. Da aber* 17 *eine Primzahl ist und weder* 10 *noch* $10^n$ *durch* 17 *teilbar ist, ist $a$ durch* 17 *teilbar.*

Das Schubfachprinzip ist gemeinsam mit vielen Weiterentwicklungen eine sehr verbreitete Beweismethode in der Mathematik. Die allgemeine Einsetzbarkeit hat aber gewissermassen auch einen Preis, denn wenn das Schubfachprinzip für einen Existenzbeweis verwendet wird, so gibt der Beweis meist keinerlei Aufschlüsse wie man denn das Objekt (im Beispiel die konkrete, durch 17 teilbare Zahl) effektiv (also abgesehen von ausprobieren) finden könnte.

### 1.6.5 Weitere Methoden

Wir werden noch vielen weiteren Methoden begegnen, die wir hier nicht auflisten möchten. Beispielsweise gibt es Aussagen, die in verschiedenen Fällen einfache (aber verschiedene) Beweise haben. Falls diese Fälle alle Möglichkeiten abdecken, haben wir den Beweis der Aussage mittels Fallunterscheidung erhalten.

Ein anderes Beispiel einer Beweismethode (vor allem aus der Kombinatorik) ist die folgende: Angenommen wir wollen die endliche Kardinalität einer Menge bestimmen. Mit Hilfe einer geschickt gewählten Bijektion von dieser Menge in eine andere kann man dieses Zählproblem an einen Ort transportieren, wo man die Antwort bereits kennt.

Wie oben bereits angedeutet, gibt es mehrere Wünsche, die man an einen Beweis stellen könnte. In erster Linie muss dieser natürlich einen Beweis darstellen, doch könnte man sich auch folgende Punkte wünschen: Der Beweis könnte eine gute Erklärung für die Aussage liefert, könnte Verallgemeinerungen zulassen oder könnte bereits als Anleitung gelesen werden, wie man aus dem Existenzbeweis einen Algorithmus zum Auffinden des gesuchten Objektes erstellen kann. Für all diese Aussagen werden wir viele Beispiele sehen. Das Auffinden eines zweiten Beweises einer Aussage ist zwar logisch gesehen unnötig (und für uns aus Zeitgründen

oft nicht möglich), kann aber mitunter diese weiteren Wünsche besser abdecken und insgesamt das Verständnis der Theorie stärken.

### 1.6.6 Beweise finden

Sie fragen sich vielleicht bereits, wie man Beweise (wie zum Beispiel für die Übungen) finden kann. Die Antwort ist mit Übung, Hartnäckigkeit und Glück. Des Weiteren empfehlen wir Ihnen, Beweise aus der Vorlesung, diesem Skript oder anderer Literatur aus dem Gedächtnis zu wiederholen. Dadurch bekommen Sie Übung und ein gewisses Gespür für die inneren Mechanismen von Beweisen. Zusätzlich erkennen Sie vielleicht, dass viele Beweise ähnliche Bauformen aufweisen und mancher Beweis, sozusagen durch die Definitionen der involvierten Objekte erzwungen werden. Letzteres können Sie aber nur bemerken, wenn Sie die bereits besprochenen Beweise und Definitionen im Gedächnis haben.

### 1.6.7 Beweise aufschreiben

Nachdem Sie die Idee für den Beweis gefunden haben, wollen Sie diesen kommunizieren und aufschreiben. Auch dazu ist (viel) Übung nötig und Sie müssen den Beweis vielleicht umstellen oder komplett neu formulieren, bevor er für andere verständlich wird (was das Ziel sein sollte). Mitunter ist die Reihenfolge der Argumente, wie man sie gefunden hat, möglicherweise komplett anders als die Reihenfolge der Argumente, wie man sie präsentieren sollte. Weiters darf ein Beweis nicht nur aus Formeln bestehen, sondern muss dieser auch die Gedanken enthalten, die diesen Formeln Sinn und Zusammenhang geben. Wir verweisen auf [SS12] und das Buch „Das ist o.B.d.A. trivial“ ([Beu09]) von Beutelspacher für weitere Tipps in diese Richtung.

An dieser Stelle möchten wir noch die Wörter „o.B.d.A.“ und „trivial“ im obigen Buchtitel kommentieren. Die Abkürzung „o.B.d.A.“steht für „ohne Beschränkung der Allgemeinheit“. Wir verwenden diese, wenn wir den Beweis einer Aussage auf den Beweis eines Spezialfalls reduzieren. Folgendes Beispiel illustriert dies:

**Beispiel 1.98** (o.B.d.A)**.** *Sei $A(n)$ die Aussage $2^{n^2} > n^2$ über ganze Zahlen $n$, die die Eigenschaft $A(n) \iff A(-n)$ für jede ganze Zahl $n$ erfüllt. Wollen wir die Wahrheit der Aussage $A(n)$ nachweisen, so reicht es anzunehmen, dass $n \geq 0$ ist, denn ein mögliches Vorzeichen von $n$ wird von der Funktion $n \in \mathbb{Z} \mapsto n^2 \in \mathbb{Z}$ absorbiert. Wir schreiben zu Beginn des Beweises also „Sei o.B.d.A. $n \geq 0$“.*

Das Wort „trivial“ (abgeleitet von lat. „trivium“) bedeutet „bekannt“ oder „allgemein bekannt“ und sollte entgegen dem üblichen Gebrauch nicht mit dem Wort „offensichtlich“ verwechselt werden. Wenn also „. . . ist trivial“ geschrieben wird, so sollte man dies als Aufforderung verstehen, zu verifizieren, wieso genau die Aussage „bekannt“ sein sollte. Dennoch werden wir und sollten Sie auf den Gebrauch dieses Wortes komplett verzichten. Selbiges betrifft Wörter wie

„offensichtlich“, „evident“, „einfach“ und „leicht“

oder Phrasen wie

„. . . ist klar“ und „Es ist leicht zu sehen, dass . . .“.

Grund dafür ist, dass diese Ausdrücke aus Sicht des Lesers als Affront aufgefasst werden können; es liegt nicht an der Verfasserin oder dem Verfasser, die Schwierigkeit einer Aussage zu beurteilen, die sie oder er selbst hingeschrieben hat. Des Weiteren kann man eine (vermeintlich) einfache Aussage oft auch in wenigen Worten erklären.

### 1.6.8 Beweise lesen

Ein wichtiger erster Schritt ist die zu beweisende Aussage zu lesen, zu verstehen und zu erkennen, was überhaupt zu beweisen ist. Wenn Sie dann den Beweis lesen, dann sollten Sie jeden Schritt hinterfragen und dabei fast so stur wie ein Computer beim Abarbeiten eines Programms vorgehen. Man kann dazu auch das Motto der Royal Society zitieren:

*„Nullus in verba“ oder „Take nobody's word for it“.*

Sollten Sie einen Schritt nicht verstehen oder nicht einsehen, wieso dieser möglich sein soll, dann fragen Sie bei Mitstudenten, bei Assistenten oder beim Professor nach, bis Sie eine zufriedenstellende Antwort bekommen haben. Bemühen Sie sich jetzt, wo die Themen noch nicht so verflochten sind, ein vollständiges Verständnis der besprochenen Begriffe und Sätze zu entwickeln, und warten Sie nicht darauf bis die Themen „interessanter“ werden. Denn dann werden diese auch schwieriger und es wird zunehmend schwieriger werden, den Einstieg zu finden.

Umgekehrt sollten Ihre Beweise auch so aufgeschrieben sein, dass auch ein sturer Leser nicht umhin kommt, die von Ihnen bewiesene Aussage zu akzeptieren. Es hilft, wenn Sie Ihren Beweis einen oder zwei Tage später nochmals lesen, da es für Sie dann wahrscheinlich leichter ist, bewusst zu vergessen, was Sie sich beim Niederschreiben gedacht haben und dann vielmehr lesen, was Sie tatsächlich niedergeschrieben haben.

### 1.6.9 Prädikatenlogik vs Umgangssprache

Wir werden unsere Beweise in der Umgangssprache (also in Deutsch anstatt in formalen Symbolen) formulieren, doch empfehlen wir Ihnen, die Übersetzung zwischen Umgangssprache und Prädikatenlogik zu üben. Wenn wir die Analogie zur Informatik etwas weiter ziehen, dann sollten Sie die Prädikatenlogik als Maschinensprache der Beweise verstehen und die Umgangssprache als die höhere Programmiersprache, die es uns Menschen leichter machen wird, Zusammenhänge zu sehen, eine Intuition für die Theorie zu entwickeln und Konversationen über den Beweis zu führen. Der Autor eines Beweises muss also „als Programmierer“ sicherstellen, dass die umgangssprachliche Formulierung ohne Zweideutigkeiten in die Prädikatenlogik übersetzt werden kann. Genauso sollte der Leser die Rolle des „sturen Computers“ spielen und überprüfen, ob der Beweis „ohne Fehler abläuft“.

Wir werden anfangs unsere Diskussionen näher an der Prädikatenlogik halten und mitunter (entgegen üblichen mathematischen Konventionen) auch die Quantoren $\forall, \exists, \exists!$ in Symbolen verwenden. Doch werden wir sehen, dass weder die Lineare Algebra noch die Analysis sinnlose Formelsammlungen in Buchstaben und den Symbolen $\neg$, $\wedge$, $\vee$, $\Longrightarrow$, $\Longleftrightarrow$, $\forall$, $\exists$, $\exists!$, $=$, $\in$,

$\subseteq$, ... bilden. Vielmehr stellen sie zwei sehr massive Gebäude mit vielen Stockwerken, interessanten Verzierungen und Erkern sowie mehreren Brücken zwischen einander und anderen Gebäuden dar. Wenn Sie das „Gebäude“ der Linearen Algebra oder das „Gebäude“ der Analysis verstehen wollen, dann müssen Sie zwar die einzelnen Bausteine (in Form von Definitionen, Lemmata und Sätzen), aber eben auch den Lageplan des Gebäudes (das wären die Zusammenhänge) kennen. Es sind diese Zusammenhänge, die in umgangsprachlich formulierten Beweisen klarer werden.

Wir wollen Ihnen am Ende des Hauptteiles dieses Kapitels noch einen Podcast der BBC empfehlen, der neben geschichtlichen Informationen auch noch eine Zusammenfassung vieler Themen dieses Kapitels bietet.

*Bemerkung.* Eigentlich beschäftigt sich obiger Podcast mit Gödel's Unvollständigkeitssatz, welcher Teil der mathematischen Logik ist. Dieses Teilgebiet der Mathematik beschäftigt sich mit der Theorie des Beweisens und Fragen wie „ist die Aussage ... aus den Axiomen ... beweisbar?“. Interessanterweise gibt es in dieser Theorie nicht bloss einen Unvollständigkeitssatz, sondern auch einen Vollständigkeitssatz. Die Auflösung dieses scheinbaren Paradox würde uns jedoch zu weit vom Thema abbringen.

## 1.7 Weitere Lernmaterialien

Wir wollen hier versuchen, Ihnen einen Überblick über dieses Kapitel zu geben und auch weitere Übungsaufgaben zu den Themen des Kapitels zu sammeln.

### 1.7.1 Verwendung des Kapitels

Wir haben in diesem Kapitel viele Themen und unter anderem Tipps, Motivationen, einige Theorie und auch noch etwas Geschichte der Mathematik besprochen. Auf Grund dieser Vielfalt wollen wir kurz noch betonen, was Sie aus diesem Kapitel eigentlich für das Folgende mitnehmen müssen: Die Themen aus den Abschnitten 1.3 und 1.4 sind grundlegend und wir werden ohne Wiederholungen alle unsere weiteren Diskussionen auf diese Abschnitte aufbauen. Vor allem sollten Sie die folgenden Begriffe so lange üben, bis Sie diese ohne Zweifel im Gedächnis haben.

- Logische Operationen, insbesondere sollten Sie alle Fälle für das Oder, die Implikation und ihre Negation ohne das Nachblättern der Wahrheitstabellen wissen.
- Quantoren und deren Verhalten bei Kombination und Negation.
- Mengenoperationen, de Morgan Gesetze.
- Begriff der Funktion und elementare Eigenschaften wie Injektivität, Surjektivität und Bijektivität, aber auch die ungenaueren Begriffe „wohldefiniert“ und „kanonisch“.
- Verhalten dieser Eigenschaften unter Verknüpfungen.
- Äquivalenzrelationen, Partitionen, Quotientenraum
- Kardinalitäten, vorläufig nur mit der Unterscheidung endlich, abzählbar unendlich oder überabzählbar.

Bei logischen Aussagen werden wir, wie Sie vielleicht schon in Abschnitt 1.6 gemerkt haben, die Notation in Zukunft etwas leichter halten. Unter anderem werden wir die Anführungszeichen „...“ bei logischen Ausdrücken weglassen und teilweise die Klammerung unterschlagen, wenn diese implizit klar ist. Zum Beispiel kann $\exists n \in \mathbb{N} : n = n^2 \implies n = 1$ nur für den Ausdruck $\exists n \in \mathbb{N} : (n = n^2 \implies n = 1)$ stehen, da ansonsten die Zahl $n$ auf der rechten Seite von $(\exists n \in \mathbb{N} : n = n^2) \implies n = 1$ nicht definiert ist.

Der Abschnitt 1.5 ist als Übersicht gedacht und Sie sollten eigentlich (vor allem für die ersten zwei Abschnitte des nächsten Kapitels) diese erste Einführung der Zahlenmengen wieder vergessen. Denn wir werden im nächsten Kapitel logischen Begriffe, Mengen und Funktionen verwenden um die reellen Zahlen axiomatisch einzuführen.

Abschnitt 1.6 und die Übungsaufgaben dieses Abschnitts sollen Ihnen helfen Ihren eigenen Zugang zu Beweisen zu finden. Schwierige Fragen, die immer wieder auftauchen, sind, „Was muss ich denn bei diesem Satz oder bei dieser Aufgabe eigentlich beweisen?“ und „Welche Aussage kann ich als gegeben annehmen?“. Dies ist in diesem Kapitel mitunter wirklich schwer

zu beantworten (und wir haben hierzu im Laufe des Kapitels auch mehrmals unsere Meinung geändert). Nach Einführung der Axiome im nächsten Kapitel wird deutlich klarer sein, was wir beweisen müssen: nämlich ausser den Axiomen alles Weitere, wobei wir aber auf bereits bewiesene Aussagen zurückgreifen dürfen. Insbesondere dürfen Sie in den wöchentlichen Übungsaufgaben die Aussagen der Vorlesung und ebenso die Aussagen des Skripts verwenden, aber keine Übungsaufgaben des Skripts und auch nur jene Seiten des Skripts, die bereits in der Vorlesung behandelt wurden. Obwohl die Axiome sehr einfach sein werden, werden wir im Laufe der Vorlesung viele komplizierte und auch überraschende Aussagen beweisen können.

### 1.7.2 Flächeninhalt

**Übung** (Allgemeinere Bereiche unter der Parabel)**.** *Berechnen Sie in Analogie zu obigem die Fläche unter der Parabel*

$$P_{a,b} = \left\{(x,y) \in \mathbb{R}^2 \mid a \leq x \leq b, 0 \leq y \leq x^2\right\},$$

*wobei $a, b \in \mathbb{R}$ zwei gegebene reelle Zahlen mit $a < b$ sind.*

**Übung.** *In dieser Übung möchten wir eine Variante illustrieren, wie man auf Lemma 1.3 schliessen kann ohne zuerst im Besitz der richtigen Formel zu sein. Wir schreiben dazu für $n \in \mathbb{N}$*

$$\begin{aligned}(n+1)^3 &= (1^3 + 2^3 + 3^3 + \ldots + (n+1)^3) - (1^3 + 2^3 + 3^3 + \ldots + n^3) \\ &= 1^3 + (2^3 - 1^3) + (3^3 - 2^3) + \ldots + ((n+1)^3 - n^3).\end{aligned}$$

*Gehen Sie nun wie folgt vor.*

- *Zeigen Sie für $a, b \in \mathbb{R}$*

  $$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

  *und erklären Sie, welche Rechenregeln Sie dabei verwenden.*

- *Verifizieren Sie die Gleichung*

  $$(n+1)^3 = (n+1) + 3(1^2 + 2^2 + \ldots + n^2) + 3(1 + 2 + \ldots + n).$$

- *Schliessen Sie auf Lemma 1.3.*

### 1.7.3 Logik

**Übung** (Draculas Bücher – aus [AE06])**.** *In der Bibliothek des Grafen Dracula gibt es keine zwei Bücher, deren Inhalt aus gleich vielen Wörtern besteht. Die Anzahl der Bücher ist die Summe der Anzahl der Wörter jedes einzelnen Buches. Des Weiteren genügen diese Aussagen, um den Inhalt mindestens eines Buches aus Draculas Bibliothek genau zu beschreiben. Was steht in diesem Buch?*

**Übung** (Vier Aussagen in Prädikatenlogik). *Wir sagen $m \in \mathbb{N}$ teilt $n \in \mathbb{N}$ falls es ein $d \in \mathbb{N}$ gibt mit $n = dm$. Beschreiben Sie die Bedeutung folgender Aussagen und bestimmen Sie, ob diese zutreffen.*

- $\forall n \in \mathbb{N}\ \exists m \in \mathbb{N}:\ m$ *teilt* $n$.
- $\exists m \in \mathbb{N}\ \forall n \in \mathbb{N}:\ m$ *teilt* $n$.
- $\forall m \in \mathbb{N}\ \exists n \in \mathbb{N}:\ m$ *teilt* $n$.
- $\exists n \in \mathbb{N}\ \forall m \in \mathbb{N}:\ m$ *teilt* $n$.

**Übung** (Allgemeinere Existenzquantoren). *Sei $X$ eine Menge. In dieser Übung wollen wir Quantoren für die Aussage, dass mehrere Elemente mit einer Eigenschaft $A(x)$ in $X$ existieren, definieren.*

*(i) Definieren Sie unter Verwendung des Existenzquantors einen neuen Quantor $\exists^{\geq 2}$, so dass die Aussage „$\exists^{\geq 2} x \in X : A(x)$“ bedeutet, dass es mindestens zwei Elemente $x$ in $X$ gibt, die die Eigenschaft $A(x)$ haben.*

*(ii) Verallgemeinern Sie (i) zu dem Quantor $\exists^{\geq n}$ für eine natürliche Zahl $n$, und verwenden Sie diese um auch den Quantor $\exists^{=n}$ zu definieren, der besagen soll, dass es genau $n$ Elemente in $X$ gibt, die die Eigenschaft $A(x)$ besitzen. (Sie dürfen entweder informell Punkte verwenden, oder formal korrekter den Funktionsbegriff, Eigenschaften von Funktionen, und die Menge $\{k \in \mathbb{N} \mid k \leq n\}$, die genau $n$ Elemente hat.)*

*(iii) Definieren Sie unter Verwendung des Existenzquantors, des Funktionsbegriffes, einer Eigenschaft von Funktionen und der natürlichen Zahlen $\mathbb{N}$ einen neuen Quantor $\exists^{\infty}$, der besagt, dass es unendlich viele Elemente in $X$ gibt, die die Eigenschaft $A(x)$ besitzen.*

### 1.7.4 Funktionen und Relationen

**Übung** (Injektive Funktionen durch Fallunterscheidungen). *Zeigen Sie folgende Behauptung: Seien $X$ und $Y$ Mengen und sei $\mathcal{P}$ eine Partition von $X$. Angenommen es ist für jedes $P \in \mathcal{P}$ eine injektive Funktion $f_P : P \to Y$ gegeben und sei $f : X \to Y$ die eindeutige Funktion mit $f|_P = f_P$ für jedes $P \in \mathcal{P}$ nach Lemma 1.67. Zeigen Sie, dass $f$ genau dann injektiv ist, falls die Mengen $f(P)$ für $P \in \mathcal{P}$ paarweise disjunkt sind.*

**Übung** (Eine Äquivalenzrelation auf dem kartesischen Produkt – aus [AE06]). *Seien $X, Y$ zwei nicht-leere Mengen, sei $\sim_X$ eine Relation auf $X$ und sei $\sim_Y$ eine Relation auf $Y$. Wir definieren damit eine Relation $\sim$ auf $X \times Y$ durch*

$$\big((x,y) \sim (x',y')\big) \iff \big((x \sim_X x') \wedge (y \sim_Y y')\big)$$

*für $(x,y),(x',y') \in X \times Y$. Zeigen Sie, dass $\sim$ genau dann eine Äquivalenzrelation auf $X \times Y$ ist, wenn $\sim_X$ eine Äquivalenzrelation auf $X$ ist und $\sim_Y$ eine Äquivalenzrelation auf $Y$ ist. Gilt dies auch, wenn eine der beiden Mengen $X, Y$ leer ist?*

**Applet** (Nichtvertauschbarkeit der Verknüpfung). *Wir betrachten zwei Funktionen $f, g : \mathbb{R} \to \mathbb{R}$, wobei $f$ nur durch den Graphen beschrieben ist und $g : x \in \mathbb{R} \mapsto g(x) = ax + b$ eine affine Funktion ist, die durch zwei Konstanten $a, b \in \mathbb{R}$ definiert wird. Durch Bewegen zweier Punkte am Graph von $g$ lassen sich $a$ und $b$ definieren. Experimentieren Sie damit um sich an die geometrische Bedeutung von den Zahlen $a$ und $b$ zu errinnern, und beobachten Sie, wie sich die beiden Funktionen $g \circ f$ und $f \circ g$ im rechten Fenster unterschiedlich verändern. Es ist nützlich sich diese Phänomene vollständig zu erklären, denn wir werden ähnlichen Verknüpfungen in unseren weiteren Überlegungen begegnen.*

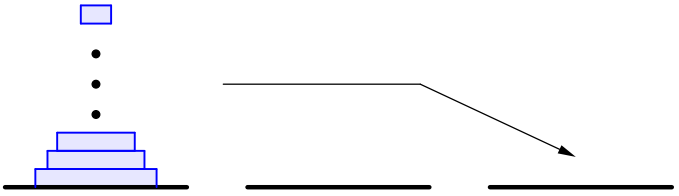
### 1.7.5 Beweismethoden

**Übung.** *Sei $n$ eine natürliche Zahl. Zeigen Sie, dass die Implikation*

$$n^2 + 17n - 13 \text{ ist gerade} \implies n \text{ ist ungerade}$$

*gilt.*

**Übung** (Türme von Hanoi). *Es seien 3 Ablageflächen gegeben. Angenommen auf der linken Ablagefläche seien $n$ Blöcke für eine natürliche Zahl $n$ aufgetürmt, wobei nie ein kleinerer Block unter einem grösseren Block liegt.*



*Zeigen Sie, dass Sie den Turm von links nach rechts umschichten können, ohne dass je ein kleinerer Block unter einem grösseren Block zu liegen kommt.*

**Übung.** *Zeigen Sie mittels vollständiger Induktion, dass für alle natürliche Zahlen $n \geq 5$ gilt $4n < 2^n$.*

**Übung.** *Zeigen Sie mittels vollständiger Induktion, dass die Ungleichung $n^2 < 3^n$ für alle natürlichen Zahlen $n$ gilt. Beschreiben Sie die Variante des Induktionsbeweises, die sie hier verwenden.*

**Übung.** *Wir färben jeden Punkt im Gitter $\mathbb{Z}^2$ mit einer von 17 verschiedenen Farben ein. Zeigen Sie, dass es ein achsenparalleles Rechteck $R$ in diesem Gitter gibt, dessen Ecken alle dieselbe Farbe besitzen.*

**Übung.** *Sei $n \in \mathbb{N}$ und sei $S$ eine Teilmenge von $\{1, \ldots, 2n\}$ mit Kardinalität $n + 1$. Zeigen Sie, dass es Elemente $a, b \in S$ gibt mit $a|b$.*

**Übung.** *Sei $T$ eine endliche Menge und seien $S_1, \ldots, S_n$ Teilmengen von $T$ mit*

$$|S_1| + \cdots + |S_n| > k|T|.$$

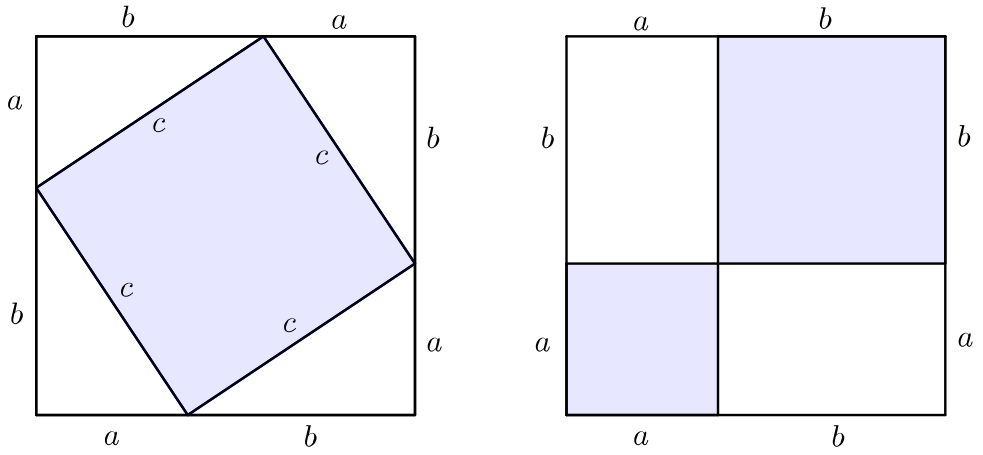
*Zeigen Sie, dass es ein Element $t \in T$ gibt, welches in mindestens $k+1$ der Mengen $S_1, \ldots, S_n$ liegt.*

### 1.7.6 Geometrische Probleme

**Übung** (Satz von Pythagoras)**.** *In dieser Übung möchten wir (auf Ihnen vielleicht bereits bekannte Weise den Satz von Pythagoras) beweisen. Wir betrachten ein rechtwinkliges Dreieck mit Katheten der Länge $a$ und $b$ und Hypotenuse der Länge $c$. Zeigen Sie, dass*

$$c^2 = a^2 + b^2.$$

*Hinweis: Betrachten Sie folgende Bilder*




*und gehen Sie dabei davon aus, dass gewisse geometrische Begriffe wie Winkel, Länge und Fläche für elementare Bereiche und intuitiv anschauliche Eigenschaften wie Invarianz der Fläche unter Verschiebung und Drehung bekannt sind.*

### 1.7.7 Übungen zu Primzahlen

Eine natürliche Zahl $p$ grösser als 1 ist **irreduzibel**, falls sie nicht als Produkt von zwei kleineren natürlichen Zahlen geschrieben werden kann. Eine natürliche Zahl $p$ grösser als 1 heisst eine **Primzahl**, falls ein Produkt $ab$ zweier natürlicher Zahlen $a, b \in \mathbb{N}$ nur dann durch $p$ teilbar ist, falls eine der beiden Zahlen durch $p$ teilbar ist.

**Übung** (Primzahlen sind irreduzibel)**.** *Zeigen Sie, dass jede Primzahl in $\mathbb{N}$ auch irreduzibel ist.*

Diese beiden Begriffe sind in der Tat für die natürlichen Zahlen äquivalent (wir werden dies nochmals etwas genauer in Abschnitt 2.2.4 besprechen) und wir werden in diesem Abschnitt irreduzible Zahlen ebenso als Primzahlen bezeichnen. Es ist eine gute Übung im Folgenden genau zu erklären welche der beiden Begriffe eigentlich verwendet wird.

**Übung** (Primfaktorzerlegung)**.** *Zeigen Sie mittels vollständiger Induktion, dass jede natürliche Zahl grösser als 1 als Produkt von Primzahlen geschrieben werden kann.*

**Übung** (Unendlich viele Primzahlen)**.** *Zeigen Sie, dass es unendliche viele Primzahlen gibt.*

**Übung** (Summen von Quadraten)**.** *In dieser Übung möchten wir (mit einer detaillierten Anleitung) zeigen, dass für jede ungerade Primzahl $p \in \mathbb{N}$ mit $p \equiv 1 \mod 4$ Zahlen $x, y \in \mathbb{N}$ mit $x^2 + y^2 = p$ existieren. Dabei folgen wir dem Artikel „A one-sentence proof that every prime $p \equiv 1 \mod 4$ is a sum of two squares.“ von Don Zagier [Zag90]. Wir zitieren:*

> *„The involution on the finite set $S = \{(x, y, z) \in \mathbb{N}^3 \mid x^2 + 4yz = p\}$ defined by*
>
> $$(x, y, z) \mapsto \begin{cases} (x + 2z, z, y - x - z) & \text{falls } x < y - z \\ (2y - x, y, x - y + z) & \text{falls } y - z < x < 2y \\ (x - 2y, x - y + z, y) & \text{falls } x > 2y \end{cases} \tag{1.9}$$
>
> *has exactly one fixed point, so $|S|$ is odd and the involution defined by $(x, y, z) \mapsto (x, z, y)$ also has a fixed point. $\square$“*

*Im Folgenden möchten wir die Details in obigem kryptischen aber auch eleganten Beweis ausarbeiten. Aber zuerst möchten wir einen Grund lieferen, wieso man obige Aussage überhaupt glauben sollte.*

*(i) Sei also $p \in \mathbb{N}$ eine ungerade Zahl, für die $x, y \in \mathbb{N}$ mit $x^2 + y^2 = p$ existieren. Zeigen Sie, dass dann $p \equiv 1 \mod 4$ ist.*

*(ii) Versuchen Sie 5, 9, 13, 17, 21 als Summe $x^2 + y^2$ von zwei Quadraten natürlicher Zahlen $x, y \in \mathbb{N}$ zu schreiben und bemerken sie, dass unter diesen Zahlen dies nur für die Primzahlen möglich ist.*

*(iii) Zeigen Sie nun, dass die obige Menge $S$ endlich ist und dass die Abbildung*

$$\tau : S \to S, \ (x, y, z) \mapsto (x, z, y)$$

*wohldefiniert ist.*

*(iv) Beweisen Sie, dass $p$ als Summe von zwei Quadraten geschrieben werden kann, wenn es für die Abbildung $\tau$ einen Punkt $a = (x, y, z) \in S$ mit $\tau(a) = a$ gibt. Ein solcher Punkt $a \in S$ nennt sich ein **Fixpunkt** von $\tau$.*

*Die Abbildung $\tau$ ist eine **Involution**, was bedeutet, dass $\tau^{\circ 2} = \mathrm{id}_S$ ist.*

*(iv) Zeigen Sie, dass eine Involution auf $S$ einen Fixpunkt besitzt, wenn $|S|$ ungerade ist. Verifizieren Sie auch, dass $|S|$ ungerade ist, wenn es eine Involution $S \to S$ mit einem eindeutig bestimmten Fixpunkt gibt.*

*(v) Betrachten Sie die Abbildung $\sigma$ aus (1.9) und verifizieren Sie folgende Behauptungen:*

- *Die Abbildung $\sigma$ ist wohldefiniert: Einerseits gilt für $(x, y, z) \in S$, dass die Fälle $x = y - z$ und $x = 2y$ nicht eintreten können, womit $\sigma$ für jedes Element in $S$ definiert ist. Andererseits ist für jedes $(x, y, z) \in S$ das Bild $\sigma((x, y, z))$ (wie oben definiert) tatsächlich ein Element von $S$.*

- *$\sigma : S \to S$ ist eine Involution.*

*(vi) Zeigen Sie, dass der einzige Fixpunkt von $\sigma$ durch $(1, 1, k)$ gegeben ist, wobei $k \in \mathbb{N}_0$ die Gleichung $p = 4k + 1$ erfüllt.*

*(vii) Kombinieren Sie obige Argumente, um die gewünschte Aussage zu zeigen.*

*Wo haben Sie verwendet, dass $p$ eine Primzahl (oder irreduzibel) ist?*

### 1.7.8 Lernkarten

Wir bieten im eSkript am Ende des Kapitels ein Tool an, das Ihnen helfen könnte, die Themen des Kapitels zu wiederholen.

# Kapitel 2

# Die reellen Zahlen

Wie bereits erwähnt wurde, werden wir uns in diesem und in allen folgenden Kapiteln an den üblichen Aufbau mathematischer Theorien halten und alle Aussagen aus den gegebenen Axiomen ableiten. In unserem Fall sind letztere die Axiome der reellen Zahlen, die wir gemeinsam mit der naiven Mengenlehre (inklusive Funktionen und Relationen) verwenden werden, um die Analysis aufzubauen. Wir werden diesen Aufbau in Etappen erledigen und nach jeder Etappe klarstellen, was wir erreicht haben und in Zukunft verwenden dürfen.

## 2.1 Die Axiome der reellen Zahlen

### 2.1.1 Körperaxiome

**Definition 2.1.** Eine Menge $\mathbb{R}$ gemeinsam mit einer Abbildung

$$+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R},\ (x, y) \mapsto x + y,$$

die wir **Addition** nennen, einer Abbildung

$$\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R},\ (x, y) \mapsto x \cdot y,$$

die wir **Multiplikation** nennen, und einer Relation $\leq$ auf $\mathbb{R}$, die wir **kleiner gleich** nennen, wird als Menge der **reellen Zahlen** bezeichnet, falls die in diesem Abschnitt 2.1 aufgelisteten 16 Axiome erfüllt sind.

**Axiome (Addition).** *Die Addition erfüllt folgende Eigenschaften:*

(1) *(Nullelement)* $\exists 0 \in \mathbb{R}\ \forall x \in \mathbb{R} : x + 0 = 0 + x = x.$

(2) *(Additives Inverses)* $\forall x \in \mathbb{R}\ \exists(-x) \in \mathbb{R} : x + (-x) = (-x) + x = 0$

(3) *(Assoziativgesetz)* $\forall x, y, z \in \mathbb{R} : (x + y) + z = x + (y + z)$

(4) *(Kommutativgesetz)* $\forall x, y \in \mathbb{R} : x + y = y + x$

An dieser Stelle kann man sich einige Fragen stellen. Beispielsweise ist nicht klar, wieso die Notation $(-x)$ für die additive Inverse eines Elements $x \in \mathbb{R}$ gerechtfertigt ist; a priori

könnte es ja mehrere additive Inverse eines Elements geben. Deswegen sollte $(-x)$ vorerst als der Name eines Elements in $\mathbb{R}$ gesehen werden und nicht als ein eindeutig bestimmtes, zu $x$ gehörendes Element. Fragen wie diese möchten wir in Kürze beantworten.

*Bemerkung.* Um formal korrekt zu sein, müsste man schreiben

(1') $\exists z \in \mathbb{R}\ \forall x \in \mathbb{R} : x + z = z + x = x$ und

(2') $\forall x \in \mathbb{R}\ \exists y \in \mathbb{R}\ \forall w \in \mathbb{R} : (x + y) + w = w + (x + y) = (y + x) + w = w + (y + x) = w.$

Dies sind die formal korrekten Version der Axiome (1) und (2): In (1') haben wir nicht ein „unbekanntes Symbol 0“ verwendet, sondern die Existenz eines Elements mit einer bestimmten Eigenschaft gefordert. Dadurch wird es klarer, dass bei der gefordeten Existenz a priori nicht klar ist, ob es nur ein oder mehrere derartige Elemente gibt. Analog haben wir in (2') die Verwendung dieses unbekannten Symbols und auch der verfrühten Notation $(-x)$ vermieden.

Wir haben bei der Formulierung der obigen Axiome versucht einen Kompromiss zwischen Lesbarkeit und formaler Korrektheit zu treffen, damit die Axiome auf den ersten Blick intuitiv Sinn machen und über jeden Zweifel erhaben sind. Formal sollten Sie die Symbole 0 und $(-x)$, wie bereits angedeutet, vorerst als seltsame aussehende Variablen interpretieren.

Alternativ hätten wir in der Formulierung der Axiome gleich zu Beginn fordern können, dass es neben der Addition auch noch ein ausgezeichnetes Element $0 \in \mathbb{R}$ und eine weitere Abbildung $- : \mathbb{R} \to \mathbb{R}$ gibt, und in den ersten beiden Axiomen die Eigenschaften dieser zusätzlichen Objekte beschreiben können. Dies widerspricht aber dem Wunsch an einem Axiomensystem minimal zu sein und nur die nötigsten Objekte, die sich nicht unter Verwendung anderer Objekte definieren lassen, einzuführen.

**Erste Folgerungen.**

(a) *Das Nullelement* 0 *(auch die Null genannt) ist durch das Axiom* (1) *eindeutig bestimmt. Insbesondere ergibt der Begriff „das Nullelement“ Sinn und es ist akzeptabel, dass* 0 *in Axiom* (2) *vorkommt, ohne dass man vorher eine Null wählen musste.* In der Tat, sind $0_1, 0_2 \in \mathbb{R}$ zwei Elemente, die die Eigenschaft in Axiom (1) erfüllen, dann gilt also

$$0_1 = 0_1 + 0_2 = 0_2.$$

(Für welche $x \in \mathbb{R}$ wurde die Eigenschaft in Axiom (1) verwendet und wie? *)*

(b) *Das Negative* $-x \in \mathbb{R}$ *ist für jedes* $x \in \mathbb{R}$ *durch die Eigenschaft* $x + (-x) = 0$ *eindeutig bestimmt. Insbesondere können wir von der additiven Inversen eines Elements sprechen und die Abbildung* $- : x \in \mathbb{R} \mapsto -x \in \mathbb{R}$ *ist wohldefiniert.* In der Tat, falls $y, z \in \mathbb{R}$ zu $x \in \mathbb{R}$ die Identitäten $x + y = x + z = 0$ erfüllen, dann gilt

$$\begin{aligned} y = y + 0 &= y + (x + z) \\ &= (y + x) + z = (x + y) + z = 0 + z = z, \end{aligned}$$

nach der Eigenschaft der Null in Axiom (1), der Annahme für $z$, dem Assoziativgesetz in Axiom (3), dem Kommutativgesetz in Axiom (4), der Annahme für $y$ und wiederum die Eigenschaft der Null in Axiom (1).

(c) *Wegen der Eindeutigkeit der additiven Inversen gilt $-(-x) = x$ für jedes $x \in \mathbb{R}$.* Denn für $x \in \mathbb{R}$ gilt $(-x) + x = 0$ nach der Definition der additiven Inversen von $x$ in Axiom (2) und damit ist nach (b) schliesslich $-(-x) = x$.

(d) *„Additives Kürzen“ ist erlaubt: Sind $x, y, z \in \mathbb{R}$ mit $x+y = x+z$, so darf man $x$ wegstreichen. (Das heisst, die Aussage $\forall x, y, z \in \mathbb{R} : x + y = x + z \implies y = z$ gilt.)* In der Tat gilt

$$\begin{aligned} y = ((-x) + x) + y &= (-x) + (x + y) \\ &= (-x) + (x + z) = ((-x) + x) + z = z, \end{aligned}$$

wobei die Eigenschaft der Null in Axiom (1), Eigenschaft der additiven Inversen in Axiom (2), das Assoziativgesetz in Axiom (3), die Annahme $x+y = x+z$, das Assoziativgesetz in Axiom (3) und nochmals die Eigenschaft in den Axiomen (2) und (1) verwendet wurden.

Nach dem Assoziativgesetz in Axiom (3) können wir für $x, y, z \in \mathbb{R}$ anstelle von $(x+y)+z$ oder $x+(y+z)$ einfach $x+y+z$ schreiben, da diese nach Axiom (3) gleich sind. Anstelle von $x+(-y)$ für $x, y \in \mathbb{R}$ schreiben wir oft auch die **Subtraktion** $x-y$ und anstelle von $(-x)+y$ auch $-x+y$.

**Wichtige Übung 2.2.** *Zeigen Sie die folgenden Regeln (unter Verwendung der Axiome* (1)-(4) *und der Folgerungen* (a)-(d)*).*

*(i) Es gilt $-0 = 0$.*

*(ii) Für alle $x, y \in \mathbb{R}$ gilt $-(x + y) = (-x) + (-y)$ (wobei wir für letzteres auch $= -x - y$ schreiben).*

*(iii) Für alle $x, y \in \mathbb{R}$ gilt $-(x - y) = -x + y$.*

*Bemerkung.* Wir sagen auch, dass die reellen Zahlen $\mathbb{R}$ gemeinsam mit der Abbildung (Verknüpfung) $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ eine **kommutative** oder **abelsche Gruppe** bilden, da die Axiome (1)-(4) gerade die Axiome einer kommutativen Gruppe bilden.

**Axiome (Multiplikation).** *Die Multiplikation erfüllt folgende Eigenschaften:*

(5) *(Einselement)* $\exists 1 \in \mathbb{R} \setminus \{0\} \;\; \forall x \in \mathbb{R} : x \cdot 1 = 1 \cdot x = x$.

(6) *(Multiplikative Inverse)* $\forall x \in \mathbb{R} \setminus \{0\} \;\; \exists (x^{-1}) \in \mathbb{R} : x \cdot (x^{-1}) = (x^{-1}) \cdot x = 1$

(7) *(Assoziativgesetz)* $\forall x, y, z \in \mathbb{R} : x \cdot (y \cdot z) = (x \cdot y) \cdot z$.

(8) *(Kommutativgesetz)* $\forall x, y \in \mathbb{R} : x \cdot y = y \cdot x$.

Des Weiteren muss bei Kombination der Addition und der Multiplikation folgendes Gesetz gelten.

**Axiome (Kompatibilität von + und ·).** *Wir verlangen*

(9) *(Distributivgesetz)* $\forall x, y, z \in \mathbb{R} : (x + y) \cdot z = (x \cdot z) + (y \cdot z)$

Wir werden sehen, dass die Axiome (5)-(9) implizieren, dass $\mathbb{R}^\times = \mathbb{R} \setminus \{0\}$ mit der Multiplikation eine abelsche Gruppe bildet, die auch die **Einheitengruppe von** $\mathbb{R}$ genannt wird. Dies wird uns insbesondere erlauben, die Folgerungen aus den Axiomen (1)-(4) auf die Multiplikation analog anzuwenden.

**Folgerungen.**

(e) *Es gilt* $0 \cdot x = x \cdot 0 = 0$ *für alle* $x \in \mathbb{R}$. Denn für ein $x \in \mathbb{R}$ gilt

$$0 \cdot x + 0 \cdot x = (0 + 0) \cdot x = 0 \cdot x = 0 \cdot x + 0$$

nach dem Distributivgesetz in Axiom (9) und der Eigenschaft der Null. Durch Wegstreichen von $0 \cdot x$ (siehe Folgerung (d)) erhalten wir $0 \cdot x = 0$. Nach dem Kommutativgesetz in Axiom (8) folgt auch $x \cdot 0 = 0$ und die Behauptung ist gezeigt.

(f) *Es gilt* $(-1) \cdot x = -x$ *für alle* $x \in \mathbb{R}$. Denn für ein $x \in \mathbb{R}$ ist

$$\begin{aligned} x + (-1) \cdot x &= 1 \cdot x + (-1) \cdot x \\ &= (1 + (-1)) \cdot x = 0 \cdot x = 0, \end{aligned}$$

was wegen Folgerung (b) die gewünschte Aussage impliziert.

(g) *Es gilt, dass für jedes* $x \in \mathbb{R}^\times$ *auch jedes Element* $(x^{-1})$ *wie in Axiom* (6) *in* $\mathbb{R}^\times$ *liegt.* Denn wäre $x \in \mathbb{R}^\times$ mit $(x^{-1}) = 0$, so würde $1 = x \cdot (x^{-1}) = x \cdot 0 = 0$ gelten, was in Axiom (5) ausgeschlossen wurde.

(h) *„Multiplikatives Kürzen“ ist erlaubt: Sei* $x \in \mathbb{R}^\times$ *und eine Gleichung der Form* $x \cdot y = x \cdot z$ *für* $y, z \in \mathbb{R}$ *gegeben. Dann darf man* $x$ *wegstreichen und es gilt* $y = z$. In der Tat ist

$$\begin{aligned} y = 1 \cdot y = ((x^{-1}) \cdot x) \cdot y &= (x^{-1}) \cdot (x \cdot y) \\ &= (x^{-1}) \cdot (x \cdot z) = ((x^{-1}) \cdot x) \cdot z = 1 \cdot z = z. \end{aligned}$$

(i) *Es gibt keine „Keine Nullteiler“: Falls* $x \cdot y = 0$ *für zwei Elemente* $x, y \in \mathbb{R}$ *gilt, dann ist* $x = 0$ *oder* $y = 0$. Denn nimmt man an, dass $x \neq 0$ ist, so folgt aus der Gleichung $x \cdot y = 0 = x \cdot 0$, die nach Folgerung (e) und der Voraussetzung gilt, dass $y = 0$ ist nach Folgerung (h).

Wegen dem Assoziativgesetz in Axiom (7) schreiben wir anstelle von $x \cdot (y \cdot z)$ oder $(x \cdot y) \cdot z$ auch $x \cdot y \cdot z$ für $x, y, z \in \mathbb{R}$. Wir verwenden im Weiteren die Regel „Punkt- vor Strichrechnung“ und lassen den Punkt in der Multiplikation oft auch weg. Insbesondere werden wir das Distributivgesetz in Axiom (9) auch in der Form $(x + y)z = xz + yz$ für alle $x, y, z \in \mathbb{R}$ schreiben.

*Bemerkung.* Die Axiome (5)-(8) (gemeinsam mit den Folgerungen (g) und (i)) machen $\mathbb{R}^\times$ ausgestattet mit (der Einschränkung) der Multiplikation $\cdot : (\mathbb{R}^\times)^2 \to \mathbb{R}^\times$ zu einer kommutativen Gruppe: Nach Folgerung (i) ist die Multiplikation wohldefiniert, nach Axiom (5) existiert ein sogenanntes neutrales Element (hier die Eins, bei der Addition war es die Null), nach Axiom (6) und Folgerung (g) hat jedes Element ein multiplikatives Inverses in $\mathbb{R} \setminus \{0\}$. Das Assoziativgesetz (resp. das Kommutativgesetz) ist wegen Axiom (7) (resp. Axiom (8)) erfüllt. Insbesondere können wir die Folgerungen (a)-(c) übernehmen.

**Folgerungen.**

(j) *Das Einselement ist durch die Eigenschaft in Axiom* (5) *eindeutig bestimmt.*

(k) *Das (multiplikatives) Inverse* $x^{-1} \in \mathbb{R}^\times$ *ist für jedes Element* $x \in \mathbb{R}^\times$ *eindeutig durch* $x \cdot x^{-1} = 1$ *bestimmt.*

(l) *Für alle* $x \in \mathbb{R}^\times$ *gilt* $(x^{-1})^{-1} = x$.

**Wichtige Übung 2.3.**

*(i) Analysieren Sie das Argument in Bemerkung 2.1.1, das Folgerungen* (j), (k) *und* (l) *beweist.*

*(ii) Leiten Sie die Folgerungen* (j),(k) *und* (l) *direkt aus den Axiomen* (5)-(7) *ab, was in diesem Fall nicht viel langsamer als die Argumentation im ersten Teil der Übung ist, aber eine klare Wiederholung der Argumente in Folgerungen* (a)-(c) *darstellt.*

**Wichtige Übung 2.4.** *Seien* $x, y, z \in \mathbb{R}$.

*(i) Zeigen Sie, dass die Identität* $(-x)(-y) = xy$. *Überprüfen Sie auch, dass* $-x \in \mathbb{R}^\times$ *und* $(-x)^{-1} = -(x^{-1})$ *gilt, falls* $x \in \mathbb{R}^\times$ *ist.*

*(ii) Zeigen Sie, dass das Distributivgesetz für die Subtraktion*

$$x(y - z) = xy - xz$$

*gilt.*

Wir verwenden oft die Schreibweise des **Quotienten** $\frac{x}{y} = xy^{-1}$ für alle **Zähler** $x \in \mathbb{R}$ und **Nenner** $y \in \mathbb{R}^\times$. Die **Inverse** $\frac{1}{y} = y^{-1}$ von $y \in \mathbb{R}^\times$ nennen wir auch den **reziproken Wert** oder den **Kehrwert** von $y$.

**Wichtige Übung 2.5** (Rechenregeln für Quotienten)**.**

*(i) Für alle* $x, z \in \mathbb{R}$ *und* $y, w \in \mathbb{R}^\times$ *gilt* $\frac{x}{y} = \frac{z}{w}$ *genau dann, wenn* $xw = yz$.

*(ii) Für alle* $x, z \in \mathbb{R}$ *und* $y, w \in \mathbb{R}^\times$ *gilt*

$$\frac{x}{y}\frac{z}{w} = \frac{xz}{yw}.$$

*(iii) Für alle $x \in \mathbb{R}$ und $y, z, w \in \mathbb{R}^\times$ gilt*

$$\frac{\frac{x}{y}}{\frac{z}{w}} = \frac{xw}{yz}.$$

*(iv) Für alle $x, z \in \mathbb{R}$ und $y, w \in \mathbb{R}^\times$ gilt*

$$\frac{x}{y} + \frac{z}{w} = \frac{xw + yz}{yw}.$$

**Wichtige Übung 2.6.** *Wir definieren $a^2 = a \cdot a$ für alle $a \in \mathbb{R}$. Zeigen Sie die Gleichungen $(a+b)^2 = a^2 + 2ab + b^2$, $(a-b)^2 = a^2 - 2ab + b^2$, und $(a+b)(a-b) = a^2 - b^2$ für alle $a, b \in \mathbb{R}$.*

Die Axiome (1)-(9) werden auch die **Körperaxiome** genannt und machen also $\mathbb{R}$ zu einem **Körper**. Diese definieren die üblichen Rechenregeln und sind damit gut als Axiome geeignet. Die Folgerungen (a)-(l), Übung 2.2 und Übungen 2.4–2.6 gelten für beliebige Körper und nicht nur für die reellen Zahlen.

**Beispiel 2.7.** *Bevor wir zu den weiteren Axiomen der reellen Zahlen kommen, wollen wir noch weiterere Beispiele von Körpern und ein Gegenbeispiel geben:*

*(i) Die rationalen Zahlen $\mathbb{Q}$, welche man aus den axiomatisch definierten natürlichen Zahlen $\mathbb{N}$ konstruieren kann (siehe Abschnitt 1.5), bilden einen Körper.*

*(ii) Die ganzen Zahlen $\mathbb{Z}$ (siehe Abschnitt 1.5) bilden keinen Körper. (Warum nicht?)*

*(iii) Aus den rationalen Zahlen lassen sich viele weitere Körper bilden, zum Beispiel*

$$\mathbb{Q}(\sqrt{2}) := \left\{ a + b\sqrt{2} \mid a, b \in \mathbb{Q} \right\}$$

*mit den natürlichen Rechenoperationen, wobei $\sqrt{2}$ eine Lösung der Gleichung $x^2 = 2$ ist. (Wir werden dies nochmals genauer besprechen, siehe Abschnitt 2.1.4.)*

*(iv) Der kleinst mögliche Körper $\mathbb{F}_2$ besteht aus der Menge $\{\mathbf{0}, \mathbf{1}\}$ gemeinsam mit den Rechenoperationen der Addition*

$$\mathbf{0} +_{\mathbb{F}_2} \mathbf{1} = \mathbf{1} +_{\mathbb{F}_2} \mathbf{0} = \mathbf{1}, \quad \mathbf{0} +_{\mathbb{F}_2} \mathbf{0} = \mathbf{1} +_{\mathbb{F}_2} \mathbf{1} = \mathbf{0}$$

*und der Multiplikation*

$$\mathbf{0} \cdot_{\mathbb{F}_2} \mathbf{1} = \mathbf{1} \cdot_{\mathbb{F}_2} \mathbf{0} = \mathbf{0} \cdot_{\mathbb{F}_2} \mathbf{0} = \mathbf{0}, \quad \mathbf{1} \cdot_{\mathbb{F}_2} \mathbf{1} = \mathbf{1}.$$

*Hier sind $\mathbf{0}, \mathbf{1}$ nicht die gewöhnlichen Zahlen (Elemente von $\mathbb{R}$), sondern zwei Elemente der neuen Menge $\mathbb{F}_2$. Die Operationen $+_{\mathbb{F}_2}$ und $\cdot_{\mathbb{F}_2}$ sind zwei neue Operationen, so dass $\mathbb{F}_2$ mit diesen beiden Operationen die Körperaxiome erfüllt (siehe Abschnitt 2.8.2).*

(v) *Der Körper $\mathbb{F}_2$ kann auch als Quotient von $\mathbb{Z}$ bezüglich der Äquivalenzrelation $\equiv$, definiert durch*

$$m \equiv n \iff n - m \text{ ist gerade}$$

*für $m, n \in \mathbb{Z}$, konstruiert werden. Diese Definition kann man für jede Primzahl $p \in \mathbb{N}$ erweitern, um damit den Körper $\mathbb{F}_p$ zu definieren (siehe wiederum Abschnitt 2.8.2).*

Obige Beispiele zeigen, dass die „üblichen Rechenoperationen“ (das wären die Axiome (1)-(9)) von vielen Zahlensystemen erfüllt werden, oder präziser formuliert, dass es viele verschiedene Körper gibt. Wir sind an diesen Körpern hier(!) nicht weiter interessiert und schliessen sie aus, indem wir weitere Axiome einführen.

### 2.1.2 Angeordnete Körper

**Axiome (Anordnung).** *Die Relation $\leq$ auf $\mathbb{R}$ erfüllt die folgenden vier Axiome*

(10) *(Reflexivität)* $\forall x \in \mathbb{R} : x \leq x$

(11) *(Antisymmetrie)* $\forall x, y \in \mathbb{R} : \big((x \leq y \wedge y \leq x) \implies x = y\big)$

(12) *(Transitivität)* $\forall x, y, z \in \mathbb{R} : \big((x \leq y \wedge y \leq z) \implies x \leq z\big)$

(13) *(Linearität)* $\forall x, y \in \mathbb{R} : \big(x \leq y \vee y \leq x\big)$

Die Axiome (10)-(12) sind die Axiome einer **Ordnung** und zusammen mit Axiom (13) bilden sie die Axiome einer **linearen** (oder auch **totalen**) **Ordnung**. Damit die Relation $\leq$ auf dem Körper $\mathbb{R}$ nützlich ist, benötigen wir die folgenden Axiome, die die Relation mit der Körperstruktur koppeln:

**Axiome (Kompatibilität von $\leq$).** *Wir verlangen*

(14) *($\leq$ und $+$)* $\forall x, y, z \in \mathbb{R} : \big(x \leq y \implies x + z \leq y + z\big)$.

(15) *($\leq$ und $\cdot$)* $\forall x, y \in \mathbb{R} : \big((0 \leq x \wedge 0 \leq y) \implies 0 \leq x \cdot y\big)$.

Wie bereits erwähnt wurde, sprechen wir $x \leq y$ als „$x$ ist **kleiner gleich** $y$“ aus. Wir definieren für $x, y \in \mathbb{R}$ auch $y \geq x$ durch $x \leq y$ und sprechen dies als „$y$ ist **grösser gleich** $x$“ aus. Weiter definieren wir $x < y$ (ausgesprochen als „$x$ ist **kleiner** als $y$“ oder „$x$ ist **echt kleiner** als $y$“) durch $x \leq y \wedge x \neq y$. Natürlich definieren wir $x > y$ durch $y < x$ und sagen „$x$ ist **grösser** als $y$“ oder „$x$ ist **echt grösser** als $y$“. Wir verwenden diese Symbole oft auch in „gleich gerichteten Ketten“; beispielsweise steht $x \leq y < z = a$ für $x \leq y \wedge y < z \wedge z = a$. Ein Element $x \in \mathbb{R}$ ist **positiv**, falls $x > 0$ gilt, und **negativ**, falls $x < 0$ gilt. Des Weiteren sagen wir ein Element $x \in \mathbb{R}$ ist **nichtnegativ** falls $x \geq 0$, beziehungsweise **nichtpositiv** falls $x \leq 0$.

VO: 02.10.17

**Folgerungen.** *Das Hinzufügen der Axiome* (10)-(15) *hat folgende Konsequenzen:*

(m) *(Trichotomie) Für alle $x, y \in \mathbb{R}$ gilt entweder $x < y$, $x = y$ oder $x > y$.* Seien $x, y \in \mathbb{R}$. Nach der Linearität in Axiom (13) gilt $x \leq y$ oder $y \leq x$. Falls $x = y$, dann können $x > y$ und $x < y$ nicht gelten (siehe obige Definitionen). Falls umgekehrt $x \neq y$, dann kann nach der Antisymmetrie in Axiom (11) nur eine der beiden Aussagen $x \leq y$ und $y \leq x$ gelten, wodurch wiederum genau eine der beiden Aussagen $x < y$ und $y < x$ gilt.

(n) *Seien $x, y, z \in \mathbb{R}$. Falls $x < y$ und $y \leq z$ ist, dann gilt auch $x < z$.* Denn wir haben $x \leq z$ nach der Transitivität in Axiom (12) und falls $x = z$ wäre, dann wäre $y \leq x$ und daher $x = y$ nach der Antisymmetrie in Axiom (11) und der Voraussetzung $x \leq y$, was aber der Annahme widerspricht. *Analog sieht man, dass $x \leq y$ und $y < z$ für $x, y, z \in \mathbb{R}$ auch $x < z$ impliziert.*

(o) *Man darf Ungleichungen folgendermassen addieren: Seien $x, y, z, w \in \mathbb{R}$ mit $x \leq y$ und $z \leq w$. Dann gilt auch $x + z \leq y + w$.* In der Tat, $x \leq y$ impliziert $x + z \leq y + z$ nach der additiven Kompatibilität in Axiom (14) und $z \leq w$ impliziert $y + z \leq y + w$ ebenso nach Axiom (14), was gemeinsam $x + z \leq y + w$ nach der Transitivität in Axiom (12) impliziert. *Analog sieht man (unter Verwendung von Folgerung* (n)*), dass für $x, y, z, w \in \mathbb{R}$ mit $x < y$ und $z \leq w$ auch $x + z < y + w$ gilt.*

(p) *Seien $y, z \in \mathbb{R}$. Dann gilt $y \leq z$ genau dann, wenn $0 \leq z - y$ gilt.* Dies folgt wiederum aus der additiven Kompatibilität in Axiom (14) durch Subtraktion resp. Addition von $y$.

(q) *Sei $x \in \mathbb{R}$. Dann gilt $x \geq 0 \iff -x \leq 0$.* Dies folgt aus (p) mit $y = -x$ und $z = 0$ gemeinsam mit Folgerung (c).

(r) *Des Weiteren ist für jedes $x \in \mathbb{R}$ das Element $x^2 = x \cdot x \geq 0$ und $x^2 > 0$, falls $x \neq 0$.* Falls $x \geq 0$ ist, so folgt die erste Aussage aus der multiplikativen Kompatibilität in Axiom (15). Falls $x \leq 0$ ist, dann ist $-x \geq 0$ nach Folgerung (q) und damit $xx = (-x)(-x) \geq 0$ nach Übung 2.4. Falls $x^2 = 0$ ist, dann gilt $x = 0$ nach Folgerung (i) und die zweite Aussage folgt.

(s) *Es gilt $0 < 1$.* Denn $1 = 1^2 \geq 0$ nach Folgerung (r) und $1 \neq 0$ nach Axiom (5).

(t) *Seien $x, y, z \in \mathbb{R}$. Falls $x \geq 0$ und $y \leq z$, dann gilt $xy \leq xz$.* Denn unter Verwendung von Folgerung (p), wonach $z - y \geq 0$, und der multiplikativen Kompatibilität in Axiom (15) gilt $xz - xy = x(z - y) \geq 0$ und damit folgt die Aussage wiederum aus Folgerung (p).

(u) *Seien $x, y, z \in \mathbb{R}$. Falls $x \leq 0$ und $y \leq z$, dann gilt $xy \geq xz$.* In der Tat ist $-x \geq 0$ nach Folgerung (q), $z - y \geq 0$ nach Folgerung (p) und somit

$$xy - xz = x(y - z) = (-x)(-(y - z)) = (-x)(z - y) \geq 0$$

nach der multiplikativen Kompatibilität in Axiom (15) und Übungen 2.4 und 2.4.

(v) *Für $x, y \in \mathbb{R}$ impliziert $0 < x \leq y$, dass $0 < y^{-1} \leq x^{-1}$.* Wir behaupten zuerst, dass $x^{-1} > 0$ ($y^{-1} > 0$ folgt analog). Denn falls nicht, dann wäre wegen der Trichotomie in Folgerung (m) und Folgerung (g) $x^{-1} < 0$. Demnach würde $1 = xx^{-1} < 0$ nach

Folgerung (t) gelten, was Folgerung (s) widerspricht. Insbesondere ist $x^{-1}y^{-1} > 0$ und es gilt

$$y^{-1} = xx^{-1}y^{-1} \leq yx^{-1}y^{-1} = x^{-1}.$$

(w) *Falls* $0 \leq x \leq y$ *und* $0 \leq z \leq w$ *für* $x, y, z, w \in \mathbb{R}$*, dann gilt auch* $0 \leq xz \leq yw$ *(siehe Übung 2.8).*

(x) *In einer Ungleichung der Form* $x + y \leq x + z$ *für* $x, y, z \in \mathbb{R}$ *darf man* $x$ *streichen, das heisst,* $y \leq z$ *folgern (siehe Übung 2.8).*

(y) *In einer Ungleichung der Form* $xy \leq xz$ *für* $x, y, z \in \mathbb{R}$ *darf man* $x$ *streichen, das heisst,* $y \leq z$ *folgern, wenn* $x > 0$ *ist (siehe Übung 2.8).*

**Übung 2.8.** *(i) Beweisen Sie die Folgerungen* (w),(x),(y)*. Was geschieht in* (y)*, wenn man die Bedingung* $x > 0$ *fallen lässt, das heisst, wenn* $x < 0$ *oder* $x = 0$*?*

*(ii) Formulieren Sie für einige der obigen Folgerungen ähnliche Versionen für die strikte Relation „<“ und beweisen Sie diese.*

Obige Axiome, Folgerungen und Aussagen in den Übungen stellen die üblichen Eigenschaften für Ungleichungen dar. Mit Hilfe dieser können wir auch Aufgaben wie in folgender Übung lösen.

**Übung 2.9.** *Zeigen Sie, dass*

$$\left\{x \in \mathbb{R} \setminus \{0\} \mid x + \frac{3}{x} + 4 \geq 0\right\} = \{x \in \mathbb{R} \setminus \{0\} \mid -3 \leq x \leq -1 \vee x > 0\}$$

Falls ein Körper (der ja per Definition die Axiome (1)-(9) erfüllt) eine Relation $\leq$ besitzt, die auch die Axiome (10)-(15) erfüllt, dann nennen wir den Körper mit der Relation einen **angeordneten** (oder **geordneten**) **Körper**.

**Beispiel 2.10** (angeordnete Körper)**.**

*(i) Es gibt keine Relation* $\leq$ *auf* $\mathbb{F}_2$*, so dass dieser einen angeordneten Körper bildet.* Nehmen wir per Widerspruch an, dass $\leq$ eine Relation auf $\mathbb{F}_2$ ist, die die Axiome $(10)-(15)$ erfüllt. Dann folgt aus $0 < 1$ und Folgerung (n) die strikte Ungleichung $0 = 0 + 0 < 1 + 1 = 0$, was einen Widerspruch darstellt.

*(ii) Die rationalen Zahlen* $\mathbb{Q}$ *sowie der Körper* $\mathbb{Q}(\sqrt{2})$ *aus Beispiel 2.7 bilden mit der üblichen Relation* $\leq$ *einen angeordneten Körper.*

### 2.1.3 Das Vollständigkeitsaxiom

Für die Analysis sind die Axiome (1)-(15) noch nicht ausreichend; Grund dafür ist, dass man vorerst noch zu viele „Lücken“ in $\mathbb{R}$ haben könnte. Wir benötigen also noch ein weiteres Axiom. Gewissermassen hat die Suche nach diesem Axiom mit den Arbeiten der Griechen

wie Pythagoras, Euklid und Archimedes begonnen, doch wurde sie erst im 19. Jahrhundert in den Arbeiten zahlreicher Mathematiker, darunter Weierstrass, Heine, Cantor und Dedekind, erfolgreich (siehe auch diesen Link).

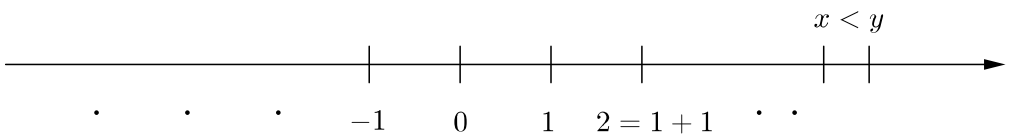
Wie wir in Kürze besprechen wollen, ist dieses Axiom trotzdem relativ leicht vorstellbar und wie wir im Laufe des Jahres sehen werden, ist es Grundlage für die ganze Analysis.

**Axiom (Vollständigkeit).** *Die reellen Zahlen erfüllen folgendes Axiom:*

(16) *Zuerst in Worten: Falls $X, Y$ zwei nicht-leere Teilmengen von $\mathbb{R}$ sind und für alle $x \in X$ und $y \in Y$ die Ungleichung $x \leq y$ gilt, dann gibt es ein $c \in \mathbb{R}$, das zwischen $X$ und $Y$ liegt in dem Sinn, als dass für alle $x \in X$ und $y \in Y$ die Ungleichung $x \leq c \leq y$ gilt. Formal:*

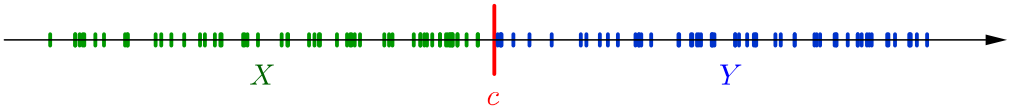
$$\forall X, Y \subseteq \mathbb{R} : \big((X \neq \emptyset \wedge Y \neq \emptyset \wedge \forall x \in X \ \forall y \in Y : x \leq y) \\ \implies (\exists c \in \mathbb{R} \forall x \in X \ \forall y \in Y : x \leq c \leq y)\big)$$

Wenn $\mathbb{R}$ die Axiome (1)–(16) erfüllt, dann sprechen wir auch von einem **vollständig angeordneten Körper**. Wir werden uns die reellen Zahlen häufig als die Punkte auf einer Geraden vorstellen, wobei wir deswegen die Gerade auch die **Zahlengerade** nennen.




Die Relation $x < y$ für $x, y \in \mathbb{R}$ interpretieren wir als „auf der Geraden liegt der Punkt $y$ rechts von dem Punkt $x$“, wobei wir „rechts“ mit einem Pfeil auf der Geraden andeuten. Wir definieren $2 = 1 + 1 > 1 > 0$, $3 = 2 + 1 > 2$ und so weiter - siehe auch den Abschnitt 2.2.

Was bedeutet in diesem Bild das Vollständigkeitsaxiom? Seien $X, Y$ nicht-leere Teilmengen von $\mathbb{R}$, so dass für alle $x \in X$ und alle $y \in Y$ die Ungleichung $x \leq y$ gilt. Dann sind alle Elemente von $X$ links von allen Elementen von $Y$ wie im nachfolgenden Bild.




Nach dem Vollständigkeitsaxiom existiert also ein $c$, das dazwischen liegt. Die Existenz der Zahl $c$ ist gewissermassen eine Versicherung, dass $\mathbb{R}$ keine „Lücken“ hat.

Es ist gut, sich die obigen Axiome und Folgerungen, aber auch alle folgenden Lemmata, Propositionen, Sätze, Theoreme und deren Beweise auf der Zahlengeraden zu veranschaulichen. Doch sollte die Zahlengerade als Motivation und zur Entwicklung einer guten Intuition, aber nicht für die Beweisführung verwendet werden.

### 2.1.4 Eine erste Anwendung der Vollständigkeit

Wir schliessen diesen Abschnitt indem wir als eine Anwendung des Vollständigkeitsaxioms die Wurzelfunktion auf $\mathbb{R}_{\geq 0} = \{x \in \mathbb{R} : x \geq 0\}$ einführen.

**Wichtige Übung 2.11** (Existenz der Wurzelfunktion)**.** *In dieser Übung wollen wir die Existenz einer bijektiven Funktion* $\sqrt{\cdot} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0},\ a \mapsto \sqrt{a}$ *mit der Eigenschaft* $\sqrt{a}^2 = a$ *für alle* $a \in \mathbb{R}_{\geq 0}$ *zeigen.*

*(i) Zeigen Sie für alle* $x, y \in \mathbb{R}_{\geq 0}$*, dass* $x < y \iff x^2 < y^2$*.*

*(ii) (Eindeutigkeit) Folgern Sie, dass es für jedes* $a \in \mathbb{R}_{\geq 0}$ *höchstens ein* $c \in \mathbb{R}_{\geq 0}$ *mit* $c^2 = a$ *gibt.*

*(iii) (Existenz) Betrachten Sie für eine reelle Zahl* $a > 0$ *die nicht-leeren Teilmengen*

$$X = \left\{x \in \mathbb{R}_{\geq 0} \mid x^2 < a\right\}, \quad Y = \left\{y \in \mathbb{R}_{\geq 0} \mid y^2 > a\right\}.$$

*Wenden Sie nun das Vollständigkeitsaxiom an, um ein* $c \in \mathbb{R}$ *mit* $x \leq c \leq y$ *für alle* $x \in X$ *und* $y \in Y$ *zu finden. Verwenden Sie, dass für alle* $\varepsilon \in \mathbb{R}$ *mit* $0 < \varepsilon < 1$ *die Aussagen* $c + \varepsilon \notin X$*,* $c - \varepsilon \notin Y$ *gelten und schliessen Sie jeweils auf* $c^2 \geq a$ *beziehungsweise* $c^2 \leq a$*.*

*Wir bezeichnen für jedes* $a \geq 0$ *die durch* $c^2 = a$ *und* $c \geq 0$ *eindeutig bestimmte reelle Zahl als* $c = \sqrt{a}$ *und sprechen von der* ***Wurzel von*** $a$*.*

*(iv) (Wachsend) Zeigen Sie für* $x, y \in \mathbb{R}_{\geq 0}$ *mit* $x < y$ *die Ungleichung* $\sqrt{x} < \sqrt{y}$*.*

*(v) (Bijektion) Zeigen Sie, dass die Wurzelfunktion von* $\mathbb{R}_{\geq 0}$ *nach* $\mathbb{R}_{\geq 0}$ *bijektiv ist.*

*(vi) (Multiplikativität) Zeigen Sie unter Verwendung von (ii), dass für alle* $x, y \in \mathbb{R}_{\geq 0}$ *gilt* $\sqrt{xy} = \sqrt{x}\sqrt{y}$*.*

*(vii) (Zwei Lösungen) Zeigen Sie, dass es für* $a > 0$ *genau zwei Lösungen der Gleichung* $x^2 = a$ *in* $x \in \mathbb{R}$ *gibt.*

### 2.1.5 Verwendung der reellen Zahlen und der Axiome

Zusammenfassend gilt, dass die Körperaxiome der reellen Zahlen die üblichen Rechenregeln und Gleichungsumformungen erlauben, wobei (wie gewohnt) Division mit Null nicht gestattet ist. Des Weiteren erfüllen die Relationen $\leq$ und $<$ die üblichen Umformungsgesetze für Ungleichungen, wobei bei Multiplikation mit negativen Zahlen die Ungleichungen natürlich umzudrehen sind. Wir werden diese Gesetze (das wären die Axiome (1)-(15), die Folgerungen (a)-(y) und die Aussagen in den Übungen) im Folgenden ohne Verweis verwenden. Das Vollständigkeitsaxiom (Axiom (16)) war bereits notwendig für den Beweis der Existenz einer Wurzelfunktion. Die wahre Bedeutung dieses Axioms werden wir hingegen erst sehen, wenn wir es für weitere Aussagen verwenden. Insbesondere werden wir bis auf Weiteres stets darauf verweisen, wenn wir es verwenden.

Wir werden häufig die Variablen $a, b, c, s, t, x, y$ verwenden um damit reelle Zahlen zu bezeichnen, werden aber im Sinne der Transparenz trotzdem immer „Sei $a \in \mathbb{R} \ldots$“ oder ähnliches schreiben.

*Bemerkung.* Wir haben in obigem die Axiome der reellen Zahlen aufgelistet. Die Tatsache, dass wir von den reellen Zahlen sprechen können, rührt daher, dass es bis auf gewisse Identifikationen nur einen angeordneten Körper gibt, der auch (16) genügt. Eine Analogie dazu findet sich im Schachspiel: Ein Schachbrett mit Schachfiguren ist nicht gleich einem anderen Schachbrett mit Schachfiguren. Für das Schachspiel ist es jedoch egal, welches Schachbrett man benutzt. Wir werden die erwähnte Eindeutigkeit etwas später genauer formulieren und auch beweisen können; für den Moment fixieren wir uns aber einen solchen Körper und nennen ihn den Körper der reellen Zahlen (wir einigen uns auf ein Schachbrett mit den dazugehörigen Figuren).

VO: 04.10.17

## 2.2 Die natürlichen Zahlen

Da wir alle unsere Diskussionen auf den Axiomen der reellen Zahlen in Abschnitt 2.1 aufbauen werden, wollen wir jetzt die natürlichen, die ganzen und die rationalen Zahlen innerhalb der reellen Zahlen finden und die wichtigsten elementaren und geometrischen Eigenschaften dieser Zahlen beweisen.

### 2.2.1 Definition der natürlichen Zahlen und vollständige Induktion

**Definition 2.12** (Induktive Teilmengen)**.** Eine Teilmenge $M \subseteq \mathbb{R}$ ist **induktiv**, falls folgende zwei Eigenschaften gelten:

(i) $1 \in M$

(ii) Für alle $x \in \mathbb{R}$ gilt $x \in M \implies x + 1 \in M$.

Beispielsweise ist $\mathbb{R}$ eine induktive Menge (gewissermassen die grösste solche). Die „kleinste“ induktive Menge sollen die natürlichen Zahlen sein.

**Definition 2.13** (Natürliche Zahlen)**.** Wir definieren die Teilmenge der **natürlichen Zahlen** $\mathbb{N} \subseteq \mathbb{R}$ als Durchschnitt aller induktiven Teilmengen von $\mathbb{R}$

$$\mathbb{N} = \bigcap_{M \subseteq \mathbb{R} \text{ induktiv}} M.$$

Aus der Definition folgt unmittelbar, dass $\mathbb{N}$ in jeder induktiven Teilmenge von $\mathbb{R}$ enthalten ist und dass $1 \in \mathbb{N}$, da jede induktive Teilmenge die Eins enthalten muss und $\mathbb{N}$ der Durchschnitt aller induktiven Teilmengen ist. Des Weiteren können wir folgern, dass für alle $n \in \mathbb{N}$ die Ungleichung $n \geq 1$ gilt. In der Tat ist die Teilmenge $\{x \in \mathbb{R} \mid x \geq 1\}$ induktiv (überprüfen Sie dies) und enthält somit $\mathbb{N}$.

**Lemma 2.14** (Kleinste induktive Menge)**.** *Die natürlichen Zahlen $\mathbb{N}$ bilden eine induktive und somit die kleinste induktive Teilmenge der reellen Zahlen.*

*Beweis.* Wir haben oben bereits gesehen, dass $1 \in \mathbb{N}$ ist. Falls nun $n \in \mathbb{N}$ ist und $M \subseteq \mathbb{R}$ eine beliebige induktive Teilmenge ist, dann gilt auch $n \in M$ (wegen der Definition von $\mathbb{N}$). Da $M$ induktiv ist, gilt $n + 1 \in M$. Da $M$ aber eine beliebige induktive Teilmenge war, liegt $n + 1$ in jeder induktiven Teilmenge und somit auch in $\mathbb{N}$ per Definition von $\mathbb{N}$. Wir haben für $\mathbb{N}$ also beide Eigenschaften einer induktiven Teilmenge nachgewiesen und das Lemma folgt. □

Wir können nun das Prinzip der vollständigen Induktion als Konsequenz unserer Definition der natürlichen Zahlen (und der Axiome der reellen Zahlen) beweisen.

**Satz 2.15** (Vollständige Induktion)**.** *Falls für eine Aussage $A(n)$ über die natürlichen Zahlen $n \in \mathbb{N}$*

- *(Induktionsanfang) $A(1)$ und*
- *(Induktionsschritt) $\forall n \in \mathbb{N} : (A(n) \implies A(n+1))$*

*gelten, dann gilt $A(n)$ für alle $n \in \mathbb{N}$.*

*Beweis.* Wir definieren $E = \{n \in \mathbb{N} \mid A(n)\}$, womit folgende Aussagen gelten.

- $1 \in E$, da $A(1)$ auf Grund des Induktionsanfanges gilt.
- Für alle $x \in \mathbb{R}$ gilt, dass $x \in E$ nach Definition $x \in \mathbb{N}$ und auf Grund des Induktionsschrittes auch $x + 1 \in E$ impliziert.

Daher ist $E$ eine induktive Menge und es folgt, dass $\mathbb{N} \subseteq E$ nach Definition von $\mathbb{N}$. Also gilt $A(n)$ für alle natürlichen Zahlen $n \in \mathbb{N}$. □

**Übung 2.16** (Peano-Axiome)**.** *Zeigen Sie, dass die oben definierte Teilmenge $\mathbb{N} \subseteq \mathbb{R}$ die Peano-Axiome (siehe Abschnitt 1.5) erfüllt, wobei $\nu : n \in \mathbb{N} \mapsto n + 1 \in \mathbb{N}$ die Nachfolgerfunktion ist.*

Wir untersuchen nun weitere algebraische und geometrische Eigenschaften von $\mathbb{N}$. Die Bedeutung der folgenden Diskussionen liegt nicht so sehr in den behaupteten Aussagen, die anschaulich klar sind. Vielmehr zeigen Sie, dass unsere Axiome von $\mathbb{R}$ und unsere Definition von $\mathbb{N}$ auch in der Lage sind, diese natürlichsten Eigenschaften von $\mathbb{N}$ zu beweisen, die wiederum Grundlage für die weiteren Diskussionen bilden.

**Lemma 2.17** (Addition und Multiplikation auf $\mathbb{N}$)**.** *Für alle $m, n \in \mathbb{N}$ gilt $m + n \in \mathbb{N}$ und $m \cdot n \in \mathbb{N}$.*

*Beweis.* Sei $A(n)$ die Aussage $\forall m \in \mathbb{N} : m + n \in \mathbb{N}$. Dann gilt $A(1)$, denn falls $m \in \mathbb{N}$, dann gilt auch $m + 1 \in \mathbb{N}$, da $\mathbb{N}$ induktiv ist wegen Lemma 2.14. Dies ist der Induktionsanfang. Für den Induktionsschritt nehmen wir also an, dass $A(n)$ für $n \in \mathbb{N}$ gilt oder in anderen Worten, dass für alle $m \in \mathbb{N}$ auch $m + n \in \mathbb{N}$ gilt. Wegen Lemma 2.14 impliziert letzteres aber auch $m + n + 1 \in \mathbb{N}$ für alle $m \in \mathbb{N}$ und wir erhalten die Aussage $A(n + 1)$. Vollständige Induktion zeigt daher $\forall n \in \mathbb{N} : A(n)$, was gerade die Aussage $\forall m, n \in \mathbb{N} : m + n \in \mathbb{N}$ ist.

Für die Multiplikation definieren wir $B(n)$ für $n \in \mathbb{N}$ als die Aussage $\forall m \in \mathbb{N} : m \cdot n \in \mathbb{N}$. Dann gilt $B(1)$, da für alle $m \in \mathbb{N}$ auch $m \cdot 1 = m \in \mathbb{N}$. Falls nun $B(n)$ für $n \in \mathbb{N}$ gilt, dann folgt aus $m \in \mathbb{N}$ auch $m \cdot n \in \mathbb{N}$ und aus dem ersten Teil des Lemmas auch

$$m \cdot (n + 1) = m \cdot n + m \in \mathbb{N}$$

Da $m$ beliebig war, gilt also $B(n) \implies B(n + 1)$ und das Lemma folgt mittels vollständiger Induktion. □

Nachdem wir im letzten Lemma einige algebraische Fragen beantwortet haben, wollen wir uns nun geometrischen Fragen zuwenden. Da $\mathbb{N}$ induktiv ist, sind 1 und $2 = 1 + 1$ in $\mathbb{N}$. Gibt es eine natürliche Zahl zwischen 1 und 2? Die negative Antwort zu dieser Frage ist in allgemeinerer Form in folgendem Lemma enthalten.

**Lemma 2.18** (Anordnung von $\mathbb{N}$)**.**

- *Für $n \in \mathbb{N}$ gilt $n = 1$ oder $n - 1 \in \mathbb{N}$.*

- *Für $m, n \in \mathbb{N}$ mit $m \leq n \leq m+1$ gilt $n = m$ oder $n = m+1$.*

*Beweis.* Für die erste Aussage zeigen wir, dass die Menge $M = \{1\} \cup \{n \in \mathbb{N} \mid n-1 \in \mathbb{N}\}$ die natürlichen Zahlen $\mathbb{N}$ enthält. In der Tat ist die Menge $M$ induktiv, da $1 \in M$ und da für $n \in M$ auch $(n+1)-1 = n \in \mathbb{N}$ und damit $n+1 \in M$ gilt. Nach Definition von $\mathbb{N}$ ist also $\mathbb{N} \subseteq M$ wie gewünscht.

Für die zweite Behauptung definieren wir für $n \in \mathbb{N}$ die Aussage $A(n)$ durch

$$\forall m \in \mathbb{N} : ((m \leq n \leq m+1) \implies n \in \{m, m+1\}).$$

Dann gilt $A(1)$, denn falls $m \in \mathbb{N}$ die Ungleichung $m \leq 1 \leq m+1$ erfüllt, dann gilt wegen $m \geq 1$ auch $m = 1 = n$.

Angenommen es gilt nun $A(n)$ für ein $n \in \mathbb{N}$ und wir wollen $A(n+1)$ zeigen. Sei also $m \in \mathbb{N}$, so dass $m \leq n+1 \leq m+1$ gilt. Falls $m = 1$ ist, dann gilt $1 \leq n+1 \leq 2 = 1+1$ und damit $n \leq 2-1 = 1$. Wegen $n \geq 1$ folgt $n = 1 = m$ und somit $n+1 = m+1$. Falls aber $m \neq 1$ ist, dann ist wegen der ersten Behauptung $m-1 \in \mathbb{N}$ und $m-1 \leq n \leq m$. Da wir aber $A(n)$ angenommen haben, gilt $n \in \{m-1, m\}$ und daher $n+1 \in \{m, m+1\}$.

Wir haben also den Induktionsanfang $A(1)$ und den Induktionsschritt $A(n) \implies A(n+1)$ für ein beliebiges $n$ gezeigt. Daher gilt $A(n)$ für alle $n \in \mathbb{N}$ und das Lemma folgt. □

Wie bereits im Abschnitt 1.6.3 kurz erwähnt haben, gibt es mehrere Versionen der vollständigen Induktion.

**Satz 2.19** (Vollständige Induktion)**.** *Falls für eine Aussage $A(n)$ über die natürlichen Zahlen $n \in \mathbb{N}$ die Aussage*

- *(Induktion)* $\forall n \in \mathbb{N} : \Big(\big(\forall k \in \mathbb{N} : (k < n \implies A(k))\big) \implies A(n)\Big)$

*erfüllt ist, dann gilt $A(n)$ für alle $n \in \mathbb{N}$*

*Beweis.* Wir definieren eine Aussage $B(n)$ für natürliche Zahlen $n \in \mathbb{N}$ durch

$$\forall k \in \mathbb{N} : k \leq n \implies A(k).$$

Mit vollständiger Induktion (siehe Satz 2.15) und der Anordnung von $\mathbb{N}$ (wie in Lemma 2.18) möchten wir nun zeigen, dass $B(n)$ für alle $n \in \mathbb{N}$ gilt. Insbesondere folgt damit, dass $A(n)$ für alle $n \in \mathbb{N}$ gilt (wieso?), was den Beweis des Satzes abschliessen wird.

Wir zeigen zuerst den Induktionsanfang, also dass die Aussage $B(1)$ gilt. Da aber $k = 1$ die einzige natürliche Zahl mit $k \leq 1$ ist, genügt es, die Aussage $A(1)$ zu verifizieren. Hierfür verwenden wir die Annahme im Satz für $n = 1$, also die Aussage

$$\big(\forall k \in \mathbb{N} : (k < 1 \implies A(k))\big) \implies A(1).$$

Da es keine natürlichen Zahlen kleiner 1 gibt, ist für jedes $k \in \mathbb{N}$ die Aussage $k < 1$ falsch, womit $(k < 1 \implies A(k))$ richtig ist. Also gilt die Vorraussetzung der im Satz angenommenen Implikation für $n = 1$, und es folgt $A(1)$ (und damit $B(1)$ wie gewünscht).

Sei nun $n \in \mathbb{N}$ gegeben. Wir wollen den Induktionsschritt $B(n) \implies B(n+1)$ beweisen. Also nehmen wir an, dass $B(n)$ bereits gilt. Die Aussage $B(n+1)$ ist durch

$$\forall k \in \mathbb{N} : k \leq n+1 \implies A(k)$$

gegeben. Für $k \in \mathbb{N}$ ist $k < n+1$ auf Grund von Lemma 2.18 äquivalent zu $k \leq n \vee k = n$. Die Aussage $B(n)$ ist damit zu

$$\forall k \in \mathbb{N} : k < n+1 \implies A(k)$$

äquivalent. Wegen der Annahme im Satz angewandt auf $n+1$ impliziert dies aber $A(n+1)$, was auf Grund obiger Äquivalenz gemeinsam mit $B(n)$ die Aussage $B(n+1)$ zeigt. Dies schliesst den Induktionsschritt und damit den Beweis des Satzes ab. □

Die vollständige Induktion in der Version von Satz 2.19 erlaubt uns im Induktionsschritt statt der Annahme, dass die Aussage bloss für die vorhergehende natürliche Zahl gilt, die stärkere Annahme, dass die Aussage bereits für alle echt kleineren natürlichen Zahlen gilt, zu verwenden.

**Übung 2.20** (Versteckter Induktionsanfang)**.** *In der Version der vollständigen Induktion in Satz 2.19 scheint es keinen Induktionsanfang zu geben. Wie kann das sein? Wo ist der Induktionsanfang versteckt?*

Wir bemerken, dass man die Induktion auch verwenden kann, um zum Beispiel für ein vorgebenes $n_0 \in \mathbb{N}$ eine Aussage für alle natürliche Zahlen $n \geq n_0$ zu zeigen. In diesem Fall würde man als Induktionsanfang die Aussage für $n = n_0$ beweisen und im Induktionsschritt für eine natürliche Zahl $n \geq n_0$ annehmen.

**Wichtige Übung 2.21** (Varianten der vollständigen Induktion)**.** *Folgern Sie aus Satz 2.15 oder aus Satz 2.19 die folgenden Varianten der vollständigen Induktion. Sei hierzu $A(n)$ eine beliebige Aussage über natürliche Zahlen $n \in \mathbb{N}$*

*(i) Angenommen die Aussagen*

- *(Induktionsanfang) $A(1)$ und $A(2)$*
- *(Induktionsschritt) $\forall n \in \mathbb{N} : (A(n) \wedge A(n+1) \implies A(n+2))$,*

*gelten, dann gilt ebenso $A(n)$ für alle $n \in \mathbb{N}$.*

*(ii) Falls für ein $n_0 \in \mathbb{N}$ die Aussagen*

- *(Induktionsanfang) $A(n_0)$*
- *(Induktionsschritt) $\forall n \in \mathbb{N} : ((n \geq n_0 \wedge A(n)) \implies A(n+1))$*

*gelten, dann gilt auch $A(n)$ für alle natürliche Zahlen $n \geq n_0$.*

**Satz 2.22** (Wohlordnung der natürlichen Zahlen)**.** *Sei $M \subseteq \mathbb{N}$ eine nicht-leere Teilmenge. Dann hat $M$ ein eindeutig bestimmtes kleinstes Element, das heisst*

$$\exists! n_0 \in M \ \forall n \in M : n \geq \ n_0.$$

Die Existenz eines kleinsten Elements zu jeder nicht-leeren Teilmenge ist etwas, was die natürlichen Zahlen auszeichnet und beispielsweise von den reellen Zahlen nicht erfüllt ist. Die Teilmenge der positiven Zahlen $\{x \in \mathbb{R} \mid x > 0\}$ oder $\mathbb{R}$ selbst sind konkrete Beispiele von Teilmengen, die kein kleinstes Element haben. (Wieso?)

*Beweis.* Die Eindeutigkeit eines solchen kleinsten Elements folgt direkt: Sind $n_0, n_0' \in M$ zwei kleinste Elemente, dann gilt $n_0' \geq n_0$, da $n_0$ ein kleinstes Element ist und $n_0 \geq n_0'$, da $n_0'$ ein kleinstes Element ist. Also gilt $n_0 = n_0'$.

Um die Existenz eines kleinsten Elements zu zeigen, verwenden wir die Kontraposition. Wir nehmen also an, dass $M$ kein kleinstes Element hat, und wollen zeigen, dass $M$ leer ist. Hierzu definieren wir für alle $n \in \mathbb{N}$ die Aussage $A(n)$ durch $n \notin M$.

Sei $n \in \mathbb{N}$. Dann bedeutet die Aussage $\forall k \in \mathbb{N} : k < n \implies A(k)$ genau, dass es unterhalb von $n$ keine Elemente in $M$ gibt. Da wir angenommen haben, dass $M$ kein kleinstes Element hat, sehen wir, dass $n$ nicht in $M$ liegen kann. Also gilt

$$\big(\forall k \in \mathbb{N} : k < n \implies A(k)\big) \implies A(n)$$

für jedes $n \in \mathbb{N}$. Die vollständige Induktion in Satz 2.19 zeigt nun, dass $A(n)$ für alle $n \in \mathbb{N}$ gilt. Damit ist $M$ die leere Menge. □

**Lemma 2.23** (Subtraktion in $\mathbb{N}$)**.** *Für alle $m, n \in \mathbb{N}$ mit $m < n$ gilt $n - m \in \mathbb{N}$.*

*Beweis.* Sei $A(n)$ für $n \in \mathbb{N}$ die Aussage

$$\forall m \in \mathbb{N} : m < n \implies n - m \in \mathbb{N}.$$

Dann gilt $A(1)$, denn es existiert kein $m \in \mathbb{N}$ mit $m < 1$. Angenommen $A(n)$ gilt für ein $n \in \mathbb{N}$ und sei $m \in \mathbb{N}$ mit $m < n + 1$. Nach Lemma 2.18 ist entweder $m = n$ oder $m < n$. Im ersten Fall gilt $(n + 1) - m = 1 \in \mathbb{N}$. Im zweiten Fall gilt $(n + 1) - m = (n - m) + 1 \in \mathbb{N}$ nach $A(n)$. Also gilt $A(n)$ für alle $n \in \mathbb{N}$ nach vollständiger Induktion. □

Wir definieren die **nicht-negativen ganzen Zahlen** als $\mathbb{N}_0 = \mathbb{N} \sqcup \{0\}$. Diese stellen auf natürliche Weise die Kardinalitäten der endlichen Mengen dar, wobei die natürlichen Zahlen die Kardinalitäten der endlichen, nicht-leeren Mengen darstellen. Summen und Produkte sind in diesem Zusammenhang auch wichtig:

**Wichtige Übung 2.24** (Kardinalität des kartesischen Produkts von endlichen Mengen)**.** *Seien $m, n \in \mathbb{N}$. Zeigen Sie per Induktion über $n$, dass das kartesische Produkt*

$$\{1, \ldots, m\} \times \{1, \ldots, n\}$$

*Kardinalität $mn$ hat. Der Ausdruck $\{1, \ldots, m\}$ ist dabei eine Abkürzung für $\{k \in \mathbb{N} \mid k \leq m\}$ und hat Kardinalität $m$.*

*Bemerkung.* Wir werden auch des öfteren eine Funktion auf $\mathbb{N}$ durch **Rekursion** definieren. Dies bedeutet, dass man die Funktion $f$ auf 1 definiert indem man $f(1)$ konkret angibt, und dann eine Rekursionsbedingung (zum Beispiel eine Formel) festlegt wie $f(n+1)$ aus $f(1), \ldots, f(n)$ bestimmt wird (wobei man $f(1), \ldots, f(n)$ als bereits definiert annimmt). Dass es höchstens eine Funktion auf $\mathbb{N}$ gibt, die beide Bedingungen erfüllt, lässt sich durch einen Induktionsbeweis schnell beweisen. (Nehmen Sie an, dass $f$ und $\tilde{f}$ sowohl $f(1) = \tilde{f}(1)$ als auch die Rekursionsbedingung erfüllen und beweisen sie $f(n) = \tilde{f}(n)$ für alle $n \in \mathbb{N}$ mittels Satz 2.19.)

Streng formal ist die Existenz etwas aufwendiger. Dazu beweist man zuerst mittels vollständiger Induktion die Aussage: "Für alle $n \in \mathbb{N}$ gibt es eine eindeutig bestimmte Funktion $f_n$ auf $\{k \in \mathbb{N} \mid k \leq n\}$, die $f_n(1) = f(1)$ und die Rekursionsbedingung erfüllt." Insbesondere gilt dann für natürliche Zahlen $m \leq n$, dass die Einschränkung der Funktion $f_n$ auf die Menge $\{k \in \mathbb{N} \mid k \leq m\}$ dieselben Gesetze wie $f_m$ erfüllt und somit gilt $f_m(k) = f_n(k)$ für alle natürlichen Zahlen $k \leq m \leq n$. Wir definieren $f$ auf $\mathbb{N}$ durch $f(k) = f_n(k)$ für $k \in \mathbb{N}$ und ein $n \in \mathbb{N}$ mit $k \leq n$ und sehen (wieder mittels vollständiger Induktion), dass diese Funktion die Rekursionbedingungen erfüllt. Man kommt nicht umhin, diesen Beweis mit dem rekursiven Algorithmus zur Berechnung von $f(n)$ für $n \in \mathbb{N}$ zu vergleichen (der ja zur Berechnung von $f(n)$ im Allgemeinen ebenso auch $f(1), f(2), \ldots, f(n-1)$ berechnen müsste).

### 2.2.2 Die ganzen Zahlen

Die **ganzen Zahlen** sind als Teilmenge von $\mathbb{R}$ durch

$$\mathbb{Z} = \mathbb{N} \sqcup \{0\} \sqcup \{-n \mid n \in \mathbb{N}\} = \mathbb{N}_0 \sqcup -\mathbb{N}$$

definiert.

**Lemma 2.25** (Addition und Multiplikation auf $\mathbb{Z}$)**.** *Die ganzen Zahlen sind unter Addition und Multiplikation abgeschlossen, das heisst, für alle $m, n \in \mathbb{Z}$ gilt $m + n \in \mathbb{Z}$ und $mn \in \mathbb{Z}$.*

*Beweis.* Für die Multiplikation sieht man dies sehr direkt: Falls $m, n \in \mathbb{N}$, dann gilt offenbar $mn = (-m)(-n) \in \mathbb{N} \subseteq \mathbb{Z}$ und $(-m)n = m(-n) = -mn \in -\mathbb{N} \subseteq \mathbb{Z}$ nach Lemma 2.17. Falls $m$ oder $n$ Null ist, gilt ebenso $mn = 0 \in \mathbb{Z}$.

Für die Addition verwenden wir die Eigenschaften von $\mathbb{N}$ in Lemma 2.17 und Lemma 2.23. Seien $m, n \in \mathbb{Z}$. Falls $m$ oder $n$ Null sind, gibt es nichts zu zeigen. Seien also $m, n \in \mathbb{N}$. Dann ist $m + n \in \mathbb{N} \subseteq \mathbb{Z}$ und $-m - n = -(m + n) \in -\mathbb{N} \subseteq \mathbb{Z}$. Falls $n > m$, dann ist $n - m \in \mathbb{N} \subseteq \mathbb{Z}$ und $-n + m = -(n - m) \in \mathbb{Z}$. Analoges gilt falls $n < m$. Falls $n = m$, ist $n - m = 0 \in \mathbb{Z}$. Dies deckt alle Möglichkeiten ab und das Lemma folgt. □

**Wichtige Übung 2.26** (Anordnung von $\mathbb{Z}$)**.** *Verallgemeinern Sie Lemma 2.18 von $\mathbb{N}$ auf $\mathbb{Z}$. Das heisst, zeigen Sie, dass für $m, n \in \mathbb{Z}$ die Ungleichung $m \leq n \leq m + 1$ ebenso $n = m$ oder $n = m + 1$ impliziert.*

### 2.2.3 Die rationalen Zahlen

Die **rationalen Zahlen** sind definiert als die Teilmenge von Quotienten

$$\mathbb{Q} = \left\{ \frac{m}{n} \mid m \in \mathbb{Z},\ n \in \mathbb{N} \right\} \subseteq \mathbb{R}.$$

**Lemma 2.27** (Rationale Zahlen)**.** *Die rationalen Zahlen bilden einen Unterkörper von $\mathbb{R}$, das heisst, für alle $r, s \in \mathbb{Q}$ gilt $-r, r+s, rs \in \mathbb{Q}$ und auch $r^{-1} \in \mathbb{Q}$, falls $r \neq 0$.*

**Wichtige Übung 2.28.** *Beweisen Sie Lemma 2.27.*

**Wichtige Übung 2.29.** *Zeigen Sie, dass sowohl die ganzen Zahlen als auch die rationalen Zahlen abzählbar unendlich sind.*

Eine reelle Zahl $x \in \mathbb{R}$ heisst **irrational**, falls $x \notin \mathbb{Q}$. An dieser Stelle könnte man sich fragen, ob es überhaupt irrationale Zahlen gibt und wenn ja, wieviele. Um die erste Frage zu beantworten, werden wir die Wurzelfunktion (siehe Übungen 2.11) verwenden, mit welcher man aus rationalen Zahlen irrationale konstruieren kann.

**Lemma 2.30** (Quadratwurzel aus 2)**.** *Die reelle Zahl $\sqrt{2}$ ist irrational. Insbesondere erfüllen die rationalen Zahlen nicht das Vollständigkeitsaxiom.*

VO: 05.10.17

Man kann Lemma 2.30 als einen Grund sehen, wieso es nicht ausreichend ist, nur rationale Zahlen zu betrachten. Eine visuelle Veranschaulichung des folgenden Beweises des Lemmas findet sich in diesem Video.

*Beweis.* Wir nehmen per Widerspruch an, dass $\sqrt{2}$ rational ist und schreiben $2 = (\frac{m}{n})^2$ für $m \in \mathbb{N}$ und das kleinst mögliche $n \in \mathbb{N}$ (dies ist nach Satz 2.22 möglich). Insbesondere gilt also $2n^2 = m^2$ und folglich

$$2(m-n)^2 = 2m^2 - 4mn + 2n^2 = 4n^2 - 4nm + m^2 = (2n-m)^2$$

Also gilt $\left(\frac{2n-m}{m-n}\right)^2 = 2$. Da $0 < m-n < m$, erhalten wir einen kleineren Nenner, der verwendet werden kann, um $\sqrt{2}$ darzustellen. Dies widerspricht der minimalen Wahl von $n$. Somit ist $\sqrt{2}$ irrational.

Für den Beweis der letzte Aussage bemerken wir zuerst, dass $\mathbb{Q}$ die Körperaxiome auf Grund von Lemma 2.27 erfüllt. Des Weiteren gelten für $\mathbb{Q}$ die Axiome der Anordnung und die Axiome der Verträglichkeit der Anordnung und Körperoperationen, da diese für $\mathbb{R}$ gelten, womit $\mathbb{Q}$ ein angeordneter Körper ist. Das einzig verbleibende Axiom ist das Vollständigkeitsaxiom. Da wir zum Beweis der Existenz der Quadratwurzel positiver reeller Zahlen nur die Axiome der reellen Zahlen in Abschnitt 2.1 verwendet haben, folgt, dass $\mathbb{Q}$ das Vollständigkeitsaxiom nicht erfüllt. □

Wir werden im Abschnitt 2.6.4 zeigen, dass die reellen Zahlen überabzählbar sind. Vergleicht man diese Tatsache mit Übung 2.29, so kommt man zum Schluss, dass es viel mehr irrationale als rationale Zahlen gibt.

### 2.2.4 Division mit Rest und Anfänge der Zahlentheorie*

VO: Abschnitt 2.2.4 im Kolloquium 09-11.10.17

**Satz 2.31** (Division mit Rest)**.** *Für alle $n \in \mathbb{N}_0$ und $d \in \mathbb{N}$ gibt es ein $q \in \mathbb{N}_0$ und ein $r \in \mathbb{N}_0$ mit $r < d$, welches wir den Rest nennen, so dass $n = qd + r$.*

*Beweis.* Für $n < d$ stimmt die Behauptung, da wir dann $q = 0$ und $r = n$ wählen können. Genauso stimmt sie für $n = d$, da wir dann $q = 1$ und $r = 0$ wählen können. Nehmen wir nun an, dass der Satz nicht zutrifft. Dann gibt es nach der Wohlordnung von $\mathbb{N}$ in Satz 2.22 ein kleinstes $n_0 \in \mathbb{N}$, für das die Division durch ein $d \in \mathbb{N}_0$ mit Rest nicht funktioniert. Nach obigem muss $n_0 > d \geq 1$ und damit auch $n_0 \geq 2$ gelten.

Insbesondere ist $n = n_0 - 1 \in \mathbb{N}$ und es gibt einen Rest $r \in \mathbb{N}_0$ mit $r < d$, so dass $n_0 - 1 = n = qd + r$ für ein $q \in \mathbb{N}_0$. Damit gilt $n_0 = qd + r + 1$. Falls $r < d - 1$, dann ist $r + 1 < d$ und $n_0$ erfüllt doch Division durch $d$ mit Rest. Falls $r = d - 1$, dann ist $d = r + 1$ und $n_0 = qd + r + 1 = qd + d = (q + 1)d + 0$ und $n_0$ erfüllt Division durch $d$ mit Rest 0. Nach der Anordnung von $\mathbb{N}$ in Lemma 2.18 erfüllt $r$ entweder $r < d - 1$ oder $r = d - 1$ und daher wurden alle Möglichkeiten für $r$ abgedeckt. Für $n_0$ ist Division durch $d$ mit Rest daher möglich, was einen Widerspruch darstellt. Also gilt der Satz. □

Wir wollen hier für Interessierte kurz andeuten, was Division mit Rest mit Begriffen wie Primzahlen, Primfaktorzerlegung, etc. zu tun hat. Da eine ausführliche Besprechung uns aber zu weit vom Thema Analysis ablenken würde, begnügen wir uns mit einer Skizze anhand einer Serie von Übungsaufgaben. Des Weiteren verweisen wir auf die Algebra 1-Vorlesung im dritten Semester des Mathematikstudiums und das Buch [RU08] für mehr Details.

Wir sagen, dass eine Zahl $d \in \mathbb{Z}$ eine Zahl $n \in \mathbb{Z}$ **teilt** und schreiben $d|n$, falls es ein $q \in \mathbb{Z}$ gibt, so dass $qd = n$. Eine natürliche Zahl $p > 1$ ist **prim** oder eine **Primzahl**, falls für alle $a, b \in \mathbb{N}$ die Implikation $p|ab \implies (p|a \vee p|b)$ zutrifft. Eine natürliche Zahl $p > 1$ heisst **irreduzibel**, falls sie nicht als Produkt $p = ab$ für $a, b \in \mathbb{N}$ mit $a > 1$ und $b > 1$ geschrieben werden kann (das heisst, ausser 1 und $p$ keine Teiler hat). Wir haben bereits in einer Übung in Abschnitt 1.7.7 gesehen, dass wir jede Zahl $n \in \mathbb{N}$ als ein Produkt von irreduziblen Zahlen darstellen können. (Die für den Beweis dieser Übung notwendige Form der vollständigen Induktion haben wir inzwischen in der Form von Satz 2.19 nachgeliefert.) Aber wie zeigt man, dass diese Produktzerlegung (bis auf die Reihenfolge der Faktoren) eindeutig bestimmt ist?

**Übung 2.32.**

*(a) Zeigen Sie, dass jede Primzahl auch irreduzibel ist.*

*(b) Nehmen Sie kurz an, dass Sie bereits wissen, dass eine Zahl irreduzibel ist genau dann, wenn sie prim ist. Zeigen Sie, dass die Primfaktorzerlegung bis auf Reihenfolge der Faktoren eindeutig bestimmt ist.*

*Der **grösste gemeinsame Teiler** zweier natürlichen Zahlen $m, n \in \mathbb{N}$ ist die grösste natürliche Zahl $d = \gcd(m, n) \in \mathbb{N}$ mit $d|m$ und $d|n$. Wir wollen zeigen, dass es $a, b \in \mathbb{Z}$ gibt mit $d = am + bn$. Dies nennt sich auch das Lemma von Bézout oder der Euklidsche Algorithmus.*

(c) *Führen Sie für $m > n$ Division von $m$ durch $n$ mit Rest durch. Sei $r \in \mathbb{N}_0$ der Rest. Zeigen Sie für $r \in \mathbb{N}$, dass $\gcd(m,n) = \gcd(n,r)$ und für $r = 0$, dass $\gcd(m,n) = n$.*

(d) *Schliessen Sie per Induktion auf die Aussage, dass der grösste gemeinsame Teiler wie oben beschrieben als Summe von Vielfachen dargestellt werden kann.*

*Wir zeigen nun, dass irreduzible Zahlen auch prim sind.*

(e) *Angenommen $p \in \mathbb{N}$ ist irreduzibel und seien $m, n \in \mathbb{N}$ mit $p|mn$, aber $p \nmid m$ (also $\neg(p|m)$). Zeigen Sie, dass es $a, b \in \mathbb{Z}$ gibt mit $1 = ap+bm$. Folgern Sie, dass $p|bmn$, dass $p|(1-ap)n$ und dass $p|n$.*

*Fassen Sie obige Diskussionen in der Form der eindeutigen Primfaktorzerlegung von natürlichen Zahlen zusammen.*

Unter Verwendung der Tatsache, dass irreduzible Zahlen prim sind (und umgekehrt), lässt es sich etwas einfacher zeigen, dass die Wurzel aus 2 (oder jeder anderen Primzahl) irrational ist. Wir überlassen das wiederum als Übung (siehe auch dieses Lied).

Wir möchten an dieser Stelle noch kurz erwähnen, dass vielleicht überraschenderweise die tiefgreifende Untersuchung von Primzahlen viele Methoden der Analysis (und auf jeden Fall alle Methoden dieser Analysis-Vorlesung) voraussetzt. Für eine Andeutung dieser Tatsache und einen historischen Exkurs verweisen wir auf den Podcast der BBC.

### 2.2.5 Verwendung der ganzen Zahlen und deren Eigenschaften

Die algebraischen und geometrischen Aussagen in diesem Abschnitt über die natürlichen, die ganzen und die rationalen Zahlen stellen bloss die Standardeigenschaften dieser Zahlen dar. Deswegen werden wir die oben bewiesenen Lemmata und Sätze im Folgenden meist ohne Referenz verwenden. Dies gilt ebenso für die vollständige Induktion in Satz 2.15.

Wir werden im Abschnitt 2.6.1 zwei weitere grundlegende Eigenschaften von $\mathbb{Z}$ respektive $\mathbb{Q}$ beweisen, die die Geometrie von $\mathbb{Z}$ und $\mathbb{Q}$ als Teilmengen von $\mathbb{R}$ beschreiben.

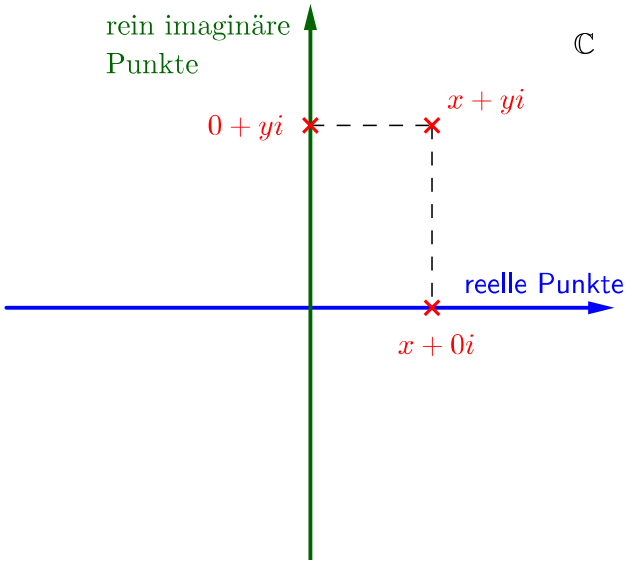
Wir werden oft die Variablen $j, k, \ell, m, n$ für natürliche oder ganze Zahlen verwenden. Weiters verwenden wir meist die Variable $r$ für rationale Zahlen.

## 2.3 Die komplexen Zahlen

Unter Verwendung der reellen Zahlen können wir die Menge der komplexen Zahlen als

$$\mathbb{C} = \mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}$$

definieren. Wir schreiben ein Element $z = (x, y) \in \mathbb{C}$ viel häufiger in der Form $z = x + y\mathrm{i}$, wobei das Symbol i als die **imaginäre Einheit** bezeichnet wird. Man beachte, dass bei dieser Identifikation $+$ vorerst als Ersatz für das Komma zu verstehen ist. Die Zahl $x \in \mathbb{R}$ wird als der **Realteil** von $z$ bezeichnet und man schreibt $x = \mathrm{Re}(z)$; die Zahl $y = \mathrm{Im}(z) \in \mathbb{R}$ ist der **Imaginärteil** von $z$. Die Elemente von $\mathbb{C}$ mit Imaginärteil 0 bezeichnet man auch als **reell** und die Elemente mit Realteil 0 als **rein imaginär**. Via der injektiven Abbildung $x \in \mathbb{R} \mapsto x + 0\mathrm{i} \in \mathbb{C}$ identifizieren wir $\mathbb{R}$ mit der Teilmenge der reellen Elemente von $\mathbb{C}$ (der „$x$-Achse“).




Die Menge $\mathbb{C}$ (inklusive deren graphische Darstellung wie oben) wird ganz im Sinne der Identifikation $\mathbb{C} = \mathbb{R}^2$ auch **komplexe Ebene** (alternativ **Gausssche Zahlenebene** oder auch **Argand-Ebene**) genannt. In der geometrischen Denkweise wird die Menge der reellen Punkte als die **reelle Achse** und die Menge der rein imaginären Punkte als die **imaginäre Achse** bezeichnet.

Wie Sie vielleicht schon erwartet haben, soll i eine Wurzel von $-1$ sein. Formal ausgedrückt, wollen wir, dass $\mathbb{C}$ einen Körper darstellt, in dem die Rechenoperationen von $\mathbb{R}$ „verallgemeinert“ werden, und dass $\mathrm{i}^2 = \mathrm{i} \cdot \mathrm{i} = -1$ gilt. Die Addition auf $\mathbb{C}$ definieren wir „komponentenweise“ durch

$$(x_1 + y_1\mathrm{i}) + (x_2 + y_2\mathrm{i}) = (x_1 + x_2) + (y_1 + y_2)\mathrm{i}$$

für $x_1, x_2, y_1, y_2 \in \mathbb{R}$. Die Multiplikation auf $\mathbb{C}$ definieren wir hingegen durch

$$(x_1 + y_1\mathrm{i}) \cdot (x_2 + y_2\mathrm{i}) = (x_1x_2 - y_1y_2) + (x_1y_2 + y_1x_2)\mathrm{i}$$

für $x_1, x_2, y_1, y_2 \in \mathbb{R}$. Insbesondere gilt $(0+1\mathrm{i})^2 = -1+0\mathrm{i}$ und die Addition und Multiplikation auf $\mathbb{C}$ erweitern die entsprechenden Operationen auf $\mathbb{R}$.

**Proposition 2.33** (Komplexe Zahlen)**.** *Mit den oben definierten Verknüpfungen definiert $\mathbb{C}$ einen Körper, den* ***Körper der komplexen Zahlen****. Hierbei ist die Null gleich $0 + 0\mathrm{i}$ und die Eins gleich $1 + 0\mathrm{i}$.*

Für die Geschichte der komplexen Zahlen verweisen wir auf den Podcast der BBC (zum Beispiel ab der 14. oder 20. Minute).

**Übung 2.34.** *Wäre $\mathbb{R}^2$ mit obiger Addition und mit der (komponentenweisen) Multiplikation definiert durch $(a, b) \times (c, d) = (ac, bd)$ für $a, b, c, d \in \mathbb{R}$ auch ein Körper? Genauer: Welche Körperaxiome gelten in diesem Fall?*

*Beweis von Proposition 2.33.* Wir verifizieren die Körperaxiome. Wie wir sehen werden, folgen die Eigenschaften der Addition auf $\mathbb{C}$ aus den Eigenschaften der Addition auf $\mathbb{R}$. Die Addition ist kommutativ: Seien $x_1, x_2, y_1, y_2 \in \mathbb{R}$. Dann gilt

$$\begin{aligned}(x_1 + y_1\mathrm{i}) + (x_2 + y_2\mathrm{i}) &= (x_1 + x_2) + (y_1 + y_2)\mathrm{i}\\ &= (x_2 + x_1) + (y_2 + y_1)\mathrm{i}\\ &= (x_2 + y_2\mathrm{i}) + (x_1 + y_1\mathrm{i}).\end{aligned}$$

Das Element $0 + 0\mathrm{i}$ ist ein (und schlussendlich also das) Nullelement der Addition, denn

$$(0 + 0i) + (x + y\mathrm{i}) = (0 + x) + (0 + y)\mathrm{i} = x + y\mathrm{i}$$

für alle $x, y \in \mathbb{R}$. Die additive Inverse eines Elements $x + y\mathrm{i}$ für $x, y \in \mathbb{R}$ ist $(-x) + (-y)\mathrm{i}$, denn

$$\begin{aligned}(x + y\mathrm{i}) + ((-x) + (-y)\mathrm{i}) &= ((-x) + (-y)\mathrm{i}) + (x + y\mathrm{i})\\ &= (x + (-x)) + (y + (-y))\mathrm{i} = 0 + 0\mathrm{i}.\end{aligned}$$

Die Addition ist assoziativ: Seien $x_i, y_i \in \mathbb{R}$ für $i \in \{1, 2, 3\}$. Dann gilt

$$\begin{aligned}&((x_1 + y_1\mathrm{i}) + (x_2 + y_2\mathrm{i})) + (x_3 + y_3\mathrm{i})\\ &\quad= ((x_1 + x_2) + (y_1 + y_2)\mathrm{i}) + (x_3 + y_3\mathrm{i})\\ &\quad= (x_1 + x_2 + x_3) + (y_1 + y_2 + y_3)\mathrm{i}\\ &\quad= \ldots = (x_1 + y_1\mathrm{i}) + ((x_2 + y_2\mathrm{i}) + (x_3 + y_3\mathrm{i})).\end{aligned}$$

Die Eigenschaften der Multiplikation fordern etwas mehr Aufwand. Wir zeigen zuerst, dass die Multiplikation kommutativ ist. Für $x_1, x_2, y_1, y_2 \in \mathbb{R}$ haben wir

$$\begin{aligned}(x_1 + y_1\mathrm{i}) \cdot (x_2 + y_2\mathrm{i}) &= (x_1x_2 - y_1y_2) + (x_1y_2 + y_1x_2)\mathrm{i}\\ &= (x_2x_1 - y_2y_1) + (x_2y_1 + y_2x_1)\mathrm{i}\\ &= (x_2 + y_2\mathrm{i}) \cdot (x_1 + y_1\mathrm{i}).\end{aligned}$$

Das Element $1 + 0\mathrm{i}$ ist ein Einselement, denn $1 + 0\mathrm{i} \neq 0 + 0i$ und für $x, y \in \mathbb{R}$ gilt

$$\begin{aligned}(1 + 0\mathrm{i}) \cdot (x + y\mathrm{i}) &= (x + y\mathrm{i}) \cdot (1 + 0\mathrm{i}) \\ &= (x \cdot 1 - y \cdot 0) + (x \cdot 0 + y \cdot 1)\mathrm{i} = x + y\mathrm{i}.\end{aligned}$$

Wir geben nun die multiplikative Inverse eines Elements $x + y\mathrm{i} \in \mathbb{C}$, wobei $x, y \in \mathbb{R}$ und $x + y\mathrm{i} \neq 0 + 0i$ (das heisst $x \neq 0$ oder $y \neq 0$), an. Wir bemerken zuerst, dass $x^2 + y^2 > 0$: Nehmen wir vorerst an, dass $x \neq 0$, dann ist $x^2 > 0$ und $y^2 \geq 0$ und damit $x^2 + y^2 > 0$. Für $y \neq 0$ gilt ebenso $x^2 \geq 0$ und $y^2 > 0$ und damit $x^2 + y^2 > 0$. Die multiplikative Inverse ist gegeben durch $\frac{x}{x^2+y^2} + \frac{-y}{x^2+y^2}\mathrm{i}$, denn

$$\begin{aligned}&(x + y\mathrm{i}) \cdot \left(\frac{x}{x^2 + y^2} + \frac{-y}{x^2 + y^2}\mathrm{i}\right) \\ &\quad = \left(x \cdot \frac{x}{x^2 + y^2} - y \cdot \frac{-y}{x^2 + y^2}\right) + \left(y \cdot \frac{x}{x^2 + y^2} + x \cdot \frac{-y}{x^2 + y^2}\right)\mathrm{i} \\ &\quad = 1 + 0\mathrm{i}\end{aligned}$$

Die verbleibenden beiden Axiome (Assoziativität der Multiplikation und Distributivität) lassen sich durch abstraktere Argumente beweisen, die aber auch etwas mehr Wissen benötigen. Wir bestätigen diese Axiome deswegen durch zwei konkrete Rechnungen.

Die Multiplikation ist assoziativ: Seien $x_i, y_i \in \mathbb{R}$ für $i \in \{1, 2, 3\}$. Nun berechnet man

$$\begin{aligned}&((x_1 + y_1\mathrm{i}) \cdot (x_2 + y_2\mathrm{i})) \cdot (x_3 + y_3\mathrm{i}) \\ &\quad = ((x_1x_2 - y_1y_2) + (x_1y_2 + y_1x_2)\mathrm{i}) \cdot (x_3 + y_3\mathrm{i}) \\ &\quad = (x_1x_2x_3 - y_1y_2x_3 - x_1y_2y_3 - y_1x_2y_3) \\ &\qquad\quad + (x_1y_2x_3 + y_1x_2x_3 + x_1x_2y_3 - y_1y_2y_3)\mathrm{i} \\ &\quad = (x_1 + y_1\mathrm{i}) \cdot ((x_2x_3 - y_2y_3) + (y_2x_3 + x_2y_3)\mathrm{i}) \\ &\quad = (x_1 + y_1\mathrm{i}) \cdot ((x_2 + y_2\mathrm{i}) \cdot (x_3 + y_3\mathrm{i}))\end{aligned}$$

Es bleibt nur noch die Distributivität: Seien also $x_i, y_i \in \mathbb{R}$ für $i \in \{1, 2, 3\}$. Dann gilt

$$\begin{aligned}&(x_1 + y_1\mathrm{i}) \cdot ((x_2 + y_2\mathrm{i}) + (x_3 + y_3\mathrm{i})) \\ &\quad = (x_1 + y_1\mathrm{i}) \cdot ((x_2 + x_3) + (y_2 + y_3)\mathrm{i}) \\ &\quad = (x_1x_2 + x_1x_3 - y_1y_2 - y_1y_3) + (y_1x_2 + y_1x_3 + x_1y_2 + x_1y_3)\mathrm{i} \\ &\quad = ((x_1x_2 - y_1y_2) + (y_1x_2 + x_1y_2)\mathrm{i}) + ((x_1x_3 - y_1y_3) + (y_1x_3 + x_1y_3)\mathrm{i}) \\ &\quad = (x_1 + y_1\mathrm{i}) \cdot (x_2 + y_2\mathrm{i}) + (x_1 + y_1\mathrm{i}) \cdot (x_3 + y_3\mathrm{i}),\end{aligned}$$

womit gezeigt wäre, dass $\mathbb{C}$ zusammen mit der oben definierten Addition und der oben definierten Multiplikation ein Körper ist. □

**Applet 2.35** (Komplexe Zahlen)**.** *Wir betrachten die Körperoperationen (Addition, Multiplikation, multiplikatives Inverse) auf den komplexen Zahlen. Die wahre geometrische Bedeutung*

*der Multiplikation und des multiplikativen Inversen lässt sich hier bereits erahnen, doch werden wir diese erst später besprechen.*

*Bemerkung* (Andere Konstruktionen der komplexen Zahlen).

(i) Wenn Sie die ersten Eigenschaften der Matrixmultiplikation kennen, lässt sich obiger Beweis deutlich vereinfachen. In diesem Fall lässt sich $x + y\mathrm{i} \in \mathbb{C}$ mit

$$\begin{pmatrix} x & -y \\ y & x \end{pmatrix} \in \mathrm{Mat}_{2,2}(\mathbb{R})$$

identfizieren. Die Multiplikation auf $\mathbb{C}$ entspricht dann der Multiplikation der Matrizen in $\mathrm{Mat}_{2,2}(\mathbb{R})$ und insbesondere folgt beispielsweise die Assoziativität der Multiplikation auf $\mathbb{C}$ aus der Assoziativität der Matrixmultiplikation. Gleiches gilt für die Distributivität. (Kommutativität der Multiplikation und die Existenz der multiplikativen Inversen müssen aber nach wie vor direkt überprüft werden, da diese beiden Eigenschaften im Allgemeinen nicht für die Matrixmultiplikation gelten.)

(ii) Ebenso lässt sich $\mathbb{C}$ aus $\mathbb{R}$ konstruieren, wenn man den Ring der Polynome mit reellen Koeffizienten kennt (siehe Abschnitt 3.2).

Wie schon zuvor angedeutet, wollen wir $\mathbb{R}$ als eine Teilmenge von $\mathbb{C}$ auffassen. Vielmehr nennt man $\mathbb{R} \subseteq \mathbb{C}$ auch einen Unterkörper, da Addition und Multiplikation auf $\mathbb{C}$ eingeschränkt auf $\mathbb{R}$ die Addition und Multiplikation auf $\mathbb{R}$ ergeben. Wir werden deswegen von nun an für alle $x \in \mathbb{R}$ kürzer $x = x + 0\mathrm{i}$ und $x\mathrm{i} = 0 + x\mathrm{i}$ schreiben. Insbesondere wollen wir auch $1 = 1 + 0\mathrm{i}$, $0 = 0 + 0\mathrm{i}$ und $\mathrm{i} = 0 + 1\mathrm{i}$ schreiben. Per Definition der Multiplikation gilt nun $\mathrm{i}^2 = -1$ wie gewünscht.

Wir wollen ebenso bemerken, dass die komplexen Zahlen keinen angeordneten Körper bilden – unabhängig davon, welche Ordnung man auf $\mathbb{C}$ wählt. Angenommen es gäbe eine Ordnung $\leq_{\mathbb{C}}$, so dass $\mathbb{C}$ mit $\leq_{\mathbb{C}}$ ein angeordneter Körper ist. In einem angeordneten Körper sollte $-1 <_{\mathbb{C}} 0$ gelten, was aber $-1 = \mathrm{i}^2 \geq_{\mathbb{C}} 0$ widerspricht (siehe Abschnitt 2.1).

An dieser Stelle möchten wir uns kurz fragen, wieso die komplexen Zahlen überhaupt von Interesse sind. Während eine der schönen Eigenschaften der reellen Zahlen deren vollständige Ordnung ist, so zeichnen sich die komplexen Zahlen unter anderem durch algebraische Schönheit aus. Auf $\mathbb{C}$ hat nicht nur die Gleichung $x^2 + 1 = 0$ eine Lösung, sondern auch jede andere Gleichung der Form $a_n x^n + ... + a_1 x + a_0 = 0$ für $n \in \mathbb{N}$ und $a_0, a_1, ..., a_n \in \mathbb{C}$ mit $a_n \neq 0$ und $n > 0$. Diese Tatsache („$\mathbb{C}$ ist algebraisch abgeschlossen“) ist Inhalt des sogenannten Fundamentalsatzes der Algebra, den wir im zweiten Semester beweisen werden. Intuitiv sollte man in Analogie zu „$\mathbb{R}$ ist vollständig, da $\mathbb{R}$ keine Lücken hat“ den Fundamentalsatz der Algebra lesen als „$\mathbb{C}$ hat algebraisch keine Lücken“.

Zum Abschluss dieses ersten Exkurses in das Reich der komplexen Zahlen wollen wir die komplexe Konjugation definieren. Diese ist im Wesentlichen nichts anderes als eine Spiegelung um die reelle Zahlengerade (und wird zum Beispiel in der Linearen Algebra in der Untersuchung von komplexen inneren Produkten unentbehrlich sein).

**Definition 2.36** (Konjugation)**.** Die **komplexe Konjugation** ist die Abbildung

$$^{-} : \mathbb{C} \to \mathbb{C},\ z = x + y\mathrm{i} \mapsto \bar{z} = x - y\mathrm{i}.$$

Im Beweis von Proposition 2.33 wurde die komplexe Konjugation indirekt schon verwendet: die multiplikative Inverse eines von Null verschiedenen Elements $x + y\mathrm{i} \in \mathbb{C}$ ist

$$(x + y\mathrm{i})^{-1} = \frac{x}{x^2 + y^2} - \frac{y}{x^2 + y^2}\mathrm{i} = \frac{x - y\mathrm{i}}{x^2 + y^2} = \frac{\overline{x + y\mathrm{i}}}{x^2 + y^2}.$$

**Lemma 2.37** (Eigenschaften der Konjugation)**.** *Die komplexe Konjugation erfüllt folgende Eigenschaften:*

*(i) Für alle $z \in \mathbb{C}$ ist $z\bar{z} \in \mathbb{R}$ und $z\bar{z} \geq 0$. Des Weiteren gilt für alle $z \in \mathbb{C}$, dass $z\bar{z} = 0$ genau dann, wenn $z = 0$.*

*(ii) Für alle $z, w \in \mathbb{C}$ gilt $\overline{z + w} = \overline{z} + \overline{w}$.*

*(iii) Für alle $z, w \in \mathbb{C}$ gilt $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$.*

*Beweis.* Wir überlassen der Leserin/dem Leser Teil $(i)$ als Übung. Seien $z = x_1 + y_1\mathrm{i}$ und $w = x_2 + y_2\mathrm{i} \in \mathbb{C}$ für $x_1, y_1, x_2, y_2 \in \mathbb{R}$. Dann gilt

$$\overline{z + w} = (x_1 + x_2) - (y_1 + y_2)\mathrm{i} = (x_1 - y_1\mathrm{i}) + (x_2 - y_2\mathrm{i}) = \overline{z} + \overline{w}$$

und

$$\overline{z \cdot w} = (x_1x_2 - y_1y_2) - (x_1y_2 + y_1x_2)\mathrm{i} = (x_1 - y_1\mathrm{i}) \cdot (x_2 - y_2\mathrm{i}) = \overline{z} \cdot \overline{w},$$

was zu zeigen war. □

**Wichtige Übung 2.38.** *Zeigen Sie (i) in Lemma 2.37.*

Wie schon angemerkt wurde, gelten die Folgerungen (a)-(l) in Abschnitt 2.1.1 für alle Körper und insbesondere auch für $\mathbb{C}$.

**Wichtige Übung 2.39.** *Zeigen Sie für alle $z, w \in \mathbb{C}$ die Rechenregeln*

$$\mathrm{Re}(z) + \mathrm{Re}(w) = \mathrm{Re}(z + w), \quad \mathrm{Im}(z) + \mathrm{Im}(w) = \mathrm{Im}(z + w)$$

*sowie*

$$\mathrm{Re}(zw) = \mathrm{Re}(z)\,\mathrm{Re}(w) - \mathrm{Im}(z)\,\mathrm{Im}(w), \quad \mathrm{Im}(zw) = \mathrm{Re}(z)\,\mathrm{Im}(w) + \mathrm{Re}(w)\,\mathrm{Im}(z).$$

**Übung 2.40.** *Zeigen Sie die Identitäten*

$$\mathrm{Re}(z) = \frac{z + \overline{z}}{2}, \quad \mathrm{Im}(z) = \frac{z - \overline{z}}{2\mathrm{i}}$$

*für alle $z \in \mathbb{C}$. Schliessen Sie insbesondere, dass $\mathbb{R} = \{z \in \mathbb{C} \mid z = \overline{z}\}$. Was bedeutet diese Gleichheit geometrisch?*

*Bemerkung.* Wie wir gesehen haben, lässt sich auf $\mathbb{C}$ keine Ordnung definieren, die zur Addition und zur Multiplikation kompatibel ist. Dennoch lässt sich auf den komplexen Zahlen Analysis betreiben, was zum Teil in diesem Kurs aber vor allem im Kurs „Funktionentheorie“ im zweiten Studienjahr des Mathematik- und Physikstudiums thematisiert wird. Grund dafür ist, dass $\mathbb{C}$ eine Verallgemeinerung des Vollständigkeitsaxiom erfüllt (welches wir erst nach etwas mehr Theorie besprechen können).

### 2.3.1 Verwendung der komplexen Zahlen

Unsere Konstruktion von $\mathbb{C}$ aus $\mathbb{R}$ mag etwas formal gewesen sein, doch muss man sich eigentlich nur merken, dass $\mathrm{i}^2 = -1$ und sonst alle gewöhnlichen Eigenschaften für die Addition und Multiplikation gelten. Sogar die Formel für das multiplikative Inverse von $z \in \mathbb{C}$ muss man nicht auswendig lernen wenn man sich stattdessen merkt, dass man den Bruch $\frac{1}{z}$ mit dem konjugierten Element $\overline{z}$ erweitert. Sie werden der komplexen Konjugation noch öfter und insbesondere in der Linearen Algebra-Vorlesung in der Diskussion von „inneren Produkten auf Vektorräumen über $\mathbb{C}$“ begegnen. Wir bemerken noch, dass wir manchmal die Variable $i$ (zum Beispiel als Indexvariable) verwenden werden. Man sollte dies allerdings vermeiden, wenn gleichzeitig komplexen Zahlen eine wesentliche Rolle in der Diskussion spielen.

Wir verwenden häufig die Variablen $z$ und $w$ für Elemente der komplexen Zahlen.

## 2.4 Intervalle und der Absolutbetrag

### 2.4.1 Intervalle

Wie bereits erwähnt, stellen wir $\mathbb{R}$ als die Zahlengerade dar. In diesem Bild entsprechen folgende Teilmengen Strecken auf dieser Geraden, wobei wir vier Möglichkeiten haben, je nachdem, ob man die Endpunkte in der Teilmenge haben will oder nicht.

**Definition 2.41** (Intervalle)**.** Seien $a, b \in \mathbb{R}$. Dann ist das **abgeschlossene Intervall** $[a, b]$ durch

$$[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\},$$

das **offene Intervall** $(a, b)$ durch

$$(a, b) = \{x \in \mathbb{R} \mid a < x < b\},$$

das **(rechts) halboffene Intervall** $[a, b)$ durch

$$[a, b) = \{x \in \mathbb{R} \mid a \leq x < b\}$$

und das **(links) halboffene Intervall** $(a, b]$ durch

$$(a, b] = \{x \in \mathbb{R} \mid a < x \leq b\}$$

definiert. Wenn das Intervall nicht-leer ist, dann wird $a$ der **linke Endpunkt**, $b$ der **rechte Endpunkt**, und $b - a$ die **Länge des Intervalls** genannt.

Wir möchten an dieser Stelle anmerken, dass beispielsweise die Intervalle $(a, b], [a, b), (a, b)$ für $a, b \in \mathbb{R}$ nicht-leer sind genau dann, wenn $a < b$, und $[a, b]$ nicht-leer ist genau dann, wenn $a \leq b$. Intervalle der Art $[a, b]$, $(a, b]$, $[a, b)$, $(a, b)$ für $a, b \in \mathbb{R}$ werden auch **endliche** oder **beschränkte Intervalle** genannt, wenn wir sie von folgenden Intervallen unterscheiden wollen.

**Definition 2.42** (Unbeschränkte Intervalle)**.** Für $a, b \in \mathbb{R}$ definieren wir die **unbeschränkten abgeschlossenen Intervalle**

$$\begin{aligned} [a, \infty) &= \mathbb{R}_{\geq a} = \{x \in \mathbb{R} \mid a \leq x\} \\ (-\infty, b] &= \mathbb{R}_{\leq b} = \{x \in \mathbb{R} \mid x \leq b\} \end{aligned}$$

und die **unbeschränkten offenen Intervalle**

$$\begin{aligned} (a, \infty) &= \mathbb{R}_{>a} = \{x \in \mathbb{R} \mid a < x\} \\ (-\infty, b) &= \mathbb{R}_{<b} = \{x \in \mathbb{R} \mid x < b\} \\ (-\infty, \infty) &= \mathbb{R} \end{aligned}$$

Statt runden Klammern werden manchmal auch umgedrehte eckige Klammern verwendet, um offene und halboffene Intervalle zu bezeichnen. Zum Beispiel findet man anstelle von $(a, b)$ für $a, b \in \mathbb{R}$ oft auch $]a, b[$ in der Literatur. Wir werden hier stets runde Klammern verwenden.

VO: 09.10.17

Der folgende Begriff wird für uns später sehr bedeutsam sein.

**Definition 2.43** (Umgebungen eines Punktes)**.** Sei $x \in \mathbb{R}$. Ein Menge, die ein offenes Intervall enthält, in dem $x$ liegt, wird auch eine **Umgebung** von $x$ genannt. Für ein $\delta > 0$ wird das offene Intervall $(x - \delta, x + \delta)$ die **$\delta$-Umgebung** von $x$ genannt.

Beispielsweise wäre also $\mathbb{Q} \cup [-1, 1]$ eine Umgebung von $0 \in \mathbb{R}$. Falls ein $y \in \mathbb{R}$ in einer $\delta$-Umgebung eines Punktes $x \in \mathbb{R}$ liegt für ein „kleines“ $\delta > 0$, so sagt man auch, dass $y$ „$\delta$-nahe“ an $x$ ist.

**Übung 2.44** (Verhalten von Intervallen unter Durchschnitt und Vereinigung)**.**

*(i) Zeigen Sie, dass ein endlicher Schnitt $\bigcap_{k=1}^{n} I_k$ von Intervallen $I_1, ..., I_n$ wieder ein Intervall ist (wobei die leere Menge auch als ein Intervall zugelassen ist). Können Sie die Endpunkte eines nicht-leeren Durchschnitts mittels der Endpunkte der ursprünglichen Intervalle beschreiben?*

*(ii) Wann ist eine Vereinigung von zwei Intervallen wieder ein Intervall? Was geschieht in diesem Fall, wenn man zwei Intervalle des selben Typs (offen, abgeschlossen, links halboffen, rechts halboffen) vereinigt?*

### 2.4.2 Der Absolutbetrag auf den reellen Zahlen

**Definition 2.45.** Der **Absolutbetrag** ist die Funktion

$$|\cdot| : \mathbb{R} \to \mathbb{R}, \quad x \mapsto |x| = \begin{cases} x & \text{falls } x \geq 0 \\ -x & \text{falls } x < 0 \end{cases} .$$

Wir betrachten zuerst einige Konsequenzen dieser Definition.

**Folgerungen.**

*(a) Für $x \in \mathbb{R}$ ist $|x| \geq 0$ und $|x| = 0$ genau dann, wenn $x = 0$.* Dies folgt aus der Trichotomie von reellen Zahlen: Für $x = 0$ gilt $|x| = 0$, für $x > 0$ gilt $|x| = x > 0$, und für $x < 0$ folgt $|x| = -x > 0$.

*(b) Es ist $|-x| = |x|$ für alle $x \in \mathbb{R}$.*

*(c) Die Absolutbetrag ist multiplikativ: $|xy| = |x||y|$ für alle $x \in \mathbb{R}$.* (Überprüfen Sie dies in den insgesamt vier Fällen, je nachdem, ob $x$, $y$ negativ sind oder nicht.)

*(d) Für alle $x \in \mathbb{R}^\times = \mathbb{R} \setminus \{0\}$ gilt $|\frac{1}{x}| = \frac{1}{|x|}$.* Dies folgt aus (c) wegen $|\frac{1}{x}||x| = 1$ für alle $x \in \mathbb{R}$.

(e) *Für alle $x, y \in \mathbb{R}$ ist $|x| \leq y$ äquivalent zu $-y \leq x \leq y$.* Denn angenommen $|x| \leq y$. Falls $x \geq 0$ dann gilt $-y \leq 0 \leq x = |x| \leq y$. Falls $x < 0$, dann ist $-y \leq -|x| = x < 0 \leq y$ und damit wiederum $-y \leq x \leq y$. Für die Umkehrung bemerken wir, dass $-y \leq x \leq y$ auch $-y \leq -x \leq y$ und somit in jedem Fall $|x| \leq y$ impliziert.

(f) *Analog ist für alle $x, y \in \mathbb{R}$ die strikte Ungleichung $|x| < y$ äquivalent zu $-y < x < y$.*

(g) *(Dreiecksungleichung) Für alle $x, y \in \mathbb{R}$ gilt*

$$|x + y| \leq |x| + |y|.$$

*Diese Ungleichung wird auch die* ***Dreiecksungleichung*** *genannt.* Sie folgt, in dem wir $-|x| \leq x \leq |x|$ und $-|y| \leq y \leq |y|$ wie in (e) addieren und anschliessend auf

$$-(|x| + |y|) \leq x + y \leq |x| + |y|$$

wiederum Eigenschaft (e) anwenden.

(h) *(umgekehrte Dreiecksungleichung) Für alle $x, y \in \mathbb{R}$ gilt*

$$\big||x| - |y|\big| \leq |x - y|.$$

Denn die Dreiecksungleichung in (g) zeigt

$$|x| \leq |x - y + y| \leq |x - y| + |y|$$

was zu $|x| - |y| \leq |x - y|$ führt. Durch Vertauschen von $x, y$ erhalten wir $|y| - |x| \leq |x - y|$. Also ist nach Eigenschaft (e) $\big||x| - |y|\big| \leq |x - y|$ wie gewünscht.

**Übung 2.46.** *Für welche $x, y \in \mathbb{R}$ gilt Gleichheit in der Dreiecksungleichung oder der umgekehrten Dreiecksungleichung?*

Für alle $x \in \mathbb{R}$ gilt $x = \operatorname{sgn}(x)|x|$, wobei $\operatorname{sgn}(x)$ das **Vorzeichen** (oder **Signum**) von $x$ ist, welches durch

$$\operatorname{sgn} : \mathbb{R} \to \{-1, 0, 1\}, \quad x \mapsto \begin{cases} 1 & \text{falls } x > 0 \\ 0 & \text{falls } x = 0 \\ -1 & \text{falls } x < 0 \end{cases}$$

definiert ist. Das Vorzeichen einer Zahlen ist also genau dann 1 (respektive $-1$), wenn die Zahl positiv (respektive negativ) ist. Die Zahl 0 ist weder positiv noch negativ und deswegen weist man ihr das „Vorzeichen Null“ zu.

**Übung 2.47** (Absolutbetrag und Quadratwurzel)**.** *Zeigen Sie für alle $x \in \mathbb{R}$ die Gleichungen $x^2 = |x|^2$ und $\sqrt{x^2} = |x|$.*

Wir bemerken noch, dass für $\delta > 0$ und $x \in \mathbb{R}$ die $\delta$-Umgebung von $x$ (siehe Definition 2.43) durch $\{y \in \mathbb{R} \mid |x - y| < \delta\}$ gegeben ist. Wir werden $|x - y|$ als den **Abstand** von $x$ zu $y$

interpretieren. Im Sinne des Wortes „Abstand“ kann man ein paar der obigen Folgerungen neu intuitiver ausdrücken. Zum Beispiel besagt (b), dass für $x, y \in \mathbb{R}$ die Gleichheit $|x-y| = |y-x|$ erfüllt ist, was also bedeutet, dass der Abstand von $x$ zu $y$ dem Abstand von $y$ zu $x$ gleich ist (wie man sich wünschen könnte). Des Weiteren werden Umgebungen einer reellen Zahl $x \in \mathbb{R}$ auch **Nachbarschaften** von $x$ genannt.

**Definition 2.48** (Offene und abgeschlossene Teilmengen)**.** Eine Teilmenge $U \subseteq \mathbb{R}$ heisst **offen** (in $\mathbb{R}$), wenn für jedes $x \in U$ ein $\varepsilon > 0$ existiert mit

$$\{y \in \mathbb{R} \mid |y - x| < \varepsilon\} = (x - \varepsilon, x + \varepsilon) \subseteq U.$$

Eine Teilmenge $A \subseteq \mathbb{R}$ heisst **abgeschlossen** (in $\mathbb{R}$), wenn ihr Komplement $\mathbb{R} \setminus A$ offen ist.

Intuitiv ausgedrückt ist eine Teilmenge offen, wenn für jeden Punkt $x$ in der Menge alle Punkte, die nahe genug an $x$ sind, wieder in der Menge liegen. Wir kennen bereits Beispiele von offenen Mengen:

**Übung 2.49** (Offene Intervalle)**.** *Zeigen Sie, dass eine Teilmenge $U \subseteq \mathbb{R}$ genau dann offen ist, wenn für jeden Punkt $x \in U$ ein offenes Intervall $I$ mit $x \in I$ und $I \subseteq U$ existiert. Schliessen Sie, dass die offenen (respektive abgeschlossenen) Intervalle auch im Sinne der obigen Definition offen (respektive abgeschlossen) sind.*

**Übung 2.50.** *Entscheiden Sie bei den folgenden Teilmengen von $\mathbb{R}$ jeweils, ob sie offen, abgeschlossen oder weder noch sind.*

- *Die Teilmengen* $\emptyset, \mathbb{N}, \mathbb{Z}, \mathbb{R}$.
- *Die Teilmengen* $[0, 1)$, $(0, 1]$ *und* $(0, 1) \cup (2, 3)$.

### 2.4.3 Der Absolutbetrag auf den komplexen Zahlen

Wir möchten nun den Absolutbetrag auf $\mathbb{C}$ so definieren, so dass dieser möglichst viele Eigenschaften des Absolutbetrags auf $\mathbb{R}$ hat und mit diesem kompatibel ist. Wir verwenden dazu die Wurzelfunktion, die in Übung 2.11 eingeführt wurde.

**Definition 2.51.** Der **Absolutbetrag** $|\cdot|$ **auf** $\mathbb{C}$ ist gegeben durch

$$|x + y\mathrm{i}| = \sqrt{x^2 + y^2}$$

für $x + y\mathrm{i} \in \mathbb{C}$.

An dieser Stelle bemerken wir, dass für $z = x + y\mathrm{i} \in \mathbb{C}$ die Summe der Quadrate $x^2 + y^2$ gerade gleich $z\overline{z}$ ist, denn

$$(x + y\mathrm{i})(x - y\mathrm{i}) = x^2 + y^2 + (xy - xy)\mathrm{i} = x^2 + y^2.$$

Somit gilt für alle $z \in \mathbb{C}$

$$|z| = \sqrt{z\overline{z}}.$$

Des Weiteren möchten wir anmerken, dass für ein $x \in \mathbb{R}$ der zu Beginn von Abschnitt 2.4.2 definierte Absolutbetrag $|x|$ und der Absolutbetrag von $x$ als Element von $\mathbb{C}$ übereinstimmen, da $\sqrt{x\overline{x}} = \sqrt{x^2} = |x|$ (vergleiche Übung 2.47). Insbesondere ist die neu eingeführte Notation nicht widersprüchlich und wir haben den Absolutbetrag von $\mathbb{R}$ auf $\mathbb{C}$ erweitert.

Wir fassen nun einige Eigenschaften des Absolutbetrags auf $\mathbb{C}$ zusammen:

**Eigenschaften des Absolutbetrags auf $\mathbb{C}$.**

*(i) (Definitheit) Für alle $z \in \mathbb{C}$ gilt $|z| \geq 0$ und $|z| = 0$ genau dann, wenn $z = 0$.*

*(ii) (Multiplikativität) Für alle $z, w \in \mathbb{C}$ gilt $|zw| = |z||w|$.*

*(iii) (Dreiecksungleichung) Für alle $z, w \in \mathbb{C}$ gilt $|z + w| \leq |z| + |w|$.*

*(iv) (Umgekehrte Dreiecksungleichung) Für alle $z, w \in \mathbb{C}$ gilt $\big||z| - |w|\big| \leq |z - w|$.*

Genauso wie auf $\mathbb{R}$ wollen wir mit Hilfe des Absolutbetrags den **Abstand** zweier Punkte $z, w \in \mathbb{C}$ als die nicht-negative Zahl $|z - w|$ auffassen. Wir bemerken noch, dass Definition 2.51 dem Satz von Pythagoras (siehe die entsprechende Übung in Abschnitt 1.7.6) entspricht. Doch haben wir dies als Definition des Absolutbetrages von $z = x + y\mathrm{i}$ gewählt, womit es (abgesehen von obigen Eigenschaften) nichts zu beweisen gibt.

*Beweis.* Zur Definitheit: Per Definition der Wurzel gilt für ein $z \in \mathbb{C}$, dass $|z| \geq 0$. Des Weiteren gilt $|z| = 0$ wegen der Injektivität der Wurzelfunktion genau dann, wenn $z\overline{z} = 0$. In Lemma 2.37 wurde jedoch gezeigt, dass $z\overline{z}$ genau dann Null ist, wenn $z$ selbst Null ist. Also folgt die Definitheit des Absolutbetrags.

Für die Multiplikativität verwenden wir die Eigenschaften der Konjugation aus Lemma 2.37 und die Multiplikativität der Wurzel (siehe Übung 2.11(vi)). Seien $z, w \in \mathbb{C}$. Dann gilt

$$|zw| = \sqrt{zw\overline{zw}} = \sqrt{z\overline{z}w\overline{w}} = \sqrt{z\overline{z}}\sqrt{w\overline{w}} = |z||w|,$$

was zu zeigen war.

Für die Dreiecksungleichung betrachten wir $z = x_1 + y_1\mathrm{i}, w = x_2 + y_2\mathrm{i} \in \mathbb{C}$. Da die Wurzelfunktion Ungleichungen zwischen positive Zahlen erhält (siehe Übung 2.11(iv)), reicht es die Ungleichung $|z + w|^2 \leq (|z| + |w|)^2$ zu zeigen. Wir berechnen

$$\begin{aligned} |z + w|^2 &= (x_1 + x_2)^2 + (y_1 + y_2)^2 \\ &= x_1^2 + x_2^2 + y_1^2 + y_2^2 + 2(x_1x_2 + y_1y_2) \\ &= |z|^2 + |w|^2 + 2(x_1x_2 + y_1y_2). \end{aligned}$$

Wie wir sehen werden, reicht es aus die Ungleichung $x_1x_2 + y_1y_2 \leq |z||w|$ zu zeigen, die auch als Cauchy-Schwarz-Ungleichung auf $\mathbb{C}$ bekannt ist. Tatsächlich gilt

$$\begin{aligned}(x_1x_2 + y_1y_2)^2 &\leq (x_1x_2 + y_1y_2)^2 + (y_1x_2 - x_1y_2)^2 \\ &= x_1^2x_2^2 + y_1^2y_2^2 + 2x_1x_2y_1y_2 + y_1^2x_2^2 + x_1^2y_2^2 - 2x_1x_2y_1y_2 \\ &= x_1^2x_2^2 + y_1^2y_2^2 + y_1^2x_2^2 + x_1^2y_2^2 \\ &= (x_1^2 + y_1^2)(x_2^2 + y_2^2) = |z|^2|w|^2,\end{aligned}$$

und daher auch $x_1x_2 + y_1y_2 \leq |x_1x_2 + y_1y_2| \leq |z||w|$. Zusammen ergibt sich

$$|z + w|^2 = |z|^2 + |w|^2 + 2(x_1x_2 + y_1y_2) \leq |z|^2 + |w|^2 + 2|z||w| = (|z| + |w|)^2$$

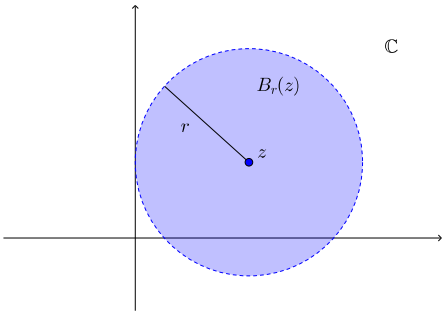
Die umgekehrte Dreiecksungleichung folgt ebenso wie im reellen Fall direkt aus der Dreiecksungleichung. □

Wie vorhin lässt sich mit Hilfe des Absolutbetrags ein Begriff von Offen- und Abgeschlossenheit einführen. Für die Definition von offenen Mengen in $\mathbb{R}$ wurden die symmetrisch um einen zuvor fixierten Punkt liegenden offenen Intervalle verwendet. In Analogie dazu definieren wir folgende Teilmengen von $\mathbb{C}$.

**Definition 2.52** (Offene Bälle)**.** Der **offene Ball** mit Radius $r > 0$ um einen Punkt $z \in \mathbb{C}$ ist die Menge

$$B_r(z) = \{w \in \mathbb{C} \mid |z - w| < r\}.$$

Der offene Ball $B_r(z)$ zu $r > 0$ und $z \in \mathbb{C}$ besteht also gerade aus jenen Punkten, die Abstand (strikt) kleiner $r$ von $z$ haben. Offene Bälle in $\mathbb{C}$ und offene Intervalle in $\mathbb{R}$ sind in folgendem Sinne kompatibel: Ist $x \in \mathbb{R}$ und $r > 0$, so ist der Schnitt des offenen Balles $B_r(x) \subseteq \mathbb{C}$ mit $\mathbb{R}$ gerade das offene, symmetrisch um $x$ liegende Intervall $(x-r, x+r)$ (wieso?).




**Wichtige Übung 2.53** (Durchschnitt von offenen Bällen)**.** *Zeigen Sie folgende Eigenschaft von Bällen: Seien $z_1, z_2 \in \mathbb{C}$, $r_1 > 0$ und $r_2 > 0$. Für jeden Punkt $z \in B_{r_1}(z_1) \cap B_{r_2}(z_2)$ existiert ein Radius $r > 0$, so dass*

$$B_r(z) \subseteq B_{r_1}(z_1) \cap B_{r_2}(z_2).$$

*Illustrieren Sie Ihre Wahl des Radius $r$ in einem Bild.*

**Definition 2.54** (Offene und abgeschlossene Teilmengen von $\mathbb{C}$)**.** Eine Teilmenge $U \subseteq \mathbb{C}$ heisst **offen** (in $\mathbb{C}$), wenn zu jedem Punkt in $U$ ein offener Ball um diesen Punkt existiert, der in $U$ enthalten ist. Formaler: Für alle $z \in U$ existiert ein Radius $r > 0$, so dass $B_r(z) \subseteq U$. Eine Teilmenge $A \subseteq \mathbb{C}$ heisst **abgeschlossen** (in $\mathbb{C}$), falls ihr Komplement $\mathbb{C} \setminus A$ offen ist.

Nach Übung 2.53 sind beispielsweise alle Bälle offen.

**Applet 2.55** (Offener Ball)**.** *Wir sehen, dass es für jeden Punkt $w$ in dem offenen Ball $B_r(z)$ um $z$ mit Radius $r$ wieder einen Radius $\varepsilon > 0$ gibt, so dass der offene Ball um $w$ mit Radius $\varepsilon$ ganz in $B_r(z)$ enthalten ist.*

Es gibt, abgesehen von den offenen Bällen, noch viele weitere, offene Teilmengen von $\mathbb{C}$. Beispielsweise ist jede Vereinigung von offenen Teilmengen offen. Zum Studium der offenen Mengen und damit verwandten Begriffen werden wir in deutlicher grösserer Allgemeinheit im zweiten Semester zurückkehren. Insbesondere wollen wir uns hier noch nicht auf eine ausführliche Diskussion einlassen.

## 2.5 Maximum und Supremum

### 2.5.1 Maximum und Minimum

**Definition 2.56** (Maximum)**.** Wir sagen, dass $x_0 = \max(X) \in \mathbb{R}$ das **Maximum** einer Teilmenge $X \subseteq \mathbb{R}$ ist, falls $x_0 \in X$ und für alle $x \in X$ die Ungleichung $x \leq x_0$ gilt.

Wir dürfen in der Tat von dem Maximum einer Teilmenge $X \subseteq \mathbb{R}$ sprechen, da es durch die Definition eindeutig bestimmt ist. Denn falls $x_0, x_0'$ beide die Eigenschaften eines Maximums erfüllen, so folgt $x_0 \leq x_0'$ (weil $x_0 \in X$ und $x_0'$ ein Maximum ist) und $x_0' \leq x_0$ (weil $x_0' \in X$ und $x_0$ ein Maximum ist) und damit $x_0 = x_0'$.

Ein abgeschlossenes Intervall $[a, b]$ mit Endpunkten $a < b$ in $\mathbb{R}$ hat $b = \max([a, b])$ als Maximum. Auch nicht-leere endliche Teilmengen und viele weitere Mengen besitzen ein Maximum. Es gibt jedoch auch Mengen, die kein Maximum besitzen. Beispielsweise hat das offene Intervall $(a, b)$ mit Endpunkten $a < b$ in $\mathbb{R}$ kein Maximum (beweisen Sie dies als Übung) - es würde sich zwar der Endpunkt $b$ anbieten, doch dieser liegt nicht in der Menge $(a, b)$ und ist also kein Kandidat für das Maximum. Wir werden in Kürze einen Begriff einführen, der auf $b$ zutrifft und gewissermassen als Ersatz für das Maximum angesehen werden kann.

Des Weiteren kann $\mathbb{R}$ (oder auch Intervalle der Form $[a, \infty), (a, \infty)$ für $a \in \mathbb{R}$) kein Maximum besitzen, da für beliebige $x \in \mathbb{R}$ die Ungleichung $x < x + 1$ gilt und damit $x$ kein Maximum sein kann.

**Definition 2.57** (Minimum)**.** Wir sagen, dass $x_0 = \min(X)$ das **Minimum** einer Teilmenge $X \subseteq \mathbb{R}$ ist, falls $x_0 \in X$ und $x \geq x_0$ für alle $x \in X$ gilt.

Die obige Diskussion lässt sich auf analoge Weise für das Minimum anwenden. Dieses ist also eindeutig bestimmt, muss aber nicht unbedingt existieren.

### 2.5.2 Supremum und Infimum

**Definition 2.58** (Beschränktheit und Schranken)**.** Eine Teilmenge $X \subseteq \mathbb{R}$ heisst **von oben beschränkt**, falls es ein $s \in \mathbb{R}$ gibt mit $x \leq s$ für alle $x \in X$. Ein solches $s \in \mathbb{R}$ nennt man in diesem Fall eine **obere Schranke** von $X$. Die Begriffe „**von unten beschränkt**“ und „**untere Schranke**“ sind analog definiert. Eine Teilmenge $X \subseteq \mathbb{R}$ heisst **beschränkt**, falls sie von oben und von unten beschränkt ist.

Wie wir bereits bemerkt haben, hat zum Beispiel das Intervall $(0, 1)$ kein Maximum. Es hat aber obere Schranken, 100 ist ein Beispiel. Natürlich ist 100 keine „gute“ obere Schranke; 10 oder auch 2 oder $\frac{3}{2}$ sind kleinere also auch „bessere“ obere Schranken. Die absolut beste obere Schranke ist aber durch 1 gegeben. Denn nach Definition von $(0, 1)$ ist 1 sicherlich eine obere Schranke und für jede obere Schranke $s$ gilt $s \geq 1$. (Falls $s < 1$ eine obere Schranke wäre, dann wäre $s \geq \frac{1}{2} > 0$, $\frac{s+1}{2} \in (0, 1)$ und $s < \frac{s+1}{2}$, wonach $s$ keine obere Schranke sein kann.)

Diese Gedanken führen gemeinsam mit dem Vollständigkeitsaxiom (Axiom (16) in Abschnitt 2.1) zu folgendem grundlegenden Begriff.

**Satz 2.59** (Supremum)**.** *Sei $X \subseteq \mathbb{R}$ eine von oben beschränkte, nicht-leere Teilmenge. Dann gibt es eine* ***kleinste obere Schranke*** *von $X$, die auch das* ***Supremum*** $\sup(X)$ *von $X$ genannt wird. Formal gelten also für $s_0 = \sup(X)$ folgende Eigenschaften:*

*(1) ($s_0$ ist eine obere Schranke)* $\forall x \in X : x \leq s_0$

*(2) ($s_0$ ist kleiner gleich jeder oberen Schranke)* $\forall s \in \mathbb{R} : ((\forall x \in X : x \leq s) \implies s_0 \leq s)$

*Äquivalenterweise kann $s_0 = \sup(X)$ auch durch (1) und die folgende Bedingung definiert werden:*

*(2') (Kleinere Zahlen sind keine oberen Schranken)* $\forall \varepsilon > 0\ \exists x \in X : x > s_0 - \varepsilon$.

Um diesen wichtigen Begriff noch etwas genauer zu beleuchten, wollen wir vor dem Beweis noch ein paar Bemerkungen machen.

- Falls das Maximum $x_0 = \max(X)$ existiert, dann ist $x_0$ eine obere Schranke von $X$ und ist vielmehr auch die kleinste obere Schranke, also $\max(X) = \sup(X)$. Denn aus $x_0 \in X$ folgt $x_0 \leq s$ für jede obere Schranke $s$ von $X$.
- Wenn das Supremum $\sup(X)$ in $X$ liegt, dann ist $\sup(X) = \max(X)$, da das Supremum eine obere Schranke ist. Also ist das Supremum eine Verallgemeinerung des Maximums einer Menge.
- Die Formulierung „kleinste obere Schranke“ ist natürlich ein Synonym für das Minimum der oberen Schranken und ist dadurch eindeutig bestimmt, falls es existiert.

*Beweis von Satz 2.59.* Nach Annahme ist $X$ nicht-leer und die Menge der oberen Schranken $Y = \{s \in \mathbb{R} \mid \forall x \in X : x \leq s\}$ ist ebenfalls nicht-leer. Des Weiteren gilt für alle $x \in X, s \in Y$ die Ungleichung $x \leq s$. Nach dem Vollständigkeitsaxiom (Axiom (16) in Abschnitt 2.1.3) folgt daher, dass es ein $c \in \mathbb{R}$ gibt, für das $x \leq c \leq s$ für alle $x \in X$ und $s \in Y$. Aus der ersten Ungleichung folgt, dass $c$ eine obere Schranke von $X$ ist. Aus der zweiten Ungleichung folgt, dass $c$ die kleinste obere Schranke von $X$ ist, und daher erfüllt $c$ sowohl (1) als auch (2). VO: 11.10.17

Wir zeigen nun, dass das Supremum auch durch (1) und (2') charakterisiert wird. Also angenommen $s_0 = \sup(X)$ und $\varepsilon > 0$, dann ist $s_0 - \varepsilon < s_0$. Daher kann $s_0 - \varepsilon$ keine obere Schranke sein und es existiert ein $x \in X$ mit $x > s_0 - \varepsilon$. Daher erfüllt $s_0$ auch (2').

Erfüllt $t_0 \in \mathbb{R}$ nun (1) und (2'), so ist $t_0$ eine obere Schranke und daher ist $s_0 \leq t_0$ nach Definition von $s_0 = \sup(X)$. Falls $s_0 < t_0$ wäre, dann wäre $s_0 = t_0 - \varepsilon$ für ein $\varepsilon > 0$. Nach der zweiten Eigenschaft von $t_0$ gäbe es ein $x \in X$ mit $x > s_0$, was der Definition von $s_0$ als (kleinste) obere Schranke widerspricht. Deswegen muss $t_0 = s_0$ gelten und $s_0$ ist eindeutig durch die Bedingungen (1) und (2') bestimmt. □

**Applet 2.60** (Supremum einer beschränkten nicht-leeren Menge)**.** *Wir betrachten eine beschränkte nicht-leere Teilmenge von $\mathbb{R}$ und zwei äquivalente Charakterisierungen des Supremums dieser Menge.*

Genauso wie auch andere Konsequenzen des Vollständigkeitsaxioms, die wir behandeln werden, ist die Existenz des Supremums in der Tat äquivalent zum Vollständigkeitsaxiom. In anderen Worten hätten wir anstelle von Axiom (16) einfach die Aussage von Satz 2.59 fordern können. Mehr dazu finden Sie im Abschnitt 2.8.2.

Für eine von unten beschränkte, nicht-leere Teilmenge $X \subseteq \mathbb{R}$ wird die grösste, untere Schranke auch das **Infimum** $\inf(X)$ von $X$ genannt. Für das Infimum gilt eine ähnliche Aussage wie in Satz 2.59:

**Übung 2.61** (Existenz des Infimums)**.** *Formulieren und beweisen Sie die analoge Aussage zu Satz 2.59 für das Infimum. Sie können dazu wie im Beweis von Satz 2.59 vorgehen oder das Supremum der Teilmenge $-X = \{-x \mid x \in X\}$ für eine von unten beschränkte, nicht-leere Teilmenge $X \subseteq \mathbb{R}$ betrachten.*

Die in obiger Übung erschienene Notation lässt sich verallgemeinern. Sei $x \in \mathbb{R}$ eine reelle Zahl und seien $A, B \subseteq \mathbb{R}$ zwei Teilmengen. Wir definieren

$$\begin{aligned} x + A &= \{x + a \mid a \in A\} \\ A + B &= \{a + b \mid a \in A,\ b \in B\} \\ xA &= \{xa \mid a \in A\} \\ AB &= \{ab \mid a \in A,\ b \in B\}\,. \end{aligned}$$

Es gelten also beispielsweise die Identitäten $x + A = \{x\} + A$, $xA = \{x\}\,A$ für alle $x \in \mathbb{R}$ und $A \subseteq \mathbb{R}$. Auch gilt $[a, b] + [c, d] = [a + c, b + d]$ für $a, b, c, d \in \mathbb{R}$ mit $a \leq b$ und $c \leq d$. (Wieso?)

**Proposition 2.62** (Supremum unter Streckung)**.** *Sei $A \subseteq \mathbb{R}$ eine nicht-leere, von oben beschränkte Teilmenge und sei $c > 0$. Dann ist $cA$ von oben beschränkt und es gilt*

$$\sup(cA) = c \sup(A).$$

Wir empfehlen Ihnen hier, sich die Aussage dieser (genauso wie der nächsten) Proposition zuerst am Begriff des Maximums zu veranschaulichen.

*Beweis.* Sei $s = \sup(A)$. Dann gilt $a \leq s$ und somit auch $ca \leq cs$ für alle $a \in A$. Da aber jedes Element von $cA$ von der Form $ca$ für ein $a \in A$ ist, erhalten wir, dass $cs$ eine obere Schranke von $cA$ ist und dass $cA$ von oben beschränkt ist.

Sei $\varepsilon > 0$. Dann existiert nach Satz 2.59 ein $a \in A$ mit $a > s - \frac{\varepsilon}{c}$, für welches die Ungleichung $ca > cs - \varepsilon$ gilt. Dies zeigt die zweite charakterisierende Eigenschaft des Supremums und wir erhalten $\sup(cA) = cs = c \sup(A)$. □

**Proposition 2.63** (Supremum unter Summen)**.** *Seien $A, B \subseteq \mathbb{R}$ zwei nicht-leere, von oben beschränkte Teilmengen von $\mathbb{R}$. Dann ist $A + B$ von oben beschränkt und es gilt*

$$\sup(A + B) = \sup(A) + \sup(B).$$

*Beweis.* Wir definieren $s_A = \sup(A)$ und $s_B = \sup(B)$. Dann gilt $a \leq s_A$ und $b \leq s_B$ für alle $a \in A$ und $b \in B$, was $a + b \leq s_A + s_B$ für alle $a \in A$ und $b \in B$ impliziert. Da aber jedes Element von $A + B$ von dieser Form ist, erhalten wir, dass $s_A + s_B$ eine obere Schranke von $A + B$ ist und dass $A + B$ von oben beschränkt ist.

Sei $\varepsilon > 0$. Dann existiert nach Satz 2.59 ein $a \in A$ mit $a > s_A - \frac{\varepsilon}{2}$ und ein $b \in B$ mit $b > s_B - \frac{\varepsilon}{2}$, was wiederum $a+b > s_A+s_B-\varepsilon$ impliziert. Dies zeigt die zweite charakterisierende Eigenschaft von $\sup(A + B)$ in Satz 2.59 und wir erhalten

$$\sup(A + B) = s_A + s_B = \sup(A) + \sup(B).$$

□

### 2.5.3 Uneigentliche Werte, Suprema und Infima

In diesem Abschnitt wollen wir die Begriffe „Supremum“ und „Infimum“ auf beliebige Teilmengen von $\mathbb{R}$ erweitern (ohne die in Abschnitt 2.5.2 getroffenen Annahmen). Dazu verwenden wir die Symbole $\infty = +\infty$ und $-\infty$, die keine reellen Zahlen darstellen. Wir definieren die **erweiterte Zahlengerade** (die auch **Zweipunktkompaktifizierung** von $\mathbb{R}$ genannt wird) durch

$$\overline{\mathbb{R}} = \mathbb{R} \sqcup \{-\infty, +\infty\}$$

und stellen uns diese als die Zahlengerade

vor. Hier haben wir den Punkt $+\infty$ rechts von $\mathbb{R}$ und den Punkt $-\infty$ links von $\mathbb{R}$ zu der Gerade hinzugefügt. Formaler formuliert: wir erweitern die Relation (Ordnung) $\leq$ auf $\overline{\mathbb{R}}$, so dass $-\infty \leq x \leq +\infty$ für alle $x \in \overline{\mathbb{R}}$ gilt, aber keine weiteren $\leq$-Relationen für die Symbole $-\infty, +\infty$ erfüllt sind. Inbesondere schreiben wir auch $-\infty < x < \infty$ für alle $x \in \mathbb{R}$.

**Übung 2.64** (Geometrie der Zweipunktkompaktifizierung)**.** *Zeigen Sie, dass die Abbildung*

$$\phi : \mathbb{R} \to (-1, 1), \quad x \mapsto \begin{cases} 1 - \frac{1}{1+x} & \textit{falls } x \geq 0 \\ -1 + \frac{1}{1-x} & \textit{falls } x < 0 \end{cases}$$

*bijektiv ist und die Ordnung erhält. Das heisst, für $x, y \in \mathbb{R}$ gilt $x < y \iff \phi(x) < \phi(y)$. Erweitern Sie $\phi$ zu einer ordnungserhaltenden Bijektion $\overline{\phi} : \overline{\mathbb{R}} \to [-1, 1]$ und erklären Sie damit das obige Bild der erweiterten Zahlengerade.*

Das Maximum und das Minimum einer Teilmenge $X \subseteq \overline{\mathbb{R}}$ ist nun wie in Abschnitt 2.5.1 definiert (falls es existiert).

Falls $X \subseteq \mathbb{R}$ nicht von oben beschränkt ist, dann definieren wir $\sup(X) = +\infty$. Falls $X$ leer ist, setzen wir $\sup(\emptyset) = -\infty$ (da jedes $x \in \mathbb{R}$ eine obere Schranke von $\emptyset$ darstellt). Analog definieren wir $\inf(\emptyset) = +\infty$ und $\inf(X) = -\infty$, falls $X \subseteq \mathbb{R}$ nicht von unten beschränkt ist.

Folgende Übungen stellen natürliche Eigenschaften von Supremum und Infimum dar. Sie sollten mindestens eine dieser Übungen ausarbeiten. Betrachten Sie hierbei die Spezialfälle, die zu einem uneigentlichen Supremum oder Infimum führen, getrennt und gehen Sie anschliessend wie im Beweis von Proposition 2.63 vor.

Weiter definieren wir für die Übungen die „Rechenregeln“

$$\begin{aligned} \infty + x = x + \infty = \infty \qquad & -\infty + x = x - \infty = -\infty \\ \infty + \infty = \infty \qquad & -\infty - \infty = -\infty \end{aligned}$$

für alle $x \in \mathbb{R}$ und

$$\begin{aligned} \infty \cdot \infty &= \infty & (-\infty) \cdot \infty &= -\infty \\ y \cdot \infty = \infty \cdot y &= \infty & (-y) \cdot \infty = \infty \cdot (-y) &= -\infty \\ \infty \cdot (-\infty) &= -\infty & (-\infty) \cdot (-\infty) &= \infty \\ y \cdot (-\infty) = (-\infty) \cdot y &= -\infty & (-y) \cdot (-\infty) = (-\infty) \cdot (-y) &= \infty \end{aligned}$$

für alle $y > 0$, wovon wir einen Teil verwenden werden. Die Ausdrücke $\infty - \infty$ und $0 \cdot \infty$ oder ähnliche bleiben wohlgemerkt aber undefiniert.

**Übung 2.65** (Eigenschaften von Supremum und Infimum unter Vereinigung)**.** *Seien $X, Y$ zwei Teilmengen von $\mathbb{R}$. Zeigen Sie, dass*

$$\sup(X \cup Y) = \max\{\sup(X), \sup(Y)\}.$$

*Formulieren und beweisen Sie eine analoge Formel für das Infimum von $X \cup Y$.*

**Übung 2.66** (Eigenschaften von Supremum und Infimum unter Summen und Produkten)**.** *Seien $A, B \subseteq \mathbb{R}$ zwei nicht-leere Teilmengen. Zeigen Sie, dass*

$$\sup(A + B) = \sup(A) + \sup(B)$$

*und dass, falls $A \subseteq \mathbb{R}_{>0}$ und $B \subseteq \mathbb{R}_{>0}$,*

$$\sup(AB) = \sup(A)\sup(B).$$

*Suchen Sie des Weiteren ähnliche Identitäten für das Infimum.*

**Übung 2.67.** *Sei $A$ eine nicht-leere Teilmenge von $\mathbb{R}$. Zeigen Sie, dass*

$$\sup|A| = \max\{\sup(A), -\inf(A)\}.$$

*Hierbei ist $|A|$ das Bild von $A$ unter dem Absolutbetrag $|\cdot|$ (als Funktion von $\mathbb{R}$ nach $\mathbb{R}$).*

### 2.5.4 Verwendung des Supremums und des Infimums

Das Supremum ist eine natürliche und notwendige Verallgemeinerung des Maximums einer Menge, da letzteres sogar für beschränkte Intervalle nicht existieren muss. Das Supremum kann

aber auch hilfreich sein in Situationen, wo das Maximum existiert. Denn falls man beweisen will, dass ein Maximum existiert, dann hat man mit dem Supremum den richtigen Kandidaten und kann den Beweis mit der Existenz des Supremums beginnen. Auf die gleiche Weise ist das Infimum einer Menge eine Verallgemeinerung des Minimums.

Es ist wichtig, dass Sie sich die charakterisierenden Eigenschaften des Supremums und Infimums einprägen, da diese Begriffe fundamentale Bausteine unserer zu entwickelnden Theorie sein werden. Zum Beispiel werden wir das Integral einer Funktion durch ein Supremum definieren (siehe Figur 1.2 und Kapitel 4).

## 2.6 Konsequenzen der Vollständigkeit

Wir haben in Abschnitt 2.1 unter Verwendung des Vollständigkeitsaxiom die Wurzelfunktion eingeführt und in Abschnitt 2.5 bereits das Vollständigkeitsaxiom verwendet um die Existenz des Supremums zu beweisen. Letzteres kann aber auch bloss als eine Umformulierung des Vollständigkeitsaxiom betrachtet werden. In diesem Abschnitt werden wir eine bereits bekannte Fragestellung und einige weitere Themen betrachten und erkennen wie nützlich das Vollständigkeitsaxiom sein kann.

### 2.6.1 Das Archimedische Prinzip

Mit Hilfe der Existenz des Supremums können wir nun das Archimedische Prinzip beweisen.

**Satz 2.68** (Das Archimedische Prinzip)**.** *Es gelten folgende Aussagen:*

*(i) Jede nicht-leere, von oben beschränkte Teilmenge von $\mathbb{Z}$ hat ein Maximum.*

*(ii) Für jedes $x \in \mathbb{R}$ existiert genau ein $n \in \mathbb{Z}$ mit $n \leq x < n+1$.*

*(iii) Für jedes $\varepsilon > 0$ existiert ein $n \in \mathbb{N}$ mit $\frac{1}{n} < \varepsilon$.*

Wir können das Archimedische Prinzip beispielsweise verwenden, um folgende Funktionen zu definieren.

- Der **ganzzahlige Anteil** $\lfloor x \rfloor$ einer Zahl $x \in \mathbb{R}$ ist die nach Satz 2.68 eindeutig bestimmte ganze Zahl $n \in \mathbb{Z}$ mit $n \leq x < n+1$. Wir erhalten also die Funktion $x \in \mathbb{R} \mapsto \lfloor x \rfloor \in \mathbb{Z}$, die auch **Abrundungsfunktion** genannt wird.

- Der **gebrochene Anteil** (oder auch **Nachkommaanteil**) ist $\{x\} = x - \lfloor x \rfloor \in [0,1)$ und wir erhalten eine Funktion $x \in \mathbb{R} \mapsto \{x\} \in [0,1)$ mit $x = \lfloor x \rfloor + \{x\}$ für alle $x \in \mathbb{R}$.

*Beweis von Satz 2.68.* Zu (i): Sei $E \subseteq \mathbb{Z}$ eine nicht-leere und (als Teilmenge von $\mathbb{R}$) von oben beschränkte Teilmenge. Nach Satz 2.59 existiert das Supremum $s_0 = \sup(E)$. Da $s_0$ die kleinste obere Schranke von $E$ ist, existiert ein $n_0 \in E$ mit $s_0 - 1 < n_0 \leq s_0$ (sonst wäre $s_0 - 1$ eine kleinere obere Schranke). Es folgt $s_0 < n_0 + 1$ und für jedes $m \in E$ gilt dann $m \leq s_0 < n_0 + 1$, woraus $m \leq n_0$ folgt (nach Lemma 2.18 und Übung 2.26). Daher ist $n_0$ das Maximum von $E$ wie in (i) behauptet.

Zu (ii): Sei $x \geq 0$ eine reelle Zahl. Dann ist $E = \{n \in \mathbb{Z} \mid n \leq x\}$ eine von oben beschränkte, nicht-leere Teilmenge von $\mathbb{Z}$ (nicht-leer, da $0 \in E$ – hier verwenden wir $x \geq 0$). Nach obigem hat $E$ ein Maximum, das heisst, es gibt ein maximales $n \in \mathbb{Z}$ mit $n \leq x$. Daraus folgt $x < n+1$ wie in (ii).

Falls $x < 0$ ist, dann können wir obigen Fall auf $-x$ anwenden und finden ein $\ell \in \mathbb{Z}$ mit $\ell \leq -x < \ell + 1$. Daraus folgt, dass es auch ein $k \in \{\ell, \ell+1\} \subseteq \mathbb{Z}$ mit $k - 1 < -x \leq k$ gibt. Für $n = -k \in \mathbb{Z}$ erhalten wir schliesslich $n \leq x < n+1$. Damit ist die Existenz in (ii) bewiesen.

Für den Beweis der Eindeutigkeit nehmen wir an, dass $n_1, n_2 \in \mathbb{Z}$ die Ungleichungen $n_1 \leq x < n_1 + 1$ und $n_2 \leq x < n_2 + 1$ gelten. Daraus folgt $n_1 \leq x < n_2 + 1$ und damit $n_1 \leq n_2$. Analog folgt $n_2 \leq n_1$, was $n_1 = n_2$ impliziert.

Zu (iii): Sei $\varepsilon > 0$ eine reelle Zahl. Dann gilt auch $\frac{1}{\varepsilon} > 0$ und es gibt nach Teil (ii) ein $n \in \mathbb{N}$ mit $\frac{1}{\varepsilon} < n$. Für dieses $n$ gilt aber auch $\frac{1}{n} < \varepsilon$, wie in (iii) behauptet wurde. □

**Übung 2.69** (Supremum von Bildmengen)**.** *Sei $A$ eine nicht-leere Teilmenge von $\mathbb{R}$. Zeigen Sie, dass im Allgemeinen $\sup(\lfloor A \rfloor) = \lfloor \sup(A) \rfloor$ nicht gilt. Hierbei ist $\lfloor A \rfloor$ das Bild von $A$ unter der Abrundungsfunktion $\lfloor \cdot \rfloor : \mathbb{R} \to \mathbb{R}$.*

Mit Hilfe des Archimedischen Prinzips können wir auch den geometrischen Zusammenhang zwischen $\mathbb{Q}$ und $\mathbb{R}$ im folgenden Korollar beschreiben. (Ein Korollar ist eine Folgerung aus einer Proposition, einem Satz oder einem Theorem.)

**Korollar 2.70** (Dichtheit von $\mathbb{Q}$ in $\mathbb{R}$)**.** *Zwischen je zwei reellen Zahlen $a, b \in \mathbb{R}$ mit $a < b$ gibt es ein $r \in \mathbb{Q}$ mit $a < r < b$.*

*Beweis.* Nach dem Archimedischen Prinzip (Satz 2.68 (iii)) existiert ein $m \in \mathbb{N}$ mit $\frac{1}{m} < b - a$. Ebenso gibt es nach dem Archimedischen Prinzip (Satz 2.68 (ii)) ein $n \in \mathbb{Z}$ mit $n - 1 \leq ma < n$ oder äquivalenterweise $\frac{n-1}{m} \leq a < \frac{n}{m}$. Insbesondere gilt $\frac{n}{m} \leq a + \frac{1}{m}$, was mit $\frac{1}{m} < b - a$ gerade

$$a < \frac{n}{m} \leq a + \frac{1}{m} < a + b - a = b$$

und damit das Korollar impliziert, wobei $r = \frac{n}{m}$ gewählt wird. □

Anders formuliert zeigt obiges Korollar, dass $\mathbb{Q}$ jede Umgebung $I$ einer reellen Zahl schneidet (das heisst, $I \cap \mathbb{Q} \neq \emptyset$), oder auch, dass wir jede reelle Zahl beliebig genau durch rationale Zahlen approximieren können. Die Eigenschaft wird auch als $\mathbb{Q}$ ist **dicht** in $\mathbb{R}$ bezeichnet und wird uns später in einem allgemeineren Kontext wiederbegegnen.

VO: 12.10.17

**Übung 2.71** (Jede reelle Zahl ist ein Supremum einer Menge von rationalen Zahlen)**.** *Zeigen Sie, dass für jedes $x \in \mathbb{R}$ das Supremum von $\{r \in \mathbb{Q} \mid r < x\}$ gerade $x$ ist.*

**Übung 2.72** (Etwas Diophantische Approximation)**.** *Sei $a \in \mathbb{R}$ eine reelle, irrationale Zahl. Betrachtet man den Beweis von Korollar 2.70 nochmals, so realisiert man, dass die Existenz einer rationalen Zahl $\frac{p}{q} \in \mathbb{Q}$ mit*

$$\left|a - \frac{p}{q}\right| \leq \frac{1}{q}$$

*gezeigt wird. Fragen zu Approximation von reellen Zahlen mit rationalen sind Fragestellungen der Diophantischen Approximation. Wir wollen hier auf elementare Weise ein stärkeres Resultat (Dirichlet's Approximationssatz) zeigen. Sei $Q \in \mathbb{N}$ eine natürliche Zahl. Zeigen Sie, dass $p \in \mathbb{Z}$ und $q \in \mathbb{N}$ mit $1 \leq q \leq Q$ existieren, die*

$$\left|a - \frac{p}{q}\right| < \frac{1}{qQ}$$

*und insbesondere $\left|a - \frac{p}{q}\right| < \frac{1}{q^2}$ erfüllen.*

### 2.6.2 Häufungspunkte einer Menge

Wie oben bereits erwähnt, kann jeder Punkt in $\mathbb{R}$ durch Punkte in $\mathbb{Q}$ approximiert werden. Allgemeiner möchten wir nun zu einer Menge $A \subseteq \mathbb{R}$ jene Punkte betrachten, denen $A$ „von aussen“ beliebig nahe kommt.

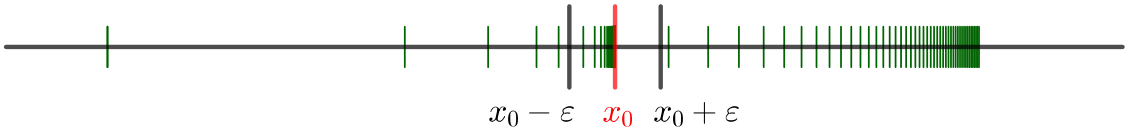
**Definition 2.73** (Häufungspunkte von Mengen)**.** Sei $A \subseteq \mathbb{R}$ und $x_0 \in \mathbb{R}$. Wir sagen, dass $x_0$ ein **Häufungspunkt der Menge** $A$ ist, falls es für jedes $\varepsilon > 0$ ein $a \in A$ gibt mit $0 < |a - x_0| < \varepsilon$.

In anderen Worten gibt es für einen Häufungspunkt $x_0$ in jeder Umgebung abgesehen von $x_0$ Punkte in $A$ (wobei es keine Rolle spielt ob $x_0$ in $A$ liegt oder nicht). Die Menge $A$ kommt also ihren Häufungspunkten „von aussen“ beliebig nahe.

**Übung 2.74** (Endliche Mengen haben keine Häufungspunkte)**.** *Zeigen Sie, dass eine endlichen Teilmenge $A \subseteq \mathbb{R}$ keine Häufungspunkte besitzt.*

Ebenso gibt es unendliche Mengen ohne Häufungspunkte. Zum Beispiel hat $A = \mathbb{Z}$ keine Häufungspunkte. In der Tat gibt es für $x_0 \in \mathbb{Z}$ kein $n \in \mathbb{Z}$ mit $0 < |n - x_0| < 1$ (auf Grund der Anordnung in Lemma 2.18 und Übung 2.26). Des Weiteren ist für jedes $x_0 \in \mathbb{R} \setminus \mathbb{Z}$ die Zahl $\varepsilon = \min\{x_0 - \lfloor x_0 \rfloor, \lfloor x_0 \rfloor + 1 - x_0\}$ positiv und erfüllt $|n - x_0| \geq \varepsilon$ für alle $n \in \mathbb{Z}$. Für beschränkte unendliche Mengen ist die Situation aber besser wie folgender Satz zeigt.

**Satz 2.75** (Existenz von Häufungspunkten)**.** *Sei $A \subseteq \mathbb{R}$ eine beschränkte unendliche Teilmenge. Dann existiert ein Häufungspunkt von $A$ in $\mathbb{R}$.*




Figur 2.1: Wir möchten in diesem Bild die Idee des Beweises von Satz 2.75 erläutern. Wir betrachten das „grösste“ $x_0 \in \mathbb{R}$, für welches links von $x_0$ (im Bild in rot) nur endliche viele Punkte von $A$ (im Bild in grün) liegen. Um genau zu sein, muss ein solches $x_0$ nicht existieren, weswegen wir $x_0$ als das Supremum über alle $x$ mit dieser Eigenschaft nehmen. Schiebt man dieses $x_0$ nun um ein kleines $\varepsilon > 0$ nach rechts auf $x_0 + \varepsilon$, so müssen unendlich viele Elemente von $A$ links von $x_0 + \varepsilon$ liegen. Umgekehrt kann man $x_0$ etwas nach links nach $x_0 - \varepsilon$ schieben, womit nur endlich viele Elemente von $A$ links von $x_0 - \varepsilon$ liegen können. Also befinden sich unendlich viele Elemente von $A$ zwischen $x_0 - \varepsilon$ und $x_0 + \varepsilon$. Da aber $\varepsilon$ beliebig war, muss $x_0$ ein Häufungspunkt von $A$ sein.

*Beweis.* Angenommen $m, M \in \mathbb{R}$ erfüllen $A \subseteq [m, M]$. Wir definieren

$$X = \{x \in \mathbb{R} \mid |A \cap (-\infty, x]| < \infty\}.$$

Dann ist $m \in X$ da $|A \cap (-\infty, m]| \leq 1$. Des Weiteren gilt $x < M$ für jedes $x \in X$, denn für $x \geq M$ ist $A \cap (-\infty, x] = A \cap (-\infty, M] = A$ eine unendliche Menge nach Annahme im Satz. Daher ist $X$ eine beschränkte, nicht-leere Teilmenge von $\mathbb{R}$, womit das Supremum $x_0 = \sup(X)$ nach Satz 2.59 existiert.

Sei nun $\varepsilon > 0$. Dann existiert ein $x \in X$ mit $x > x_0 - \varepsilon$, was zeigt, dass $A \cap (-\infty, x_0 - \varepsilon]$ eine endliche Menge ist, da

$$A \cap (-\infty, x_0 - \varepsilon] \subseteq A \cap (-\infty, x]$$

gilt. Des Weiteren gilt $x_0 + \varepsilon \notin X$ auf Grund der Definition von $x_0$. Damit ist die Kardinalität von $A \cap (-\infty, x_0 + \varepsilon]$ unendlich. Es folgt, dass

$$A \cap (x_0 - \varepsilon, x_0 + \varepsilon] = (A \cap (-\infty, x_0 + \varepsilon]) \setminus (A \cap (-\infty, x_0 - \varepsilon])$$

eine unendliche Menge ist und abgesehen von möglicherweise $x_0, x_0 + \varepsilon$ noch weitere Punkte besitzen muss. Da $\varepsilon > 0$ beliebig war, sehen wir, dass $x_0$ ein Häufungspunkt der Menge $A$ ist. □

Insbesondere erkennen wir im Beweis eine stärkere Aussage für den gefundenen Häufungspunkt $x_0$ der Menge $A$, nämlich dass für alle $\varepsilon > 0$ der Durchschnitt $A \cap (x_0 - \varepsilon, x_0 + \varepsilon)$ unendlich ist. Dies stellt eine alternative Definition des Begriffs dar.

**Übung 2.76** (Alternative Charakterisierung von Häufungspunkten)**.** *Sei $A \subseteq \mathbb{R}$ und $x_0 \in \mathbb{R}$. Zeigen Sie, dass $x_0$ genau dann ein Häufungspunkt der Menge $A$ ist, wenn für jedes $\varepsilon > 0$ der Durchschnitt von $A$ mit der $\varepsilon$-Umgebung $(x_0 - \varepsilon, x_0 + \varepsilon)$ unendlich viele Punkte enthält.*

Die Existenz eines Häufungspunkt in Satz 2.75 kann man als ein Schubfachprinzip der Analysis auffassen: Anstatt einer echten Übereinstimmung wie im regulären Schubfachprinzip für endliche Mengen (wie in Abschnitt 1.6.4) erlauben wir näherungsweise Übereinstimmungen wie in der Definition eines Häufungspunktes und haben einen Punkt gefunden mit dem unendlich viele Punkte der Menge „fast übereinstimmen“. Wir werden noch andere Sätze kennenlernen, bei denen ein beschränktes, abgeschlossenes Intervall ähnliche Eigenschaften wie eine endliche Menge haben wird.

### 2.6.3 Intervallschachtelungsprinzip

Der Durchschnitt von ineinander geschachtelten, nicht-leeren Intervallen, das heisst, Intervallen $I_1 \supseteq I_2 \supseteq I_3 \subseteq \cdots$ in $\mathbb{R}$, die kleiner werden, muss nicht unbedingt nicht-leer sein. Zum Beispiel gilt

$$\bigcap_{n=1}^{\infty} [n, \infty) = \emptyset,$$
$$\bigcap_{n=1}^{\infty} \left(0, \frac{1}{n}\right) = \emptyset$$

auf Grund des Archimedischen Prinzip in Satz 2.68. Für abgeschlossene und beschränkte Intervalle ist die Situation aber deutlich besser. Dies ist nochmals eine Konsequenz des Vollständigkeitsaxioms.

**Satz 2.77** (Intervallschachtelungsprinzip)**.** *Sei für jedes $n \in \mathbb{N}$ ein nicht-leeres, abgeschlossenes, beschränktes Intervall $I_n = [a_n, b_n]$ gegeben, so dass für alle natürlichen Zahlen $m \leq n$ die Inklusion $I_m \supseteq I_n$ oder äquivalenterweise die Ungleichungen $a_m \leq a_n \leq b_n \leq b_m$ gelten. Dann ist der Durchschnitt*

$$\bigcap_{n=1}^{\infty} I_n = \left[\sup\{a_n \mid n \in \mathbb{N}\}, \inf\{b_n \mid n \in \mathbb{N}\}\right]$$

*nicht-leer.*

Sollte die Aussage verwirrend sein, überzeugen Sie sich doch zuerst davon, dass

$$[a_1, b_1] \cap [a_2, b_2] = [\max\{a_1, a_2\}, \min\{b_1, b_2\}]$$

für beliebige $a_1, a_2, b_1, b_2 \in \mathbb{R}$ gilt.

*Beweis von Satz 2.77.* Nach Annahme gilt für natürliche Zahlen $\ell, m, n$ mit $\ell \leq m \leq n$ die Ungleichung

$$a_\ell \leq a_m \leq a_n \leq b_n \leq b_m \leq b_\ell.$$

Insbesondere ist $b_m$ eine obere Schranke von $\{a_k \mid k \in \mathbb{N}\}$, woraus

$$\sup\{a_k \mid k \in \mathbb{N}\} \leq b_m$$

folgt. Da $m \in \mathbb{N}$ beliebig war, sehen wir nun, dass $\overline{a} = \sup\{a_k \mid k \in \mathbb{N}\} \in \mathbb{R}$ eine untere Schranke von $\{b_m \mid m \in \mathbb{N}\}$ ist. Daher hat letztere Menge ein Infimum $\overline{b} \in \mathbb{R}$ und wir erhalten

$$\overline{a} = \sup\{a_k \mid k \in \mathbb{N}\} \leq \inf\{b_m \mid m \in \mathbb{N}\} = \overline{b}.$$

Insbesondere ist der Schnitt $\bigcap_{n=1}^{\infty} I_n$ nicht-leer, da er zum Beispiel $\overline{a}$ enthält. Für $x \in \mathbb{R}$ gilt nun die Abfolge von Äquivalenzen

$$\begin{aligned} x \in \bigcap_{n=1}^{\infty} [a_n, b_n] &\iff \forall n \in \mathbb{N} : a_n \leq x \leq b_n \\ &\iff \big(\forall n \in \mathbb{N} : a_n \leq x\big) \wedge \big(\forall n \in \mathbb{N} : x \leq b_n\big) \\ &\iff \overline{a} \leq x \wedge x \leq \overline{b}, \end{aligned}$$

womit $\bigcap_{n=1}^{\infty} [a_n, b_n] = [\overline{a}, \overline{b}]$ gilt und der Satz folgt. □

**Applet 2.78** (Intervallschachtelung)**.** *Bei Vergrösserung sehen wir, dass die Intervalle immer wieder weitere Intervalle enthalten. Man kann sich vorstellen, wie dies unbeschränkt weitergeht und der Durchschnitt in dem betrachteten Fall aus der Menge mit nur einem Punkt besteht. (Allerdings können wir dies hier nicht unbeschränkt beobachten, da geogebra nach einigen Vergrösserungen auf die Grenzen der Rechengenauigkeit stösst.)*

**Übung 2.79** (Charakterisierung von Intervallen)**.**

*(i) Zeigen Sie, dass eine Teilmenge $I \subseteq \mathbb{R}$ genau dann ein Intervall ist, wenn für alle $x, y \in I$ und alle $z \in \mathbb{R}$ die Implikation $(x \leq z \leq y) \implies z \in I$ gilt.*

*(ii) Schliessen Sie daraus, dass ein beliebiger Schnitt $\bigcap_{I \in \mathcal{I}} I$ von Intervallen $I \in \mathcal{I}$ ein Intervall ist.*

**Übung 2.80** (Zusammenziehende Intervalle)**.** *Seien $I_n = [a_n, b_n]$ für $n \in \mathbb{N}$ wie im Satz 2.77 und nehmen Sie zusätzlich an, dass $\inf\{b_n - a_n | n \in \mathbb{N}\} = 0$. Intuitiv heisst dies also, dass die betrachteten Intervalle immer kürzer werden. Zeigen Sie, dass in diesem Fall $\bigcap_{n=1}^{\infty} I_n$ aus nur einem Punkt besteht.*

### 2.6.4 Überabzählbarkeit

In diesem Teilabschnitt möchten wir eine klassische Anwendung des Intervallschachtelungsprinzips präsentieren, nämlich den Beweis für die Überabzählbarkeit der reellen Zahlen.

**Korollar 2.81** (Überabzählbarkeit von $\mathbb{R}$)**.** *Die Teilmenge $[0, 1] \subseteq \mathbb{R}$ (und daher auch $\mathbb{R}$) ist überabzählbar.*

*Beweis.* Wir wollen den Beweis in der Form eines Spiels zwischen den Spielern Alice und Bob darstellen. Bob behauptet (fälschlicherweise), dass $[0, 1]$ abzählbar ist und darf in jedem seiner Züge (von abzählbar vielen Spielzügen) ein weiteres Element auflisten und Alice glaubt ihm nicht. Er hat das Ziel alle Elemente von $[0, 1]$ am Ende aufgelistet zu haben und Alice hat das Ziel ein Element zu finden, das Bob nicht aufgelistet hat. Als Gegenzug zur Wahl des jeweiligen Elements von Bob, darf Alice ihr zuletzt konstruiertes Intervall (beginnend mit $[0, 1]$) verkleinern. Wir beschreiben eine Gewinnstrategie für Alice, die bei Befolgung und Anwendung von Satz 2.77 zu einem neuen Element in $[0, 1]$ führt, welches von Bob nicht aufgelistet wurde. Das heisst, Alice kann das Spiel immer gewinnen, was beweist, dass Bob nicht Recht hat.

Sei $n \in \mathbb{N} \mapsto x_n \in [0, 1]$ eine beliebige Funktion, die die Züge von Bob vollständig beschreibt. Wir beschreiben nun die Strategie von Alice, die rekursiv Intervalle $I_n = [a_n, b_n]$ definiert, so dass diese $x_n \notin I_n$ erfüllen und den Bedingungen in Satz 2.77 genügen. Für $n = 1$ definieren Alice $I_1$ durch Fallunterscheidung mittels

$$I_1 = [a_1, b_1] = \begin{cases} \left[\frac{2}{3}, 1\right] & \text{falls } x_1 \in [0, \frac{1}{2}], \\ \left[0, \frac{1}{3}\right] & \text{falls } x_1 \in (\frac{1}{2}, 1]. \end{cases}$$

Dies stellt sicher, dass $I_1$ ein abgeschlossenes Intervall ist, dass $x_1$ nicht enthält. Grob gesagt wählt Alice also $I_1$ als das linke Drittel von $[0, 1]$, wenn $x_1$ in der rechten Hälfte liegt, und als das rechte Drittel, wenn $x_1$ in der linken Hälfte liegt.

Angenommen Alice hat bereits $I_n = [a_n, b_n]$ für ein $n \in \mathbb{N}$ definiert und Bob hat in seinem darauffolgenden Spielzug das Element $x_{n+1} \in [0, 1]$ gewählt. Alice definiert nun

$$I_{n+1} = [a_{n+1}, b_{n+1}] = \begin{cases} I_n & \text{falls } x_{n+1} \notin I_n, \\ \left[b_n - \frac{1}{3}(b_n - a_n), b_n\right] & \text{falls } x_{n+1} \in [a_n, a_n + \frac{1}{2}(b_n - a_n)], \\ \left[a_n, a_n + \frac{1}{3}(b_n - a_n)\right] & \text{falls } x_{n+1} \in (a_n + \frac{1}{2}(b_n - a_n), b_n] \end{cases}$$

das heisst, sie verwendet also das vorherige Intervall $I_n$, wenn es $x_{n+1}$ nicht enthält und das rechte (resp. linke) Drittel von $I_n$, wenn $x_{n+1}$ in der linken (resp. rechten) Hälfte von $I_n$ liegt. Per Konstruktion ist insbesondere $I_{n+1} \subseteq I_n$ und $x_{n+1} \notin I_{n+1}$ erfüllt.

Nach dem Intervallschachtelungsprinzip ist der Durchschnitt $I = \bigcap_{n=1}^{\infty} I_n \subseteq [0,1]$ nichtleer, sagen wir $x \in I$. Dann ist $x \neq x_n$ für alle $n \in \mathbb{N}$, da $x_n \notin I_n$ nach Konstruktion von $I_n$ und $x \in I \subseteq I_n$. Alice hat also ein neues Element von $[0,1]$ gefunden, das von Bob nicht aufgelistet wurde.

Formal gesehen, zeigt obiges Argument, dass eine beliebige Abbildung $\mathbb{N} \to [0,1]$ nicht surjektiv sein kann und das Korollar folgt. $\square$

### 2.6.5 Die Cantor-Menge*

VO: 16.10.17, Rest des Kapitels ist Eigenlektüre

Wir definieren rekursiv Teilmengen des Intervalles $[0,1]$ durch

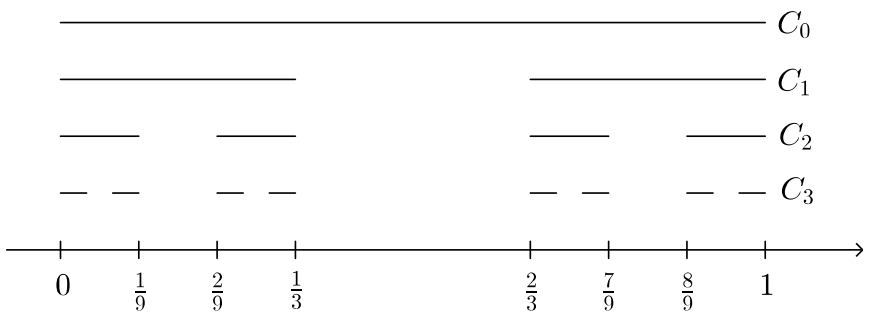
$$C_0 = [0,1], \quad C_{n+1} = \frac{1}{3}C_n \cup \left(\frac{1}{3}C_n + \frac{2}{3}\right).$$

für $n \in \mathbb{N}_0$. Beispielsweise gilt also

$$\begin{aligned} C_1 &= \frac{1}{3}[0,1] \cup \left(\frac{1}{3}[0,1] + \frac{2}{3}\right) = \left[0, \frac{1}{3}\right] \cup \left(\left[0, \frac{1}{3}\right] + \frac{2}{3}\right) = \left[0, \frac{1}{3}\right] \sqcup \left[\frac{2}{3}, 1\right] \\ C_2 &= \left[0, \frac{1}{9}\right] \sqcup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left(\left(\left[0, \frac{1}{9}\right] \sqcup \left[\frac{2}{9}, \frac{1}{3}\right]\right) + \frac{2}{3}\right) \\ &= \left[0, \frac{1}{9}\right] \sqcup \left[\frac{2}{9}, \frac{3}{9}\right] \sqcup \left[\frac{6}{9}, \frac{7}{9}\right] \sqcup \left[\frac{8}{9}, 1\right]. \end{aligned}$$

**Definition 2.82** (Cantor-Menge)**.** Die **Cantor-Menge** ist der Durchschnitt $C = \bigcap_{n=0}^{\infty} C_n$.

Ziel dieses Teilabschnitts ist es, die Cantor-Menge und ihre schönen „fraktalen“ Eigenschaften auszukundschaften. Dazu wollen wir zuerst die Mengen $C_n$ besser verstehen. Grob gesagt wollen wir die für folgendes Bild notwendigen Aussagen treffen:




**Erste Feststellungen.**

*(i) (Abfallende Mengen) Für jedes $n \in \mathbb{N}_0$ gilt $C_{n+1} \subseteq C_n$.* Dies folgt aus einem kurzen Induktionsargument. Für $n = 0$ wurde die Aussage oben bewiesen. Angenommen, dass

$C_{n+1} \subseteq C_n$ für ein $n \in \mathbb{N}_0$. Dann gilt

$$\frac{1}{3}C_{n+1} \subseteq \frac{1}{3}C_n \subseteq C_{n+1}, \quad \frac{1}{3}C_{n+1} + \frac{2}{3} \subseteq \frac{1}{3}C_n + \frac{2}{3} \subseteq C_{n+1}$$

und also $C_{n+2} \subseteq C_{n+1}$. Somit folgt die Aussage per vollständiger Induktion.

*(ii) (Disjunktheit) Die Mengen $\frac{1}{3}C_n$ und $\left(\frac{1}{3}C_n + \frac{2}{3}\right)$ sind disjunkt für jedes $n \in \mathbb{N}_0$. Insbesondere gilt $C_{n+1} = \frac{1}{3}C_n \sqcup \left(\frac{1}{3}C_n + \frac{2}{3}\right)$ für jedes $n \in \mathbb{N}_0$.* Bei genauerer Betrachtung sieht man sogar, dass $\frac{1}{3}C_n$ in $[0, \frac{1}{3}]$ und $\left(\frac{1}{3}C_n + \frac{2}{3}\right)$ in $[\frac{2}{3}, 1]$ liegt. Dies impliziert die behauptete Disjunktheit und folgt sofort aus $C_n \subseteq [0, 1]$.

*(iii) (Zerlegung in Intervalle) Für $n \in \mathbb{N}_0$ besteht $C_{n+1}$ aus genau doppelt so vielen disjunkten, abgeschlossenen Intervallen wie $C_n$.* Dies folgt aus (ii).

*(iv) (Von der Zerlegung von $C_n$ zur Zerlegung von $C_{n+1}$) Sei $I$ eines der Intervalle in der Zerlegung von $C_n$ und schreibe $I = [a, b]$. Dann sind $\left[a, a + \frac{b-a}{3}\right]$ und $\left[b - \frac{b-a}{3}, b\right]$ zwei Intervalle in der Zerlegung von $C_{n+1}$ und*

$$[a, b] \cap C_{n+1} = \left[a, a + \tfrac{b-a}{3}\right] \sqcup \left[b - \tfrac{b-a}{3}, b\right]$$

*Insbesondere ist die Länge eines Intervalles in $C_{n+1}$ ein Drittel der Länge eines Intervalles in $C_n$.* Für $n = 0$ (und auch $n = 1$) haben wir dies bereits direkt nachgerechnet. Wenn dies bereits für $C_n$ und $C_{n+1}$ bekannt ist, dann gilt die Aussage auch für die verkleinerten Versionen $\frac{1}{3}C_{n+1} \subseteq \frac{1}{3}C_n$ und $(\frac{1}{3}C_{n+1} + \frac{2}{3}) \subseteq (\frac{1}{3}C_n + \frac{2}{3})$, was wiederum die Aussage für $C_{n+1}$ und $C_{n+2}$ ergibt.

Wir wenden uns nun der Cantor-Menge $C$ zu. Wir möchten einem Punkt in $C$ eine Art „Adresse“ zuweisen. Dazu definieren wir rekursiv die Adresse eines Punktes $x \in C$ wie folgt. Setze

$$a_1(x) = \begin{cases} \mathrm{l} & \text{falls } x \in [0, \frac{1}{3}], \\ \mathrm{r} & \text{falls } x \in [\frac{2}{3}, 1], \end{cases} \qquad I_1(x) = \begin{cases} \left[0, \frac{1}{3}\right] & \text{falls } a_1(x) = \mathrm{l}, \\ \left[\frac{2}{3}, 1\right] & \text{falls } a_1(x) = \mathrm{r}. \end{cases}$$

Wir weisen $x$ also die erste Adresse l zu, falls $x$ im linken Drittel von $[0, 1]$ liegt und r, falls $x$ im rechten Drittel von $[0, 1]$ liegt. Kennen wir die erste Adresse $a_1(x)$ von $x$, so ist $I_1(x)$ das entsprechende Intervall, in dem $x$ liegt. Wir fahren genauso fort: Die zweite Adresse von $x$ ist definiert durch

$$a_2(x) = \begin{cases} \mathrm{l} & \text{falls } x \text{ im linken Drittel von } I_1(x) \text{ liegt} \\ \mathrm{r} & \text{falls } x \text{ im rechten Drittel von } I_1(x) \text{ liegt} \end{cases}$$

Des Weiteren ist $I_2(x)$ das linke Drittel von $I_1(x)$, wenn $a_2(x) = \mathrm{l}$, und sonst das rechte Drittel. Genauso fährt man fort, um für jedes $k \in \mathbb{N}_0$ die $k$-te Adresse $a_k(x)$ von $x$ zu erhalten. Wir erhalten somit eine Liste von Adressen oder genauer eine Abbildung

$$a(x) : \mathbb{N}_0 \to \{\mathrm{l}, \mathrm{r}\}, \ k \mapsto a_k(x).$$

Die Menge der Abbildung $\mathbb{N}_0 \to \{\mathrm{l}, \mathrm{r}\}$ schreiben wir als $\{\mathrm{l}, \mathrm{r}\}^{\mathbb{N}_0}$. Unter dem Strich haben wir also die Funktion

$$f : x \in C \mapsto a(x) \in \{\mathrm{l}, \mathrm{r}\}^{\mathbb{N}_0}$$

konstruiert, die einem Element in der Cantor-Menge ihre Adressen zuweist. Folgende Beschreibung der Cantor-Menge folgt aus dem Intervallschachtelungsprinzip (Satz 2.77).

**Korollar 2.83** (Cantor-Menge)**.** *Die oben konstruierte Abbildung $f : C \to \{\mathrm{l}, \mathrm{r}\}^{\mathbb{N}_0}$ ist eine Bijektion. Insbesondere ist also $C \sim \{\mathrm{l}, \mathrm{r}\}^{\mathbb{N}_0} \sim \{0, 1\}^{\mathbb{N}_0} \sim \mathcal{P}(\mathbb{N}_0)$ und die Cantor-Menge ist überabzählbar.*

Wir bemerken, dass die letzte Bijektion in obigen Satz auf Grund von Übung 1.76 gilt und die Überabzählbarkeit daher aus Cantor's Diagonalargument (Theorem 1.74) folgt.

*Beweis.* Wir zeigen zuerst, dass $f$ surjektiv ist. Sei also $a \in \{\mathrm{l}, \mathrm{r}\}^{\mathbb{N}_0}$ eine Liste von Adressen. Um Verwirrungen zu vermeiden schreiben wir $a_k$ anstelle von $a(k)$ für $k \in \mathbb{N}_0$. Genau wie oben (und ähnlich wie im Beweis von Korollar 2.81) diktieren diese Adressen eine Liste von Intervalle $I_1 \supseteq I_2 \supseteq \ldots$. Präziser formuliert ist

$$I_1 = \begin{cases} \left[0, \frac{1}{3}\right] & \text{falls } a_1 = \mathrm{l}, \\ \left[\frac{2}{3}, 1\right] & \text{falls } a_1 = \mathrm{r} \end{cases}$$

und rekursiv ist $I_{n+1}$ das linke Teilintervall von $I_n \cap C_{n+1}$ falls $a_{n+1} = \mathrm{l}$ und ansonsten das rechte Teilintervall. Nach dem Intervallschachtelungsprinzip (Satz 2.77) ist $\bigcap_{k=1}^{\infty} I_k$ nicht-leer und nach Konstruktion in der Cantor-Menge enthalten. Ist $x$ ein Element dieses Schnittes, dann gilt $I_k(x) = I_k$ und insbesondere $a_k(x) = a_k$ für alle $k \in \mathbb{N}_0$. Also ist $a(x) = a$ und $f$ ist surjektiv.

Zur Injektivität von $f$. Seien $x, y \in C$ mit $a(x) = a(y)$. Dann gilt also auch $I_k(x) = I_k(y)$ für alle $k \in \mathbb{N}_0$. Die Punkte $x, y$ liegen beide im Schnitt $\bigcap_{k=1}^{\infty} I_k(x) = \bigcap_{k=1}^{\infty} I_k(y)$. Wir behaupten, dass die Länge von einem Intervall $I_n$ in der Zerlegung von $C_n$ kleiner gleich $\frac{1}{n}$ ist. Nach dem Archimendischen Prinzip in Satz 2.68 und Übung 2.80 besteht also $\bigcap_{k=1}^{\infty} I_k(x)$ aus einem Punkt und somit gilt $x = y$.

Die Behauptung gilt sicherlich für $n = 0$, da $I_1 = [0, 1]$ Länge $1 = \frac{1}{1}$ hat. Angenommen sie gilt für $n \in \mathbb{N}_0$. Dann hat ein Intervall $I_{n+1}$ in der Zerlegung von $C_{n+1}$ ein Drittel der Länge eines Intervalles in der Zerlegung von $C_n$ und es gilt

$$\text{Länge}(I_{n+1}) \leq \frac{1}{3n} \leq \frac{1}{n+1}.$$

Dies schliesst den Beweis der Behauptung ab und impliziert somit den Satz. □

Wir stellen die Cantor-Menge im folgenden Applet dar und werden sie einige Male am Rande für weitere „fraktale“ Konstruktionen verwenden. Unser Hauptinteresse wird aber bei „glatten“ Objekten und weniger bei derartigen „fraktalen“ Objekten liegen.

**Applet 2.84** (Selbstähnlichkeit der Cantor-Menge)**.** *Wir stellen hier die Cantor-Menge dar, wobei Sie die Cantor-Menge vergrössern und verschieben können. Sie werden bemerken, dass die Cantor-Menge selbstähnlich ist, da sie den Vergrösserungsfaktor nur an der Beschriftung nicht aber an der Form der dargestellten Teilmenge feststellen können.*

## 2.7 Modelle und Eindeutigkeit der Menge der reellen Zahlen*

Wir wollen in diesem Abschnitt andeuten, warum die Axiome der reellen Zahlen die reellen Zahlen im Wesentlichen eindeutig festlegen. Damit dürfen wir ohne Weiteres eine beliebige Version der reellen Zahlen betrachten oder je nach Zusammenhang auch verschiedene Vorstellungen von reellen Zahlen haben. Zuerst wollen wir aber einige mögliche Modelle der reellen Zahlen ansprechen.

VO: wurde am 18.10. kurz besprochen.

### 2.7.1 Ebene Geometrie und die Zahlengerade

Falls wir von der zweidimensionalen euklidischen Geometrie ausgehen, so können wir $\mathbb{R}$ definieren indem wir eine Gerade $g$ in der Ebene gemeinsam mit zwei verschiedenen ausgezeichneten Punkten $P_0, P_1 \in g$ auswählen. Addition und Multiplikation wird dann mit geometrischen Konstruktionen (siehe folgendes Applet) so definiert, dass in $\mathbb{R} = g$ die Punkte $P_0$ und $P_1$ die Rolle von 0 und 1 in $\mathbb{R}$ übernehmen.

**Applet 2.85** (Definition mittels der Zahlengerade)**.** *Wir deuten an, wie man mittels parallelen Geraden die Addition und mittels dem Strahlensatz die Multiplikation auf der Zahlengerade definieren kann.*

### 2.7.2 Dezimalbrüche

Eine übliche Vorstellung der reellen Zahlen wird durch (im Allgemeinen nicht abbrechenden und vorzeichenbehafteten) Dezimalbrüche gegeben. Die Addition und Multiplikation von Dezimalbrüchen sind durch die bekannten Algorithmen gegeben. Allerdings muss die formale Beschreibung dieser Algorithmen auch nicht abbrechende Dezimalbrüche erlauben und die reellen Zahlen $\mathbb{R}$ müssen in diesem Zusammenhang als Quotientenraum von der Menge der Dezimalbrüche modulo einer Äquivalenzrelation definiert werden. Denn zum Beispiel stellen $1.00\ldots$ und $0.99\ldots$ dieselbe reelle Zahl dar, und es gibt unendlich viele reelle Zahlen mit zwei Dezimalbruchentwicklungen. Daher muss vor Besprechung der Axiome diese Äquivalenzrelation genau definiert werden und auch gezeigt werden, dass die Algorithmen wohldefinierte Abbildungen auf dem Quotientenraum $\mathbb{R}$ definieren.

Insgesamt ist die korrekte Konstruktion mit dieser Methode überraschend aufwendig. Des Weiteren gibt es keinen guten Grund nur Dezimalbrüche zu betrachten und nicht auch andere Basen zu erlauben (zum Beispiel Binärdarstellungen von Zahlen). Letzteres wirft aber die Frage auf, ob denn vielleicht manche Eigenschaften der reellen Zahlen davon abhängen, ob man 10 Symbole oder eine andere Anzahl verwendet. Die Eindeutigkeit der reellen Zahlen verneint diese Frage. Wir werden auf Dezimalbruchentwicklungen im Kapitel 6 nochmals zu sprechen kommen.

### 2.7.3 Dedekind-Schnitte

Formal einfacher ist die Konstruktion der reellen Zahlen ausgehend von den rationalen Zahlen durch sogenannte Dedekind-Schnitte, welche nach dem deutschen Mathematiker Dedekind

(1831-1916) benannt sind. Hier definiert man $\mathbb{R}$ als die Familie aller von oben beschränkten Teilmengen $A \subseteq \mathbb{Q}$ ohne Maximum, welche die Eigenschaft

$$\forall a \in A : \forall s \in \mathbb{Q} : (s \leq a \implies s \in A)$$

erfüllen (Strahlen in $\mathbb{Q}$). Der Hinweis, warum dies funktionieren sollte, ist in Übung 2.71 enthalten. Die rationale Zahl $r \in \mathbb{Q}$ identifiziert man mit der Teilmenge $\{s \in \mathbb{Q} \mid s < r\}$ und das Verhalten von Suprema unter Summenbildung in Proposition 2.63 erklärt, warum die Definition der Summe von „den Zahlen“ $A$ und $B$ durch $A + B = \{a + b \mid a \in A, b \in B\}$ gegeben ist. Die Definition der Multiplikation erfordert das Betrachten mehrerer Fälle und ist etwas aufwendiger (siehe auch Übung 2.66).

### 2.7.4 Vervollständigung der rationalen Zahlen

Es gibt eine wichtige Methode, wie man ausgehend von einem Raum mit einer Abstandsfunktion (einem sogenannten „metrischen Raum“) einen grösseren „vollständigen Raum“ definieren kann. Dies erfordert etwas mehr Theorie, die wir im Laufe des ersten Semesters besprechen werden, kann auch auf die Menge $\mathbb{Q}$ der rationalen Zahlen angewendet werden und liefert in diesem Fall eine Konstruktion der reellen Zahlen $\mathbb{R}$. Auch hier wird $\mathbb{R}$ als Quotientenraum einer grösseren Menge modulo einer Äquivalenzrelation definiert. Wir werden diesen Existenzbeweis später ausführlich besprechen.

### 2.7.5 Definition mittels Steigungen

Der schweizer Mathematiker A'Campo (geb. 1941) hat in diesem Jahrhundert (2003) eine weitere kuriose Konstruktion gefunden (siehe [A'C03]). Für dies nennen wir eine Abbildung $f : \mathbb{Z} \to \mathbb{Z}$ quasi-linear, falls sie die Eigenschaft

$$\big| \{f(m+n) - f(m) - f(n) \mid m, n \in \mathbb{Z}\} \big| < \infty$$

hat. Wir bezeichnen die Menge der quasi-linearen Abbildungen mit $\mathcal{Q}$ und die Menge der Abbildungen von $\mathbb{Z}$ nach $\mathbb{Z}$ mit endlichem Bild mit $\mathcal{K}$. Wir bemerken, dass $\mathcal{K} \subseteq \mathcal{Q}$ und zum Beispiel $\mathrm{id}_{\mathbb{Z}} \in \mathcal{Q} \setminus \mathcal{K}$. Man kann nun $\mathbb{R}$ als den Quotientenraum von $\mathcal{Q}$ definieren, wobei $f_1, f_2 \in \mathcal{Q}$ equivalent sind, wenn die Funktion

$$f_1 - f_2 : m \in \mathbb{Z} \to f_1(m) - f_2(m) \in \mathbb{Z}$$

endliches Bild besitzt, das heisst, in $\mathcal{K}$ liegt. Im Sinne der Algebra ist $\mathcal{Q}$ bereits eine Gruppe bezüglich „punktweiser Addition“, $\mathcal{K} \subseteq \mathcal{Q}$ ist eine Untergruppe, und $\mathbb{R} = \mathcal{Q}/\mathcal{K}$. Überraschenderweise ist die Multiplikation der reellen Zahlen $[f_1], [f_2] \in \mathcal{Q}/\mathcal{K}$ durch $[f_1 \circ f_2]$ und die 1 durch $[\mathrm{id}_{\mathbb{Z}}]$ gegeben. Eine rationale Zahl $r \in \mathbb{Q}$ wird in diesem Modell durch die Äquivalenzklasse $[f_r]$ der Funktion

$$f_r : m \in \mathbb{Z} \mapsto \lfloor rm \rfloor \in \mathbb{Z}$$

dargestellt. Gewissermassen wird hier eine reelle Zahl durch die „durchschnittliche Steigung“ einer quasi-linearen Abbildung definiert.

Sie sollten nicht zu viel Zeit dazu verwenden, diese Konstruktion vollständig zu verstehen, denn wir werden im Laufe des ersten Semesters mittels einigen Übungen ein besseres Verständnis für diesen Zugang erhalten. Vor allem aber wollten wir damit und auch mit den anderen, obigen Modellen Ihnen klarstellen, dass es nicht nur eine Art und Weise gibt, wie man die reellen Zahlen finden kann. Dies unterstreicht die Wichtigkeit der Frage nach der Eindeutigkeit.

### 2.7.6 Eindeutigkeit

Wie auch in anderen Situationen ist der Beweis der Eindeutigkeit einfacher als der Beweis der Existenz. Er enthält aber trotzdem gewisse Einblicke, die weiterentwickelt auch zum Existenzbeweis beitragen.

**Satz 2.86** (Eindeutigkeit der reellen Zahlen)**.** *Die Axiome von $\mathbb{R}$ in Abschnitt 2.1 legen die reellen Zahlen bis auf Isomorphie fest. Genauer formuliert gilt folgende Aussage:*

*Sei $\mathbb{R}'$ eine weitere Menge, auf der eine Addition, eine Multiplikation und eine kleinergleich-Relation definiert sind, so dass alle Axiome der reellen Zahlen erfüllt sind (das heisst, $\mathbb{R}'$ ist ein weiterer vollständiger angeordneter Körper). Wir bezeichnen mit $0' \in \mathbb{R}'$ das Nullelement in $\mathbb{R}'$ und mit $1' \in \mathbb{R}'$ das Einselement in $\mathbb{R}'$. Dann existiert eine bijektive Abbildung $\Phi : \mathbb{R} \to \mathbb{R}'$, so dass folgende Eigenschaften erfüllt sind.*

*(i)* $\Phi(0) = 0'$ *und* $\Phi(1) = 1'$.

*(ii)* *($\Phi$ ist additiv)* $\forall x, y \in \mathbb{R} : \Phi(x + y) = \Phi(x) + \Phi(y)$

*(iii)* *($\Phi$ ist multiplikativ)* $\forall x, y \in \mathbb{R} : \Phi(xy) = \Phi(x)\Phi(y)$

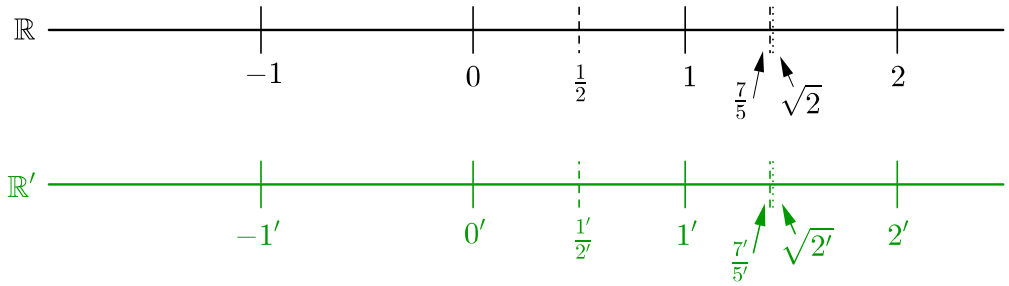
*(iv)* *($\Phi$ ist ordnungserhaltend)* $\forall x, y \in \mathbb{R} : (x \leq y \iff \Phi(x) \leq \Phi(y))$

Dieser Satz ist befriedigend, da es wegen ihm nicht darauf ankommt, welches Modell der reellen Zahlen man untersucht oder welche Konstruktion ausgehend aus den rationalen (oder auch aus den natürlichen) Zahlen wir verwenden, um die reellen Zahlen zu finden. Eine Abbildung $\Phi$ wie in Satz 2.86 nennen wir eine **Isomorphie** (von angeordneten Körpern), denn sie ist eine Bijektion, die alle Strukturen von $\mathbb{R}$ auf die entsprechenden Strukturen von $\mathbb{R}'$ abbildet.

Wie bereits bemerkt, könnte man obigen Satz auch mit Schach vergleichen, wo es ebenso nicht darauf ankommt, ob die Figuren aus Glas, Holz, Plastik oder Metall sind. Denn sobald man die verschiedenen Figuren richtig erkannt hat, besteht eine klare Korrespondenz (eine Isomorphie) zwischen den Figuren in dem Schachspiel aus Glas und dem Schachspiel aus Holz, und weiter kann man mit den einen genau dasselbe machen wie mit den anderen (die Isomorphie erhält alle möglichen Schachzüge).

Wir möchten im Folgenden einen Beweis zur Existenz einer solchen Abbildung $\Phi$ skizzieren und überlassen Interessierten die Verifikation einiger Details (siehe Übung 2.88). Wie wir

sehen werden besteht der Beweis gewissermassen aus einer Wiederholung vieler fundamentaler Themen aus diesem Kapitel. Etwas überspitzt lässt sich die Idee des Beweises wie folgt darstellen:




Figur 2.2: Wir definieren $\Phi$ zuerst so auf $0, 1, 2, ...$, dass $0$ auf $0'$, $1$ auf $1'$, $2$ auf $2'$ und so weiter abgebildet wird. Dann erweitern wir $\Phi$ auf alle Brüche so, dass beispielsweise $\frac{1}{2}$ auf $\frac{1'}{2'}$ abgebildet wird, und schliesslich verwenden wir das Supremum von Teilmengen von $\mathbb{Q}$ um $\Phi$ auf ganz $\mathbb{R}$ zu definieren. Dadurch erhalten wir, dass $\mathbb{R}'$ dasselbe ist wie $\mathbb{R}$, nur halt bloss in grün.

Des Weiteren ist folgende Übung ein hilfreicher erster Schritt für den Eindeutigkeitsbeweis.

**Übung 2.87.** *Sei $\varphi : \mathbb{R} \to \mathbb{R}$ eine bijektive Funktion, die $x < y \iff \varphi(x) < \varphi(y)$ für alle $x, y \in \mathbb{R}$ erfüllt. Zeigen Sie, dass $\varphi$ genau dann die Identitätsabbildung ist, wenn $\varphi(x) = x$ für alle $x \in \mathbb{Q}$ gilt.*

*Beweisskizze.* Angenommen $\mathbb{R}$ und $\mathbb{R}'$ sind zwei Mengen von reellen Zahlen, die jeweils die Axiome in Abschnitt 2.1 und damit auch deren Folgerungen in den Abschnitten 2.1, 2.2, 2.5 und 2.6.1 erfüllen. Seien $\mathbb{N}, \mathbb{N}_0, \mathbb{Z}, \mathbb{Q}$ respektive $\mathbb{N}', \mathbb{N}_0', \mathbb{Z}', \mathbb{Q}'$ die in $\mathbb{R}$ respektive $\mathbb{R}'$ konstruierten natürlichen Zahlen, ganzen Zahlen und rationalen Zahlen. Wir werden diese verwenden, um die gewünschte Abbildung in mehreren Schritten zu definieren.

*Definition von $\Phi$ auf $\mathbb{N}_0$.* Wir verwenden Rekursion, um eine Abbildung $\Phi_0 : \mathbb{N}_0 \to \mathbb{N}_0'$ mit „guten“ Eigenschaften zu definieren. In der Tat setzen wir $\Phi_0(0) = 0'$, $\Phi_0(1) = 1'$ und mittels Rekursion für alle $n \in \mathbb{N}$

$$\Phi_0(n+1) = \Phi_0(n) + 1'.$$

Mittels vollständiger Induktion in $\mathbb{N}$ kann man nun zeigen, dass $\Phi_0$ dadurch eindeutig festgelegt ist und dass die Regeln

$$\Phi_0(m+n) = \Phi_0(m) + \Phi_0(n)$$
$$\Phi_0(mn) = \Phi_0(m)\Phi_0(n)$$

für alle $m, n \in \mathbb{N}_0$ gelten. Die Verifikation dieser Regeln überlassen wir als Übung und ist sehr ähnlich dem Beweis von Lemma 2.17, nach dem $\mathbb{N}$ unter Addition und Multiplikation abgeschlossen ist. Des Weiteren gilt

$$\Phi_0(m) = 0 \iff m = 0$$

Insbesondere implizieren obigen Eigenschaften, dass $\Phi_0$ injektiv ist (wieso?). Da das Bild von $\Phi_0$ sowohl $0'$ als auch $1'$ enthält und induktiv ist, ist $\Phi_0$ auch surjektiv und damit bijektiv. Des Weiteren gilt (wegen Lemma 2.23) für $m, n \in \mathbb{N}_0$ auch

$$m \leq n \iff (\exists k \in \mathbb{N}_0 : n = m + k),$$

was auf Grund der obigen Eigenschaften von $\Phi_0$ auch

$$m \leq n \iff (\exists k' \in \mathbb{N}_0' : \Phi_0(n) = \Phi_0(m) + k') \iff \Phi_0(m) \leq \Phi_0(n)$$

impliziert.

*Definition von $\Phi$ auf $\mathbb{Z}$.* Für $n \in \mathbb{Z}$ definieren wir

$$\Phi_{\mathbb{Z}}(n) = \begin{cases} \Phi_0(n) \in \mathbb{N}_0' & \text{falls } n \in \mathbb{N}_0 \\ -\Phi_0(-n) \in -\mathbb{N}' & \text{falls } n \in -\mathbb{N} \end{cases}$$

und erhalten dadurch eine Abbildung $\Phi_{\mathbb{Z}} : \mathbb{Z} \to \mathbb{Z}'$, die ebenso obige Eigenschaften erfüllt.

*Definition von $\Phi$ auf $\mathbb{Q}$.* Für $m \in \mathbb{Z}$ und $n \in \mathbb{N}$ setzen wir $\Phi_{\mathbb{Q}}(\frac{m}{n}) = \frac{\Phi_0(m)}{\Phi_0(n)} \in \mathbb{Q}'$. Somit erhalten wir eine Abbildung $\Phi_{\mathbb{Q}} : \mathbb{Q} \to \mathbb{Q}'$, die noch immer alle gewünschten Eigenschaften erfüllt. Insbesondere gilt $a \leq b \iff \Phi_{\mathbb{Q}}(a) \leq \Phi_{\mathbb{Q}}(b)$ zuerst für $a, b \in \mathbb{Z}$ und dann (wegen $\Phi_{\mathbb{Q}}(na) = \Phi_{\mathbb{Q}}(n)\Phi_{\mathbb{Q}}(a)$ für $a \in \mathbb{Q}$ und $n \in \mathbb{N}$) auch für $a, b \in \mathbb{Q}$.

*Definition von $\Phi$ auf $\mathbb{R}$.* Für jedes $x \in \mathbb{R}$ gilt wegen dem Korollar 2.70 des Archimedischen Prinzip (Satz 2.68) und Übung 2.71

$$x = \sup \{r \in \mathbb{Q} \mid r < x\}.$$

Deswegen drängt sich die Definition

$$\Phi(x) = \sup{}' \{\Phi_{\mathbb{Q}}(r) \mid r \in \mathbb{Q},\ r < x\} = \sup{}' \Phi_{\mathbb{Q}}(\{r \in \mathbb{Q} \mid\ r < x\})$$

für alle $x \in \mathbb{R}$ auf, wobei $\sup'$ das Supremum auf den in $\mathbb{R}'$ von oben beschränkten Teilmengen bezeichnet. Dazu gehört auch die Teilmenge $\{\Phi_{\mathbb{Q}}(r) \mid r \in \mathbb{Q},\ r < x\}$, da eine obere Schranke durch $\Phi_{\mathbb{Q}}(\lfloor x \rfloor + 1)$ gegeben ist (wieso?). Für $x \in \mathbb{Q}$ gilt nun

$$\Phi(x) = \sup{}' \{\Phi_{\mathbb{Q}}(r) \mid r \in \mathbb{Q},\ r < x\} = \sup{}' \{\Phi_{\mathbb{Q}}(r) \mid r \in \mathbb{Q},\ r < x\} = \Phi_{\mathbb{Q}}(x)$$

da $\Phi_{\mathbb{Q}}$ die Ordnung erhält. Für $x \leq y$ gilt

$$\{r \in \mathbb{Q} \mid r < x\} \subseteq \{r \in \mathbb{Q} \mid r < y\},$$
$$\Phi_{\mathbb{Q}}(\{r \in \mathbb{Q} \mid r < x\}) \subseteq \Phi_{\mathbb{Q}}(\{r \in \mathbb{Q} \mid r < y\})$$

und damit

$$\Phi(x) = \sup{}' \{\Phi_{\mathbb{Q}}(r) \mid r \in \mathbb{Q},\ r < x\} \leq \sup{}' \{\Phi_{\mathbb{Q}}(r) \mid r \in \mathbb{Q},\ r < y\} = \Phi(y).$$

Für $x < y$ existieren $s, t \in \mathbb{Q}$ mit $x \leq s < t \leq y$ nach Korollar 2.70, was

$$\Phi(x) \leq \Phi(s) < \Phi(t) \leq \Phi(y)$$

und damit $\Phi(x) < \Phi(y)$ impliziert. Insbesondere ist $\Phi$ injektiv. Für $x, y \in \mathbb{R}$ gilt (siehe Übung 2.88d) )

$$\{r \in \mathbb{Q} \mid r < x\} + \{s \in \mathbb{Q} \mid s < y\} = \{t \in \mathbb{Q} \mid t < x + y\}\,, \tag{2.1}$$

was gemeinsam mit dem Verhalten des Supremums unter Summenbildung von Mengen in Proposition 2.63 und der Additivität von $\Phi_{\mathbb{Q}}$ gerade $\Phi(x + y) = \Phi(x) + \Phi(y)$ impliziert. Insbesondere gilt wegen $\Phi(0) = 0$ auch $\Phi(-x) = -\Phi(x)$ für alle $x \in \mathbb{R}$. Auch gilt für alle $x, y \in \mathbb{R}_{>0}$, dass

$$\{r \in \mathbb{Q} \mid 0 < r < x\} \cdot \{s \in \mathbb{Q} \mid 0 < s < y\} = \{t \in \mathbb{Q} \mid 0 < t < xy\} \tag{2.2}$$

und damit $\Phi(x)\Phi(y) = \Phi(xy)$ für alle $x, y \in \mathbb{R}_{>0}$ wegen der Multiplikativität von $\Phi_{\mathbb{Q}}$ und Übung 2.66. Da wir aber bereits wissen, dass $\Phi(-x) = -\Phi(x)$ für alle $x \in \mathbb{R}$, gilt $\Phi(x)\Phi(y) = \Phi(xy)$ für alle $x, y \in \mathbb{R}$.

Wir müssen noch zeigen, dass $\Phi$ bijektiv ist. Durch Vertauschen von $\mathbb{R}$ mit $\mathbb{R}'$ erhalten wir analog eine Funktion $\Psi : \mathbb{R}' \to \mathbb{R}$, die unter anderem $x' < y' \iff \Psi(x') < \Psi(y')$ für alle $x', y' \in \mathbb{R}'$ erfüllt. Die Verknüpfung $\Psi \circ \Phi : \mathbb{R} \to \mathbb{R}$ genügt $x < y \iff \Psi \circ \Phi(x) < \Psi \circ \Phi(y)$ für alle $x, y \in \mathbb{R}$. Nach Konstruktion ist $\Psi \circ \Phi$ auf $\mathbb{Q}$ die Identitätsabbildung und damit $\Psi \circ \Phi = \mathrm{id}_{\mathbb{R}}$ nach Übung 2.87. Analog zeigt man, dass $\Phi \circ \Psi = \mathrm{id}_{\mathbb{R}'}$ gilt, womit die Abbildung $\Phi$ bijektiv sein muss.

Überprüfen Sie nun, dass wir damit alle notwendigen Eigenschaften von $\Phi$ gezeigt haben. □

**Übung 2.88.** *Beweisen Sie die in obiger Beweisskizze unterlassenen Details:*

*(i) Zeigen Sie die in obigen Beweis behaupteten Eigenschaften von $\Phi_0$.*

*(ii) Stellen Sie eine Liste von Eigenschaften der Abbildung $\Phi_{\mathbb{Q}}$ auf (in Analogie zu $\Phi_0$) und beweisen Sie diese.*

*(iii) Zeigen Sie, dass die Einschränkung von $\Phi$ auf $\mathbb{Q}$ mit $\Phi_{\mathbb{Q}}$ übereinstimmt.*

*(iv) Beweisen Sie die Gleichungen (2.1) und (2.2).*

*(v) Zeigen Sie, dass jede additive und multiplikative Abbildung $\phi : \mathbb{Q} \to \mathbb{Q}$ mit $\phi(0) = 0$ und $\phi(1) = 1$ die Identitätsabbildung ist.*

**Übung 2.89.** *Zeigen Sie die Eindeutigkeit der Abbildung $\Phi$ in Satz 2.86.*

## 2.8 Weitere Lernmaterialien

### 2.8.1 Verwendung des Kapitels

Die Themen dieses Kapitels stellen den Anfang unserer Entwicklung der Analysis dar und sind aus diesem Grunde für das Folgende fundamental. Wie bereits erwähnt werden wir die üblichen Eigenschaften der reellen, natürlichen, ganzen, rationalen und komplexen Zahlen (inklusive der Konjugation komplexer Zahlen) im Folgenden ohne Verweise verwenden. Es ist auch nicht notwendig, die Beweise der elementaren Aussagen in Abschnitt 2.1 auswendig zu lernen. Manche der Beweise in Abschnitt 2.2 sind auch etwas zu formal, als dass sie für das Folgende von grosser Bedeutung sein werden. Für ein fundiertes Verständnis der Induktion sind die besprochenen Varianten der Induktion samt Beweise wichtig und auch die Beweise der algebraischen und geometrischen Aussagen stellen eine gute Übung dar. In Abschnitt 2.4 haben wir einige Ihnen wahrscheinlich bekannte Definition ausgesprochen, doch werden auch die Ihnen wahrscheinlich neuen Begriffe „offen“ und „abgeschlossen“ zunehmend an Bedeutung gewinnen.

Die Kernthemen dieses Kapitels sind hingegen in folgender Liste enthalten.

- Das Vollständigkeitsaxiom in Abschnitt 2.1.3.
- Existenz und Eigenschaften des Supremums und Infimums in Abschnitt 2.5 (inbesondere beispielsweise die Unterscheidung von Maximum und Supremum).
- Korollare der Vollständigkeit in Abschnitt 2.6: Das Archimedische Prinzip (Satz 2.68), die Existenz von Häufungspunkten für beschränkte unendliche Mengen (Satz 2.75), das Intervallschachtelungsprinzip (Satz 2.77), und die Überabzählbarkeit von $\mathbb{R}$ in Korollar 2.81.

Diese Themen und deren Beweismethoden sind von zentraler Bedeutung für das Folgende und Sie werden weitere Vorlesungsstunden besser verstehen, wenn Sie diese Kernthemen bereits im Gedächnis und auf Abruf bereit haben.

Im Laufe dieses Kapitels haben wir auch bereits einige grundlegende Funktionen eingeführt, welche wir ohne Verweis und mit den üblichen Eigenschaften in Zukunft wieder benötigen werden.

- Die Körperoperationen auf $\mathbb{R}$ oder $\mathbb{C}$: Addition, Subtraktion, Multiplikation, Division.
- Das Quadrieren $(\cdot)^2$ auf $\mathbb{R}$ oder $\mathbb{C}$.
- Die Quadratwurzel $\sqrt{\cdot} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$.
- Der Absolutbetrag $|\cdot|$ auf $\mathbb{R}$ oder $\mathbb{C}$.
- Die Vorzeichenfunktion $\operatorname{sgn}(\cdot)$ auf $\mathbb{R}$.
- Der ganzzahlige Anteil $\lfloor \cdot \rfloor : \mathbb{R} \to \mathbb{Z}$.
- Der Nachkommaanteil $\{\cdot\} : \mathbb{R} \to [0, 1)$.

- Das Maximum $\max(x, y) = \max(\{x, y\})$ und das Minimum $\min(x, y) = \min(\{x, y\})$ zweier reeller Zahlen $x, y \in \mathbb{R}$ ergeben sich als Spezialfälle von Maximum und Minimum der Menge $\{x, y\}$ (welche auf Grund einer Fallunterscheidung basierend auf die Trichotomie reeller Zahlen immer existieren).

### 2.8.2 Weitere Übungsaufgaben

**Übung** (Parallelogrammidentität). *Zeigen Sie für alle $z, w \in \mathbb{C}$ die Gleichung*

$$|z + w|^2 + |z - w|^2 = 2(|z|^2 + |w|^2).$$

**Übung** (Mittelsenkrechte). *Seien $w_1, w_2 \in \mathbb{C}$ zwei verschiedene Punkte. Erklären und beweisen Sie, wieso die Teilmenge $\{z \in \mathbb{C} \mid |z - w_1| = |z - w_2|\}$ eine Gerade ist. Eine Gerade ist dabei eine Teilmenge der Form $\{a + tv \mid t \in \mathbb{R}\}$ für $a, v \in \mathbb{C}$.*

**Übung** (Körper mit zwei Elementen). *Zeigen Sie, dass die Menge $\mathbb{F}_2$ mit den in Übung 2.7 definierten Operationen einen Körper mit zwei Elementen bildet. Wieso gibt es keinen Körper mit nur einem Element?*

**Übung** (Körper von Primzahlordnung). *Sei $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ gegeben wie in Abschnitt 1.5.2 durch Kongruenzen modulo einer Primzahl $p$. Wie in Übung 1.93 gezeigt wurde, sind wohldefinierte Operationen der Addition und Multiplikation auf $\mathbb{F}_p$ durch*

$$(a + p\mathbb{Z}) + (b + p\mathbb{Z}) = (a + b) + p\mathbb{Z}$$
$$(a + p\mathbb{Z}) \cdot (b + p\mathbb{Z}) = (ab) + p\mathbb{Z}$$

*für $a + p\mathbb{Z}, b + p\mathbb{Z} \in \mathbb{F}_p$ definiert. Wir möchten in dieser Übung zeigen, dass $(\mathbb{F}_p, +, \cdot)$ ein Körper ist.*

*(i) Zeigen Sie, dass $\mathbb{F}_p$ mit diesen beiden Operationen allen Körperaxiomen bis auf (6) genügt, wobei das Nullelement durch $0 + p\mathbb{Z}$ und das Einselement durch $1 + p\mathbb{Z}$ gegeben ist.*

*(ii) Verifizieren Sie mit Division durch Rest, dass $\mathbb{F}_p$ genau $p$ Elemente hat (und insbesondere endlich ist).*

*(iii) Zeigen Sie, dass jedes Element $a + p\mathbb{Z} \neq 0 + p\mathbb{Z}$ von $\mathbb{F}_p$ eine multiplikative Inverse besitzt. Betrachten Sie dazu die Multiplikation mit diesem Element auf $\mathbb{F}_p$ und überprüfen Sie zuerst, dass diese injektiv (und damit auch surjektiv) ist.*

*(iv) Zeigen Sie, dass es keine Ordnung auf $\mathbb{F}_p$ gibt, die $\mathbb{F}_p$ zu einem angeordnetem Körper macht.*

*Somit haben wir für jede Primzahl $p$ ein Körper dieser Kardinalität konstruiert. Aus ähnlichen Gründen wie auch für $\mathbb{C}$ lässt sich auf keinem dieser Körper eine Ordnung definieren, die diese*

*zu geordneten Körper macht. Wir bemerken auch, dass sich für jede Primzahlpotenz wie zum Beispiel* 4 *oder* 9 *ein Körper definieren lässt; siehe nächstes Kapitel.*

**Übung.** *Entscheiden Sie bei den folgenden Teilmengen von* $\mathbb{C}$ *jeweils, ob sie offen, abgeschlossen oder weder noch sind.*

- *Die Zahlenmengen* $\emptyset, \mathbb{N}, \mathbb{Z}, \mathbb{R}, \mathbb{C}$.
- *Die Teilmenge der komplexen Zahlen mit Absolutbetrag Eins.*
- *Das Rechteck* $\{z \in \mathbb{C} \mid a < \mathrm{Re}(z) < b,\ c < \mathrm{Im}(z) < d\}$ *für* $a, b, c, d \in \mathbb{R}$ *mit* $a < b$ *und* $c < d$.

**Übung** (Topologie auf $\mathbb{R}$ und $\mathbb{C}$)**.** *Sei* $\mathcal{T}$ *die Menge der offenen Teilmengen von* $\mathbb{C}$. *Zeigen Sie, dass folgende Eigenschaften erfüllt sind.*

- $\emptyset \in \mathcal{T}$ *und* $\mathbb{C} \in \mathcal{T}$.
- *Für* $U_1, ..., U_n \in \mathcal{T}$ *ist* $\bigcap_{i=1}^n U_i \in \mathcal{T}$.
- *Für eine Kollektion* $\mathcal{U} \subseteq \mathcal{T}$ *gilt* $\bigcup_{U \in \mathcal{U}} U \in \mathcal{T}$.

*In Worten ausgedrückt sind also endliche Schnitte und beliebige Vereinigungen von offenen Mengen offen. Die analoge Aussage gilt für die offenen Teilmengen von* $\mathbb{R}$. *Was gilt für abgeschlossene Mengen?*

In Abschnitt 2.6.1 haben wir bereits beschrieben, was Dichtheit der rationalen Zahlen in $\mathbb{R}$ bedeutet. Allgemeiner sagt man, dass eine Teilmenge $A \subseteq \mathbb{R}$ **dicht** ist, wenn für jedes offene, nicht-leere Intervall $I \subseteq \mathbb{R}$ der Schnitt $I \cap A$ nicht-leer ist.

**Übung** (Charakterisierung von Dichtheit)**.** *Zeigen Sie, dass folgende Aussagen über eine Teilmenge* $A \subseteq \mathbb{R}$ *äquivalent sind.*

*(i)* $A$ *ist dicht.*

*(ii) Die Menge der Häufungspunkte von* $A$ *ist gleich* $\mathbb{R}$.

*(iii) Jede abgeschlossene Menge, die* $A$ *enthält, ist gleich* $\mathbb{R}$.

**Übung** (Dichtheit der irrationalen Zahlen)**.** *Zeigen Sie, dass die Menge* $\mathbb{R} \setminus \mathbb{Q}$ *der irrationalen Zahlen dicht liegt in* $\mathbb{R}$.

**Übung.** *Berechnen Sie die Häufungspunkte folgender Teilmengen von* $\mathbb{R}$.

$$\left\{\tfrac{1}{n} \mid n \in \mathbb{N}\right\}, \quad (0,1), \quad \left\{\tfrac{1}{1-r} \mid r \in (-1,1)\right\}$$

**Übung** (Supremum als Häufungspunkt)**.** *Sei* $A \subseteq \mathbb{R}$ *eine von oben beschränkte Teilmenge. Zeigen Sie, dass* $A$ *ein Maximum besitzt oder das Supremum von* $A$ *ein Häufungspunkt der Menge* $A$ *ist.*

**Übung** (Überabzählbare Mengen haben Häufungspunkte). *Sei $A \subseteq \mathbb{R}$ überabzählbar (aber möglicherweise unbeschränkt). Zeigen Sie, dass dann $A$ einen Häufungspunkt besitzt.*

**Übung.** *Finden Sie für jedes $n \in \mathbb{N}$ ein Intervall $I_n = [a_n, b_n]$ mit rationalen Endpunkten $a_n, b_n \in \mathbb{Q}$ wie in obigem Satz, so dass $\bigcap_{n=1}^{\infty} I_n = \{\sqrt{2}\}$ gilt. Schliessen Sie daraus, dass das Intervallschachtelungsprinzip in $\mathbb{Q}$ nicht erfüllt ist. (Hierbei ist ein Intervall in $\mathbb{Q}$ definiert als der Durchschnitt von $\mathbb{Q}$ mit einem reellen Intervall mit rationalen Endpunkten.)*

**Übung** (Das Vollständigkeitsaxiom und das Supremum). *Zeigen Sie, dass Satz 2.59 zum Vollständigkeitsaxiom (Axiom* (16)*) äquivalent ist. Genauer formuliert: zeigen Sie, dass die Axiome eines angeordneten Körpers (das wären Axiome* (1)–(15)*) gemeinsam mit der Aussage in Satz 2.59 das Vollständigkeitsaxiom (Axiom* (16)*) implizieren.*

**Übung** (Weitere Formen des Vollständigkeitsaxioms). *Zeigen Sie in Analogie zu obiger Übung, dass*

*(i) der Satz über die Existenz von Häufungspunkten sowie*

*(ii) das Intervallschachtelungsprinzip zusammen mit dem Archimedischen Prinzip*

*(unter Annahme der Axiome eines angeordneten Körpers* (1)–(15)*) äquivalent sind zum Vollständigkeitsaxiom.*

**Übung.** *Zeigen Sie, dass jede offene Teilmenge von $\mathbb{R}$ überabzählbar ist.*

**Übung** (Multiplikation mit 3 auf der Cantor-Menge). *Zeigen Sie, dass die Abbildung*

$$m_3 : C_{n+1} \to C_n, \quad x \mapsto \begin{cases} 3x & \text{falls } x \in [0, \frac{1}{3}], \\ 3(x - \frac{2}{3}) & \text{falls } x \in [\frac{2}{3}, 1]. \end{cases}$$

*wohldefiniert ist. Intuitiv sagt uns die Abbildung $m_3$ also, dass $C_{n+1}$ aus zwei Hälften besteht, die jeweils aussehen wie kontrahierte Kopien von $C_n$. (Wieso?)*

**Übung** (Rechtecksschachtelungprinzip in $\mathbb{C}$). *Wir bezeichnen eine Menge der Form*

$$R = [a, b] \times [c, d] = \{z = x + y\mathrm{i} \mid x \in [a, b],\ y \in [c, d]\}$$

*als ein abgeschlossenes beschränktes Rechteck. Beweisen Sie folgendes Rechtecksschachtelungsprinzip in $\mathbb{C}$: Seien $R_n$ für jedes $n \in \mathbb{N}$ ein abgeschlossenes beschränktes Rechteck so dass $R_m \supseteq R_n$ für $m \leq n$. Dann ist der abzählbare Durchschnitt $\bigcap_{n=1}^{\infty} R_n$ nicht-leer.*

**Übung** (Häufungspunkte in $\mathbb{C}$). *Sei $A \subseteq \mathbb{C}$ und $z_0 \in \mathbb{C}$. Dann heisst $z_0$ ein Häufungspunkt von der Menge $A$ falls es zu jedem $\varepsilon > 0$ ein $a \in A$ gibt mit $0 < |a - z_0| < \varepsilon$. Sei nun $A$ eine unendliche und beschränkte (das heisst, es existiert $M > 0$ mit $A \subseteq B_M(0)$) Teilmenge. Zeigen Sie, dass ein Häufungspunkt der Menge $A$ in $\mathbb{C}$ existiert.*

*Eine kurze Anleitung: Auf Grund der Beschränktheit der Menge $A$ existiert ein $D > 0$ so dass $A \subseteq [-D, D] \times [-D, D]$. Sie können für den Beweis zuerst obiges Rechtecksschachtelungsprinzip beweisen und dann verwenden. Alternativ können Sie den Beweis von Satz 2.75*

*adaptieren: definieren Sie*

$$\begin{aligned} X &= \{x \in \mathbb{R} \mid \left|A \cap ([-\infty, x] \times \mathbb{R})\right| < \infty\} \\ x_0 &= \sup X \\ Y &= \{y \in \mathbb{R} \mid \exists \varepsilon > 0 : \left|A \cap ([x_0 - \varepsilon, x_0 + \varepsilon] \times (-\infty, y])\right| < \infty\} \\ y_0 &= \sup Y \end{aligned}$$

*und zeigen Sie, dass* $x_0 + y_0\mathrm{i}$ *ein Häufungspunkt ist.*

### 2.8.3 Multiple-Choice Fragen

In folgendem Fragen-Typ können Sie aus mehreren verschiedenen Lösungsvarianten auswählen (Multiple-Choice). Es ist dabei nicht unbedingt nur eine Lösung richtig.

**Übung.** *Sei* $A \subseteq \mathbb{R}$ *und* $x_0 \in \mathbb{R}$. *Welche der folgenden Aussagen sind äquivalent zur Aussage, dass* $x_0$ *ein Häufungspunkt von* $A$ *ist?*

*(1)* $\forall \varepsilon > 0\ \exists! a \in A : 0 < |a - x_0| < \varepsilon$.

*(2)* $\forall \varepsilon > 0\ \exists a \in A : 0 < |a - x_0| < \varepsilon$.

*(3)* $\exists \varepsilon_0 > 0\ \forall \varepsilon \in (0, \varepsilon_0)\ \exists a \in A : 0 < |a - x_0| < \varepsilon$.

*(4)* $\forall \varepsilon > 1\ \exists a \in A : 0 < |a - x_0| < \varepsilon$.

**Übung.** *Welche der folgenden Punkte liegen in der Cantor-Menge?*

*(1)* $\frac{2}{3}$

*(2)* $\frac{7}{16}$

*(3)* $\frac{8}{27}$

*(4)* $\frac{4}{15}$

**Übung** (Challenge)**.** *Gibt es eine Kollektion* $\{A_t \mid t \in \mathbb{R}\}$ *von Teilmengen von* $\mathbb{N}$ *mit der Eigenschaft* $A_t \subsetneq A_{t'}$ *für alle* $t < t'$ *in* $\mathbb{R}$ *und* $\bigcup_{t \in \mathbb{R}} A_t = \mathbb{N}$*?*

### 2.8.4 Lernkarten

Für das Lernen der Themen dieses Kapitels könnten wieder die Lernkarten helfen. Des Weiteren können Sie auch die Zusammenhänge der einzelnen Sätze dieses Kapitels mit Hilfe eines Graphen wiederholen. (Die Darstellung wird zufällig generiert, und es empfiehlt sich die Knoten des Graphen etwas logischer zu platzieren.)

# Kapitel 3

# Funktionen und die reellen Zahlen

Wir haben im letzten Kapitel die reellen Zahlen eingeführt und ihre grundlegenden Eigenschaften besprochen. In diesem Kapitel werden wir einige weitere elementare Konstruktionen und reellwertige oder komplexwertige Funktionen betrachten. Des Weiteren werden wir für solche Funktionen erste Eigenschaften definieren und das Vollständigkeitsaxiom (in der Form der Existenz des Supremums) dazu verwenden, grundlegendes Wissen über „stetige“ Funktionen auf Intervallen zu erarbeiten.

## 3.1 Summen und Produkte

Sei $n \in \mathbb{N}$ und seien $a_1, \ldots, a_n \in \mathbb{C}$ oder $a_1, \ldots, a_n$ Elemente eines Vektorraums $V$ (wie zum Beispiel $\mathbb{R}^d$ für ein $d \geq 1$). Wir wollen hier für eine natürliche Zahl $n \in \mathbb{N}$ die **Summe** von $a_1$ bis $a_n$, also

$$\sum_{j=1}^{n} a_j = a_1 + \ldots + a_n,$$

besprechen und formal korrekt definieren.

Vom formalen Standpunkt her gesehen ist $j \mapsto a_j \in V$ eine Funktion, (die oft auch durch eine konkrete Formel gegeben sein wird und) deren Definitionsbereich die Menge

$$\{j \in \mathbb{N} \mid 1 \leq j \leq n\}$$

enthalten muss. Wir können $\sum_{i=1}^{n} a_j$ rekursiv definieren durch

$$\sum_{j=1}^{1} a_j = a_1 \text{ und } \sum_{j=1}^{k+1} a_j = \left(\sum_{j=1}^{k} a_j\right) + a_{k+1}$$

für $k \in \{1, \ldots, n-1\}$. Diese Definition entspricht einem einfachen rekursiven Algorithmus, um die Summe $\sum_{i=1}^{n} a_j$ zu berechnen. Allgemeiner ist die Summe $\sum_{i=m}^{n} a_j$ für ganze Zahlen

$m, n$ ebenso rekursiv durch

$$\sum_{j=m}^{n} a_j = \begin{cases} 0 & \text{falls } m > n, \\ a_m & \text{falls } m = n \text{ und} \\ \left(\sum_{j=m}^{n-1} a_j\right) + a_n & \text{falls } m < n \end{cases}$$

definiert. Wir werden $a_j$ als die **Summanden** und $j$ als den **Index** der Summe $\sum_{j=1}^{n} a_j$ bezeichnen.

Falls nun $m, n$ ganze Zahlen und $a_m, \ldots, a_n \in \mathbb{C}$ sind, dann können wir auch das **Produkt** $\prod_{j=m}^{n} a_j$ von $a_m$ bis $a_n$ rekursiv durch

$$\prod_{j=m}^{n} a_j = \begin{cases} 1 & \text{falls } m > n, \\ a_m & \text{falls } m = n \text{ und} \\ \left(\prod_{j=m}^{n-1} a_j\right) \cdot a_n & \text{falls } m < n \end{cases}$$

definieren. Wir werden $a_j$ als die **Faktoren** und $j$ als den **Index** des Produkts $\prod_{j=m}^{n} a_j$ bezeichnen.

Der Index $j$ in der Summe $\sum_{j=1}^{n} a_j$ oder dem Produkt $\prod_{j=1}^{n} a_j$ hat ausserhalb der Summe oder dem Produkt keinerlei Bedeutung; er ist sozusagen eine interne Variable für das rekursive Teilprogramm und wird von einem Programm, welches das Teilprogramm aufruft, nicht gesehen. Insbesondere gilt

$$\sum_{j=m}^{n} a_j = \sum_{k=m}^{n} a_k = \sum_{\ell=m}^{n} a_\ell$$

und analog für das Produkt. Manchmal werden wir auch eine **Indexverschiebung** anwenden, wie zum Beispiel in

$$\sum_{j=m}^{n} a_j = \sum_{k=m-1}^{n-1} a_{k+1} = \sum_{\ell=m+1}^{n+1} a_{\ell-1}. \tag{3.1}$$

Dies lässt sich direkt mittels vollständiger Induktion beweisen (siehe die folgende Übung), doch wollen wir bemerken, dass es leicht ist, sich diese Formeln zu merken. Statt diese auswendig zu lernen, überprüfen Sie einfach bei Auftreten von Indexverschiebungen dieser Form bei beiden Summen, ob jeweils diesselben Ausdrücke für den ersten und den letzten Summanden auftreten.

**Übung 3.1** (Indexverschiebung)**.** *Beweisen Sie Gleichung* (3.1) *und eine analoge Formel für das Produkt mittels vollständiger Induktion.*

Der einfachste Fall einer Funktion $j \mapsto a_j$ ist der Fall der konstanten Funktion $a_j = z$ für ein $z$ und für alle $j$. In diesem Fall ergibt sich die Summe zu $\sum_{j=1}^{n} z = nz$ für alle $n \in \mathbb{N}$ und $z \in \mathbb{C}$ (oder $z$ in einem Vektorraum). Im Falle des Produkts erhalten wir aber die Definition

der **Potenzfunktion** für $z \in \mathbb{C}$ und $n \in \mathbb{N}$

$$z^n = \prod_{j=1}^{n} z,$$

die somit rekursiv durch

$$z^1 = z \text{ und } z^{n+1} = z^n z$$

für alle $n \in \mathbb{N}$ definiert ist. Wir nennen $z$ die **Basis** und $n$ den **Exponenten**. Wir erweitern diese Definition durch

$$z^0 = 1 \tag{3.2}$$

für alle $z \in \mathbb{C}$ (insbesondere[1] für $z = 0$) und

$$z^{-n} = (z^n)^{-1}$$

für alle $z \in \mathbb{C}^\times := \mathbb{C} \setminus \{0\}$ und $n \in \mathbb{N}$. Die nächste Übung zeigt, dass die so definierte Potenzfunktion die üblichen Rechenregeln erfüllt.

**Wichtige Übung 3.2.** *Beweisen Sie* $(zw)^m = z^m w^m$, $z^{m+n} = z^m z^n$ *und* $(z^m)^n = z^{mn}$ *zuerst für alle* $z, w \in \mathbb{C}$ *und* $m, n \in \mathbb{N}_0$ *mit vollständiger Induktion und dann für alle* $z, w \in \mathbb{C}^\times$ *und* $m, n \in \mathbb{Z}$.

*Bemerkung.* Formal gesehen sollten wir auch alle weiteren Rechenregeln in diesem Abschnitt mit vollständiger Induktion beweisen. Da uns diese Beweise aber sehr wenig lehren, werden wir darauf verzichten.

### 3.1.1 Rechenregeln für die Summe

Die Summe erfüllt für gegebene ganze Zahlen $m, n$ mit $m \leq n$ die Gleichungen

$$\sum_{k=m}^{n} (a_k + b_k) = \sum_{k=m}^{n} a_k + \sum_{k=m}^{n} b_k$$

und

$$\sum_{k=m}^{n} (ca_k) = c \sum_{k=m}^{n} a_k,$$

wobei $a_m, \ldots, a_n$, $b_m, \ldots, b_n$ in einem reellen (respektive komplexen) Vektorraum $V$ liegen und $c \in \mathbb{R}$ (respektive $c \in \mathbb{C}$) ein Skalar ist. (Die erste Eigenschaft ist eine Mischung aus

[1] Da $z \in \mathbb{C} \mapsto z^0$ die konstante Funktion mit Wert 1 darstellt, wäre es sehr eigenartig (und im nächsten Abschnitt bei der Diskussion von Polynomen extrem störend), wenn wir diese Funktion für $z = 0$ undefiniert lassen oder mit einem anderen Wert versehen. Trotzdem ist der Ausdruck $0^0$ undefiniert, wenn dieser losgelöst von der Diskussion der Potenzfunktion $z \in \mathbb{C} \mapsto z^n$ für $n = 0$ auftritt.

Assoziativgesetz und Kommutativgesetz für die Addition, und die zweite Eigenschaft ist eine Verallgemeinerung des Distributivgesetzes.)

Diese beiden Eigenschaften (die Summe wird auf die Summe und das skalare Vielfache auf das skalare Vielfache abgebildet) werden auch als **Linearität der Abbildung** $\sum$ bezeichnet, wobei $\sum$ auf dem Vektorraum $V^{\{m,\ldots,n\}}$ der Funktionen von $\{m,\ldots,n\}$ nach $V$ definiert ist, den Vektorraum $V$ als Zielbereich besitzt, und $(a_m,\ldots,a_n) \in V^{\{m,\ldots,n\}}$ auf $\sum_{k=m}^n a_k$ abbildet. Wie der Name sagt, wird Linearität ausführlicher in der Linearen Algebra besprochen. Es handelt sich dabei aber auch um eine wichtige Eigenschaft für die Analysis, welche also häufig auftreten.

Des Weiteren gilt die Formel für die **Teleskopsumme**

$$\begin{aligned}\sum_{k=m}^n (a_{k+1}-a_k) &= (\cancel{a_{m+1}} - a_m) + (a_{m+2} - \cancel{a_{m+1}}) + \ldots + (\cancel{a_n} - a_{n-1}) + (a_{n+1} - \cancel{a_n}) \\ &= a_{n+1} - a_m\end{aligned}$$

wobei $a_m,\ldots,a_{n+1}$ in einem reellen oder einem komplexen Vektorraum liegen. Formaler argumentiert gilt

$$\begin{aligned}\sum_{k=m}^n (a_{k+1}-a_k) &= \sum_{k=m}^n a_{k+1} - \sum_{k=m}^n a_k = \sum_{j=m+1}^{n+1} a_j - \sum_{k=m}^n a_k \\ &= \left(a_{n+1} + \sum_{j=m+1}^n a_j\right) - \left(a_m + \sum_{k=m+1}^n a_k\right) = a_{n+1} - a_m\end{aligned}$$

wie bereits behauptet. Die Formel für die Teleskopsumme lässt sich zur Abel-Summationsformel verallgemeinern, welche überraschend viele Anwendungen in der Analysis und Zahlentheorie findet.

**Übung 3.3** (Abel-Summation)**.** *Seien $a_1,\ldots,a_n,b_1,\ldots,b_n \in \mathbb{C}$. Wir setzen $A_k = \sum_{j=1}^k a_j$ für $k \in \mathbb{N}_0$ mit $k \leq n$. Zeigen Sie die Abel-Summationsformel*

$$\sum_{k=1}^n a_k b_k = A_n b_n + \sum_{k=1}^{n-1} A_k (b_k - b_{k+1}).$$

*Verwenden Sie dazu die Gleichung $a_k = A_k - A_{k-1}$ für alle $k \in \mathbb{N}$ mit $k \leq n$. Wenden Sie des Weiteren die Abel-Summation auf die Summe $\sum_{k=1}^{2n} \frac{(-1)^k}{k}$ an.*

Anstelle der Dreiecksungleichung werden wir oft auch folgende verallgemeinerte Dreiecksungleichung für Summen verwenden.

**Wichtige Übung 3.4** (Verallgemeinerte Dreiecksungleichung)**.** *Zeigen Sie, dass für alle Zahlen $a_1,\ldots,a_n \in \mathbb{C}$ die Ungleichung*

$$\left|\sum_{i=1}^n a_i\right| \leq \sum_{i=1}^n |a_i|.$$

*gilt.*

Manchmal wollen wir in einer Summe einen gewissen Summanden getrennt betrachten und dazu die Summe aufteilen. Dies kann dann zum Beispiel für $1 \leq k \leq n$ die Form

$$\sum_{j=1}^{n} a_j = \sum_{j=1}^{k} a_j + \sum_{j=k+1}^{n} a_j = \sum_{j=1}^{k-1} a_j + a_k + \sum_{j=k+1}^{n} a_j$$

annehmen. In dem Spezialfall $k = 1$ sollte dies aber mit $a_1 + \sum_{j=2}^{n} a_j$ und in dem Spezialfall $k = n$ mit $\sum_{j=1}^{n-1} a_j + a_n$ übereinstimmen, was auf Grund unserer Definitionen $\sum_{j=1}^{0} a_j = 0$ und $\sum_{j=n+1}^{n} a_j = 0$ in der Tat gilt. (Der formale Beweis erfolgt wiederum mit Induktion nach $n \geq k$.)

### 3.1.2 Rechenregeln für das Produkt

Für ganze Zahlen $m \leq n$ und $a_m, \ldots, a_n, b_m, \ldots, b_n \in \mathbb{C}$ gilt

$$\prod_{k=m}^{n} (a_k b_k) = \Big(\prod_{k=m}^{n} a_k\Big)\Big(\prod_{k=m}^{n} b_k\Big).$$

Insbesondere ist für alle $c \in \mathbb{C}$

$$\prod_{k=m}^{n} (c a_k) = c^{n-m+1}\Big(\prod_{k=m}^{n} a_k\Big).$$

Des Weiteren gilt für alle $a_m, \ldots, a_n$ die Formel für das **Teleskopprodukt**

$$\begin{aligned}\prod_{k=m}^{n} \frac{a_{k+1}}{a_k} &= \Big(\prod_{k=m}^{n} a_{k+1}\Big)\Big(\prod_{k=m}^{n} \frac{1}{a_k}\Big) = \Big(\prod_{k=m+1}^{n+1} a_k\Big)\Big(\prod_{k=m}^{n} \frac{1}{a_k}\Big) \\ &= a_{n+1}\Big(\prod_{k=m+1}^{n} a_k\Big)\Big(\prod_{k=m+1}^{n} \frac{1}{a_k}\Big)\frac{1}{a_m} = \frac{a_{n+1}}{a_m}.\end{aligned}$$

**Lemma 3.5** (Bernoulli'sche Ungleichung)**.** *Für alle reellen Zahlen $a \geq -1$ und $n \in \mathbb{N}_0$ gilt $(1+a)^n \geq 1 + na$.*

*Beweis.* Wir verwenden vollständige Induktion. Für $n = 0$ haben wir $(1+a)^n = 1 = 1 + na$. Angenommen die Ungleichung $(1+a)^n \geq 1 + na$ gilt für ein $n \in \mathbb{N}_0$. Nach Annahme an $a$ ist $a \geq -1$, was in Kombination mit der Annahme an $n$

$$\begin{aligned}(1+a)^{n+1} &= (1+a)^n(1+a) \\ &\geq (1+na)(1+a) = 1 + na + a + na^2 \\ &\geq 1 + (n+1)a\end{aligned}$$

ergibt und damit den Induktionsschritt zeigt. Das Lemma folgt. □

**Übung 3.6** (Archimedisches Prinzip für Potenzen)**.** *Verwenden Sie die Bernoulli'sche Ungleichung und das Archimedische Prinzip (Satz 2.68), um folgende Aussage zu beweisen. Für alle $x, y \in \mathbb{R}$ mit $x > 1$ existiert ein $n \in \mathbb{N}_0$, so dass $x^n \geq y$.*

**Übung 3.7** (Zifferndarstellungen natürlicher Zahlen)**.** *Sei $q \in \mathbb{N}$ eine natürliche Zahl. Zeigen Sie, dass sich jede natürliche Zahl $m$ als Summe der Form $m = \sum_{k=0}^{\ell} a_k q^k$ schreiben lässt wobei $\ell \in \mathbb{N}_0$ und die Koeffizienten $a_0, \ldots, a_\ell \in \mathbb{N}_0 \cap [0, q-1]$. Diese Aussage kennen Sie schon für $q = 10$ wegen der Dezimaldarstellung natürlicher Zahlen und vielleicht auch für $q = 2$ wegen der Binärdarstellung. Für ein allgemeines $q$ spricht man auch von der $q$-nären Darstellung.*

### 3.1.3 Die geometrische Summe

In diesem kurzen Abschnitt möchten wir folgende, vermutlich schon bekannte und für uns später sehr wichtige Formel beweisen.

**Proposition 3.8** (Geometrische Summenformel)**.** *Sei $n \in \mathbb{N}_0$ und $q \in \mathbb{C}$. Dann gilt*

$$\sum_{k=0}^{n} q^k = \begin{cases} n+1 & \text{falls } q = 1 \\ \frac{q^{n+1}-1}{q-1} & \text{falls } q \neq 1 \end{cases}.$$

Der direkte (aber sicher nicht eleganteste) Beweis verwendet vollständige Induktion:

*Beweis.* Für $q = 1$ ist $q^k = 1$ für alle $k \in \mathbb{N}_0$ und die Aussage folgt aus den Eigenschaften der Summe. Sei nun $q \neq 1$. Für $n = 0$ gilt $\sum_{k=0}^{0} q^k = q^0 = 1 = \frac{q-1}{q-1}$, was also den Induktionsanfang zeigt. Angenommen die Formel in der Proposition gilt bereits für $n$. Dann ist

$$\sum_{k=0}^{n+1} q^k = \sum_{k=0}^{n} q^k + q^{n+1} = \frac{q^{n+1}-1}{q-1} + q^{n+1} = \frac{q^{n+1}-1}{q-1} + \frac{q^{n+2}-q^{n+1}}{q-1} = \frac{q^{n+2}-1}{q-1},$$

womit der Induktionsschritt gezeigt ist und die Proposition folgt. □

Wir laden Sie dazu ein, in folgender Übung einen eleganteren Beweis zu finden.

**Übung 3.9** (Geometrische Summenformel)**.** *Verwenden Sie eine Teleskopsumme um die geometrische Summenformel (Proposition 3.8) für $q \neq 1$ zu beweisen.*

**Übung 3.10** (Eindeutigkeit der Ziffernentwickung natürlicher Zahlen)**.** *Zeigen Sie, dass die $q$-näre Darstellung einer natürlichen Zahl in Übung 3.7 eindeutig bestimmt ist. Das heisst, für jedes $m \in \mathbb{N}$ mit $m = \sum_{k=0}^{\ell} a_k q^k$ und $a_\ell \neq 0$ sind $\ell \in \mathbb{N}_0$ und die Koeffizienten $a_0, \ldots, a_\ell \in \mathbb{N}_0 \cap [0, q-1]$ eindeutig durch $m$ bestimmt.*

VO: 18.10.17

## 3.2 Polynome

**Definition 3.11** (Polynomfunktionen)**.** Eine **Polynomfunktion** auf $\mathbb{C}$ ist eine Funktion der Form

$$f : z \in \mathbb{C} \mapsto \sum_{k=0}^{n} a_k z^k \in \mathbb{C}$$

für $n \in \mathbb{N}_0$ und $a_0, \ldots, a_n \in \mathbb{C}$. Die Zahlen $a_0, \ldots, a_n \in \mathbb{C}$ heissen die **Koeffizienten** von $f$. Das grösste $k \in \{0, \ldots, n\}$ mit $a_k \neq 0$ ist der **Grad** $\deg(f)$ der Polynomfunktion $f$ und $a_{\deg(f)}$ ist der **Leitkoeffizient** oder **führende Koeffizient** von $f$. Falls kein solches $k$ existiert, das heisst, falls $f$ die Polynomfunktion $z \in \mathbb{C} \mapsto 0z^0 = 0 \in \mathbb{C}$ ist, so nennt man die Polynomfunktion die **Null** und setzt den Grad auf $-\infty$. Eine Polynomfunktion der Form $z \in \mathbb{C} \mapsto a_0 z^0 \in \mathbb{C}$ für $a_0 \in \mathbb{C}$ wird auch **konstant** genannt und kurz mit $a_0$ bezeichnet. Eine Polynomfunktion mit Grad $\leq 1$ wird **affin** oder **linear** genannt. Eine Polynomfunktion der Form $z \in \mathbb{C} \mapsto a_k z^k \in \mathbb{C}$ für $k \in \mathbb{N}_0$ heisst ein **Monom**. Wir sagen, dass eine Polynomfunktion **reell** ist, wenn die Koeffizienten reell gewählt werden können. Wir werden eine reelle Polynomfunktion auch mit der zugehörigen Funktion von $\mathbb{R}$ nach $\mathbb{R}$ identifizieren.

Wir bemerken, dass die Definition (3.2) für die Definition einer Polynomfunktion gewissermassen notwendig ist, dann falls wir $z^0$ bei $z = 0$ nicht definiert hätten, dann wäre eine Polynomfunktion bei $z = 0$ nicht definiert.

Polynomfunktionen lassen sich auf natürliche Weise addieren und multiplizieren.

**Wichtige Übung 3.12** (Ringstruktur der Polynomfunktionen)**.** *Seien $f, g$ zwei Polynomfunktionen. Wir möchten hier unter anderem zeigen, dass die **Summe** von $f$ und $g$ definiert durch*

$$f + g : z \in \mathbb{C} \mapsto f(z) + g(z) \in \mathbb{C}$$

*und das **Produkt***

$$f \cdot g : z \in \mathbb{C} \mapsto f(z)g(z) \in \mathbb{C}$$

*von $f$ und $g$ wieder Polynomfunktionen sind.*

*(i) Angenommen die Koeffizienten von $f$ sind $a_0, \ldots, a_m \in \mathbb{C}$ und die Koeffizient von $g$ sind $b_0, \ldots, b_n \in \mathbb{C}$. Setzen Sie $a_j = 0$ für alle $j > m$ und $b_j = 0$ für alle $j > n$. Zeigen Sie, dass $f + g$ durch*

$$z \in \mathbb{C} \mapsto (f + g)(z) = \sum_{j=0}^{\max\{m,n\}} (a_j + b_j) z^j$$

*gegeben ist und insbesondere wieder eine Polynomfunktion ist.*

*(ii) Zeigen Sie, dass $f \cdot g$ durch*

$$z \in \mathbb{C} \mapsto f \cdot g(z) = \sum_{k=0}^{n+m} \left( \sum_{j=0}^{k} a_j b_{k-j} \right) z^k$$

*gegeben ist und insbesondere wieder eine Polynomfunktion ist.*

*Zeigen Sie damit auch die Ungleichungen*

$$\deg(f+g) \leq \max\{\deg(f), \deg(g)\}, \quad \deg(f \cdot g) = \deg(f) + \deg(g)$$

*und dass im ersten Fall Gleichheit gilt, falls* $\deg(f) \neq \deg(g)$*. Folgern Sie auch, dass der Leitkoeffizient von* $f \cdot g$ *gerade das Produkt der Leitkoeffizienten von* $f$ *und* $g$ *ist.*

Wir bemerken, dass die Axiome eines Ringes eine Teilmenge der Axiome eines Körper sind und insbesondere für einen Ring nicht gefordert wird, dass ein multiplikatives Inverses existiert. Für eine präzise Definition verweisen wir auf [SS12] und auf [AE06]. Die ganzen Zahlen und auch die Menge der Polynomfunktionen über $\mathbb{R}$ bilden einen Ring. Dies bringt uns zu einer weiteren Definition.

**Definition 3.13** (Polynome)**.** Sei $\mathbb{K}$ ein beliebiger Körper. Ein **Polynom** $f$ über $\mathbb{K}$ ist ein formaler Ausdruck der Form $\sum_{k=0}^{n} a_k T^k$ für $n \in \mathbb{N}_0$ und Koeffizienten $a_0, \ldots, a_n \in \mathbb{K}$. Hierbei ist $T$ ein Symbol, das man auch als **Variable** bezeichnet und das verwendet wird, um die Koeffzienten von einander zu trennen. Wir schreiben auch $T^1 = T$ und $aT^0 = a$ für alle $a \in \mathbb{K}$. Weiter darf ein Summand der Form $0T^k$ für $k \in \mathbb{N}_0$ aus der Summe entfernt werden. Wir definieren den **Polynomring** $\mathbb{K}[T]$ als die Menge der Polynome über $\mathbb{K}$ in der Variablen $T$ mit Addition und Multiplikation gegeben durch die Formeln in Übung 3.12 und verwenden ebenso die Begriffe Grad, Koeffizient, etc. wie in Definition 3.11 für Polynome[2].

Sie fragen sich jetzt vielleicht, und zwar mit Recht, was denn der Unterschied zwischen Definition 3.11 und Definition 3.13 sei (abgesehen davon, dass wir in letzterer allgemeinere Körper zugelassen haben und einen anderen Buchstaben für die Variable verwendet haben). In der Tat, jedem Polynom $f \in \mathbb{C}[T]$ können wir eine Polynomfunktion $z \in \mathbb{C} \mapsto f(z) \in \mathbb{C}$ zuordnen. Diese Zuordnung ist nach Definition des Begriffs Polynomfunktion surjektiv, doch ist nicht klar, ob die Polynomfunktion ihre Koeffizienten eindeutig bestimmt. Insbesondere ist nicht klar, ob es nicht vielleicht zwei Darstellung der gleichen Polynomfunktion mit verschiedenen Koeffizienten gibt und ob nicht vielleicht der Grad der Polynomfunktion von dieser Darstellung abhängt.

**Beispiel 3.14** (Polynome auf endlichen Körpern)**.** *Wir betrachten für den Körper* $\mathbb{F}_2$ *mit zwei Elementen die Polynomfunktionen*

$$f : a \in \mathbb{F}_2 \mapsto a^3 + a + 1 \in \mathbb{F}_2, \quad g : a \in \mathbb{F}_2 \mapsto 1 \in \mathbb{F}_2.$$

*Bei* $0 \in \mathbb{F}_2$ *gilt* $f(0) = 1 = g(0)$ *und an der Stelle* $1 \in \mathbb{F}_2$ *gilt* $f(1) = 1 + 1 + 1 = 1$ *und* $g(1) = 1$*. Insbesondere gilt* $f = g$*, obwohl* $f$ *und* $g$ *nicht durch die gleichen Koeffizienten gegeben sind. Wir unterscheiden die Polynome* $T^3 + T + 1 \in \mathbb{F}_2[T]$ *(mit Grad* $3$*) und* $1 \in \mathbb{F}_2[T]$

[2]Noch formaler ist für jedes $n \in \mathbb{N}_0$ ein Polynom $f(T) = \sum_{k=0}^{n} a_k T^k \in \mathbb{K}[T]$ mit der endlichen Liste von Koeffizienten $(a_0, \ldots, a_n) \in \mathbb{K}^{n+1}$ zu identifizieren, wobei des Weiteren die Identifikation $(a_0, \ldots, a_n) = (a_0, \ldots, a_n, 0)$ verwendet wird – siehe die formale Konstruktion in einer Übung im Abschnitt 3.9. Die Schreibweise als Summe ist hübscher und bei weitem natürlicher für die Definition der Multiplikation in $\mathbb{K}[T]$.

*(mit Grad 0), obwohl die zugehörigen Polynomfunktionen identisch sind (womit es für diese Polynomfunktion keinen wohldefinierten Grad gibt).*

Dieses Beispiel zeigt, dass die oben erwähnte Unterscheidung zwischen einer Polynomfunktion und einem Polynom für gewisse Körper notwendig ist. Wir werden hier zeigen, dass für $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ die Zuordnung zwischen Polynome (welche per Definition eineindeutig einer Liste von Koeffizienten mit wohldefiniertem Grad entsprechen) und der zugehörigen Polynomfunktion bijektiv ist. Wir beweisen dies mittels einer weiteren wichtigen Eigenschaft von Polynomfunktionen.

**Proposition 3.15** (Wachstum von Polynomfunktionen und Eindeutigkeit der Koeffizienten)**.** *Sei $f(T) \in \mathbb{C}[T]$ ein nicht-konstantes Polynom. Dann gibt es zu jeder positiven reellen Zahl $M > 0$ eine reelle Zahl $R \geq 1$, so dass für alle $z \in \mathbb{C}$ mit $|z| \geq R$ auch $|f(z)| \geq M$ gilt. Insbesondere ist die Zuordnung, die jedem Polynom $f(T) \in \mathbb{C}[T]$ die zugehörige Polynomfunktion $z \in \mathbb{C} \mapsto f(z) \in \mathbb{C}$ zuweist, bijektiv. Dies gilt analog ebenso für reelle Polynome $f(T) \in \mathbb{R}[T]$ und reelle Polynomfunktionen $x \in \mathbb{R} \mapsto f(x) \in \mathbb{R}$.*

Intuitiv formuliert besagt die Proposition, dass ein nicht-konstantes Polynom bei „grossen“ $z \in \mathbb{C}$ auch grosse Werte annimmt (gross ist im Absolutbetrag zu verstehen). Auf Grund der zweiten Aussage in obiger Proposition werden wir in der Analysis in Zukunft die Begriffe Polynom und Polynomfunktion nicht mehr unterscheiden und für eine Polynomfunktion $f$ auch $f \in \mathbb{C}[T]$ schreiben.

*Beweis.* Sei $f(T) = a_nT^n + a_{n-1}T^{n-1} + \ldots + a_1T + a_0 \in \mathbb{C}[T]$ mit $a_n \neq 0$ und $n \geq 1$. Wir definieren $q(T) \in \mathbb{C}[T]$ durch $q(T) = a_{n-1}T^{n-1} + \ldots + a_1T + a_0$, womit $f(T) = a_nT^n + q(T)$. Nun behaupten wir, dass die Polynomfunktion $q(z)$ „langsamer wächst als“ $a_nz^n$ und schätzen also $|q(z)|$ für $z \in \mathbb{C}$ mit $|z| \geq 1$ nach oben ab:

$$\begin{aligned}
|q(z)| &= |a_{n-1}z^{n-1} + \ldots + a_1z + a_0| \\
&\leq |a_{n-1}z^{n-1}| + \ldots + |a_1z| + |a_0| \\
&= |a_{n-1}||z^{n-1}| + \ldots + |a_1||z| + |a_0| \\
&\leq (|a_{n-1}| + \ldots + |a_1| + |a_0|)|z|^{n-1} = A|z|^{n-1},
\end{aligned}$$

wobei wir $A = |a_{n-1}| + \ldots + |a_1| + |a_0|$ gesetzt haben und $|z| \geq 1$ in der Form $|z|^k \leq |z|^{n-1}$ für $k \in \{0, \ldots, n-1\}$ verwendet haben. Mit der umgekehrten Dreiecksungleichung (siehe Abschnitt 2.4.2) und $f(z) = a_nz^n + q(z)$ gilt somit

$$\begin{aligned}
|f(z)| \geq |a_nz^n| - |q(z)| &\geq |a_n||z|^n - A|z|^{n-1} \\
&= (|a_n||z| - A)|z|^{n-1} \geq |a_n||z| - A,
\end{aligned}$$

falls $|a_n||z| - A \geq 0$ oder äquivalenterweise $|z| \geq \frac{A}{|a_n|}$. Sei nun $M > 0$ beliebig. Dann wählen wir

$$R = \max\left\{1, \frac{A}{|a_n|}, \frac{A+M}{|a_n|}\right\}.$$

Falls nun $z \in \mathbb{C}$ die Ungleichung $|z| \geq R$ erfüllt, dann gilt $|z| \geq 1$ und $|z| \geq \frac{A}{|a_n|}$, wonach obige Ungleichungen ergeben

$$|f(z)| \geq |a_n||z| - A \geq |a_n|\frac{A+M}{|a_n|} - A = M,$$

was die erste Behauptung der Proposition beweist.

Angenommen $f_1(T), f_2(T) \in \mathbb{C}[T]$ sind zwei Polynome, die $f_1(z) = f_2(z)$ für alle $z$ in $\mathbb{C}$ erfüllen. Dann hat das Polynom $g(T) = f_1(T) - f_2(T)$ die Eigenschaft, dass $g(z) = 0$ für alle $z \in \mathbb{C}$ gilt. Falls der Grad des Polynoms $g(T)$ grösser gleich Eins ist, widerspricht dies dem ersten Teil der Proposition. Also ist $g(T)$ konstant, womit $g(T) = 0$ gelten muss und daher sind die Polynome $f_1(T)$ und $f_2(T)$ identisch (d.h. sie haben denselben Grad und dieselben Koeffizienten). Diesen Beweis kann man ebenso für reelle Polynome und die zugehörigen Polynomfunktionen von $\mathbb{R}$ nach $\mathbb{R}$ durchführen. $\square$

### 3.2.1 Polynomdivision

Wie wir in Satz 2.31 gesehen haben, existiert auf $\mathbb{N}$ eine Division von $n$ durch $d$ mit Rest gegeben durch $n = qd + r$. Dabei ist der Rest $r$ strikt kleiner (bezüglich $\leq$) als $d$. Division mit Rest gibt es auch für Polynome. Hier hat der Rest bei der Divison von $f$ durch $d$ einen kleineren Grad als $d$. Wir illustrieren dies an einem Beispiel und verschieben die allgemeine Aussage auf die nächste Übung.

**Beispiel 3.16.** *Seien $f, d$ die durch $f(x) = 3x^4 - 2x^2 + 5$, $d(x) = x^2 + 1$ für alle $x \in \mathbb{C}$ gegebenen Polynome. Wir behaupten, dass Polynome $q$ und $r$ existieren, so dass $f = q \cdot d + r$ mit $\deg(r) < \deg(d) = 2$. Dazu wählen wir zuerst ein Polynom $q_1$ von der Form $q_1(x) = 3x^2$ für alle $x \in \mathbb{C}$, denn dann ist der Grad von $q \cdot d$ vier und $r_1(x) = f(x) - q_1(x)d(x) = -5x^2 + 5$ hat einen strikt kleineren Grad als $f$. Wir wenden das gleiche Prinzip nochmals auf $r_1$ an und betrachten das Polynom $q_2$ gegeben durch $q_2(x) = -5$ für alle $x \in \mathbb{C}$. Dann gilt $r_1(x) - q_2(x)d(x) = 10$ für alle $x \in \mathbb{C}$. Insbesondere hat das Polynom $r_1 - q_2 \cdot d$ einen strikt kleineren Grad als das Polynom $d$ (nämlich Null); wir setzen somit $r = r_1 - q_2 \cdot d$. Dann gilt*

$$r = r_1 - q_2 \cdot d = f - q_1 \cdot d - q_2 \cdot d = f - (q_1 + q_2) \cdot d.$$

*Wenn wir $q = q_1 + q_2$ setzen, haben wir also $f = q \cdot d + r$ mit $\deg(r) = 0 < 2 = \deg(d)$ wie gewünscht. Im Gymnasium wurde das Vorgehen vielleicht durch folgendes Diagramm dargestellt:*

$$\begin{array}{l}
(3x^4 - 2x^2 + 5) : (x^2 + 1) = 3x^2 - 5 \\
\underline{-3x^4 - 3x^2} \\
\qquad\quad - 5x^2 + 5 \\
\qquad\quad \underline{5x^2 + 5} \\
\qquad\qquad\quad 10
\end{array}$$

Für uns wird folgende Übung wichtig sein, wobei für konkrete Rechnungen später nicht nur die Existenz der Division mit Rest notwendig sein wird, sondern auch eine gewisse Rechenfertigkeit mit Polynomen und der Division mit Rest (wie in Beispiel 3.16) vorausgesetzt sein wird.

**Wichtige Übung 3.17** (Division mit Rest)**.** *Zeigen Sie folgende Version von Division mit Rest: Falls $d$ ein Polynom verschieden von Null ist, dann gibt es für jedes Polynom $f$ zwei eindeutig bestimmte Polynome $q, r$ mit* $\deg(r) < \deg(d)$ *und* $f = q \cdot d + r$.

Für zwei Polynome $f, d \in \mathbb{C}[z]$ mit $d \neq 0$ wird die Funktion $z \mapsto \frac{f(z)}{d(z)}$ als eine **rationale Funktion** bezeichnet, diese hat als natürlichen Definitionsbereich die Menge $\mathbb{C} \setminus \{z \in \mathbb{C} \mid d(z) = 0\}$. Die Division mit Rest hat für rationale Funktion $\frac{f(z)}{d(z)}$ die Bedeutung, dass man letztere in der Form $q(z) + \frac{r(z)}{d(z)}$ schreiben kann, wobei der Grad von $r$ aber kleiner als der Grad von $d$ ist. Dies ist nützlich, da das Polynom $q(z)$ und ebenso die rationale Funktion $\frac{r(z)}{d(z)}$ oft einfacher zu behandeln sind.

### 3.2.2 Nullstellen und Interpolation

Beim Betrachten eines expliziten Polynoms (und auch sonst) interessiert man sich oft für sehr spezifische Punkte, die Nullstellen des Polynoms. Eine **Nullstelle** eines Polynoms $f$ ist eine Zahl $z_1 \in \mathbb{C}$ mit $f(z_1) = 0$. In Abschnitt 2.3 wurde bereits erwähnt, dass nach dem Fundamentalsatz der Algebra jede Gleichung der Form $a_n x^n + \ldots + a_1 x + a_0 = 0$ für $a_0, \ldots, a_n \in \mathbb{C}$ mit $a_n \neq 0$ und $n > 0$ eine Lösung über $\mathbb{C}$ besitzt. Äquivalent dazu ist, dass jedes nicht-konstante Polynom eine Nullstelle hat.

**Übung 3.18** (Ein Spezialfall des Fundamentalsatzes der Algebra)**.** *Verwenden Sie die Wurzelfunktion aus Übung 2.11, um zu zeigen, dass jedes Polynom von Grad* 1 *und jedes reelle Polynom (also mit reellen Koeffizienten) von Grad* 2 *eine Nullstelle in* $\mathbb{C}$ *besitzt. Geben Sie dabei die Nullstellen explizit an.*

An dieser Stelle muss man anmerken, dass der Beweis der allgemeinen Aussage deutlich komplizierter ist. Für einen geschichtlichen Exkurs bezüglich Nullstellen eines Polynoms vom Grad 3 (also kubische Gleichungen) verweisen wir nochmals auf den Podcast der BBC (von der 14. Minute bis zur 21. Minute). Des Weiteren existiert für Polynome von Grad grösser gleich 5 im Allgemeinen keine explizite Formel für die Nullstellen. (Diese letzte Aussage ist Teil der Galois-Theorie, die auch in der Algebra-Vorlesung des 2. Studienjahres des Mathematikstudiums behandelt wird.) In Analogie zu ganzen Zahlen sagen wir, dass ein Polynom $d$ ein Polynom $f$ **teilt** falls es ein Polynom $q$ gibt mit $f = q \cdot d$. In folgender Übung interessieren wir uns für die Anzahl der Nullstellen.

**Wichtige Übung 3.19** (Anzahl Nullstellen eines Polynoms)**.** *Zeigen Sie, für ein beliebiges Polynom $f(T) \in \mathbb{C}[T]$ und eine komplexe Zahl $z \in \mathbb{C}$, dass das Polynom $T - z$ genau dann $f$ teilt, wenn $f$ bei $z$ eine Nullstelle hat. Schliessen Sie daraus, dass $f$ höchstens $n$ verschiedene Nullstellen in $\mathbb{C}$ besitzt, falls $f$ nicht gerade gleich Null ist.*

In Hinblick auf Übung 3.19 sagen wir auch, dass eine Nullstelle $z \in \mathbb{C}$ von $f(T) \in \mathbb{C}[T]$ **Vielfachheit** $k \in \mathbb{N}$ hat, falls $(T-z)^k$ das Polynom $f$ teilt, aber $(T-z)^{k+1}$ das Polynom $f$ nicht teilt.

**Wichtige Übung 3.20** (Koeffizientenvergleich)**.** *Sei $n \in \mathbb{N}_0$ und seien $f, g$ zwei Polynome mit Grad kleiner gleich $n$. Angenommen $f$ und $g$ stimmen auf mehr als $n$ Punkten überein (das heisst, $f(z) = g(z)$ gilt für mehr als $n$ Punkte $z \in \mathbb{C}$). Zeigen Sie, dass dies $f = g$ impliziert und insbesondere, dass die Grade und Koeffizienten von $f$ und $g$ übereinstimmen.*

Falls $n \in \mathbb{N}$ und $z_1, \ldots, z_n \in \mathbb{C}$ verschiedene Punkte und $w_1, \ldots, w_n \in \mathbb{C}$ beliebige Werte sind, dann kann man ein Polynom $f(z) \in \mathbb{C}[z]$ finden mit $f(z_k) = w_k$ für $k \in \{1, \ldots, n\}$. Auf Grund des Koeffizientenvergleich in Übung 3.20 ist dieses auch eindeutig durch diese $2n$ komplexen Zahlen bestimmt. Das Auffinden eines solchen Polynoms wird auch als **Lagrange Polynominterpolation** bezeichnet. Unter Verwendung der Summennotation und einer leicht adaptierten Produktnotation können wir dieses Polynom auch konkret durch

$$f(z) = \sum_{k=1}^{n} w_k \prod_{\substack{j=1 \\ j \neq k}}^{n} (z_k - z_j)^{-1} \prod_{\substack{j=1 \\ j \neq k}}^{n} (z - z_j)$$

angeben (was ohne Verwendung dieser Notation extrem unangenehme Ausdrücke liefern würde). In der Tat können wir für $k = 1, \ldots, n$ die Polynome

$$q_k(z) = \prod_{\substack{j=1 \\ j \neq k}}^{n} (z - z_j)$$
$$p_k(z) = q_k(z_k)^{-1} q_k(z)$$

definieren, wobei wir $q_k(z_k) = \prod_{j=1, j \neq k}^{n} (z_k - z_j) \neq 0$ für die Definition des Polynoms $p_k(z)$ verwendet haben. Setzt man nun die Werte $z = z_\ell$ für $\ell = 1, \ldots, n$ in $p_k$ ein, so sieht man

$$p_k(z_\ell) = \begin{cases} 1 & \text{falls } \ell = k, \\ q_k(z_\ell) = 0 & \text{falls } \ell \neq k \end{cases}$$

und damit

$$f(z_\ell) = \sum_{j=1}^{n} w_k p_k(z_\ell) = w_\ell.$$

Für uns werden diese Formeln nicht besonders wichtig sein, doch zeigen sie sehr deutlich den Vorteil der Summen- und Produktnotation auf. Des Weiteren zeigt obige Diskussion, dass für $n$ vorgegebene und paarweise verschiedene Zahlen $z_1, \ldots, z_n \in \mathbb{C}$ die Polynome $p_1, \ldots, p_n$ eine Basis des Vektorraums $\{f \in \mathbb{C}[z] \mid \deg(f) \leq n-1\}$ bilden, die eben bei diesem Interpolationsproblem der eher üblichen Basis $1, z, \ldots, z^{n-1}$ vorzuziehen ist.

**Applet 3.21** (Polynominterpolation)**.** *Wir stellen in diesem Applet die Polynom-Interpolation grafisch dar, wobei sie bis zu $n = 7$ Punkte $x_1, x_2, \ldots, x_n$ verwenden können, um ein Polynom*

*zu definieren. Nach einigen Experimenten sieht man bereits Nachteile der Polynominterpolation. Wenn man bespielsweise die Werte $x_1 < x_2$ nahe an einander wählt aber $y_1$ und $y_2$ nicht so nahe, so ergeben sich schnell grosse Koeffizienten des Interpolationspolynoms, was bespielsweise zwischen $x_3$ und $x_4$ mit $x_3 < x_4$ unerwartete Auswirkungen für die Funktionswerte des Polynoms hat.*

Wir erwähnen noch, dass viele Aussagen (mit Ausnahme des Fundamentalsatzes der Algebra), die wir zuvor für den Körper $\mathbb{C}$ formuliert haben, auch für reelle Polynome und den Polynomring $\mathbb{R}[x]$ in analoger Weise richtig sind.

**Übung 3.22.** *Verallgemeinern Sie in der richtigen Notation die Übungen 3.12, 3.17, 3.19 und 3.20 für einen beliebigen Körper $\mathbb{K}$.*

*Bemerkung.* Der Grund, wieso man oft auch den hier erklärten, formalen Standpunkt der Unterscheidung von Polynomfunktionen und Polynomen einnimmt, ist zum einen, dass man damit auch endliche Körper gescheit behandeln kann und zum anderen, dass die formale Herangehensweise geeigneter ist für algebraische Konstruktionen. Wir wollen ein wichtiges Beispiel dazu erwähnen.

Die komplexen Zahlen lassen sich als Äquivalenzklassen von $\mathbb{R}[x]$ bezüglich der Äquivalenzrelation

$$f \sim g \iff (x^2+1) \text{ teilt } f-g$$

auffassen. In der Tat bezeichnen wir die Äquivalenzklasse von $x$ einfach durch i und erhalten aus den Definitionen $\mathrm{i}^2 = -1$. In dieser Konstruktion kann man schnell eine Addition und eine Multiplikation auf $\mathbb{C}$ definieren und erhält die Körperstruktur auf $\mathbb{C}$, ohne dass man dabei Kommutativität, Assoziativität, und Distributivität verifizieren müsste, da diese Eigenschaften bereits auf $\mathbb{R}[x]$ gelten. (Wieso gelten diese auf $\mathbb{R}[x]$ und wieso gelten diese dann auch für den Quotientenraum? Können Sie den Beweis dieser Eigenschaften auf die Axiome zurückführen?) Auf diese Art und Weise kann man viele weitere Körper aus $\mathbb{K}[x]$ für einen Körper $\mathbb{K}$ erhalten (siehe wiederum die Algebra-Vorlesungen des zweiten Jahres des Mathematik-Studiums).

### 3.2.3 Algebraische und transzendente Zahlen

Eine Zahl $\alpha \in \mathbb{C}$ heisst **algebraisch**, falls es ein von Null verschiedenes Polynom $f \in \mathbb{Q}[x]$ gibt mit $f(\alpha) = 0$. Beispielsweise sind i und $\sqrt{2}$ algebraisch, denn $x^2+1$ hat i als Nullstelle und $x^2-2$ hat $\sqrt{2}$ als Nullstelle. Des Weiteren ist jede rationale Zahl algebraisch. Die Menge $\overline{\mathbb{Q}}$ der algebraischen Zahlen wird auch der **algebraische Abschluss** von $\mathbb{Q}$ genannt und ist (wie wir hier nicht zeigen wollen) ein Unterkörper von $\mathbb{C}$.

Nicht-algebraische Zahlen nennt man **transzendent**. Interessanterweise sind die meisten Zahlen transzendent, wie die nächste Übung zeigt. Beispiele von transzendenten Zahlen werden wir allerdings erst später angeben können.

**Übung 3.23.** *Zeigen Sie, dass der Polynomring $\mathbb{Q}[x]$ abzählbar unendlich ist. Schliessen Sie, dass der algebraische Abschluss $\overline{\mathbb{Q}}$ von $\mathbb{Q}$ abzählbar unendlich ist.*

VO: 19.10.17

## 3.3 Die Fakultät und der Binomialsatz

### 3.3.1 Fakultät

**Definition 3.24** (Fakultät)**.** Die Funktion $n \in \mathbb{N}_0 \mapsto n! \in \mathbb{N}$ ist definiert durch

$$0! = 1, \; n! = \prod_{k=1}^{n} k.$$

Die Zahl $n!$ wird als $n$-**Fakultät** oder $n$-**Faktorielle** bezeichnet.

Insbesondere (nämlich per Definition des Produkts) gilt also die rekursive Formel

$$(n+1)! = (n!) \cdot (n+1)$$

für alle $n \in \mathbb{N}_0$. Wir werden dieser Funktion in vielen weiteren Funktionen und Ausdrücken begegnen. Sie hat jedoch auch für sich gesehen eine (kombinatorische) Bedeutung.

**Lemma 3.25** (Kardinalität der Menge der Permutationen einer endlichen Menge)**.** *Für $n \in \mathbb{N}$ ist $n!$ die Kardinalität der Menge $\mathcal{S}_n$ der bijektiven Abbildungen $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ (auch **Permutationen** von $\{1, \ldots, n\}$ genannt).*

Intuitiv ausgedrückt gibt es also genau $n!$ verschiedene Möglichkeiten die Menge $\{1, \ldots, n\}$ zu sortieren oder auch $n!$ Möglichkeiten für die verschiedenen Reihenfolgen, wenn alle $n$ nummerierte Bälle zufällig aus einer Urne gezogen werden.

*Bemerkung.* Zu $n \in \mathbb{N}$ bildet die Menge $\mathcal{S}_n$ zusammen mit der Verknüpfung von Elementen $(\sigma, \tau) \in \mathcal{S}_n^2 \mapsto \sigma \circ \tau \in \mathcal{S}_n$ eine Gruppe (die **symmetrische Gruppe**).

*Beweis von Lemma 3.25.* Wir beweisen die Aussage per Induktion. Für $n = 1$ gibt es genau eine (bijektive) Abbildung $\{1\} \to \{1\}$, was den Induktionsanfang darstellt.

Angenommen die Aussage des Lemmas gilt bereits für ein $n \in \mathbb{N}$. Wir betrachten nun eine Permutation $\sigma$ von $\{1, \ldots, n+1\}$. Falls $\sigma(n+1) = n+1$ gilt, so erhalten wir mittels Einschränkung auf $\{1, \ldots, n\}$ eine bijektive Abbildung $\sigma' : k \in \{1, \ldots, n\} \to \{1, \ldots, n\}$ in $\mathcal{S}_n$. Umgekehrt können wir für jedes $\sigma' \in \mathcal{S}_n$ eine Fortsetzung $\sigma \in \mathcal{S}_{n+1}$ mit $\sigma(n+1) = n+1$ definieren. Daher wissen wir also per Induktionsannahme, dass es $n!$ Abbildungen $\sigma \in \mathcal{S}_{n+1}$ mit $\sigma(n+1) = n+1$ gibt. Wir bezeichnen die Menge aller solcher Permutation in $\mathcal{S}_{n+1}$ mit

$$H = \{\sigma \in \mathcal{S}_{n+1} \mid \sigma(n+1) = n+1\},$$

so dass $|H| = n!$.

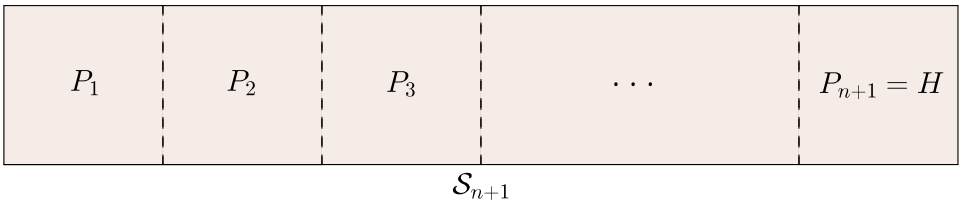
Die Menge $\mathcal{S}_{n+1}$ lässt sich wie folgt partitionieren:

$$\mathcal{S}_{n+1} = \bigsqcup_{k=1}^{n+1} P_k \text{ mit } P_k = \{\tau \in \mathcal{S}_{n+1} \mid \tau(n+1) = k\}$$

für $k = 1, \ldots, n+1$. Wir behaupten nun, dass die Mengen $P_k$ auf der rechten Seite alle Kardinalität $n!$ haben (für $k = n+1$ ist dies bereits bekannt, da $P_{n+1} = H$). Dies impliziert

$$|\mathcal{S}_{n+1}| = (n!) \cdot (n+1) = (n+1)!$$

und damit nach vollständiger Induktion das Lemma.




Figur 3.1: Wir haben $\mathcal{S}_{n+1}$ in $n+1$ Teile zerlegt und beweisen, dass jedes Element dieser Partition genau $n! = |H|$ Elemente enthält. Dazu konstruieren wir für jedes $k \in \{0, \ldots, n\}$ eine Bijektion $P_k \to P_{n+1} = H$.

Sei $k \in \{1, \ldots, n\}$ fix und sei $\delta_k : \{1, \ldots, n+1\} \to \{1, \ldots, n+1\}$ die bijektive Abbildung, die $k$ und $n+1$ vertauscht und sonst wirkungslos ist, das heisst, die Abbildung

$$\delta_k : \{1, \ldots, n+1\} \to \{1, \ldots, n+1\}\,, \ \ell \mapsto \begin{cases} k & \text{falls } \ell = n+1 \\ n+1 & \text{falls } \ell = k \\ \ell & \text{falls } \ell \notin \{k, n+1\}\,. \end{cases}$$

Falls nun $\tau \in \mathcal{S}_{n+1}$ die Eigenschaft $\tau(n+1) = k$ hat, dann bildet $\sigma = \delta_k \circ \tau \in \mathcal{S}_{n+1}$ das Element $n+1$ auf $n+1$ ab. Die Abbildung

$$\Phi : \tau \in P_k \mapsto \delta_k \circ \tau \in H = \{\sigma \in \mathcal{S}_{n+1} \mid \tau(n+1) = n+1\}$$

ist somit wohldefiniert. Sie ist auch bijektiv, da die Abbildung

$$\sigma \in H = P_{n+1} \mapsto \delta_k \circ \sigma = \tau \in P_k$$

eine Inverse darstellt. (Wieso?) Dies beweist die obige Behauptung, was den Beweis des Induktionsschritts abschliesst. □

### 3.3.2 Binomialkoeffizienten

Für $n, k \in \mathbb{N}_0$ mit $0 \leq k \leq n$ definieren wir den **Binomialkoeffizienten** $\binom{n}{k}$, als „$n$ über $k$“ ausgesprochen, durch

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}.$$

Ersetzen wir $k$ bei gleichbleibendem $n$ im Binomialkoeffizienten durch $n-k$, so vertauschen sich bloss die beiden Ausdrücke im Nenner und wir erhalten

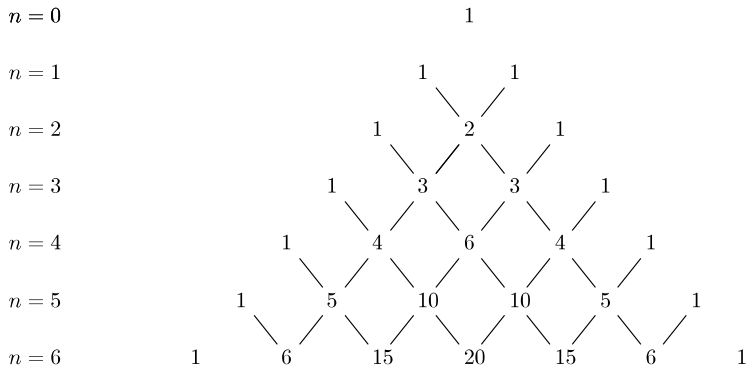
$$\binom{n}{k} = \binom{n}{n-k}$$

für alle $k, n \in \mathbb{N}_0$ mit $0 \leq k \leq n$.

**Proposition 3.26** (Additionseigenschaft der Binomialkoeffizienten)**.** *Für alle $n, k \in \mathbb{N}_0$ mit $0 \leq k \leq n-1$ gilt $\binom{n}{0} = \binom{n}{n} = 1$ und*

$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}. \tag{3.3}$$

*Insbesondere ist $\binom{n}{k} \in \mathbb{N}$ für alle $n, k \in \mathbb{N}_0$ mit $0 \leq k \leq n$.*




Figur 3.2: Die Binomialkoeffizienten lassen sich auch bildlich im sogenannten **Pascal Dreieck** festhalten, wobei in der $n$-ten Zeile die Zahlen $\binom{n}{0}, \binom{n}{1}, \binom{n}{2}, \ldots, \binom{n}{n}$ stehen und die diagonalen Striche die Additionsformel (3.3) in Proposition 3.26 andeuten.

*Beweis.* Wir verwenden die Definition der Binomialkoeffizienten und erhalten

$$\binom{n}{0} = \binom{n}{n} = \frac{n!}{0!\ n!} = 1$$

sowie

$$\begin{aligned}
\binom{n}{k} + \binom{n}{k+1} &= \frac{n!}{k!(n-k)!} + \frac{n!}{(k+1)!\ (n-k-1)!} \\
&= \frac{(k+1)n!}{(k+1)!\ (n-k)!} + \frac{(n-k)n!}{(k+1)!\ (n-k)!} \\
&= \frac{(k+1+n-k)n!}{(k+1)!\ ((n+1)-(k+1))!} = \binom{n+1}{k+1}
\end{aligned}$$

durch Erweiterung mit $k+1$ beziehungsweise $n-k$.

Die letzte Aussage ergibt sich aus den ersten beiden Aussagen und Induktion nach $n$. In der Tat entsteht die jeweils nächste Zeile im Pascal Dreieck (siehe Bild 3.2), indem man an den Rändern jeweils eine 1 und dazwischen die Summen aus der vorherigen Zeile niederschreibt. □

**Wichtige Übung 3.27** (Kombinatorische Bedeutung der Binomialkoeffizienten)**.** *Beweisen Sie, dass die Zahl* $\binom{n}{k}$ *für* $n, k \in \mathbb{N}_0$ *mit* $0 \leq k \leq n$ *die möglichen Resultate bei der Auswahl einer* $k$*-elementigen Teilmenge von* $\{1, \ldots, n\}$ *darstellt. Formaler ausgedrückt: Zeigen Sie, dass es genau* $\binom{n}{k}$ *Teilmengen von* $\{1, \ldots, n\}$ *gibt, die* $k$ *Elemente besitzen. Berechnen Sie* $\binom{42}{6}$ *(beziehungsweise* $\binom{49}{6}$ *falls Sie aus Deutschland stammen oder* $\binom{45}{6}$ *falls Sie aus Österreich stammen).*

Obige kombinatorische Bedeutung liefert sowohl eine Interpretation als auch einen kombinatorischen Beweis von (3.3): Angenommen wir haben $n+1$ Bälle mit den Zahlen 1 bis $n+1$ beschriftet, doch hat der letzte Ball $n+1$ die Farbe rot und alle anderen die Farbe blau. Sei nun $0 \leq k \leq n-1$. Für jede der $\binom{n+1}{k+1}$ Möglichkeiten $k+1$ Bälle aus den $n+1$ Bällen auszuwählen fällt nun zuerst auf, ob der rote Ball ausgewählt wurde oder nicht. In dem ersten Fall haben wir den roten Ball gemeinsam mit einer der $\binom{n}{k}$ Auswahlmöglickeiten für $k$ Bälle aus den $n$ blauen Bällen. Im zweiten Fall sehen wir eine der $\binom{n}{k+1}$ Auswahlmöglichkeiten für $k+1$ Bälle aus den $n$ blauen Bällen. Dies definiert also eine Bijektion zwischen den $\binom{n+1}{k+1}$ Auswahlmöglichkeiten für $k+1$ aus $n+1$ und der (disjunkten) Vereinigung der $\binom{n}{k}$ Auswahlmöglichkeiten für $k$ aus $n$ und den $\binom{n}{k+1}$ Auswahlmöglichkeiten für $k+1$ aus $n$, wodurch (3.3) nochmals bewiesen wird.

### 3.3.3 Der binomische Lehrsatz

**Satz 3.28** (Binomischer Lehrsatz)**.** *Für* $w, z \in \mathbb{C}$ *und* $n \in \mathbb{N}_0$ *gilt*

$$(w+z)^n = \sum_{k=0}^{n} \binom{n}{k} w^{n-k} z^k.$$

Die Geschichte des binomischen Lehrsatz ist zum einen ziemlich verzweigt und zum anderen ziemlich lang. Die erste Formulierung eines Spezialfalls des binomischen Lehrsatzes wird bereits Euklid zugeschrieben. Der binomische Lehrsatz und die Binomialkoeffizienten waren den Hindus im ersten Jahrtausend vermutlich bekannt, die erste, relativ exakte Formulierung geht wahrscheinlich auf den persischen Mathematiker Al-Karaji zurück, der den Lehrsatz mit einer groben Variante von vollständiger Induktion bewies. Wir verweisen unter anderem auf [Coo49] für mehr Details.

*Beweis des binomischen Lehrsatzes.* Wir verwenden vollständige Induktion über $n$. Für $n = 0$ gilt die Aussage, da

$$(w+z)^0 = 1 = \sum_{k=0}^{0} 1 w^{0-k} z^k.$$

Angenommen die Aussage des Satzes gilt für ein $n \in \mathbb{N}_0$. Dann erhalten wir

$$\begin{aligned}
(w+z)^{n+1} = (w+z)^n(w+z) &= \left(\sum_{k=0}^{n} \binom{n}{k} w^{n-k} z^k\right)(w+z) \\
&= \sum_{k=0}^{n} \binom{n}{k} w^{n+1-k} z^k + \sum_{k=0}^{n} \binom{n}{k} w^{n-k} z^{k+1} \\
&= w^{n+1} + \sum_{j=1}^{n} \binom{n}{j} w^{n+1-j} z^j + \sum_{k=0}^{n-1} \binom{n}{k} w^{n-k} z^{k+1} + z^{n+1} \\
&= w^{n+1} + \sum_{k=0}^{n-1} \binom{n}{k+1} w^{n-k} z^{k+1} + \sum_{k=0}^{n-1} \binom{n}{k} w^{n-k} z^{k+1} + z^{n+1} \\
&= w^{n+1} + \sum_{k=0}^{n-1} \binom{n+1}{k+1} w^{(n+1)-(k+1)} z^{k+1} + z^{n+1} \\
&= w^{n+1} + \sum_{\ell=1}^{n} \binom{n+1}{\ell} w^{n+1-\ell} z^{\ell} + z^{n+1} \\
&= \sum_{\ell=0}^{n+1} \binom{n+1}{\ell} w^{n+1-\ell} z^{\ell}
\end{aligned}$$

unter Verwendung von (zwei) Indexverschiebungen und der Additionsformel (3.3). □

**Übung 3.29** (Summe der Binomialkoeffizienten)**.** *Zeigen Sie für jedes $n \in \mathbb{N}_0$ die Identitäten*

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n, \quad \sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0.$$

**Übung 3.30** (Eine weitere Summe)**.** *Verwenden Sie den binomischen Lehrsatz, um die Identität*

$$\sum_{m=k}^{n} \binom{m}{k} \binom{n}{m} = \binom{n}{k} 2^{n-k}$$

*für alle $k \leq n$ zu beweisen. Dies verallgemeinert die erste Aussage aus Übung 3.29, wieso? Können Sie einen kombinatorischen Beweis für diese Aussage finden?*

### 3.3.4 Eine Summe von Binomialkoeffizienten*

**Für unseren Aufbau war Proposition 3.31 zwar nötig um $\int_a^b x^n \, \mathrm{d}x = [\frac{1}{n+1} x^{n+1}]_a^b$ zu beweisen, doch lässt sich dies nach Besprechung des Hauptsatzes aus der Ableitungsregel für $x^n$ viel schneller beweisen.*

In diesem Teilabschnitt wollen wir mit Hilfe der Binomialkoeffizienten Lemma 1.3 für beliebige Potenzen verallgemeinern.

**Proposition 3.31** (Summe von Potenzen)**.** *Für jedes $d \in \mathbb{N}_0$ gibt es rationale Konstanten $c_0, \ldots, c_d \in \mathbb{Q}$, so dass*

$$\sum_{k=1}^{n} k^d = \frac{1}{d+1} n^{d+1} + c_d n^d + \ldots + c_1 n + c_0$$

*für alle $n \in \mathbb{N}$ gilt.*

Anders ausgedrückt ist die Summe der $d$-ten Potenzen $1^d + \cdots + n^d$ gleich den Funktionswerten eines Polynoms mit Grad $d+1$ und Leitkoeffizient $\frac{1}{d+1}$ ausgewertet bei $x = n$. Wir werden dies im Kapitel 4 verwenden, um Flächeninhalte unter Graphen von Polynomen zu berechnen.

Per Definition erzeugen die Funktionen $1, x, x^2, \ldots$ den Vektorraum $\mathbb{R}[x]$ der Polynome und bilden in der Tat sogar eine Basis von $\mathbb{R}[x]$. (Wieso?) Für das Problem in obiger Proposition ist diese Basis jedoch nicht besonders gut geeignet. Stattdessen ist die Basis $p_0(x), p_1(x), \ldots \in \mathbb{C}[x]$ definiert durch

$$\begin{aligned} p_0(x) &= 1, \\ p_1(x) &= x, \\ p_2(x) &= \frac{x(x-1)}{2} \end{aligned}$$

und allgemeiner durch

$$p_d(x) = \frac{1}{d!} \prod_{j=0}^{d-1} (x - j) \tag{3.4}$$

für alle $d \in \mathbb{N}_0$ deutlich besser dafür geeignet. Wir haben diesen Zugang gewählt, da wir im Zuge des Beweises eine kombinatorische Bedeutung solcher Summen kennenlernen werden. Es wäre auch möglich Proposition 3.31 schneller zu beweisen, indem man die Teleskopsumme für die Summanden $a_n = (n+1)^d - n^d$ verwendet.

**Übung 3.32** (Eine besser geeignete Basis)**.** *Zeigen Sie, dass die in Gleichung* (3.4) *definierten Polynome $p_0, p_1, p_2, \ldots$ tatsächlich eine Basis von $\mathbb{R}[x]$ bilden.*

**Proposition 3.33** (Summe von Binomialkoeffizienten)**.** *Für jedes $d \in \mathbb{N}$ ist $p_d(x)$ ein Polynom mit rationalen Koeffizienten vom Grad $d$, welches Leitkoeffizient $\frac{1}{d!}$ und Nullstellen $0, \ldots, d-1$ besitzt. Für jedes $d \in \mathbb{N}_0$ gilt des Weiteren $p_d(n) = \binom{n}{d}$ für alle $n \geq d$ (siehe auch das Bild 3.3) und wir haben die Summenformel*

$$\sum_{k=0}^{n} p_d(k) = p_{d+1}(n+1)$$

*für alle $n \in \mathbb{N}$.*

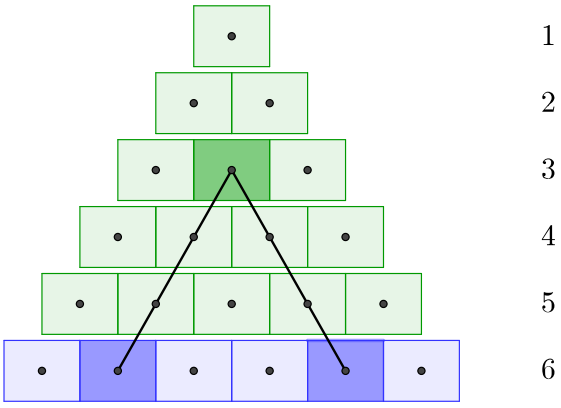
Die ersten beiden Aussagen ergeben sich direkt aus (3.4). Für $n, d \in \mathbb{N}_0$ mit $n \geq d$ gilt des Weiteren

$$\begin{aligned} p_d(n) &= \frac{\prod_{j=0}^{d-1}(n-j)}{d!} \cdot 1 \\ &= \frac{\prod_{j=0}^{d-1}(n-j)\prod_{j=d}^{n-1}(n-j)}{d!\ \prod_{k=1}^{n-d} k} = \frac{n!}{d!\ (n-d)!} = \binom{n}{d}. \end{aligned}$$

Im Spezialfall $d = 1$ können wir die Summenformel in Proposition 3.33 geometrisch interpretieren. Zuerst bemerken wir, dass $p_1(x) = x$ und somit gilt

$$\sum_{k=0}^{n} k = \frac{n(n+1)}{2} = p_2(n+1) = \binom{n+1}{2}.$$

Nun betrachten wir folgendes Bild für $n = 5$.




Je zwei Quadrate in der sechsten Zeile bestimmen also eindeutig ein Quadrat in den oberen (hier fünf) Zeilen und umgekehrt. Daher gilt $1+2+3+4+5 = \binom{6}{2}$, da die linke Seite die Anzahl Quadrate in den oberen 5 Zeilen ist und die rechte Seite die Anzahl Paare von Quadraten in der sechsten Zeile ist. Allgemeiner zeigt man analog, dass $1 + 2 + \ldots + n = \binom{n+1}{2}$.

Im Allgemeinen wollen wir einen kombinatorischen Beweis der Summenformel in Proposition 3.33 geben, wobei wir hierfür die Aussage in Übung 3.27 verwenden werden, welche den Binomialkoeffizienten eine kombinatorische Bedeutung gibt.

*Beweis von Proposition 3.33.* Seien $d, n \in \mathbb{N}_0$. Falls $n \in \{0, \ldots, d-1\}$ liegt, gilt $p_d(k) = 0$ für alle $k \in \{0, \ldots, n\}$ und damit auch

$$p_{d+1}(n+1) = 0 = \sum_{k=0}^{n} p_d(k).$$

Wir dürfen also $d \leq n$ annehmen. Sei $\mathcal{P}$ die Menge aller $(d+1)$-elementigen Teilmengen von $\{1, \ldots, n+1\}$. Da $A \in \mathcal{P}$ genau $d+1$ Elemente und nur natürliche Zahlen enthält, gilt

$\max(A) \geq d+1$. Wir verwenden dies um

$$\mathcal{P} = \bigsqcup_{\ell=d+1}^{n+1} \mathcal{P}_\ell$$

in die Teilmengen $\mathcal{P}_\ell = \{A \in \mathcal{P} \mid \max(A) = \ell\}$ für $\ell \in \{d+1, \ldots, n+1\}$ zu partitionieren, so dass

$$\binom{n+1}{d+1} = |\mathcal{P}| = \sum_{\ell=d+1}^{n+1} |\mathcal{P}_\ell|.$$

Für $\ell \in \{d+1, \ldots, n+1\}$ und eine Teilmenge $A \in \mathcal{P}_\ell$ gilt nach Definition $\ell \in A$ und $A \subseteq \{1, \ldots, \ell\}$. Entfernen wir von diesem $A$ das Element $\ell$, so erhalten wir eine $d$-elementige Teilmenge $A \setminus \{\ell\}$ von $\{1, \ldots, \ell-1\}$, welche eindeutig von $A$ bestimmt ist und gemeinsam mit $\ell$ die Menge $A$ eindeutig bestimmt. Zusammenfassend bildet die Abbildung $A \mapsto A \setminus \{\ell\}$ also eine Bijektion zwischen $\mathcal{P}_\ell$ und den $d$-elementigen Teilmengen von $\{1, \ldots, \ell-1\}$. Nach Übung 3.27 ergibt sich somit $|\mathcal{P}_\ell| = \binom{\ell-1}{d} = p_d(\ell-1)$ für alle $\ell = d+1, \ldots, n+1$ und daher

$$p_{d+1}(n+1) = \binom{n+1}{d+1} = \sum_{\ell=d+1}^{n+1} \binom{\ell-1}{d} = \sum_{k=d}^{n} \binom{k}{d} = \sum_{k=0}^{n} p_d(k),$$

da $p_d(0) = \cdots = p_d(d-1) = 0$. □

**Applet 3.34** (Kombinatorischer Beweis)**.** *Es wird für kleine Wahlen von $n$ und $d$ die in obigem kombinatorischem Beweis gefundene Bijektion dargestellt. Falls Sie nicht genau sehen, was die Animation darstellt, dann sollten Sie zuerst den Beweis lesen.*

Wenn wir auf die sehr befriedigende kombinatorische Interpretation der Summenformel in Proposition 3.33 verzichten wollen, so können wir die Summenformel auch sehr einfach im Pascal-Dreieck sehen und mit Induktion beweisen.

$d = 3$

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | | | | | |
| 1 | 1 | 0 | 0 | 0 | | | | |
| 1 | 2 | 1 | 0 | 0 | 0 | | | |
| 1 | 3 | 3 | 1 | 0 | 0 | | | |
| 1 | 4 | 6 | 4 | 1 | 0 | 0 | | |
| 1 | 5 | 10 | 10 | 5 | 1 | 0 | | |
| 1 | 6 | 15 | 20 | 15 | 6 | 1 | 0 | |
| 1 | 7 | 21 | 35 | 35 | 21 | 7 | 1 | |

Figur 3.3: Illustration der Summenformel in Proposition 3.33 für $d = 3$. In Rosa ersichtlich sind $p_3(k)$ für $k \in \{0, 1, 2, 3, 4, 5, 6\}$, die summiert gerade $p_4(7) = \binom{7}{4} = 35$ ergeben, wie in Grün markiert.

**Übung 3.35** (Summe der Binomialkoeffizienten mittels Induktion)**.** *Beweisen Sie die letzte Aussage in Proposition 3.33 mittels vollständiger Induktion.*

*Beweis von Proposition 3.31.* Wir beweisen mittels vollständiger Induktion nach $d \in \mathbb{N}_0$, dass es rationale Zahlen $c_d, \dots, c_0 \in \mathbb{Q}$ gibt, für die

$$\sum_{k=1}^{n} k^d = \frac{1}{d+1} n^{d+1} + c_d n^d + \cdots + c_0 n^0 \tag{3.5}$$

für alle natürlichen Zahlen $n \in \mathbb{N}$ gilt. (Sämtliche Konstanten in (3.5) hängen von $d$ aber nicht von $n$ ab.) Sei also $d \in \mathbb{N}_0$, so dass (3.5) bereits für alle $d' \in \mathbb{N}_0$ mit $d' < d$ bewiesen wurde (wir verwenden also die Induktionsvariante in Satz 2.19 für $\mathbb{N}_0$), und sei $n \in \mathbb{N}$. Per Definition von $p_{d+1}(T)$ (siehe (3.4)) ist $p_{d+1}(T)$ ein Polynom mit rationalen Koeffizienten und Leitkoeffizient $\frac{1}{(d+1)!}$ und es gilt

$$p_{d+1}(n+1) = \frac{1}{(d+1)!}(n+1)n(n-1)\dots(n-d)$$

für alle $n \in \mathbb{N}$. Wir multiplizieren dies mit $d!$ und erhalten nach Ausmultiplizieren sowie nach Proposition 3.33, dass

$$d! \sum_{k=0}^{n} p_d(k) = \frac{1}{d+1} n^{d+1} + c'_d n^d + \cdots + c'_0 n^0$$

für gewisse $c'_d, \dots, c'_0 \in \mathbb{Q}$ und alle $n \in \mathbb{N}$ gilt. Ebenso können wir aber $p_d$ ausmultiplizieren und erhalten analog

$$d!\ p_d(k) = k^d + c''_{d-1} k^{d-1} + \cdots + c''_0 k^0$$

für gewisse Konstanten $c''_{d-1}, \dots, c''_0 \in \mathbb{Q}$ und alle $k \in \mathbb{N}$. Dies ergibt

$$\sum_{k=0}^{n} k^d = \frac{1}{d+1} n^{d+1} + c'_d n^d + \cdots + c'_0 n^0 - \sum_{j=0}^{d-1} c''_j \sum_{k=0}^{n} k^j$$

Nach Induktionsvoraussetzung ist aber $\sum_{k=0}^{n} k^j$ gleich $f_j(n)$ für ein Polynom $f_j(T)$ mit Grad $j + 1 \leq d$ und mit rationalen Koeffizienten. Daher ist auch

$$c'_d n^d + \cdots + c'_0 n^0 - \sum_{j=0}^{d-1} c''_j \sum_{k=0}^{n} k^j$$

gleich $g(n)$ für ein Polynom $g(T)$ mit Grad kleiner gleich $d$ mit rationalen Koeffizienten. (Wieso?) Dies beweist den Induktionsschritt und damit die Proposition. □

VO: 23.10.17

## 3.4 Reellwertige Funktionen

Für eine beliebige, nicht-leere Menge $D$ definieren wir die Menge der **$\mathbb{R}$-wertigen** oder **reellwertigen** Funktionen auf $D$ als

$$\mathcal{F}(D) = \mathbb{R}^D = \{f \mid f : D \to \mathbb{R}\}.$$

Die Menge $\mathcal{F}(D)$ bildet einen Vektorraum über $\mathbb{R}$, wobei Addition und skalare Multiplikation punktweise gegeben sind durch

$$\begin{aligned}(f_1 + f_2)(x) &= f_1(x) + f_2(x),\\ (af_1)(x) &= \alpha f_1(x)\end{aligned}$$

für $f_1, f_2 \in \mathcal{F}(D)$, $a \in \mathbb{R}$ und alle $x \in D$. Funktionen in $\mathcal{F}(D)$ lassen sich sogar multiplizieren und zwar durch

$$(f_1 f_2)(x) = f_1(x) f_2(x)$$

für alle $f_1, f_2 \in \mathcal{F}(D)$ und $x \in D$. (Die Menge $\mathcal{F}(D)$ bildet einen kommutativen Ring mit Eins.) Wir sagen, dass $x \in D$ eine **Nullstelle** von $f \in \mathcal{F}(D)$ ist, falls $f(x) = 0$ gilt. Die **Nullstellenmenge** von $f$ ist durch $\{x \in D \mid f(x) = 0\}$ definiert.

**Übung 3.36** (Nullstellenmenge eines Produkts)**.** *Seien $N_1, N_2 \subseteq D$ die Menge der Nullstellen von $f_1 \in \mathcal{F}(D)$ beziehungsweise $f_2 \in \mathcal{F}(D)$. Was ist die Nullstellenmenge von $f_1 f_2$?*

Wir definieren auch eine $\leq$-Relation (tatsächlich eine Ordnung) auf $\mathcal{F}(D)$ durch

$$f_1 \leq f_2 \iff \forall x \in D : f_1(x) \leq f_2(x)$$

für $f_1, f_2 \in \mathcal{F}(D)$. Wir sagen, dass $f \in \mathcal{F}(D)$ **nicht-negativ** ist, falls $f \geq 0$ gilt

**Übung 3.37** (Ordnung auf $\mathcal{F}(D)$)**.** *Zeigen Sie, dass die oben definierte Relation $\leq$ eine Ordnung ist und dass diese Ordnung genau dann linear ist, wenn $D$ aus genau einem Punkt besteht.*

### 3.4.1 Beschränktheit

In diesem und im nächsten Abschnitt möchten wir jeweils einen wichtigen Begriff zu reellwertigen Funktionen einführen.

**Definition 3.38** (Beschränktheit von Funktionen)**.** Sei $D$ eine nicht-leere Menge und sei $f : D \to \mathbb{R}$ eine Funktion. Wir sagen, dass die Funktion $f$

- **von oben beschränkt** ist, falls die Wertemenge $f(D)$ von oben beschränkt ist,
- **von unten beschränkt** ist, falls die Wertemenge $f(D)$ von unten beschränkt ist,
- **beschränkt** ist, falls $f$ von oben und von unten beschränkt ist.

Wir bemerken, dass eine Teilmenge $A \subseteq \mathbb{R}$ genau dann beschränkt ist, wenn es ein $M \in \mathbb{R}$ gibt, so dass $|y| \leq M$ für alle $y \in A$. Daher ist eine Funktion $f \in \mathcal{F}(D)$ genau dann beschränkt, wenn es ein $M \in \mathbb{R}$ gibt, so dass für alle $x \in D$ gilt $|f(x)| \leq M$.

**Übung 3.39** (Der Unterraum der beschränkten Funktionen)**.** *Sei $D$ eine nicht-leere Menge und sei $\mathcal{B}(D) \subseteq \mathcal{F}(D)$ die Menge der beschränkten Funktionen von $D$ nach $\mathbb{R}$. Zeigen Sie, dass $\mathcal{B}(D)$ ein Unterraum von $\mathcal{F}(D)$ bildet und dass zu $f_1, f_2 \in \mathcal{B}(D)$ auch $f_1 \cdot f_2 \in \mathcal{B}$ liegt.*

### 3.4.2 Monotonie

In Folgendem wollen wir annehmen, dass die Menge $D$ (der Definitionsbereich der Funktionen, die wir betrachten wollen) eine nicht-leere Teilmenge von $\mathbb{R}$ ist.

**Definition 3.40** (Monotonieeigenschaften)**.** Eine Funktion $f : D \to \mathbb{R}$ ist

- **monoton wachsend**, falls $\forall x, y \in D : x \leq y \implies f(x) \leq f(y)$,
- **streng monoton wachsend**, falls $\forall x, y \in D : x < y \implies f(x) < f(y)$,
- **monoton fallend**, falls $\forall x, y \in D : x \leq y \implies f(x) \geq f(y)$,
- **streng monoton fallend**, falls $\forall x, y \in D : x < y \implies f(x) > f(y)$,
- **monoton**, falls $f$ monoton wachsend oder monoton fallend ist,
- **streng monoton**, falls $f$ streng monoton wachsend oder streng monoton fallend ist.

Eine streng monotone Funktion ist per Definition auch monoton; die Bezeichnung „streng“ ist also passend. Wir betrachten nun ein paar elementare Beispiele.

**Beispiel 3.41.**

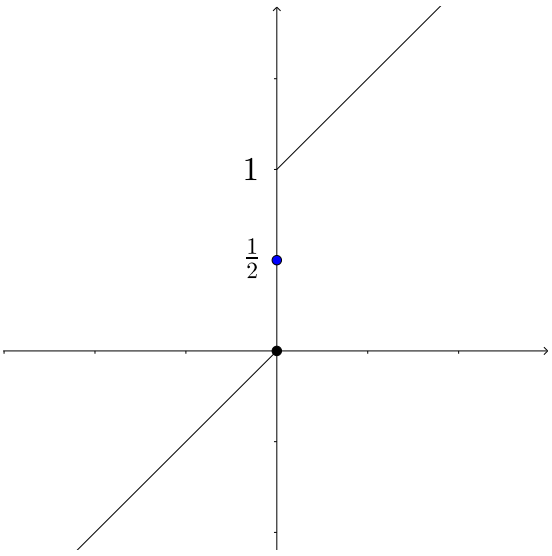
- *Die Funktion $x \in \mathbb{R}_{\geq 0} \mapsto x^2 \in \mathbb{R}$ ist streng monoton wachsend. Die Funktion $x \in \mathbb{R}_{\leq 0} \mapsto x^2 \in \mathbb{R}$ ist jedoch streng monoton fallend und $x \in \mathbb{R} \mapsto x^2 \in \mathbb{R}$ hat keine Monotonieeigenschaft. Dies zeigt, dass Monotonie vom Definitionsbereich abhängt.*
- *Die Funktion $x \in \mathbb{R} \mapsto x^3 \in \mathbb{R}$ ist streng monoton wachsend.*
- *Allgemeiner gilt: Für ein $n \in \mathbb{N}$ ist die Funktion $x \in \mathbb{R} \mapsto x^n \in \mathbb{R}$ ist genau dann (streng) monoton, wenn $n$ ungerade ist.*
- *Die Abrundungsfunktion $x \in \mathbb{R} \mapsto \lfloor x \rfloor \in \mathbb{R}$ ist monoton wachsend, aber nicht streng monoton wachsend (wieso?).*
- *Eine konstante Funktion $x \in \mathbb{R} \mapsto c \in \mathbb{R}$ für ein (festes) $c \in \mathbb{R}$ ist sowohl monoton fallend als auch monoton wachsend.*

**Übung 3.42.** *Beweisen Sie die Behauptungen in Beispiel 3.41.*

Eine streng monotone Funktion ist stets injektiv. Sie braucht jedoch nicht surjektiv zu sein. Beispielsweise ist die Funktion

$$\mathbb{R} \to \mathbb{R},\ x \mapsto \begin{cases} x+1 & \text{falls } x \geq 0 \\ x & \text{falls } x < 0 \end{cases}$$

streng monoton wachsend, aber nicht surjektiv ($\frac{1}{2}$ liegt nicht im Bild).




Verlangt man jedoch von einer streng monotonen Funktion, dass sie „nicht springt“, so kann man in vielen Fällen Surjektivität zeigen. Wir haben bereits ein Beispiel dafür gesehen als wir die Existenz der Wurzelfunktion zeigten (siehe Übung 2.11). Den dazu notwendigen Begriff des „Nicht-Springens“ besprechen wir im nächsten Abschnitt.

**Übung 3.43** (Alternative Charakterisierung von Monotonie und strenger Monotonie)**.** *Sei $D \subseteq \mathbb{R}$ eine Teilmenge mit $|D| \geq 3$ und $f : D \to \mathbb{R}$ eine Funktion. Betrachten Sie die folgenden Aussagen über $f$, beschreiben Sie ihre Bedeutung in Worten, und geben Sie einen Beweis von drei Ihrer Behauptungen.*

*(i)* $\forall x, y \in D : x < y \iff f(x) < f(y)$

*(ii)* $\forall x, y \in D : x \leq y \iff f(x) \leq f(y)$

*(iii)* $\forall x, y, z \in D : x < y < z \implies \big(f(x) < f(y) < f(z) \vee f(x) > f(y) > f(z)\big)$

*(iv)* $\forall x, y, z \in D : x < y < z \implies \big(f(x) \leq f(y) \leq f(z) \vee f(x) \geq f(y) \geq f(z)\big)$

*(v)* $\forall x, y, z \in D : x \leq y \leq z \implies \big(f(x) < f(y) < f(z) \vee f(x) > f(y) > f(z)\big)$

*(vi)* $\forall x, y, z \in D : x \leq y \leq z \implies \big(f(x) \leq f(y) \leq f(z) \vee f(x) \geq f(y) \geq f(z)\big)$

**Übung 3.44** (Monotonie unter Summen und Produkten)**.** *Sei $D \subseteq \mathbb{R}$ eine Teilmenge und seien $f, f_1, f_2 \in \mathcal{F}(D)$ streng monoton wachsend. Zeigen Sie, dass $f_1 + f_2 \in \mathcal{F}(D)$ streng monoton wachsend ist und dass für $a \in \mathbb{R}$ die Funktion $af \in \mathcal{F}(D)$ streng monoton wachsend*

*ist, falls* $a > 0$ *und streng monoton fallend ist, falls* $a < 0$*. Zeigen Sie, dass* $f_1 f_2$ *streng monoton wachsend ist, falls* $f_1(x), f_2(x) > 0$ *für alle* $x \in D$*.*

**Applet 3.45** (Monotonie von Einschränkungen)**.** *Bei vielen aber nicht allen Funktion* $f$ *(definiert auf Teilmengen von* $\mathbb{R}$*) erhält man eine monotone Funktion mittels Einschränkung von* $f$ *auf kleinere Intervalle.*

## 3.5 Stetigkeit

Sei $D \subseteq \mathbb{R}$ eine nicht-leere Teilmenge.

**Definition 3.46** (Stetigkeit)**.** Sei $f : D \to \mathbb{R}$ eine Funktion. Wir sagen, dass $f$ **stetig bei einem Punkt** $x_0 \in D$ ist, falls es für alle $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $x \in D$ die Implikation

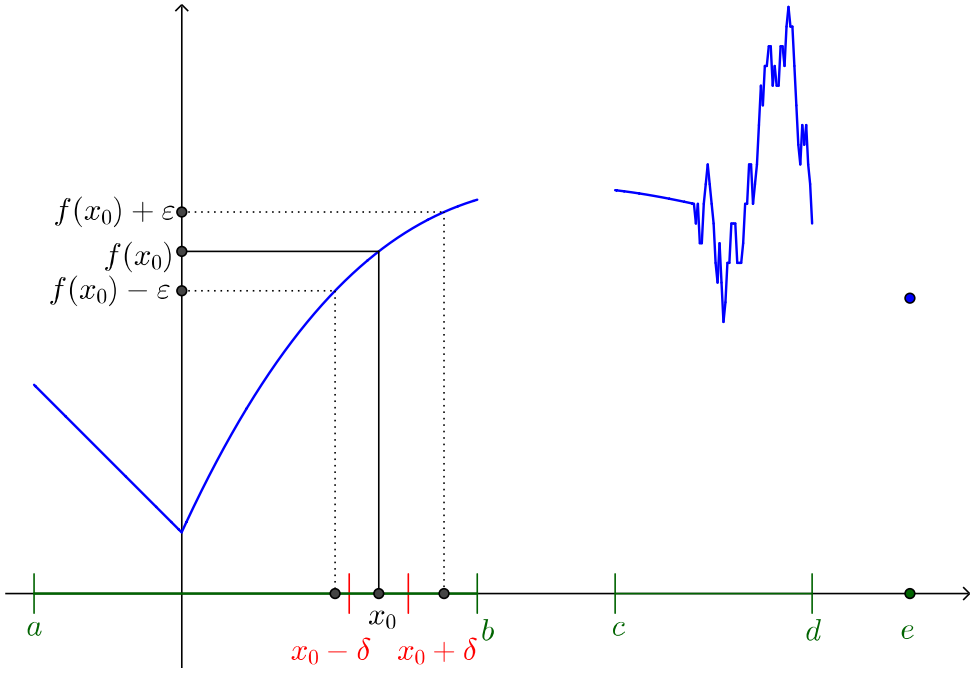
$$|x - x_0| < \delta \implies |f(x) - f(x_0)| < \varepsilon$$

gilt. Die Funktion $f$ ist **stetig**, falls sie bei jedem Punkt in $D$ stetig ist. Formal ist Stetigkeit von $f$ also durch

$$\forall x_0 \in D : \forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x \in D : |x - x_0| < \delta \implies |f(x) - f(x_0)| < \varepsilon$$

definiert.

Diese Definition ist vielleicht auf den ersten Blick überraschend, wird für uns aber sehr fundamental sein.[3] Wir wollen im Folgenden wieder einige Funktionen und deren Graphen untersuchen, damit wir ein Gespür für die Definition der Stetigkeit erhalten können.




Figur 3.4: Bei dieser Funktion $f$ mit einem „kontinuierlichen“ Graphen sehen wir, dass $f$ bei $x_0$ stetig ist. Egal wie klein man $\varepsilon > 0$ wählt, kann man sich gut vorstellen, dass für ein geeignetes $\delta > 0$ für alle $x$, die $\delta$-nahe bei $x_0$ liegen, auch $f(x)$ $\varepsilon$-nahe an $f(x_0)$ ist. Die Funktion ist sogar auf dem ganzen Definitionsbereich $[a, b] \cup [c, d] \cup \{e\}$ stetig.

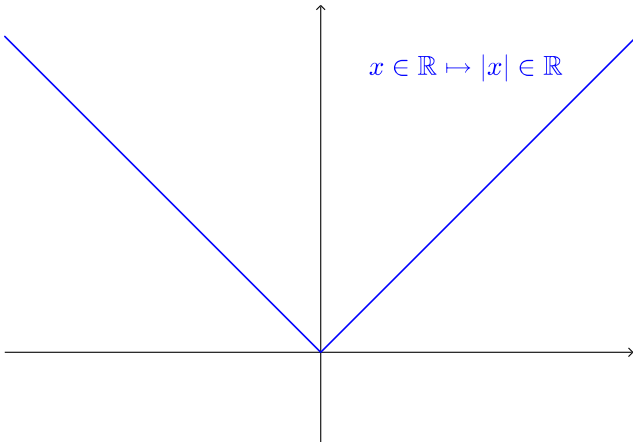
**Applet 3.47** (Stetigkeit)**.** *Wir betrachten eine Funktion, die an den meisten (aber nicht allen) Punkten des Definitionsbereichs stetig ist.*

[3] Das Wort „stetig“ und daraus abgeleitete Wörter kommen in der vollständigen Version von diesem Skript über 850 mal vor, was die Wichtigkeit des Begriffs gewissermassen quantifiziert.

In den meisten Situationen wird $D$ für uns ein Intervall $I$ sein. Wir untersuchen nun ein paar konkrete Beispiele.

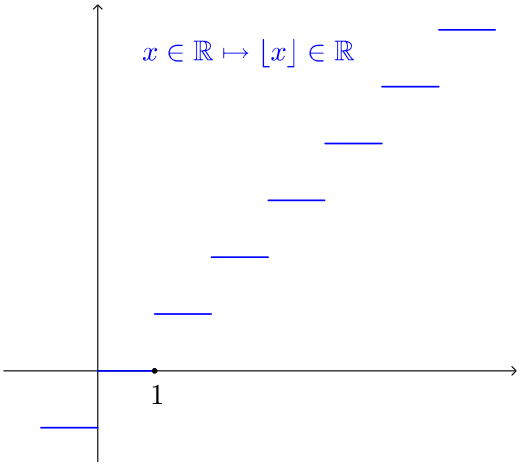
**Beispiel 3.48.**

- *Die konstante Funktion $x \in \mathbb{R} \mapsto c \in \mathbb{R}$ für ein $c \in \mathbb{R}$ ist stetig. (Wieso?).*
- *Die lineare Funktion $x \in \mathbb{R} \mapsto ax \in \mathbb{R}$ für ein $a \in \mathbb{R}$ ist stetig.* Ist $a = 0$, so ist die Funktion konstant und somit stetig. Sei also $a \in \mathbb{R}^\times$. Ist $x_0 \in \mathbb{R}$ und $\varepsilon > 0$, dann gilt $|ax - ax_0| = |a||x - x_0|$ für alle $x \in \mathbb{R}$. Betrachtet man also die Wahl $\delta = \frac{\varepsilon}{|a|}$ und ein $x \in \mathbb{R}$ mit $|x - x_0| < \delta$, so ist $|ax - ax_0| = |a||x - x_0| < |a|\delta = \varepsilon$.
- *Die Funktion $x \in \mathbb{R} \mapsto |x| \in \mathbb{R}$ ist stetig.*




Ist $x_0 \in \mathbb{R}$ und $\varepsilon > 0$ beliebig, so gilt wegen der umgekehrten Dreiecksungleichung $\big||x| - |x_0|\big| \leq |x - x_0|$ für alle $x \in \mathbb{R}$ . Damit kann man mit der Wahl $\delta = \varepsilon$ erreichen, dass wenn $|x - x_0| < \delta$ auch $\big||x| - |x_0|\big| \leq |x - x_0| < \delta = \varepsilon$ für alle $x \in \mathbb{R}$ gilt, was die Stetigkeit zeigt.

- *Die Funktion $x \in \mathbb{R} \mapsto \lfloor x \rfloor \in \mathbb{R}$ ist bei Punkten in $\mathbb{Z}$ nicht stetig.*




Ist $n \in \mathbb{Z}$, dann gilt für jedes $\delta > 0$, dass $|(n - \frac{1}{2}\delta) - n| < \delta$ und

$$\big|\lfloor n - \tfrac{1}{2}\delta \rfloor - \lfloor n \rfloor\big| = \lfloor n \rfloor - \lfloor n - \tfrac{1}{2}\delta \rfloor \geq n - (n - 1) = 1.$$

Damit ist die Stetigkeitsbedingung bei $n \in \mathbb{Z}$ für $\varepsilon < 1$ nicht erfüllt.

Wir möchten an dieser Stelle kurz anmerken, dass es im Beweis der Stetigkeit einer Funktion reicht, $\varepsilon > 0$ „klein genug“ zu betrachten. Genauer können wir annehmen, dass $\varepsilon \leq \varepsilon_0$ für ein $\varepsilon_0 > 0$. (Wieso?)

Man könnte an dieser Stelle natürlich auch die Stetigkeit weiterer Funktionen beweisen, wie man in folgender Übung verifizieren kann.

**Übung 3.49** (Quadrat und Quadratwurzel)**.** *Zeigen Sie, dass die Funktionen $x \in \mathbb{R} \mapsto x^2 \in \mathbb{R}$ und $x \in \mathbb{R}_{>0} \mapsto \sqrt{x} \in \mathbb{R}_{>0}$ stetig sind.*

Da wir jedoch nicht jedes Mal „von Hand“ überprüfen möchten, ob eine Funktion stetig ist, wenden wir uns nun allgemeineren Aussagen zu. Eine erste solche wollen wir Ihnen als Übung überlassen.

**Wichtige Übung 3.50** (Einschränkung von stetigen Funktionen)**.** *Sei $D \subseteq \mathbb{R}$ ein Teilmenge und $f : D \to \mathbb{R}$ stetig. Sei $D'$ ein Teilmenge von $D$. Zeigen Sie die Stetigkeit von $f|_{D'}$.*

**Proposition 3.51** (Stetigkeit unter Addition und Multiplikation von Funktionen)**.** *Sei $D \subseteq \mathbb{R}$. Falls $f_1, f_2 : D \to \mathbb{R}$ Funktionen sind, die bei einem Punkt $x_0 \in D$ stetig sind, dann sind auch $f_1 + f_2$, $f_1 \cdot f_2$ und $af_1$ für $a \in \mathbb{R}$ stetig bei $x_0$. Insbesondere bildet die Menge der stetigen Funktionen*

$$C(D) = \{f \in \mathcal{F}(D) \mid f \text{ ist stetig}\}$$

*einen Unterraum des Vektorraums $\mathcal{F}(D)$.*

*Beweis.* Angenommen $f_1, f_2 \in \mathcal{F}(D)$ sind bei $x_0 \in D$ stetig und sei $\varepsilon > 0$. Dann existieren $\delta_1, \delta_2 > 0$, so dass für alle $x \in D$ gilt

$$\begin{aligned} |x - x_0| < \delta_1 &\implies |f_1(x) - f_1(x_0)| < \frac{\varepsilon}{2} \\ |x - x_0| < \delta_2 &\implies |f_2(x) - f_2(x_0)| < \frac{\varepsilon}{2}. \end{aligned}$$

Wir setzen $\delta = \min\{\delta_1, \delta_2\} > 0$ und erhalten

$$\begin{aligned} |x - x_0| < \delta \implies |(f_1 + f_2)(x) - (f_1 + f_2)(x_0)| \\ &\leq |f_1(x) - f_1(x_0)| + |f_2(x) - f_2(x_0)| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

Da $\varepsilon > 0$ beliebig war, erhalten wir, dass $f_1 + f_2$ bei $x_0 \in D$ stetig ist.

Das Argument für $f_1 f_2$ ist ähnlich, aber etwas komplizierter. Wir beginnen mit der Abschätzung

$$\begin{aligned} |f_1(x)f_2(x) - f_1(x_0)f_2(x_0)| &= |f_1(x)f_2(x) - f_1(x_0)f_2(x) + f_1(x_0)f_2(x) - f_1(x_0)f_2(x_0)| \\ &\leq |f_1(x)f_2(x) - f_1(x_0)f_2(x)| + |f_1(x_0)f_2(x) - f_1(x_0)f_2(x_0)| \\ &= |f_1(x) - f_1(x_0)||f_2(x)| + |f_1(x_0)||f_2(x) - f_2(x_0)| \end{aligned}$$

für $x \in D$ unter Verwendung der Dreiecksungleichung. Sei $\varepsilon > 0$ und wähle $\delta_1 > 0$ und $\delta_2 > 0$, so dass für $x \in D$

$$\begin{aligned} |x - x_0| < \delta_1 &\implies |f_1(x) - f_1(x_0)| < \frac{\varepsilon}{2(|f_2(x_0)| + 1)} \\ |x - x_0| < \delta_2 &\implies |f_2(x) - f_2(x_0)| < \min\left\{1, \frac{\varepsilon}{2(|f_1(x_0)| + 1)}\right\} \end{aligned}$$

erfüllt sind. Dann gilt für ein $x \in D$ mit $|x - x_0| < \delta = \min\{\delta_1, \delta_2\}$, dass

$$|f_2(x)| = |f_2(x) - f_2(x_0) + f_2(x_0)| \leq |f_2(x) - f_2(x_0)| + |f_2(x_0)| < 1 + |f_2(x_0)|$$

und damit

$$|f_1(x) - f_1(x_0)||f_2(x)| < \frac{\varepsilon}{2(|f_2(x_0)| + 1)}(1 + |f_2(x_0)|) = \frac{\varepsilon}{2}.$$

Für das zweite Argument gilt ebenso

$$|f_1(x_0)||f_2(x) - f_2(x_0)| \leq |f_1(x_0)|\frac{\varepsilon}{2(|f_1(x_0)| + 1)} < \frac{\varepsilon}{2}.$$

Gemeinsam erhalten wir $|f_1(x)f_2(x) - f_1(x_0)f_2(x_0)| < \varepsilon$ wie gewünscht. Die Aussage über $af_1$ für $a \in \mathbb{R}$ folgt mit Obigem und der Tatsache, dass die konstante Funktion $x \in \mathbb{R} \mapsto a \in \mathbb{R}$ stetig ist. Insbesondere ist $C(D)$ auch nicht leer, da die konstante Nullfunktion in $C(D)$ liegt, und somit ist $C(D)$ ein Unterraum von $\mathcal{F}(D)$. □

Wir präsentieren eine direkte Anwendung dieser Proposition.

**Korollar 3.52.** *Polynome sind stetig, das heisst,* $\mathbb{R}[x] \subseteq C(\mathbb{R})$.

*Beweis.* Wie wir in Beispiel 3.48 gesehen haben, ist die Funktion $p(x) = x$ stetig und konstante Funktionen sind stetig. Vollständige Induktion und Proposition 3.51 zeigen, dass jedes Polynom $\sum_{k=0}^{n} a_k x^k \in \mathbb{R}[x]$ stetig ist. □

Eine weitere Art und Weise, wie man zeigen kann, dass eine Funktion stetig ist, ist, dass man sie als Verknüpfung von stetigen Funktionen darstellt.

**Proposition 3.53** (Stetigkeit unter Verknüpfung)**.** *Seien* $D_1, D_2 \subseteq \mathbb{R}$ *zwei Teilmengen und sei* $x_0 \in D_1$. *Angenommen* $f : D_1 \to D_2$ *ist eine bei* $x_0$ *stetige Funktion und* $g : D_2 \to \mathbb{R}$ *ist eine bei* $f(x_0)$ *stetige Funktion. Dann ist* $g \circ f : D_1 \to \mathbb{R}$ *bei* $x_0$ *stetig. Insbesondere ist die Verknüpfung von stetigen Funktionen wieder stetig.*

Es folgt damit, dass zum Beispiel die Abbildung $x \in D \mapsto |f(x)| \in \mathbb{R}$ für eine beliebige stetige Funktion $f : D \to \mathbb{R}$ stetig ist (nach Beispiel 3.48).

*Beweis.* Sei $\varepsilon > 0$. Dann existiert wegen der Stetigkeit von $g$ bei $f(x_0)$ ein $\eta > 0$, so dass für alle $y \in D_2$

$$|y - f(x_0)| < \eta \implies |g(y) - g(f(x_0))| < \varepsilon.$$

Da $\eta > 0$ ist und $f$ bei $x_0$ stetig ist, gibt es aber auch ein $\delta > 0$, so dass für alle $x \in D_1$

$$|x - x_0| < \delta \implies |f(x) - f(x_0)| < \eta.$$

Zusammen ergibt sich (für $y = f(x) \in f(D_1) \subseteq D_2$), dass für alle $x \in D_1$

$$|x - x_0| < \delta \implies |f(x) - f(x_0)| < \eta \implies |g(f(x)) - g(f(x_0))| < \varepsilon.$$

gilt. Dies beweist auch die letzte Aussage, da $x_0$ ein beliebiger Punkt in $D_1$ war. □

**Übung 3.54** (Stetigkeit von Quotienten)**.** *Zeigen Sie, dass die Funktion $x \in \mathbb{R}^\times \mapsto \frac{1}{x} \in \mathbb{R}^\times$ stetig ist. Schliessen Sie, dass Funktionen der Art $x \in D \mapsto \frac{f(x)}{g(x)} \in \mathbb{R}$ stetig sind, wenn $D \subseteq \mathbb{R}$ eine Teilmenge und $f : D \to \mathbb{R},\ g : D \to \mathbb{R}^\times$ stetige Funktionen sind. (Beachte, dass $g$ auf $D$ keine Nullstelle haben darf.)*

Eine andere Art und Weise, wie man stetige Funktionen konstruieren kann, ist durch „zusammenkleben“, wie wir in folgender Übung diskutieren wollen.

**Wichtige Übung 3.55** (Stetige Funktionen durch Fallunterscheidung)**.** *Seien zwei Intervalle durch $I_1 = [a, b],\ I_2 = [b, c] \subseteq \mathbb{R}$ gegeben mit $a < b < c$ und seien $f_1 : I_1 \to \mathbb{R}$ und $f_2 : I_2 \to \mathbb{R}$ stetige Funktionen. Zeigen Sie, dass die Funktion*

$$f : [a, c] \to \mathbb{R},\ x \mapsto \begin{cases} f_1(x) & \text{falls } x \in [a, b) \\ f_2(x) & \text{falls } x \in [b, c] \end{cases}$$

*genau dann stetig ist, wenn $f_1(b) = f_2(b)$ gilt.*

Der Vollständigkeit halber erwähnen wir noch folgende Charakterisierung von Stetigkeit, auf die wir später wieder zu sprechen kommen werden und die beispielsweise eine grundlegende Definition der Topologie-Vorlesung im vierten Semester des Mathematikstudums darstellt.

**Übung 3.56** (Stetigkeit über offene Mengen)**.** *Sei $I \subseteq \mathbb{R}$ ein offenes Intervall. Zeigen Sie, dass $f$ genau dann stetig ist, wenn für jede offene Menge $U \subseteq \mathbb{R}$ auch $f^{-1}(U)$ offen ist.*

Ist eine Funktion $f$ stetig, so ist das Verhalten von $f$ in kleinen Umgebungen eines Punktes oft vorhersehbar.

**Übung 3.57** (Lokale Eigenschaften von stetigen Funktionen)**.** *Sei $D \subseteq \mathbb{R}$ eine Teilmenge und $f : D \to \mathbb{R}$ eine Funktion. Zeigen Sie die folgenden Aussagen.*

*a) (Lokal beschränkt) Wenn $f$ bei $x_0 \in D$ stetig ist, dann gibt es eine $\delta$-Umgebung $U$ von $x_0$ und ein $M > 0$, so dass $|f(x)| \leq M$ für alle $x \in D \cap U$.*

*b) (Lokal das gleiche Vorzeichen) Wenn $f$ bei $x_0 \in D$ stetig ist und $f(x_0) \neq 0$ ist, dann gibt es eine $\varepsilon$-Umgebung $U$ von $x_0$, so dass $f(x)f(x_0) > 0$ für alle $x \in D \cap U$ (das heisst, $f(x)$ und $f(x_0)$ haben dasselbe Vorzeichen).*

**Übung 3.58** (Challenge – „Fast überall“ Stetigkeit von monotonen Funktionen)**.** *In dieser Übung möchten wir zeigen, dass es zu einer monotonen Funktion $f$ auf einem Intervall $[a, b]$ mit $a < b$ höchstens abzählbar viele Punkte geben kann, bei denen $f$ nicht stetig ist (sogenannte Unstetigkeitsstellen). Gehen Sie dazu wie folgt vor: Sei $A \subseteq [a, b]$ die Menge der Unstetigkeitsstellen von $f$.*

*(i) Sei $x \in A$. Wir setzen*

$$f_-(x) = \sup\left\{f(x') \mid x' \in [a, b],\ x' < x\right\},\ f_+(x) = \inf\left\{f(x') \mid x' \in [a, b],\ x' > x\right\}.$$

*Zeigen Sie, dass $f_-(x) < f_+(x)$. Wählen Sie anschliessend eine rationale Zahl $g(x)$ in $(f_-(x), f_+(x))$. (Wir nehmen hier an, dass wir beliebig oft eine Wahl treffen können.)*

*(ii) Zeigen Sie, dass $g : x \in A \mapsto g(x) \in \mathbb{Q}$ injektiv ist und schliessen Sie auf die Aussage.*

### 3.5.1 Komplex-wertige Funktionen

Für eine Menge $D$ können wir in Analogie zum reellen Vektorraum $\mathcal{F}(D)$ auch den komplexen Vektorraum

$$\mathcal{F}_{\mathbb{C}}(D) = \{f \mid f : D \to \mathbb{C}\}$$

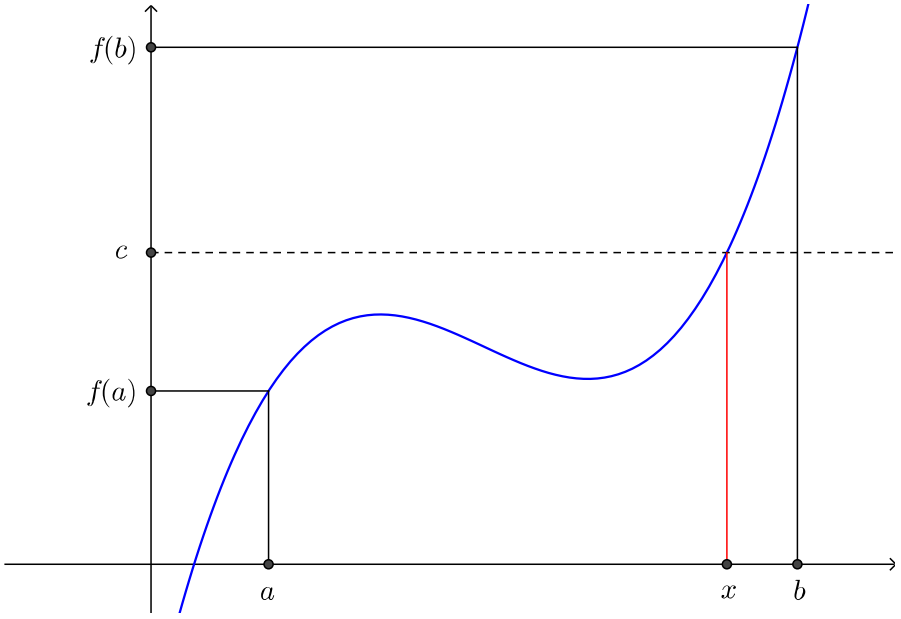
der **$\mathbb{C}$-wertigen Funktionen** oder **komplex-wertigen** Funktionen Funktionen auf $D$ definieren. Weiter sagen wir, dass eine Funktion $f : D \to \mathbb{C}$ **beschränkt** ist, falls die reellwertige Funktion $x \in D \mapsto |f(x)|$ beschränkt ist. Ein Punkt $x \in D$ ist eine **Nullstelle** einer Funktion $f : D \to \mathbb{C}$, falls $f(x) = 0$ gilt.

Dieser Abschnitt lässt sich ohne grosse Änderung auf $\mathbb{C}$-wertige Funktionen auf einer Teilmenge $D \subseteq \mathbb{C}$ übertragen. Wir empfehlen Ihnen diesen Abschnitt nochmals zu lesen, aber diesmal komplex-wertige Funktionen auf Teilmengen $D \subseteq \mathbb{C}$ zu erlauben, wobei $\varepsilon > 0$ und $\delta > 0$ weiterhin reelle Zahlen darstellen. Wir werden später nochmals in grösserer Allgemeinheit auf den Begriff der Stetigkeit zu sprechen kommen. Wir bemerken noch, dass die Monotonieeigenschaften in Abschnitt 3.4.2 hingegen kein Analog für komplex-wertige Funktionen besitzen (da auf $\mathbb{C}$ keine natürliche Ordnung gegeben ist).

## 3.6 Der Zwischenwertsatz

In diesem Abschnitt wollen wir einen fundamentalen Satz beweisen, der die Heuristik, dass der Graph einer stetigen Funktion auf einem Intervall „eine durchgehende Kurve“ darstellt, formalisiert. Wir sagen, dass eine reelle Zahl $c$ **zwischen zwei reellen Zahlen** $x_1, x_2$ **liegt**, falls $x_1 \leq c \leq x_2$ oder $x_2 \leq c \leq x_1$ gilt. Wir sagen $c$ liegt **echt zwischen** $x_1$ **und** $x_2$ falls $x_1 < c < x_2$ oder $x_2 < c < x_1$ ist.

**Satz 3.59** (Zwischenwertsatz)**.** *Sei $I \subseteq \mathbb{R}$ ein Intervall, $f : I \to \mathbb{R}$ eine stetige Funktion und $a, b \in I$. Für jedes $c \in \mathbb{R}$ zwischen $f(a)$ und $f(b)$ gibt es ein $x \in \mathbb{R}$ zwischen $a$ und $b$, so dass $f(x) = c$ gilt.*




Figur 3.5: Der Graph einer stetigen Funktion kann keine Sprünge machen und die Funktion nimmt alle Werte zwischen $f(a)$ und $f(b)$ an.

Wie wir sehen werden, verwendet der Beweis die Existenz des Supremums (und damit das Vollständigkeitsaxiom).

*Beweis.* Wir nehmen ohne Beschränkung der Allgemeinheit an, dass $a < b$ und $f(a) \leq f(b)$ gilt (falls $f(a) > f(b)$ ist, betrachtet man zuerst $-f$ und bemerkt, dass die Aussage des Satzes für $-f$ die Aussage des Satzes für $f$ impliziert).

Sei nun $c \in [f(a), f(b)]$. Falls $c = f(a)$ oder $c = f(b)$ gilt, sind wir fertig. Also angenommen $c \in (f(a), f(b))$. Wir definieren

$$X = \{x \in [a, b] \mid f(x) \leq c\}$$

und bemerken, dass $a \in X$ und $X \subseteq [a, b]$, wodurch $X$ nicht-leer und von oben beschränkt ist. Nach Satz 2.59 existiert daher $x_0 = \sup(X) \in [a, b]$. Wir werden nun die Stetigkeit von $f$ bei $x_0$ verwenden, um zu zeigen, dass $f(x_0) = c$.

Für jedes $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für alle $x \in [a, b]$ gilt

$$|x - x_0| < \delta \implies |f(x) - f(x_0)| < \varepsilon. \tag{3.6}$$

Angenommen $f(x_0) < c$. Dann folgt $x_0 < b$ wegen $f(b) > c$ und $x_0 \in [a, b]$. Wir wenden nun die Stetigkeit von $f$ bei $x_0$ an und finden für $\varepsilon = c - f(x_0) > 0$ ein $\delta > 0$ wie in (3.6). Da $x_0 < b$ ist, existiert ein $x \in (x_0, x_0 + \delta) \cap [a, b]$. Für dieses $x$ gilt dann

$$f(x) = f(x_0) + (f(x) - f(x_0)) < f(x_0) + c - f(x_0) = c.$$

Also muss $x$ in $X$ liegen, was aber $\sup(X) = x_0 < x$ widerspricht.

Angenommen $f(x_0) > c$. Dann folgt $x_0 > a$ wegen $f(a) < c$. Wir verwenden wieder die Stetigkeit von $f$ bei $x_0$ und finden zu $\varepsilon = f(x_0) - c$ ein $\delta > 0$ mit der Eigenschaft in (3.6). Für $x \in (x_0 - \delta, x_0) \cap [a, b]$ gilt dadurch

$$f(x) = f(x_0) + (f(x) - f(x_0)) > f(x_0) - (f(x_0) - c) = c,$$

wodurch $x \notin X$ und daher $(x_0 - \delta, x_0) \cap [a, b] \cap X = \emptyset$. Also ist $x_0 - \delta$ eine obere Schranke von $X$, was aber $x_0 = \sup(X)$ widerspricht. Daher gilt $f(x_0) = c$ und der Satz folgt. □

**Übung 3.60** (Flugreise)**.** *In dieser Übung möchten wir eine Anwendung des Zwischenwertsatzes aus dem Alltag präsentieren. Angenommen Sie fliegen von Zürich nach Lima. Zeigen Sie, dass Sie dabei den Äquator überqueren müssen.*

Wir wenden uns nun formaleren Anwendungen des Zwischenwertsatzes zu. Der Zwischenwertsatz ist beispielsweise nützlich, wenn man versucht Nullstellen von stetigen Funktionen zu finden oder wenn man versucht, „Löcher“ im Bild einer Funktion auszuschliessen.

**Übung 3.61** (Nullstellen von Polynomen)**.** *Zeigen Sie, dass jedes reelle Polynom von ungeradem Grad eine (reelle) Nullstelle besitzt. Gehen Sie dazu wie folgt vor: Sei $f = \sum_{j=0}^{n} a_j x^j$ in $\mathbb{R}[x]$ mit $a_n \neq 0$ und $n$ ungerade.*

**Übung 3.62** (Injektivität und Monotonie)**.** *Sei $I$ ein nicht-leeres Intervall und $f : I \to \mathbb{R}$ eine stetige, injektive Abbildung. Zeigen Sie, dass $f$ streng monoton ist.*

In den folgenden zwei Übungen möchten wir auf einen weiteren Begriff hinweisen, der eng in Verbindung zum Zwischenwertsatz steht. Im zweiten Semester werden wir auf diesen Zusammenhang zurückkehren.

**Übung 3.63** (Zusammenhängende Teilmengen von $\mathbb{R}$)**.** *Wir nennen eine Teilmenge $M \subseteq \mathbb{R}$* ***zusammenhängend****, wenn es keine zwei offenen Mengen $U, V \subseteq \mathbb{R}$ mit*

$$(U \cap M) \sqcup (V \cap M) = M, \ U \cap M \neq \emptyset, \ V \cap M \neq \emptyset \tag{3.7}$$

*gibt. Intuitiv ist die Menge $M$ also zusammenhängend, wenn sie sich nicht durch offene Mengen auseinanderreissen lässt. Ziel dieser Übung ist es zu zeigen, dass die zusammenhängenden Teilmengen von $\mathbb{R}$ gerade die Intervalle sind.*

*(i) Sei $M \subseteq \mathbb{R}$ eine Teilmenge, die nicht ein Intervall ist. Zeigen Sie unter Verwendung von Übung 2.79, dass $M$ nicht zusammenhängend ist.*

*(ii) Sei nun $I \subseteq \mathbb{R}$ ein nicht-leeres Intervall und $U, V \subseteq \mathbb{R}$ offen wie in Gleichung (3.7) für $M = I$. Seien $u \in U \cap I$ und $v \in V \cap I$ und ohne Beschränkung der Allgemeinheit $u < v$. Betrachten Sie die Menge $S = \{s \mid [u, s] \subseteq U\}$ und zeigen Sie in Analogie zum Beweis des Zwischenwertsatzes, dass das Supremum von $S$ weder in $U$ noch in $V$, aber in $I$ liegen muss.*

**Übung 3.64** (Zwischenwertsatz via Zusammenhang)**.** *Zeigen Sie den Zwischenwertsatz in folgenden Schritten. Sei $I = [a, b]$ ein Intervall und $f : I \to \mathbb{R}$ stetig.*

*(i) (Charakterisierung von Stetigkeit) Zeigen Sie für jede offene Menge $U \subseteq \mathbb{R}$, dass $f^{-1}(U)$ von der Form $U' \cap I$ für eine offene Menge $U' \subseteq \mathbb{R}$ ist.*

*(ii) Zeigen Sie, dass das Bild von $f$ zusammenhängend ist.*

*(iii) Schliessen Sie auf den Zwischenwertsatz unter Verwendung von Übung 3.63 und (ii).*

VO: 25.10.17

## 3.7 Der Satz über die Umkehrabbildung

In diesem Teilabschnitt wollen wir nun zeigen, dass jede stetige, streng monotone Abbildung eine inverse Abbildung (mit denselben schönen Eigenschaften) besitzt.

**Satz 3.65** (Umkehrsatz). *Sei $I$ ein Intervall und $f : I \to \mathbb{R}$ eine stetige, streng monotone Funktion. Dann ist $f(I) \subseteq \mathbb{R}$ wieder ein Intervall und die Abbildung $f : I \to f(I)$ hat eine stetige, streng monotone inverse Abbildung $f^{-1} : f(I) \to I$. Falls $I = [a, b]$ für reelle Zahlen $a < b$, dann gilt des Weiteren, dass $f(I)$ die Endpunkte $f(a)$ und $f(b)$ hat.*

Tatsächlich ist $f^{-1}$ streng monoton wachsend (respektive streng monoton fallend), wenn $f$ streng monoton wachsend (respektive streng monoton fallend) ist. (Wieso?). Bevor wir uns dem Beweis zuwenden, wollen wir eine Anwendung dieses Satzes betrachten.

**Beispiel 3.66** (Existenz von Wurzeln höherer Ordnung). *Sei $n \in \mathbb{N}$. Dann ist die Funktion $x \in [0, \infty) \mapsto x^n \in [0, \infty)$ streng monoton wachsend und surjektiv. Um Surjektivität zu sehen betrachten wir ein beliebiges $c \in [0, \infty)$. Nach der Bernoullischen Ungleichung (Lemma 3.5) gilt $(c+1)^n > nc \geq c$, womit $c$ zwischen $0 = 0^n$ und $(c+1)^n$ liegt. Aus dem Zwischenwertsatz (Satz 3.59) folgt nun, dass es ein $x$ zwischen $0$ und $c+1$ gibt, für das $x^n = c$ ist.*

*Nach dem Umkehrsatz (Satz 3.65) existiert eine stetige, streng monoton wachsende Umkehrabbildung*

$$x \in [0, \infty) \mapsto \sqrt[n]{x} \in [0, \infty),$$

*die die **n-te Wurzel** genannt wird. Des Weiteren definieren wir für $x \in [0, \infty)$ und $m, n \in \mathbb{N}$*

$$x^{\frac{m}{n}} = \sqrt[n]{x^m}$$

*und für $x \in (0, \infty)$ und $m, n \in \mathbb{N}$ auch $x^{-\frac{m}{n}} = (x^{\frac{m}{n}})^{-1}$*

**Wichtige Übung 3.67** (Rechenregeln für rationale Potenzen). *Zeigen Sie die Rechenregeln (in Analogie zu Übung 3.2)*

$$(xy)^{\frac{m}{n}} = x^{\frac{m}{n}} y^{\frac{m}{n}}, \quad x^{\frac{m_1}{n_1} + \frac{m_2}{n_2}} = x^{\frac{m_1}{n_1}} x^{\frac{m_2}{n_2}}, \quad (x^{\frac{m_1}{n_1}})^{\frac{m_2}{n_2}} = x^{\frac{m_1}{n_1} \frac{m_2}{n_2}}$$

*für positive Basen $x, y \in (0, \infty)$ und rationale Exponenten $\frac{m}{n}, \frac{m_1}{n_1}, \frac{m_2}{n_2} \in \mathbb{Q}$.*

*Beweis von Satz 3.65.* Ohne Beschränkung der Allgemeinheit können wir annehmen, dass $f$ streng monoton wachsend ist (sonst ersetzt man $f$ mit $-f$). Wir bemerken zuerst, dass die Funktion $f : I \to f(I)$ bijektiv ist, da sie (per Definition) surjektiv ist und auf Grund der strengen Monotonie auch injektiv ist. Somit existiert eine (eindeutig bestimmte) Umkehrabbildung $g = f^{-1} : f(I) \to I$, welche auch streng monoton wachsend sein muss: Da $f$ streng monoton wachsend ist, gilt (siehe Übung 3.43)

$$x_1 < x_2 \iff f(x_1) < f(x_2)$$

für alle $x_1, x_2 \in I$, was zu

$$f^{-1}(y_1) < f^{-1}(y_2) \iff y_1 < y_2$$

für alle $y_1, y_2 \in f(I)$ äquivalent ist.

Wir möchten nun zeigen, dass $f(I)$ auch ein Intervall ist und nehmen dazu vorerst an, dass $I = [a, b]$ ein abgeschlossenenes, beschränkten Intervall für zwei reelle Zahlen $a < b$ ist. Auf Grund der Monotonie-Annahme gilt $f(x) \in [f(a), f(b)]$ für alle $x \in [a, b]$. Nach dem Zwischenwertsatz (Satz 3.59) ist auch $f([a, b]) = [f(a), f(b)]$ und damit ist $f(I)$ insbesondere ein Intervall.

Sei nun $I$ ein beliebiges[4] Intervall in $\mathbb{R}$ mit Endpunkten $a < b$ in $\overline{\mathbb{R}}$. Wir definieren nun die Punkte $c = \inf f(I) \in \overline{\mathbb{R}}$ und $d = \sup f(I) \in \overline{\mathbb{R}}$ und behaupten, dass $(c, d)$ in $f(I)$ enthalten ist. Sei $y \in (c, d)$. Dann gibt es nach Definition von $c = \inf f(I)$ und wegen $c < y$ ein $x_- \in I$ mit $c \leq f(x_-) < y$. Ebenso gibt es nach Definition von $d = \sup f(I)$ und wegen $y < d$ ein $x_+ \in I$ mit $y < f(x_+) \leq d$. Nach dem Zwischenwertsatz (Satz 3.59) ist also $y \in f(I)$ und die Behauptung folgt. Wir haben damit insbesondere gezeigt, dass $f(I)$ ein Intervall ist.

Falls der linke Endpunkt $a$ von $I$ zu $I$ gehört, dann ist $c = f(a) = \inf f(I) = \min f(I)$. Falls $a$ nicht zu $I$ gehört, dann gibt es zu jedem $x \in I$ ein Element $x_- \in I$ mit $x_- < x$, was wiederum wegen der strengen Monotonie impliziert, dass $f(I)$ kein Minimum besitzt (da es zu jedem $y \in f(I)$ ein Element $y_- \in f(I)$ mit $y_- < y$ gibt). Das heisst, der linke Endpunkt von $I$ gehört zu $I$ genau dann, wenn $c$ zum Intervall $f(I)$ gehört. Dasselbe gilt auch für den rechten Endpunkt.

Wir wollen nun zeigen, dass $g = f^{-1}$ stetig ist. Sei also $y_0 \in f(I)$ und $\varepsilon > 0$. Wir definieren den Punkt $x_0 = g(y_0)$.

Falls $y_0 < d$ und damit auch $x_0 < b$ ist, dann gibt es einen Punkt $x_+$ in $(x_0, x_0 + \varepsilon)$ mit $x_+ < b$. Wir definieren $y_+ = f(x_+) > y_0$ und erhalten für alle $y \in f(I)$

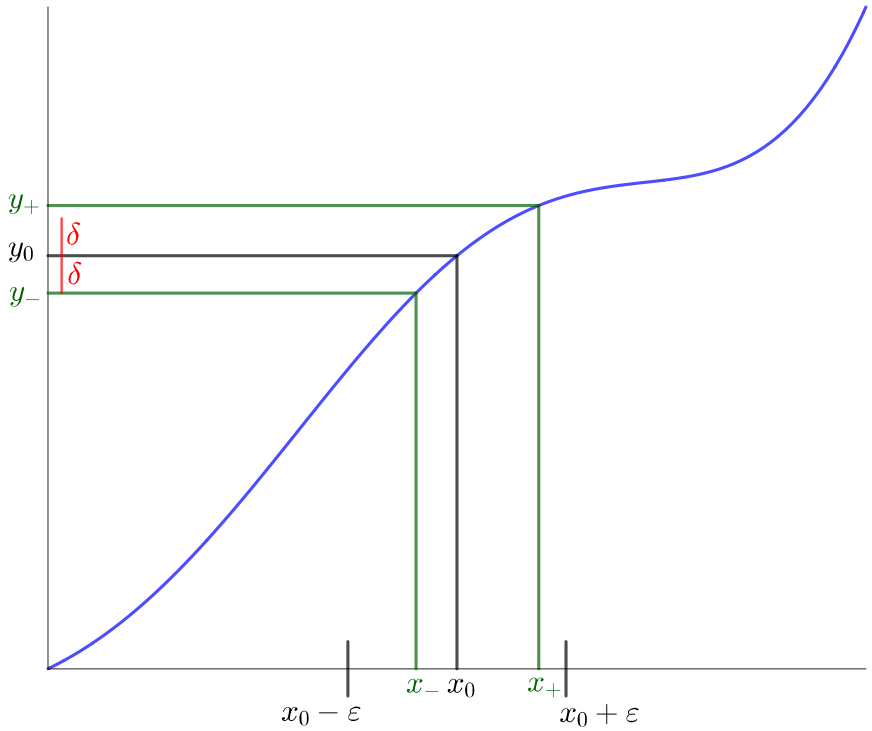
$$y_0 \leq y < y_+ \implies f^{-1}(y_0) \leq f^{-1}(y) < f^{-1}(y_+) = x_+ < f^{-1}(y_0) + \varepsilon$$

oder auch

$$y_0 \leq y < y_+ \implies |f^{-1}(y) - f^{-1}(y_0)| < \varepsilon. \tag{3.8}$$

Falls $a \in I$ und $y_0 = c = f(a)$ gilt, ist dies bereits die Stetigkeitsbedingung bei $y_0$ für $\varepsilon$ und die Wahl $\delta = y_+ - y_0$. In der Tat falls $y \in f(I)$ und $|y - y_0| < \delta$ ist, so folgt $y \geq c = f(a) = y_0$ aus $c = \min(f(I))$ und damit $y_0 \leq y < y_+$ aus der Definition von $\delta$.

[4] Es gibt noch weitere 3 Fälle von beschränkten und weitere 5 Fälle von unbeschränkten Intervallen.




Falls $y_0 > c$ und damit auch $x_0 > a$ ist, dann gibt es einen Punkt $x_- \in (x_0 - \varepsilon, x_0)$ mit $x_- > a$. Wir definieren $y_- = f(x_-) < y_0$ und erhalten wie zuvor für alle

$$y_- < y \leq y_0 \implies f^{-1}(y_0) - \varepsilon < x_- = f^{-1}(y_-) < f^{-1}(y) \leq f^{-1}(y_0)$$

oder auch

$$y_- < y \leq y_0 \implies |f^{-1}(y) - f^{-1}(y_0)| < \varepsilon. \tag{3.9}$$

Falls $b \in I$ und $y_0 = f(b)$ ist dies wiederum die Stetigkeitsbedingung bei $y_0$ für $\varepsilon$ und $\delta = y_0 - y_-$.

Für einen beliebigen Punkt $y_0 \in (a, b)$ setzen wir $\delta = \min\{y_+ - y_0, y_0 - y_-\}$ und können die Gleichungen (3.8) und (3.9) kombinieren zu

$$|y - y_0| < \delta \implies |f^{-1}(y) - f^{-1}(y_0)| < \varepsilon$$

für alle $y \in f(I)$, was zu beweisen war. □

### 3.7.1 Wurzeln aus natürlichen Zahlen

Wir wollen hier nochmals betonen, dass wann immer wir das Supremum verwenden, wir implizit auch das Vollständigkeitsaxiom (Axiom (16) in Abschnitt 2.1) verwenden. Dies ist insbesondere für den Zwischenwertsatz (Satz 3.59) und damit auch für den Umkehrsatz (Satz 3.65) der Fall. Manifestieren tut sich diese Tatsache darin, dass Wurzeln eher selten rationale Zahlen liefern.

**Lemma 3.68.** *Seien $m, k \in \mathbb{N}$. Die $m$-te Wurzel $\sqrt[m]{k}$ ist genau dann rational, wenn sie eine ganze Zahl ist.*

Dieses Lemma vereinfacht unter anderem in konkreten Fällen die Verifikation, ob eine ganze Zahl eine rationale Wurzel hat oder nicht. Grund dafür ist beispielsweise, dass Primzahlen keine $m$-te Wurzel in $\mathbb{Z}$ haben können. (Wieso?)

*Beweis.* Angenommen $\sqrt[m]{k} = \frac{p}{q} \in \mathbb{Q}$ für zwei natürliche Zahlen $p, q \in \mathbb{N}$. Nach Kürzen mit dem grössten gemeinsamen Teiler können wir annehmen, dass $\frac{p}{q}$ durchgekürzt ist oder äquivalent dazu, dass $p$ und $q$ teilerfremd sind. Dann ist aber auch $k = \left(\frac{p}{q}\right)^m = \frac{p^m}{q^m}$ ein durchgekürzter Bruch, denn jeder Primfaktor von $p^m$ (resp. $q^m$) ist ein Primfaktor von $p$ (resp. $q$) und somit sind $p^m$ und $q^m$ teilerfremd. Nach Annahme ist aber $\frac{p^m}{q^m} = k \in \mathbb{Z}$, was $q^m \mid p^m$, $q^m = 1$ und also $q = 1$ impliziert. Dann ist $k = p^m$ und $\sqrt[m]{k} = p \in \mathbb{Z}$. Die Umkehrung folgt aus $\mathbb{Z} \subseteq \mathbb{Q}$. $\square$

Wir sehen also, dass $\sqrt{2}, \sqrt{3}, \sqrt{5}, ..., \sqrt[3]{2}, \sqrt[3]{3}, \sqrt[3]{4}, \sqrt[3]{5}, \sqrt[3]{6}, \sqrt[3]{7}, \sqrt[3]{9}, ...$ alles irrationale Zahlen sind, und haben somit eine Kollektion von (konkreten) irrationalen Zahlen zur Verfügung. Genauer gilt folgendes Kriterion.

**Übung 3.69.** *Sei $k \in \mathbb{N}$ und $m \in \mathbb{N}$. Zeigen Sie, dass $\sqrt[m]{k}$ genau dann rational ist, wenn jeder Primfaktor $p$ in der Primfaktorzerlegung von $k$ die Eigenschaft $p^\ell | k$ und $p^{\ell+1} \nmid k$ für eine durch $m$ teilbare natürliche Zahl $\ell$ hat.*

## 3.8 Stetige Funktionen auf kompakten Intervallen

In diesem Abschnitt möchten wir zeigen, dass stetige Funktionen auf beschränkten, abgeschlossenen Intervallen – sogenannten **kompakten** Intervallen – besondere Eigenschaften besitzen.

### 3.8.1 Beschränktheit

**Satz 3.70** (Beschränktheit)**.** *Seien $a, b \in \mathbb{R}$ zwei reelle Zahlen mit $a < b$ und sei $f : [a, b] \to \mathbb{C}$ eine stetige Funktion. Dann ist $f$ beschränkt. Das heisst, es existiert ein $M \in \mathbb{R}$ mit $|f(x)| \leq M$ für alle $x \in [a, b]$.*

*Beweis.* Wir definieren zuerst die Teilmenge

$$X = \left\{ t \in [a, b] \mid f|_{[a,t]} \text{ ist beschränkt} \right\}.$$

Da $[a, a] = \{a\}$ gilt, liegt $a \in X$, womit $X \subseteq [a, b]$ eine nicht-leere, beschränkte Teilmenge von $\mathbb{R}$ ist. Nach dem Satz über das Supremum (Satz 2.59) existiert daher das Supremum $s_0 = \sup(X)$ von $X$. Des Weiteren muss $s_0 \in [a, b]$ liegen, da zum einen $a \in X$ liegt und zum anderen $X$ in $[a, b]$ enthalten ist und somit $b$ eine obere Schranke ist.

Wir verwenden die Stetigkeit von $f$ bei $s_0$ für $\varepsilon = 1$, wonach es ein $\delta > 0$ gibt, so dass für alle $x \in [a, b]$ die Implikation

$$|x - s_0| < \delta \implies |f(x) - f(s_0)| < 1$$

gilt. Wir definieren $t_0 = \max\{a, s_0 - \delta\}$ und $t_1 = \min\{b, s_0 + \delta\}$, womit

$$|f(x)| \leq |f(x) - f(s_0)| + |f(s_0)| < 1 + |f(s_0)| \tag{3.10}$$

für alle $x \in (t_0, t_1)$.

Da $s_0 - \delta$ keine obere Schranke von $X$ ist, gibt es ein $t \in X$ mit $t > s_0 - \delta$. Daher ist $f|_{[a,t]}$ beschränkt und es existiert ein $M_0 > 0$ mit $|f(x)| \leq M_0$ für alle $x \in [a, t]$.

Es gilt $t \geq t_0 = \max\{a, s_0 - \delta\}$ und daher $[a, t] \cup (t_0, t_1) = [a, t_1)$. Auf Grund von Gleichung (3.10) und der Wahl von $M_0$ gilt somit

$$|f(x)| \leq \max\{M_0, 1 + |f(s_0)|, |f(t_1)|\}$$

für alle $x \in [a, t_1]$. Wir schliessen, dass $t_1 \in X$ liegt und $t_1 \leq s_0$ gilt. Da aber $t_1 = \min\{b, s_0 + \delta\}$ per Definition, muss $t_1 = b$ sein. Also ist $f$ auf $[a, b]$ beschränkt. □

Es drängt sich bei obiger Beweismethode (die wir bereits das dritte Mal angewendet haben) der Vergleich mit einem Induktionsbeweis auf. Wir wollen dies kurz in Worte fassen auch wenn dies bloss eine Analogie darstellt und keineswegs formal als Ersatz von obiger Beweisführung angesehen werden kann. Wir beginnen das Argument damit zu zeigen, dass $a \in X$ ist, was dem Induktionsanfang entspricht. Danach zeigen wir für $t \in X$ mit $t < b$, dass es eine grössere

Zahl gibt, die ebenso in $X$ liegt. Dies entspricht gewissermassen dem Induktionsschritt. Wir hoffen auf diese Weise $b \in X$ zu beweisen, was den Beweis abschliessen würde. Allerdings ist auf diese Weise nicht erklärt ob wir in endlich vielen Schritten $b$ erreichen können. Deswegen benötigen wir die Existenz des Supremums um den Fall auszuschliessen, dass wir zwar immer grössere Elemente in $X$ finden aber vielleicht irgendwo in $(a, b)$ „steckenbleiben“ und nie nach $b$ gelangen.

**Wichtige Übung 3.71** (Gegenbeispiele)**.**

*(i) Finden Sie eine unbeschränkte, stetige Funktion auf einem beschränkten Intervall.*

*(ii) Finden Sie eine unbeschränkte, stetige Funktion auf einem abgeschlossenen Intervall.*

*(iii) Finden Sie eine unbeschränkte Funktion auf einem kompakten Intervall, die nur in einem einzigen Punkt unstetig ist.*

### 3.8.2 Maximum und Minimum

Sei $D \subseteq \mathbb{R}$ eine Teilmenge und $f \in \mathcal{F}(D)$ eine reellwertige Funktion auf $D$. Wir sagen, dass $f$ das **Maximum in** $x_{\max} \in D$ **annimmt**, falls $f(x) \leq f(x_{\max})$ für alle $x \in D$. Analog **nimmt** $f$ ein **Minimum in** $x_{\min} \in D$ **an**, falls $f(x) \geq f(x_{\min})$ für alle $x \in D$. Wir bezeichnen $f(x_{\max})$ als das **Maximum** von $f$ (auf $D$) und $f(x_{\min})$ als das **Minimum** von $f$ (auf $D$). Wir wollen nun zeigen, dass stetige Funktion auf einem kompaktem Intervall stets ihr Minimum und ihr Maximum annehmen.

**Korollar 3.72** (Annahme des Maximums und des Minimums)**.** *Seien $a, b \in \mathbb{R}$ zwei reelle Zahlen mit $a < b$ und sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion. Dann nimmt $f$ sowohl das Maximum als auch das Minimum an.*

*Beweis.* Nach Satz 3.70 ist $f$ beschränkt, womit nach Satz 2.59 das Supremum $S = \sup f([a, b])$ existiert. Wir nehmen nun indirekt an, dass $f(x) < S$ für alle $x \in [a, b]$ gilt, das heisst, dass die Funktion $f$ ihr Maximum nicht annimmt. Dadurch ist

$$F : [a, b] \to (0, \infty),\ x \mapsto \frac{1}{S - f(x)}$$

eine wohldefinierte Funktion. Diese ist nach Proposition 3.53 stetig. Nach Satz 3.70 ist $F$ also beschränkt, womit ein $M > 0$ mit

$$\frac{1}{S - f(x)} = F(x) \leq M$$

für alle $x \in [a, b]$ existiert. Somit gilt

$$\frac{1}{M} \leq S - f(x)$$

oder anders ausgedrückt

$$f(x) \leq S - \frac{1}{M}$$

für alle $x \in [a, b]$. Letzteres widerspricht aber der Definition von $S$ als das Supremum von $f([a, b])$. Daher existiert ein $x_{\max} \in [a, b]$ mit $f(x_{\max}) = \sup f([a, b]) = \max f([a, b])$.

Durch Anwendung des obigen Arguments auf $-f$ ergibt sich ebenso, dass das Minimum von $f$ angenommen wird. □

**Übung 3.73.** *Nimmt jede stetige Funktion $f$ auf dem offenen Intervall $(0, 1)$ ihr Maximum an?*

### 3.8.3 Gleichmässige Stetigkeit

Ein zweiter, grundlegender Satz über stetige Funktionen auf kompakten Intervallen verwendet folgende Verstärkung des Begriffs der $\varepsilon$-$\delta$-Stetigkeit von Definition 3.46.

**Definition 3.74** (Gleichmässige Stetigkeit)**.** Eine reellwertige Funktion $f$ auf einer nichtleeren Teilmenge $D \subseteq \mathbb{R}$ heisst **gleichmässig stetig**, falls es für alle $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $x, y \in D$ gilt

$$|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon.$$

In anderen Worten wollen wir genauso wie bei der Definition von Stetigkeit für jedes $\varepsilon > 0$ ein $\delta > 0$ finden. Diesmal soll jedoch das gewählte $\delta > 0$ nur von $\varepsilon$ abhängen und nicht noch zusätzlich von $x \in D$. Dies entspricht der Vertauschung eines Allquantors mit einem Existenzquantor. Wie wir in Abschnitt 1.3.2 besprochen haben, ergibt dies im Allgemeinen eine inäquivalente Aussage.

**Applet 3.75** (Gleichmässige Stetigkeit)**.** *Wir sehen eine gleichmässig stetige Funktion und können rechts durch getrennte Vergrösserung der Achsen (mit Shift und Maus oder mit zwei Fingern) sowohl $\varepsilon > 0$ als auch $\delta > 0$ verändern.*

**Übung 3.76.** *Zeigen Sie, dass das Polynom $f(x) = x^2$ stetig, aber nicht gleichmässig stetig ist auf $\mathbb{R}$. Verifizieren Sie des Weiteren, dass die Einschränkung von $f$ auf $[0, 1]$ gleichmässig stetig ist.*

**Applet 3.77** (Keine gleichmässige Stetigkeit)**.** *Wir sehen zwei bekannte aber nicht gleichmässig stetige Funktionen und können rechts durch getrennte Vergrösserung der Achsen (mit Shift und Maus oder mit zwei Finger) sowohl $\varepsilon > 0$ als auch $\delta > 0$ verändern.*

Betrachtet man aber nur stetige Funktionen auf kompakten Intervallen, wie wir es hier tun wollen, so befindet man sich in einer ganz anderen Situation.

**Satz 3.78** (Heine, gleichmässige Stetigkeit)**.** *Sei $[a, b]$ ein kompaktes Intervall für $a < b$ und $f : [a, b] \to \mathbb{C}$ eine stetige Funktion. Dann ist $f$ gleichmässig stetig.*

VO: 26.10.17

Wir verwenden im Beweis dieses Satzes nochmals dieselbe Methode wie schon in den Beweisen von dem Satz über die Existenz von Häufungspunkten (Satz 2.75), dem Zwischenwertsatz (Satz 3.59) und dem Satz über die Beschränktheit von stetigen Funktionen auf kompakten Intervallen (Satz 3.70).

*Beweis.* Sei $\varepsilon > 0$. Wir definieren die Teilmenge

$$X = \left\{ t \in [a,b] \;\middle|\; \exists \delta > 0 \, \forall x_1, x_2 \in [a,t] : |x_1 - x_2| < \delta \implies |f(x_1) - f(x_2)| < \varepsilon \right\}$$

von $[a,b]$. In Worten ausgedrückt ist $X$ die Menge der Endpunkte $t \in [a,b]$, für die ein uniformes $\delta > 0$ existiert für die Einschränkung $f|_{[a,t]}$. Wir möchten also zeigen, dass $b \in X$ liegt.

Wir bemerken zuerst, dass $a \in X$ ist, da für $x_1, x_2 \in [a,a] = \{a\}$ sogar $|f(x_1) - f(x_2)| = 0$ gilt und somit jedes $\delta > 0$ gewählt werden kann. Also ist $X$ nicht-leer. Da $X$ in $[a,b]$ enthalten ist und somit beschränkt ist, existiert nach dem Satz über das Supremum (Satz 2.59) das Supremum $s_0 = \sup(X)$ von $X$. Wir bemerken zuerst, dass $t \in X$ und $t' \in [a,t]$ auch $t' \in X$ impliziert (das $\delta$ zu $t$ erfüllt auch die nötige Eigenschaft für $t'$). Daher gelten die Inklusionen

$$[a, s_0) \subseteq X \subseteq [a, s_0]. \tag{3.11}$$

Wir behaupten nun, dass $s_0 = b \in X$ gilt.

Nach Stetigkeit von $f$ bei $s_0 \in [a,b]$ existiert ein $\delta_1 > 0$, so dass für alle $x \in [a,b]$ die Implikation

$$|x - s_0| < \delta_1 \implies |f(x) - f(s_0)| < \frac{\varepsilon}{2}$$

gilt. Für $x_1, x_2 \in [a,b] \cap (s_0 - \delta_1, s_0 + \delta_1)$ gilt damit nach der Dreiecksungleichung

$$|f(x_1) - f(x_2)| \leq |f(x_1) - f(s_0)| + |f(s_0) - f(x_2)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \tag{3.12}$$

Auf Grund von (3.11) liegt $t_0 = \max\left\{a, s_0 - \frac{1}{2}\delta_1\right\}$ in $X$ und daher existiert ein $\delta_0 > 0$ mit

$$\forall x_1, x_2 \in [a, t_0] : |x_1 - x_2| < \delta_0 \implies |f(x_1) - f(x_2)| < \varepsilon. \tag{3.13}$$

Wir definieren $t_1 = \min\left\{b, s_0 + \frac{1}{2}\delta_1\right\}$ sowie $\delta = \min\left\{\delta_0, \frac{1}{2}\delta_1\right\}$ und behaupten, dass für diese Zahlen

$$\forall x_1, x_2 \in [a, t_1] : |x_1 - x_2| < \delta \implies |f(x_1) - f(x_2)| < \varepsilon \tag{3.14}$$

gilt.

Für den Beweis dieser Behauptung nehmen wir also Punkte $x_1, x_2 \in [a, t_1]$ mit $|x_1 - x_2| < \delta$. Nun unterscheiden wir zwei Fälle.

- Angenommen $|x_1 - s_0| \leq \frac{1}{2}\delta_1$ oder $|x_2 - s_0| \leq \frac{1}{2}\delta_1$. Wir gehen ohne Beschränkung der Allgemeinheit von ersterem aus. Dann gilt nach der Dreiecksungleichung

  $$|x_2 - s_0| \leq |x_2 - x_1| + |x_1 - s_0| < \delta + \tfrac{1}{2}\delta_1 \leq \tfrac{1}{2}\delta_1 + \tfrac{1}{2}\delta_1 = \delta_1$$

  und somit $|f(x_1) - f(x_2)| < \varepsilon$ nach (3.12).

- Angenommen $|x_1 - s_0| > \frac{1}{2}\delta_1$ und $|x_2 - s_0| > \frac{1}{2}\delta_1$. Da auch $x_j \leq t_1 \leq s_0 + \frac{1}{2}\delta_1$ für $j \in \{1, 2\}$ gilt, folgt $x_j \leq s_0 - \frac{1}{2}\delta_1 \leq t_0$ für $j \in \{1, 2\}$ und insbesondere $x_1, x_2 \in [a, t_0]$. Nach Gleichung (3.13) gilt also auch in diesem Fall $|f(x_1) - f(x_2)| < \varepsilon$.

Dies beweist die Behauptung, womit auch $t_1 \in X$ gilt. Da aber $s_0$ das Supremum von $X$ ist und kleiner gleich $t_1 = \min\left\{b, s_0 + \frac{1}{2}\delta_1\right\}$ ist, muss $t_1 = s_0$ sein. Dies ist per Definition von $t_1$ aber nur dann möglich, wenn $s_0 = b$ ist, womit wir $b = s_0 = t_1 \in X$ gezeigt haben. Das heisst, für $\varepsilon > 0$ existiert ein $\delta > 0$, welches für alle $x_1, x_2 \in [a, b]$ die Implikation

$$|x_1 - x_2| < \delta \implies |f(x_1) - f(x_2)| < \varepsilon$$

erfüllt. Da $\varepsilon > 0$ beliebig war, beweist dies die gleichmässige Stetigkeit von $f$. □

**Übung 3.79.** *Gilt die Aussage von Satz 3.78 auch für stetige Funktionen auf dem offenen Intervall* $(0, 1)$*?*

Es gibt weitere, interessante Beweise von Satz 3.78. In der folgenden Übung möchten wir illustrieren, wie Satz 3.78 mit einer stetigen Wahl von $\delta$ zusammenhängen kann und dass eine solche Wahl (unabhängig vom Definitionsbereich in $\mathbb{R}$) existiert.

**Übung 3.80.** *Sei $D \subseteq \mathbb{R}$ eine Teilmenge und sei $f : D \to \mathbb{R}$ eine stetige Funktion. Sei $\varepsilon > 0$. Nach Definition der Stetigkeit gibt es für jedes $x_0 \in D$ ein $\delta_{x_0} > 0$, so dass*

$$|x - x_0| < \delta_{x_0} \implies |f(x) - f(x_0)| < \varepsilon. \tag{3.15}$$

*für alle $x \in D$ gilt. Wir möchten nun eine Funktion $\delta : D \to \mathbb{R}_{>0}, x_0 \mapsto \delta_{x_0}$ konstruieren, welche stetig ist. Für jeden Punkt $x_0 \in D$ definieren wir dazu*

$$\delta_{x_0} = \sup\left\{\delta' \in (0, 1] \mid \forall x, y \in (x_0 - \delta', x_0 + \delta') : |f(x) - f(y)| < \varepsilon\right\}.$$

*(i) Zeigen Sie, dass die Menge rechts in obiger Gleichung nicht-leer ist und $\delta_{x_0}$ somit wohldefiniert ist.*

*(ii) Zeigen Sie, dass die Abbildung $x_0 \in D \mapsto \delta_{x_0} \in (0, 1]$ stetig ist und dass für jedes $x_0 \in D$ die Zahl $\delta_{x_0}$ die Implikation* (3.15) *erfüllt.*

*Sei nun $D = [a, b]$. Verwenden Sie die oben konstruierte Funktion $x_0 \in [a, b] \mapsto \delta_{x_0} \in (0, 1]$ und Korollar 3.72, um Satz 3.78 zu beweisen.*

Nach Satz 3.78 bildet jede stetige Funktion auf einem kompakten Intervall ein Beispiel einer gleichmässig stetigen Funktion. Weitere Beispiele (auch auf allgemeineren Teilmengen von $\mathbb{R}$) kann man mittels folgendem Begriff finden.

**Übung 3.81** (Lipschitz-Stetigkeit)**.** *In dieser Übung möchten wir einen weiteren Stetigkeitsbegriff diskutieren.*

*a) Sei $D \subseteq \mathbb{R}$ eine Teilmenge. Wir nennen eine reellwertige Funktion $f$ auf $D$* ***Lipschitz-stetig****, falls ein $L \geq 0$ existiert mit $|f(x) - f(y)| \leq L|x - y|$ für alle $x, y \in D$. Geben Sie ein,*

*zwei Beispiele von Lipschitz-stetigen Funktionen und zeigen Sie, dass eine Lipschitz-stetige Funktion auch gleichmässig stetig ist.*

b) *Zeigen Sie, dass die Wurzelfunktion* $x \in [0, 2] \mapsto \sqrt{x}$ *zwar gleichmässig stetig, aber nicht Lipschitz-stetig ist.*

c) *Zeigen Sie, dass die Wurzelfunktion* $x \in [1, \infty) \mapsto \sqrt{x}$ *Lipschitz-stetig und gleichmässig stetig ist.*

d) *Folgern Sie, dass die Wurzelfunktion* $x \in [0, \infty) \mapsto \sqrt{x}$ *gleichmässig stetig ist.*

## 3.9 Weitere Lernmaterialien

### 3.9.1 Verwendung des Kapitels

Wir werden die Notationen $\sum_{k=m}^{n} a_n$ und $\prod_{k=m}^{n} a_n$ und das Verhalten dieser, zum Beispiel unter Indexverschiebung, immer häufiger benötigen. Ebenso sind Polynome, die Fakultät, der Binomialsatz und die Begriffe der Monotonie und Stetigkeit für alles Weitere von fundamentaler Bedeutung, weswegen diese Begriffe und die ersten Resultate für diese Begriffe in Zukunft meist ohne Verweis auf die jeweiligen Definitionen oder Sätze verwendet werden.

Der Zwischenwertsatz (Satz 3.59) ist ein wichtiges Resultat. Vor allem aber ist er ein wichtiger Bestandteil unseres Beweises von dem Satz über den Umkehrsatz (Satz 3.65), welchen wir später für die korrekte Konstruktion vieler Funktionen verwenden werden. Insbesondere erlaubt uns letzterer die Funktionen $x \in [0, \infty) \mapsto x^{\frac{1}{m}}$ für jedes $m \in \mathbb{N}$ und $x \in (0, \infty) \mapsto x^r$ für jedes $r \in \mathbb{Q}$ zu definieren. Wir werden diese und alle dazugehörigen Potenzregeln in Übung 3.67 in Zukunft ohne Verweis verwenden.

Die Resultate aus Abschnitt 3.8 (also der Satz über die Beschränktheit und die gleichmässige Stetigkeit) werden bereits im nächsten Kapitel Bedeutung erhalten. Wie wir später sehen werden, sind diese Resultate Spezialfälle von allgemeineren Aussage für stetige Funktionen auf sogenannten „kompakten metrischen Räumen“. Mittlerweile sollten Sie logisch geschult sein und den Unterschied (vergleiche Beispiele 1.6 und 1.7) in den Definitionen von Stetigkeit und gleichmässiger Stetigkeit klar erkennen, weswegen Sie auch den Satz über die gleichmässige Stetigkeit besonders schätzen sollten. Wir wollen noch betonen, dass diese Unterscheidung keine Spitzfindigkeit darstellt.

### 3.9.2 Weitere Übungsaufgaben

**Übung.** *Sei $n \in \mathbb{N}$ und seien $v_1, \ldots, v_n$ Elemente eines komplexen Vektorraums $V$. Finden Sie einen vereinfachten Ausdruck für die Doppelsumme*

$$\sum_{j=1}^{n} \sum_{k=j+1}^{n} (v_j - v_k).$$

**Übung** (Formale Definition des Polynomrings)**.** *Das Ziel dieser Aufgabe ist, den Ring der Polynome über einem beliebigen Körper formal zu definieren. Im Folgenden ist $\mathbb{K}$ ein beliebiger Körper und $\mathbb{K}[X]$ bezeichnet die Teilmenge* der schliesslich verschwindenden *Funktionen in $\mathbb{N}_0 \to \mathbb{K}$, das heisst,*

$$\mathbb{K}[X] = \{f : \mathbb{N}_0 \to \mathbb{K} \mid \exists N \in \mathbb{N}_0 \ \forall n \in \mathbb{N}_0 : n \geq N \implies f(n) = 0\}.$$

*Des Weiteren definieren wir Operationen $+$ und $\cdot$ auf $\mathbb{K}[X]$ durch*

$$(f+g)(n) = f(n) + g(n)$$
$$(f \cdot g)(n) = \sum_{k=0}^{n} f(k)g(n-k).$$

*für alle $n \in \mathbb{N}_0$ und $f, g \in \mathbb{K}[X]$.*

*(i) Zeigen Sie, dass $\mathbb{K}[X]$ mit den oben definierten Operationen einen kommutativen Ring bildet.*

*(ii) Wir fassen $\mathbb{K}$ als eine Teilmenge von $\mathbb{K}[X]$ auf, indem wir $a \in \mathbb{K}$ mit der Funktion $n \in \mathbb{N}_0 \mapsto a\mathbb{1}_{\{0\}}(n)$ identifizieren. Zeigen Sie, dass $0 \in \mathbb{K}$ eine Null und $1 \in \mathbb{K}$ eine Eins des Ringes $\mathbb{K}[X]$ ist.*

*(iii) Für alle $k \in \mathbb{N}_0$ sei $X^k \in \mathbb{K}[X]$ die Abbildung gegeben durch*

$$X^k(n) = \begin{cases} 1 & \text{falls } n = k \\ 0 & \text{sonst} \end{cases}$$

*für alle $n \in \mathbb{N}_0$. Zeigen Sie, dass sich jedes Element $f \in \mathbb{K}[X]$ als eindeutig bestimmten Ausdruck der Form*

$$f = \sum_{n=0}^{N} a_n X^n$$

*für ein $N \in \mathbb{N}$ und Zahlen $a_0, \ldots, a_n \in \mathbb{K}$ mit $a_N \neq 0$ schreiben lässt.*

*(iv) Vergleichen Sie die obigen Definitionen zur Definition des Polynomrings in Definition 3.13.*

**Übung** (Ein Körper mit neun Elementen)**.** *Wir möchten in dieser Übung einen Körper mit neun Elementen konstruieren und folgen dabei der Bemerkung am Ende von Abschnitt 3.2.2.*

*(i) Zeigen Sie, dass das Polynom $f(x) = x^2 + x + 2$ über dem Körper $\mathbb{F}_3$ keine Nullstelle besitzt.*

*Wir betrachten nun den Polynomring $\mathbb{F}_3[x]$ und die Relation $g_1 \sim g_2 \iff f$ teilt $(g_1 - g_2)$.*

*(ii) Zeigen Sie, dass $\sim$ eine Äquivalenzrelation ist. Sei $\mathbb{K} = \mathbb{F}_3[x]/\!\sim$ der dazugehörige Quotientenraum.*

*(iii) Zeigen Sie, dass die Operationen*

$$[g_1]_\sim + [g_2]_\sim = [g_1 + g_2]_\sim$$
$$[g_1]_\sim \cdot [g_2]_\sim = [g_1 \cdot g_2]_\sim$$

*wohldefiniert sind und aus $\mathbb{K}$ einen Körper mit neun Elementen machen.*

**Übung** (Zwei Identitäten für Binomialkoeffizienten)**.** *Seien $k, n \in \mathbb{N}_0$ mit $1 \leq k \leq n$. Zeigen Sie die Identitäten*

$$\binom{n}{k} = \frac{n+1-k}{k}\binom{n}{k-1}, \quad \binom{n-1}{k} - \binom{n-1}{k-1} = \frac{n-2k}{n}\binom{n}{k}.$$

**Übung** ($\mathbb{R}$-wertige Funktionen auf einer Zweipunktmenge)**.** *Sei $D$ eine Menge bestehend aus 2 Elementen. Zeigen Sie, dass es einen Isomorphismus von Vektorräumen $F(D) \cong \mathbb{R}^2$ gibt. Induzieren Sie durch diese Bijektion eine Ordnung auf $\mathbb{R}^2$ und beschreiben Sie diese (beispielsweise duch Beschreibung welche Elemente grösser als $(0,0)$ und welche kleiner als $(0,0)$ sind).*

**Übung** (Dimension von $F(D)$)**.** *Sei $D$ eine nicht-leere Menge. Zeigen Sie, dass $F(D)$ genau dann endlich-dimensional ist, wenn $D$ endlich ist und dass in diesem Fall die Dimension gerade $|D|$ ist.*

**Übung** (Das Multinomialtheorem)**.** *In dieser Übung möchten wir in Analogie zum binomischen Lehrsatz (Satz 3.28) Ausdrücke der Form $(z_1 + \ldots + z_d)^n$ für komplexe Zahlen $z_1, \ldots, z_d$ und $n \in \mathbb{N}_0$ untersuchen. Wir betrachten dazu sogenannte Multiindizes $\boldsymbol{\alpha} \in \mathbb{N}_0^d$. Ist $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d$ ein Multiindex, $n = \sum_{k=1}^d \alpha_k$ und $\boldsymbol{z} = (z_1, \ldots, z_d) \in \mathbb{C}^d$, so definieren wir*

$$\binom{n}{\boldsymbol{\alpha}} = \frac{n!}{(\alpha_1!)\ldots(\alpha_n!)}$$

*sowie*

$$\boldsymbol{z}^{\boldsymbol{\alpha}} = z_1^{\alpha_1} \ldots z_d^{\alpha_d}.$$

*Zeigen Sie für alle $\boldsymbol{z} = (z_1, \ldots, z_d) \in \mathbb{C}^d$ und $n \in \mathbb{N}_0$ das Multinomialgesetz*

$$(z_1 + \ldots + z_d)^n = \sum_{\boldsymbol{\alpha} \in \mathbb{N}_0^d \colon \alpha_1 + \ldots + \alpha_d = n} \binom{n}{\boldsymbol{\alpha}} \boldsymbol{z}^{\boldsymbol{\alpha}}.$$

**Übung** (Eigenschaften komplexwertiger Funktionen)**.** *Sei $D \subseteq \mathbb{C}$ eine nicht-leere Teilmenge.*

*(i) Definieren Sie den Begriff der Stetigkeit (in einem Punkt in $D$) für Funktionen $D \to \mathbb{C}$.*

*(ii) Zeigen Sie, dass eine Funktion $f : D \to \mathbb{C}$ genau dann in $x_0 \in D$ stetig ist, wenn die Funktionen $\mathrm{Re}(f) : D \to \mathbb{R},\ x \mapsto \mathrm{Re}(f(x))$ und $\mathrm{Im}(f) : D \to \mathbb{R},\ x \mapsto \mathrm{Im}(f(x))$ in $x_0$ stetig sind.*

*(iii) Formulieren Sie das Analogon von Proposition 3.51 für komplexwertige Funktionen und beweisen Sie es (zum Beispiel unter Verwendung von (ii) oder direkt).*

*(iv) Formulieren und beweisen Sie Proposition 3.53 für komplexwertige Funktionen.*

**Übung** (Formalisierung der Nicht-Stetigkeit)**.** *Sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ eine Funktion. Drücken Sie die Aussagen „$f$ ist nicht stetig“ und „$f$ ist nicht stetig bei einem Punkt*

*$x_0 \in I$" in Prädikatenlogik aus. Zeigen Sie damit, dass die Funktion*

$$\mathbb{R} \to \mathbb{R},\ x \mapsto \begin{cases} x+1 & \text{falls } x \geq 0 \\ x & \text{falls } x < 0 \end{cases}$$

*aus dem Teilabschnitt 3.4.2 nicht stetig ist.*

**Übung** (Dirichlet-Funktion)**.** *Zeigen Sie, dass die charakteristische Funktion*

$$\mathbb{1}_{\mathbb{Q}\cap[0,1]} : [0,1] \to [0,1],\ x \mapsto \begin{cases} 1 & \text{falls } x \in \mathbb{Q} \\ 0 & \text{falls } x \notin \mathbb{Q} \end{cases}$$

*an keinem Punkt in $[0,1]$ stetig ist.*

**Übung** (Lineare Abschätzung bei $x_0$)**.** *Sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ eine Funktion. Angenommen es existiert zu $x_0 \in I$ eine Konstante $L_{x_0} \geq 0$, so dass für alle $x \in I$ gilt $|f(x) - f(x_0)| \leq L_{x_0}|x - x_0|$. Zeigen Sie, dass $f$ stetig bei $x_0$ ist.*

**Übung.** *Sei $D \subseteq \mathbb{R}$ eine Teilmenge und seien $f_1, f_2 \in C(D)$. Zeigen Sie, dass dann auch die Funktionen*

$$\max(f_1, f_2) : D \to \mathbb{R},\ x \mapsto \max(f_1(x), f_2(x))$$
$$\min(f_1, f_2) : D \to \mathbb{R},\ x \mapsto \min(f_1(x), f_2(x))$$

*stetig sind.*

**Übung** (Kompakter Träger)**.** *Wir sagen, dass eine Funktion $f : \mathbb{R} \to \mathbb{R}$ einen kompakten Träger hat, falls ein $M > 0$ existiert mit $f(x) = 0$ für alle $x \in \mathbb{R}$ mit $|x| > M$. Sei nun $f : \mathbb{R} \to \mathbb{R}$ eine stetige Funktion mit kompaktem Träger. Zeigen Sie, dass $f$ gleichmässig stetig und beschränkt ist.*

**Übung** (Offene und abgeschlossene Intervalle)**.** *In dieser Übung möchten wir zeigen, dass sich das offene $(0,1)$ Intervall vom abgeschlossenen $[0,1]$ Intervall zwar von der Kardinalität her nicht unterscheiden, aber von der Ordnung her sehr wohl.*

*(i) Finden Sie eine Bijektion $f : [0,1] \to (0,1)$.*

*(ii) Zeigen Sie, dass keine stetige, bijektive Abbildung $[0,1] \to (0,1)$ existieren kann.*

**Übung.** *Zeigen Sie, dass die Abbildung $x \in \mathbb{R} \mapsto x^7 + x^5 + x^3 + x \in \mathbb{R}$ bijektiv ist (ohne zu versuchen, eine Formel für die inverse Abbildung anzugeben).*

**Übung.** *Beweisen Sie Satz 3.70 und Korollar 3.72 mit Hilfe des Intervallschachtelungsprinzips in Satz 2.77.*

### 3.9.3 Lernkarten

Sie können wiederum die Lernkarten oder den Graphen für Ihre Wiederholung der Themen des Kapitels verwenden.

# Kapitel 4

# Das Riemann-Integral

Wir werden in diesem Kapitel die Idee von Abschnitt 1.1 aufgreifen und diese mit Hilfe des Supremums und des Infimums (also implizit des Vollständigkeitsaxioms) aus Kapitel 2 und der $\sum$-Notation aus Kapitel 3 zum Begriff des Riemann-Integrals ausbauen.

Leser fragen sich vielleicht, warum wir hier schon das Integral besprechen, obwohl wir die Ableitung noch nicht besprochen haben. Es gibt viele Wege, die zum Ziel führen, und wir könnten in der Tat ebenso das Integral nach der Ableitung einführen. Für diese Reihenfolge sprechen die folgenden Argumente:

- Flächeninhalte wurden bereits seit der Antike untersucht und (teilweise) berechnet. Die Ableitung hat eine kürzere Geschichte und ist eigentlich ein schwierigeres Konzept als das Integral.

- Auch vom rein mathematischen Gesichtspunkt gesehen, ist das Integral viel einfacher. Wie wir hier sehen werden, erfüllt das Integral einige sehr nette Eigenschaften (zum Beispiel Monotonie und eine verallgemeinerte Dreiecksungleichung), welche keine Entsprechung für die Ableitung haben. Wir werden später diese netten Eigenschaften des Integrals verwenden, um gewisse Aussagen für die Ableitung zu zeigen. Da wir in dieser Vorlesung die Analysis nach ihren inneren Strukturen aufbauen wollen, spricht dies dafür das Integral zuerst zu besprechen.

- Der Zusammenhang zwischen Ableitung und Integral ist eines der Hauptziele für dieses erste Semester in Analysis. Wir hoffen, dass die frühe Einführung des Integrals die Wichtigkeit dieses Zusammenhangs weiter betont.

## 4.1 Treppenfunktionen und deren Integral

Im Folgenden seien $a < b$ zwei reelle Zahlen und $[a, b] \subseteq \mathbb{R}$ das kompakte Intervall.

### 4.1.1 Zerlegungen

**Definition 4.1** (Zerlegung)**.** Eine **Zerlegung** (oder **Unterteilung**) $\mathfrak{Z}$ von $[a, b]$ ist gegeben durch endlich viele Punkte

$$a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b$$

mit $n \in \mathbb{N}$. Die Punkte $x_0, \ldots, x_n \in [a, b]$ werden die Teilungspunkte der Zerlegung genannt. Wir schreiben $\mathfrak{Z} = \{a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b\}$.

Formal gesehen ist eine Zerlegung also eine endliche Teilmenge unseres Intervalls $[a, b]$, die $a$ und $b$ enthält, gemeinsam mit einer Auflistung ihrer Elemente durch eine streng monotone Funktion $k \in \{0, \ldots, n\} \mapsto x_k$. (Die Aufzählung ist eindeutig durch die Teilmenge bestimmt, da wir die Forderung $x_0 < x_1 < \ldots < x_n$ stellen). Eine Zerlegung induziert auch eine spezielle Art von Partition, nämlich

$$\mathcal{P}(\mathfrak{Z}) = \big\{\{a\}, (x_0, x_1), \{x_1\}, \ldots, (x_{n-1}, x_n), \{b\}\big\},$$

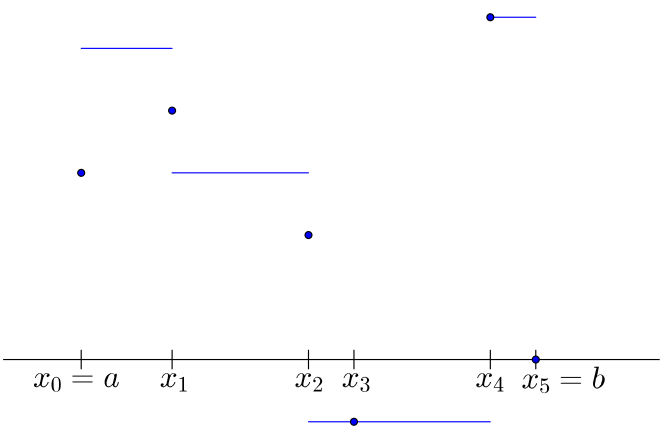
die fortan implizit in den Diskussionen verwendet wird.

**Definition 4.2** (Treppenfunktion)**.** Eine Funktion $f : [a, b] \to \mathbb{R}$ ist eine **Treppenfunktion** (abgekürzt $\mathcal{TF}$), falls es eine Zerlegung $\mathfrak{Z} = \{a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b\}$ gibt, so dass es für jedes $k \in \{1, \ldots, n\}$ eine Zahl $c_k \in \mathbb{R}$ gibt mit

$$\forall x \in (x_{k-1}, x_k) : f(x) = c_k.$$

Eine Treppenfunktion soll also konstant sein auf den Intervallen in der Partition $\mathcal{P}(\mathfrak{Z})$. Die Intervalle $(x_{k-1}, x_k)$ für $k \in \{1, \ldots, n\}$ heissen auch **Konstanzintervalle** der Treppenfunktion $f$ und $\mathfrak{Z}$ heisst eine **Zerlegung in Konstanzintervalle** von $f$. Die Zahlen $c_1, \ldots, c_n$ nennen wir **Konstanzwerte** von $f$ bezüglich $\mathfrak{Z}$.

Beispielsweise sind konstante Funktionen auch Treppenfunktionen.




Figur 4.1: Der Graph einer Treppenfunktion auf dem Intervall $[a, b]$. Die blauen Punkte deuten die Funktionswerte bei den Punkten $x_0, \ldots, x_5$ an.

**Definition 4.3.** Seien $\mathfrak{Z}, \mathfrak{Z}'$ zwei Zerlegungen von $[a, b]$. Wir sagen, dass $\mathfrak{Z}'$ **feiner als** $\mathfrak{Z}$ ist, falls jeder Teilungspunkt von $\mathfrak{Z}$ ein Teilungspunkt von $\mathfrak{Z}'$ ist. Die **gemeinsame Verfeinerung** zweier Zerlegungen $\mathfrak{Z}$ und $\mathfrak{Z}'$ ist die Zerlegung, deren Menge von Teilungspunkten durch die Vereinigung der Menge der Teilungspunkte von $\mathfrak{Z}$ und von $\mathfrak{Z}'$ gegeben ist.

### 4.1.2 Das Integral einer Treppenfunktion

**Definition 4.4.** Sei $f : [a, b] \to \mathbb{R}$ eine Treppenfunktion und $\mathfrak{Z} = \{a = x_0 < \ldots < x_n = b\}$ eine Zerlegung von $[a, b]$ in Konstanzintervalle von $f$. Seien $c_1, \ldots, c_n$ die Konstanzwerte von $f$ bezüglich $\mathfrak{Z}$. Dann definieren wir

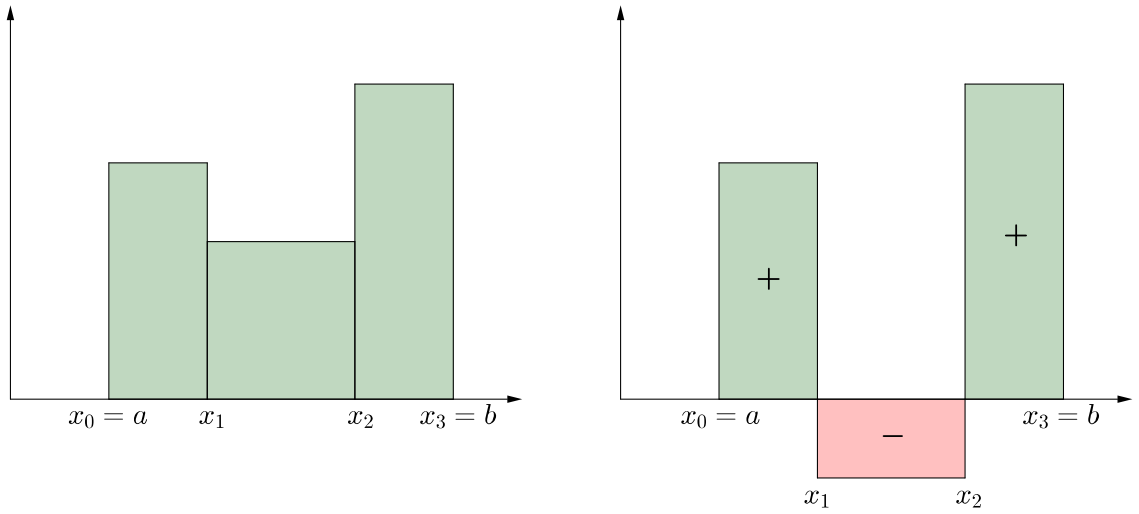
$$I(f, \mathfrak{Z}) = \sum_{k=1}^{n} c_k(x_k - x_{k-1}) = \sum_{k=1}^{n} c_k \Delta x_k,$$

wobei $\Delta x_k = (x_k - x_{k-1})$ für die Länge des $k$-ten Konstanzintervalls in der Zerlegung $\mathfrak{Z}$ für $k = 1, \ldots, n$ steht.

Für eine nicht-negative Treppenfunktion $f \geq 0$ interpretieren wir $I(f, \mathfrak{Z})$ als Flächeninhalt der Ordinatenmenge

$$\big\{(x, y) \in \mathbb{R}^2 \mid a \leq x \leq b,\ 0 \leq y \leq f(x)\big\}$$

und im Allgemeinen als vorzeichenbehafteter Nettoflächeninhalt (siehe Bild unten).




Figur 4.2: Die Zahl $I(f, \mathfrak{Z})$ kann die Summe der Flächeninhalten von Rechtecken über der $x$-Achse oder eine Differenz von Flächeninhalten sein, wenn die Treppenfunktion positive und negative Werte auf den Konstanzintervallen annimmt.

**Lemma 4.5** (Unabhängigkeit von Zerlegung in Konstanzintervalle)**.** *Sei $f : [a, b] \to \mathbb{R}$ eine Treppenfunktion und $\mathfrak{Z} = \{a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b\}$ eine Zerlegung in Konstanzintervalle von $f$. Dann hängt $I(f, \mathfrak{Z})$ nicht von den Funktionswerten $f(x_k)$ für $k \in \{1, \ldots, n\}$ und nicht von der Wahl der Zerlegung $\mathfrak{Z}$ in Konstanzintervalle der Funktion $f$ ab.*

Dieses Lemma wird uns erlauben, den Wert $I(f, \mathfrak{Z})$ als das Integral von $f$ zu definieren.

*Beweis.* Seien $f, f'$ zwei Treppenfunktionen auf $[a, b]$ mit derselben Zerlegung in Konstanzintervalle $\mathfrak{Z} = \{a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b\}$. Falls nun $f(x) = c_k = f'(x)$ für alle $x \in (x_{k-1}, x_k)$ und alle $k \in \{1, \ldots, n\}$, dann gilt

$$I(f, \mathfrak{Z}) = \sum_{k=1}^{n} c_k(x_k - x_{k-1}) = I(f', \mathfrak{Z})$$

und die erste Behauptung im Lemma folgt.

Sei nun $f$ eine Treppenfunktion auf $[a, b]$ und sowohl $\mathfrak{Z}$ als auch $\mathfrak{Z}'$ Zerlegungen in Konstanzintervalle von $f$. Die zweite Behauptung des Lemmas besagt $I(f, \mathfrak{Z}) = I(f, \mathfrak{Z}')$. (Zum Beispiel könnte $\mathfrak{Z} = \{x_0 = a < x_1 < x_2 < x_4 < x_5 = b\}$ und $\mathfrak{Z}' = \{x_0 = a < x_1 < x_2 < x_3 < x_4 < x_5 = b\}$ wie in Figur 4.1 sein.)

Wir beweisen diese Behauptung in drei Schritten. Im ersten Schritt nehmen wir an, dass $\mathfrak{Z}'$ feiner als $\mathfrak{Z} = \{a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b\}$ ist und bloss einen zusätzlichen Trennungspunkt $y \in (x_{\ell-1}, x_\ell)$ für ein $\ell \in \{1, \ldots, n\}$ hat. Unter Verwendung der Abkürzung $\Delta x_k = (x_k - x_{k-1})$ für die Länge des $k$-ten Teilintervalls von $\mathfrak{Z}$ für $k \in \{1, \ldots, n\}$ erhalten wir

$$\begin{aligned} I(f, \mathfrak{Z}) &= \sum_{k=1}^{\ell-1} c_k \Delta x_k + c_\ell(x_\ell - x_{\ell-1}) + \sum_{k=\ell+1}^{n} c_k \Delta x_k \\ &= \sum_{k=1}^{\ell-1} c_k \Delta x_k + c_\ell(x_\ell - y) + c_\ell(y - x_{\ell-1}) + \sum_{k=\ell+1}^{n} c_k \Delta x_k \\ &= I(f, \mathfrak{Z}'), \end{aligned}$$

da $f(x) = c_\ell$ für alle $x \in (x_{\ell-1}, y) \cup (y, x_\ell) \subseteq (x_{\ell-1}, x_\ell)$.

Mittels vollständiger Induktion nach $|\mathfrak{Z}' \setminus \mathfrak{Z}|$ folgt aus obigem Fall, dass $I(f, \mathfrak{Z}) = I(f, \mathfrak{Z}')$, falls $\mathfrak{Z}'$ feiner als $\mathfrak{Z}$ ist. In der Tat kann man eine Liste von Zerlegungen finden, die mit $\mathfrak{Z}$ beginnt, mit $\mathfrak{Z}'$ endet, und in der die nächste jeweils einen Punkt mehr besitzt als die vorhergehende Zerlegung in der Liste.

Falls nun $\mathfrak{Z}, \mathfrak{Z}'$ beliebige Zerlegungen in Konstanzintervalle von $f$ sind, dann können wir die gemeinsame Verfeinerung $\mathfrak{Z}''$ betrachten und erhalten aus dem vorherigen Fall

$$I(f, \mathfrak{Z}) = I(f, \mathfrak{Z}'') = I(f, \mathfrak{Z}'),$$

was den Beweis des Lemmas abschliesst. □

**Definition 4.6.** Für eine Treppenfunktion $f : [a, b] \to \mathbb{R}$ definieren wir das **Integral der Treppenfunktion** $f$ als

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^b f\,\mathrm{d}x = I(f, \mathfrak{Z}),$$

wobei $\mathfrak{Z}$ eine Zerlegung in Konstanzintervalle von $f$ ist.

Nach Lemma 4.5 hängt diese Definition des Integrals nicht von der Wahl der Zerlegung ab.

Wir bemerken auch, dass das Symbol $\int$ für ein stilisiertes $S$ steht und damit an den Zusammenhang zu einer Summe erinnert. Des Weiteren ist die Variable $x$ in der Notation $\int f(x)\,\mathrm{d}x$ eine interne Variable für die Notation des Integrals (genauso wie die Variable $k$ in der Summe $\sum_{k=1}^{n} c_k \Delta x_k$), die ausserhalb des Integrals keine Bedeutung hat (und, um vorprogrammierte Verwirrungen zu vermeiden, auch keine haben sollte).

**Lemma 4.7** (Linearität des Integrals von Treppenfunktionen)**.** *Die nicht-leere Menge*

$$\mathcal{TF}([a,b]) = \{f \in \mathcal{F}([a,b]) \mid f \text{ ist eine Treppenfunktion}\}$$

*der Treppenfunktionen auf dem Intervall $[a,b]$ ist ein Unterraum des Vektorraums $\mathcal{F}([a,b])$ der reellwertigen Funktionen auf $[a,b]$. Des Weiteren ist die Abbildung $\int : \mathcal{TF}([a,b]) \to \mathbb{R}$ linear. Das heisst, für alle $f, g \in \mathcal{TF}([a,b])$ und $s \in \mathbb{R}$ ist $f+g \in \mathcal{TF}([a,b])$, $sf \in \mathcal{TF}([a,b])$ und es gilt*

$$\int_a^b (f+g)\,\mathrm{d}x = \int_a^b f\,\mathrm{d}x + \int_a^b g\,\mathrm{d}x,$$
$$\int_a^b (sf)\,\mathrm{d}x = s\int_a^b f\,\mathrm{d}x.$$

*Beweis.* Falls $\mathfrak{Z}_f$ eine Zerlegung in Konstanzintervalle von $f$ und $\mathfrak{Z}_g$ eine Zerlegung in Konstanzintervalle von $g$ ist, dann existiert eine gemeinsame Verfeinerung

$$\mathfrak{Z} = \{a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b\}$$

von $\mathfrak{Z}_f$ und $\mathfrak{Z}_g$. Dies ist eine Zerlegung in Konstanzintervalle von $f$ und $g$. Seien $c_1, \ldots, c_n$ respektive $d_1, \ldots, d_n \in \mathbb{R}$ die Konstanzwerte von $f$ respektive $g$ bezüglich der Zerlegung $\mathfrak{Z}$, das heisst, es gilt

$$\forall x \in (x_{k-1}, x_k) : f(x) = c_k \text{ und } g(x) = d_k \tag{4.1}$$

für alle $k \in \{1, \ldots, n\}$. Insbesondere ergibt dies für alle $k \in \{1, \ldots, n\}$

$$\forall x \in (x_{k-1}, x_k) : f(x) + g(x) = c_k + d_k \text{ und } (sf)(x) = sc_k$$

und wir erhalten $f+g, sf \in \mathcal{TF}([a,b])$. Des Weiteren gilt

$$\begin{aligned}
\int_a^b (f+g)\,\mathrm{d}x = I(f+g, \mathfrak{Z}) &= \sum_{k=1}^{n} (c_k + d_k)\Delta x_k \\
&= \sum_{k=1}^{n} c_k \Delta x_k + \sum_{k=1}^{n} d_k \Delta x_k \\
&= I(f, \mathfrak{Z}) + I(g, \mathfrak{Z}) \\
&= \int_a^b f\,\mathrm{d}x + \int_a^b g\,\mathrm{d}x
\end{aligned}$$

und ebenso

$$\begin{aligned}\int_a^b (sf)dx = I(sf, \mathfrak{Z}) &= \sum_{k=1}^n sc_k \Delta x_k \\ &= s\sum_{k=1}^n c_k \Delta x_k = sI(f, \mathfrak{Z}) = s\int_a^b f dx.\end{aligned}$$

□

**Lemma 4.8** (Monotonie des Integrals von Treppenfunktionen)**.** *Sind $f, g \in \mathcal{TF}([a,b])$ zwei Treppenfunktionen mit $f \leq g$. Dann gilt*

$$\int_a^b f \,\mathrm{d}x \leq \int_a^b g \,\mathrm{d}x.$$

*Insbesondere impliziert $f \in \mathcal{TF}([a,b])$ und $f \geq 0$, dass $\int_a^b f \,\mathrm{d}x \geq 0$.*

*Beweis.* Wie schon im Beweis des letzten Lemmas können wir für $f, g \in \mathcal{TF}([a,b])$ eine gemeinsame Zerlegung $\mathfrak{Z} = \{a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b\}$ in Konstanzintervalle finden. Wir schreiben wieder $c_1, \ldots, c_n, d_1, \ldots, d_n$ für die Konstanzwerte von $f$ resp. $g$ bezüglich $\mathfrak{Z}$ (wie in Gleichung (4.1)). Falls nun $f \leq g$ (also $f(x) \leq g(x)$ für alle $x \in [a,b]$) ist, dann ist $c_k \leq d_k$ für alle $k \in \{1, \ldots, n\}$ und wir erhalten

$$\int_a^b f \,\mathrm{d}x = I(f, \mathfrak{Z}) = \sum_{k=1}^n c_k \Delta x_k \leq \sum_{k=1}^n d_k \Delta x_k = I(g, \mathfrak{Z}) = \int_a^b g \,\mathrm{d}x.$$

Die zweite Aussage folgt aus der ersten angewendet auf 0 und $f$. □

Durch genauere Betrachtung des obigen Beweises oder Lemma 4.5 sieht man sogar, dass die Ungleichung $f \leq g$ auf den durch eine Zerlegung gegebenen offenen Intervallen für die Konklusion $\int_a^b f \,\mathrm{d}x \leq \int_a^b g \,\mathrm{d}x$ ausreichend ist.

**Übung 4.9** (Integral von „zusammengeklebten“ Treppenfunktionen)**.** *Seien $[a,b], [b,c]$ zwei beschränkte und abgeschlossene Intervalle und sei $f_1 \in \mathcal{TF}([a,b])$ und $f_2 \in \mathcal{TF}([b,c])$. Zeigen Sie, dass die Funktion*

$$f : [a,c] \to \mathbb{R}, \quad x \mapsto \begin{cases} f_1(x) & \text{falls } x \in [a,b) \\ f_2(x) & \text{falls } x \in [b,c] \end{cases}$$

*eine Treppenfunktion auf $[a,c]$ ist und geben Sie eine Zerlegung in Konstanzintervalle von $f$ an. Beweisen Sie anschliessend, dass das Integral von $f$ gegeben ist durch*

$$\int_a^c f \,\mathrm{d}x = \int_a^b f_1 \,\mathrm{d}x + \int_b^c f_2 \,\mathrm{d}x.$$

*Zeigen Sie des Weiteren, dass jede Treppenfunktion auf $[a,c]$ von obiger Form ist.*

VO: 30.10.17

## 4.2 Definition des Riemann-Integrals

Wie schon im letzten Abschnitt betrachten wir im Folgenden Funktionen auf einem kompakten Intervall $[a, b] \subseteq \mathbb{R}$ zu reellen Zahlen $a < b$.

Wir bemerken, dass Treppenfunktionen beschränkt sind, da sie endliche Wertemengen haben. Des Weiteren ist eine reellwertige Funktion $f$ genau dann beschränkt, wenn es Treppenfunktionen $u, o \in \mathcal{TF}([a, b])$ gibt, die $u \leq f \leq o$ erfüllen. In der Tat, falls $u \leq f \leq o$ für gewisse Treppenfunktionen $u, o$ gilt, dann ist $f([a, b])$ von oben durch das Maximum von $o([a, b])$ beschränkt und von unten durch das Minimum von $u([a, b])$ beschränkt. (Wieso?). Umgekehrt können wir konstante Treppenfunktionen $u, o \in \mathcal{TF}([a, b])$ verwenden, falls $f$ beschränkt ist.

**Definition 4.10.** Sei $f \in \mathcal{F}([a, b])$ beschränkt. Dann definieren wir die (nicht-leere) Menge der **Untersummen** durch

$$\mathcal{U}(f) = \left\{ \int_a^b u \, \mathrm{d}x \mid u \in \mathcal{TF}([a, b]) \text{ und } u \leq f \right\}$$

und die (nicht-leere) Menge der **Obersummen** durch

$$\mathcal{O}(f) = \left\{ \int_a^b o \, \mathrm{d}x \mid o \in \mathcal{TF}([a, b]) \text{ und } f \leq o \right\}.$$

Für $u, o \in \mathcal{TF}([a, b])$ mit $u \leq f \leq o$ gilt nach Lemma 4.8 auch

$$\int_a^b u \, \mathrm{d}x \leq \int_a^b o \, \mathrm{d}x.$$

Jede Untersumme ist also kleiner gleich jeder Obersumme. Äquivalenterweise ist jede Obersumme $\int_a^b o \, dx$ eine obere Schranke der nicht-leeren Menge der Untersummen und daher ist

$$\sup \mathcal{U}(f) \leq \int_a^b o \, \mathrm{d}x,$$

da das Supremum die kleinste obere Schranke ist. Insbesondere ist $\sup \mathcal{U}(f)$ eine untere Schranke der Menge der Obersummen $\mathcal{O}(f)$ und es gilt

$$\sup \mathcal{U}(f) \leq \inf \mathcal{O}(f), \tag{4.2}$$

da das Infimum die grösste untere Schranke ist.

**Definition 4.11** (Riemann-Integrierbarkeit)**.** Für eine beschränkte Funktion $f \in \mathcal{F}([a, b])$ wird $\underline{I}(f) = \sup \mathcal{U}(f)$ das **untere Integral** von $f$ und $\overline{I}(f) = \inf \mathcal{O}(f)$ das **obere Integral** von $f$ genannt. Die Funktion $f$ heisst **Riemann-integrierbar**, oder kurz **R-integrierbar**, falls $\underline{I}(f) = \overline{I}(f)$. In diesem Fall wird dieser gemeinsame Wert das Riemann-Integral

$$\int_a^b f \, \mathrm{d}x = \underline{I}(f) = \overline{I}(f)$$

genannt. Des Weiteren definieren wir

$$\mathcal{R}([a,b]) = \big\{ f \in \mathcal{F}([a,b]) \mid f \text{ ist Riemann-integrierbar} \big\}.$$

Wir bezeichnen $a$ als die **untere** und $b$ als die **obere Integrationsgrenze** und die Funktion als den **Integrand** für das Integral $\int_a^b f \,\mathrm{d}x$.

Wir haben hier den Zugang von Darboux für die Definition des Riemann-Integrals gewählt; im nächsten Kapitel werden wir aber auch kurz die sogenannten Riemann-Summen besprechen, die von Riemann als Ausgangspunkt seiner Definition verwendet wurden. Es gibt neben diesen beiden äquivalenten Definitionen noch weitere, die wir nicht besprechen werden.

Falls $f \in \mathcal{F}([a,b])$ nicht-negativ (das heisst, es gilt $f \geq 0$), beschränkt und Riemann-integrierbar ist, dann interpretieren wir die Zahl $\int_a^b f \,\mathrm{d}x$ als den Flächeninhalt der Menge

$$\big\{ (x,y) \in \mathbb{R}^2 \mid a \leq x \leq b,\ 0 \leq y \leq f(x) \big\}.$$

**Proposition 4.12** (Charakterisierungen der Riemann-Integrierbarkeit)**.** *Sei $f \in \mathcal{F}([a,b])$ beschränkt. Folgende Bedingungen sind äquivalent:*

*(i) $f$ ist Riemann-integrierbar.*

*(ii) Es existiert höchstens eine (oder auch genau eine) reelle Zahl $I$, die die Ungleichungen*

$$\int_a^b u \,\mathrm{d}x \leq I \leq \int_a^b o \,\mathrm{d}x$$

*für alle $u, o \in \mathcal{TF}([a,b])$ mit $u \leq f \leq o$ erfüllt.*

*(iii) Für alle $\varepsilon > 0$ existieren $u, o \in \mathcal{TF}([a,b])$ mit $u \leq f \leq o$, so dass $\int_a^b (o-u) \,\mathrm{d}x < \varepsilon$.*

Der dritte Punkt in obiger Proposition bedeutet intuitiv, dass $f$ sich zwischen zwei Treppenfunktionen „einquetschen“ lässt, so dass deren Differenz im Mittel (geometrisch formuliert, der Flächeninhalt zwischen den beiden Treppenfunktionen) klein ist.

*Beweis.* Angenommen $f$ ist Riemann-integrierbar wie in (i). Wir wollen (iii) zeigen. Sei also $\varepsilon > 0$. Dann existiert (wegen der zweiten Charakterisierung des Supremums in Satz 2.59) ein $u \in \mathcal{TF}([a,b])$ mit $u \leq f$ und $\int_a^b u \,\mathrm{d}x > \underline{I}(f) - \frac{\varepsilon}{2}$. Genauso existiert ein $o \in \mathcal{TF}([a,b])$ mit $o \geq f$ und $\int_a^b o \,\mathrm{d}x < \overline{I}(f) + \frac{\varepsilon}{2}$. Da $\underline{I}(f) = \overline{I}(f)$ nach Voraussetzung folgt nun mit Lemma 4.7

$$\begin{aligned} \int_a^b (o-u) \,\mathrm{d}x &= \int_a^b o \,\mathrm{d}x - \int_a^b u \,\mathrm{d}x \\ &< \overline{I}(f) + \frac{\varepsilon}{2} - \Big( \underline{I}(f) - \frac{\varepsilon}{2} \Big) = \varepsilon \end{aligned}$$

wie in (iii) behauptet.

Angenommen $f \in \mathcal{F}([a,b])$ ist beschränkt und erfüllt die Aussage in (iii). Wir wollen (ii) zeigen und nehmen also an, dass $I_1, I_2 \in \mathbb{R}$ die Ungleichungen

$$\int_a^b u \,\mathrm{d}x \leq I_1 \leq \int_a^b o \; dx$$
$$\int_a^b u \,\mathrm{d}x \leq I_2 \leq \int_a^b o \; dx$$

für alle $u, o \in \mathcal{TF}([a,b])$ mit $u \leq f \leq o$ erfüllen. Für ein beliebiges $\varepsilon > 0$ können wir wegen (iii) $u, o \in \mathcal{TF}([a,b])$ finden, so dass die obigen Ungleichungen kombiniert zu

$$I_1 - I_2 \leq \int_a^b o \,\mathrm{d}x - \int_a^b u \,\mathrm{d}x < \varepsilon$$

und

$$I_2 - I_1 \leq \int_a^b o \,\mathrm{d}x - \int_a^b u \,\mathrm{d}x < \varepsilon$$

führen. Daher ist $|I_2 - I_1| < \varepsilon$ für alle $\varepsilon > 0$ und es muss $I_1 = I_2$ gelten. Dies zeigt, dass es höchstens eine Zahl $I \in \mathbb{R}$ gibt, die die Ungleichung in (ii) erfüllt.

Angenommen (ii) gilt. Wir behaupten, dass die Ungleichungen dann von genau einer Zahl erfüllt werden und dass $f$ Riemann-integrierbar ist. In der Tat gilt nach Gleichung (4.2), dass

$$\int_a^b u \,\mathrm{d}x \leq \sup \mathcal{U}(f) = \underline{I}(f) \leq \overline{I}(f) = \inf \mathcal{O}(f) \leq \int_a^b o \,\mathrm{d}x$$

für alle $u, o \in \mathcal{TF}([a,b])$ mit $u \leq f \leq o$. Das heisst, dass sowohl $\overline{I}(f)$ wie auch $\underline{I}(f)$ die Ungleichungen in (ii) erfüllen. Nach Voraussetzung (von (ii)) folgt $\overline{I}(f) = \underline{I}(f)$ und damit, dass $f$ Riemann-integrierbar ist.

Wir haben gesehen, dass die Implikationen (i) $\Longrightarrow$ (iii), (iii) $\Longrightarrow$ (ii) und (ii) $\Longrightarrow$ (i) gelten, also folgt die Proposition. □

**Applet 4.13** (Unter- und Obersummen)**.** *Wir sehen den Graph einer Funktion, können die betrachtete Zerlegung verfeinern (mit dem Punkt +) und dann (mit den Pfeilen) sowohl bessere Untersummen also auch besser Obersummen zu der Funktion finden. Können Sie die optimalen Unter- und Obersummen zu einer Zerlegung in 5 Intervalle finden? Nach einigen Experimenten sollten Sie davon überzeugt sein, dass die betrachtete Funktion Riemann-integrierbar ist – dies wird aus den späteren Sätzen dieses Kapitels recht schnell folgen.*

Gut zu wissen ist, dass das Riemann-Integral eine Verallgemeinerung des Integrals von Treppenfunktionen darstellt und in diesem Sinne auch einfach vom Riemann-Integral einer Treppenfunktion gesprochen werden kann.

**Übung 4.14** (Zur Wohldefiniertheit)**.** *Sei $t \in \mathcal{TF}([a,b])$ eine Treppenfunktion. Zeigen Sie, dass $t$ Riemann-integrierbar ist und dass das Riemann-Integral von $t$ gleich dem Integral von $t$ als Treppenfunktion ist.*

**Übung 4.15** (Integral der Parabelfunktion)**.** *Wiederholen Sie den Beweis von Proposition 1.1 und zeigen Sie (in der Sprache dieses Abschnitts), dass $f : x \in [0,1] \mapsto x^2 \in \mathbb{R}$ Riemann-integrierbar ist mit $\int_0^1 x^2 \,\mathrm{d}x = \frac{1}{3}$. Verifizieren Sie an dieser Stelle auch, dass*

$$\mathcal{U}(f) = \big(-\infty, \tfrac{1}{3}\big),\ \mathcal{O}(f) = \big(\tfrac{1}{3}, \infty\big).$$

Die Charakterisierung (iii) in Proposition 4.12 ist unter anderem dann nützlich, wenn man von spezifischen Funktionen die Riemann-Integrierbarkeit zeigen will. Ihre Bedingungen lassen sich sogar noch abschwächen, was wir in folgender Übung diskutieren wollen.

**Wichtige Übung 4.16** (Betrachten spezieller Ober- und Untersummen)**.** *Sei $f \in \mathcal{F}([a,b])$ eine beschränkte Funktion und sei $T_U$ eine Menge von Treppenfunktionen mit $u \leq f$ für alle $u \in T_U$ und $T_O$ eine Menge von Treppenfunktionen mit $f \leq o$ für alle $o \in T_O$. Angenommen für jedes $\varepsilon > 0$ existieren $u \in T_U$ und $o \in T_O$ mit*

$$\int_a^b (o - u) \,\mathrm{d}x < \varepsilon.$$

*Zeigen Sie, dass $f$ Riemann-integrierbar ist und*

$$\begin{aligned}\int_a^b f \,\mathrm{d}x &= \sup\left\{\int_a^b u \,\mathrm{d}x \mid u \in T_U\right\} \\ &= \inf\left\{\int_a^b o \,\mathrm{d}x \mid o \in T_O\right\}.\end{aligned}$$

**Beispiel 4.17** (Eine nicht-Riemann-integrierbare Funktion)**.** *Wir betrachten wieder die sogenannte Dirichlet-Funktion, das heisst, die charakteristische Funktion*

$$f = \mathbb{1}_{\mathbb{Q}\cap[0,1]} : [0,1] \to [0,1],\ x \mapsto \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}.$$

*Die Behauptung ist, dass diese nicht Riemann-integrierbar ist. Dazu berechnen wir das untere und das obere Integral von $f$. Sei $o \in \mathcal{TF}([0,1])$ mit $f \leq o$. Sei*

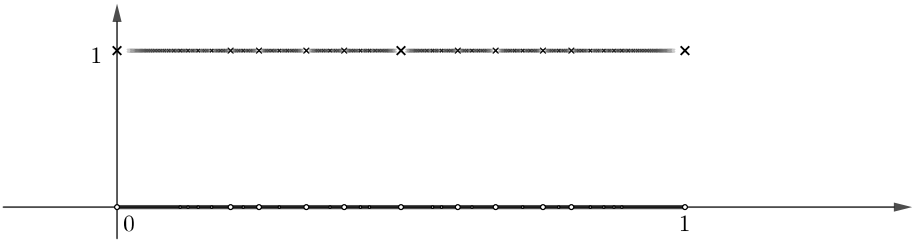
$$\mathfrak{Z} = \{0 = x_0 < x_1 < \ldots < x_n = 1\}$$

*eine Zerlegung in Konstanzintervalle von $o$. Sei $k \in \{1, \ldots, n\}$ und $c_k \in \mathbb{R}$ mit $o(x) = c_k$ für alle $x \in (x_{k-1}, x_k)$. Da $\mathbb{Q}$ dicht in $\mathbb{R}$ ist (siehe Korollar 2.70), existiert ein $x \in (x_{k-1}, x_k)$ mit $x \in \mathbb{Q}$. Wegen $f \leq o$ gilt $1 = f(x) \leq o(x) = c_k$. Somit gilt*

$$\int_0^1 o(x) \,\mathrm{d}x = \sum_{k=1}^n c_k(x_k - x_{k-1}) \geq \sum_{k=1}^n (x_k - x_{k-1}) = x_n - x_0 = 1$$

*unter Verwendung von Teleskopsummen. Damit ist das obere Integral von $f$ durch $1$ gegeben, da die Treppenfunktion mit konstantem Wert $1$ Integral $1$ hat und $o$ beliebig war. Ähnlich (siehe Übung 4.18) zeigt man, dass das untere Integral von $f$ durch $0$ gegeben ist. Somit ist $f$ nicht Riemann-integrierbar.*

*Es ist etwas schwierig den Graphen der Dirichlet-Funktion zu zeichnen (vor allem da für die meisten Computerprogramme alle Zahlen rational sind). Wir wollen dies aber trotzdem versuchen, wobei die verschiedenen Kreuze die Funktionswerte der ersten rationalen Zahlen andeuten.*




**Übung 4.18.** *Zeigen Sie, dass die Funktion $f$ aus Beispiel 4.17 unteres Integral* 0 *hat.*

## 4.3 Erste Integrationsgesetze

Wie schon zuvor betrachten wir hier Funktionen und den Begriff des Riemann-Integrals auf einem kompakten Intervall $[a, b] \subseteq \mathbb{R}$ für $a < b$. Wir möchten nun Eigenschaften des Riemann-Integrals nachweisen, die zu den Eigenschaften des Integrals von Treppenfunktionen (genauer Lemma 4.7 und Lemma 4.8) analog sind.

### 4.3.1 Linearität

**Satz 4.19** (Linearität des Riemann-Integrals)**.** *Die Menge*

$$\mathcal{R}([a, b]) = \{f \in \mathcal{F}([a, b]) \mid f \text{ ist Riemann-integrierbar}\}$$

*der Riemann-integrierbaren Funktionen auf* $[a, b]$ *bildet einen Unterraum von* $\mathcal{F}([a, b])$ *und das Integral ist eine lineare Funktion auf* $\mathcal{R}([a, b])$*. Das heisst, für* $f_1, f_2, f \in \mathcal{R}([a, b])$ *und* $s \in \mathbb{R}$ *ist* $f_1 + f_2, sf \in \mathcal{R}([a, b])$ *und*

$$\int_a^b (f_1 + f_2)(x)\,\mathrm{d}x = \int_a^b f_1(x)\,\mathrm{d}x + \int_a^b f_2(x)\,\mathrm{d}x,$$
$$\int_a^b (sf)(x)\,\mathrm{d}x = s\int_a^b f(x)\,\mathrm{d}x.$$

Im Beweis werden wir folgendes allgemeines Prinzip mehrmals anwenden. Falls $A \subseteq B$ nicht-leere Teilmengen von $\mathbb{R}$ sind und $B$ von oben beschränkt ist, dann ist $\sup(B)$ eine obere Schranke von $A$ und daher $\sup(A) \leq \sup(B)$ (nach Definition des Supremums). Analog gilt $\inf(A) \geq \inf(B)$, falls $B$ von unten beschränkt ist.

*Beweis.* Sei $f \in \mathcal{R}([a, b])$ und $s \geq 0$. Für Treppenfunktionen $u, o \in \mathcal{TF}([a, b])$ mit $u \leq f \leq o$ gilt somit $su \leq sf \leq so$. Mit $s\int_a^b u(x)\,\mathrm{d}x = \int_a^b su(x)\,\mathrm{d}x$ und $s\int_a^b o(x)\,\mathrm{d}x = \int_a^b so(x)\,\mathrm{d}x$ nach Lemma 4.7 folgt $s\mathcal{U}(f) \subseteq \mathcal{U}(sf)$ und $s\mathcal{O}(f) \subseteq \mathcal{O}(sf)$. In der Tat ist

$$\begin{aligned} s\mathcal{U}(f) &= \left\{ s\int_a^b u(x)\,\mathrm{d}x \mid u \in \mathcal{TF}([a, b]),\ u \leq f \right\} \\ &= \left\{ \int_a^b su(x)\,\mathrm{d}x \mid u \in TF([a, b],\ su \leq sf \right\} \end{aligned}$$

eine Teilmenge von $\mathcal{U}(sf)$ und analog für $s\mathcal{O}(f) \subseteq \mathcal{O}(sf)$. Aus der Bemerkung vor dem Beweis folgt also

$$\sup(s\mathcal{U}(f)) \leq \sup(\mathcal{U}(sf)) = \underline{I}(sf) \leq \overline{I}(sf) = \inf(\mathcal{O}(sf)) \leq \inf(s\mathcal{O}(f)).$$

Nach Proposition 2.62 ist jedoch

$$s\underline{I}(f) = s\sup(\mathcal{U}(f)) = \sup(s\mathcal{U}(f)) \leq \underline{I}(sf) \leq \overline{I}(sf) \leq \inf(s\mathcal{O}(f)) = s\inf(\mathcal{O}(f)) = s\overline{I}(f).$$

Da aber $f$ Riemann-integrierbar ist und somit $\underline{I}(f) = \overline{I}(f) = \int_a^b f(x)\,\mathrm{d}x$ erfüllt ist, gilt in obiger Abschätzung (wegen Gleichheit der kleinsten und der grössten Zahl) überall Gleichheit und wir schliessen

$$\underline{I}(sf) = \overline{I}(sf) = s\int_a^b f(x)\,\mathrm{d}x.$$

Damit ist $sf$ Riemann-integrierbar mit Integral $s\int_a^b f(x)\,\mathrm{d}x$. Ist $s < 0$, so kehren sich in obigem alle Abschätzungen, die $s$ beinhalten, um (zum Beispiel gilt $so \leq sf \leq su$) und man erhält vollkommen analog die gewünschte Aussage (siehe Übung 4.20).

Wir zeigen nun Additivität des Integrals. Seien also $f_1, f_2 \in \mathcal{R}([a,b])$ zwei Riemann-integrierbare Funktionen auf $[a,b]$ und $u_1, u_2, o_1, o_2 \in \mathcal{TF}([a,b])$ Treppenfunktionen mit

$$\begin{aligned} u_1 &\leq f_1 \leq o_1, \\ u_2 &\leq f_2 \leq o_2. \end{aligned}$$

Dann ist auch $u_1 + u_2 \leq f_1 + f_2 \leq o_1 + o_2$, was gemäss Lemma 4.7

$$\begin{aligned} \mathcal{U}(f_1) + \mathcal{U}(f_2) &\subseteq \mathcal{U}(f_1 + f_2), \\ \mathcal{O}(f_1) + \mathcal{O}(f_2) &\subseteq \mathcal{O}(f_1 + f_2) \end{aligned} \tag{4.3}$$

zur Folge hat. Des Weiteren gilt nach Proposition 2.63, dass

$$\begin{aligned} \sup(\mathcal{U}(f_1) + \mathcal{U}(f_2)) &= \sup(\mathcal{U}(f_1)) + \sup(\mathcal{U}(f_2)) = \underline{I}(f_1) + \underline{I}(f_2) \\ &= \int_a^b f_1(x)\,\mathrm{d}x + \int_a^b f_2(x)\,\mathrm{d}x \end{aligned}$$

nach Riemann-Integrierbarkeit von $f_1$ und $f_2$ und ebenso

$$\begin{aligned} \inf(\mathcal{O}(f_1) + \mathcal{O}(f_2)) &= \inf(\mathcal{O}(f_1)) + \inf(\mathcal{O}(f_2)) = \overline{I}(f_1) + \overline{I}(f_2) \\ &= \int_a^b f_1(x)\,\mathrm{d}x + \int_a^b f_2(x)\,\mathrm{d}x. \end{aligned}$$

Gemeinsam mit der Bemerkung vor dem Beweis ergibt sich nun wiederum

$$\begin{aligned} \int_a^b f_1(x)\,\mathrm{d}x + \int_a^b f_2(x)\,\mathrm{d}x &= \sup(\mathcal{U}(f_1) + \mathcal{U}(f_2)) \\ &\leq \sup(\mathcal{U}(f_1 + f_2)) \\ &= \underline{I}(f_1 + f_2) \leq \overline{I}(f_1 + f_2) \\ &= \inf(\mathcal{O}(f_1 + f_2)) \\ &\leq \inf(\mathcal{O}(f_1)) + \inf(\mathcal{O}(f_2)) \\ &= \int_a^b f_1(x)\,\mathrm{d}x + \int_a^b f_2(x)\,\mathrm{d}x \end{aligned}$$

Dies zeigt

$$\underline{I}(f_1 + f_2) = \overline{I}(f_1 + f_2) = \int_a^b f_1(x)\,\mathrm{d}x + \int_a^b f_2(x)\,\mathrm{d}x$$

und insbesondere Riemann-Integrierbarkeit von $f_1 + f_2$. Wir haben also die Linearität des Riemann-Integrals bewiesen. □

**Übung 4.20** (Negative Vielfache)**.** *Formulieren Sie den Fall $s < 0$ im obigen Beweis aus.*

**Übung 4.21.** *Zeigen Sie, dass Gleichheit in* (4.3) *(siehe obigen Beweis) nicht erfüllt sein muss.*

**Übung 4.22** (Ändern bei einem Punkt)**.** *Sei $f \in \mathcal{R}([a,b])$ Riemann-integrierbar. Sei $f^* \in \mathcal{F}([a,b])$ eine Funktion, die erhalten wurde, indem der Wert von $f$ an nur einem Punkt in $[a,b]$ abgeändert wurde. Zeigen Sie, dass $f^*$ Riemann-integrierbar ist und das gleiche Riemann-Integral wie $f$ hat.*

### 4.3.2 Monotonie

Für $f \in \mathcal{F}([a,b])$ definieren wir Funktionen $f^+, f^-, |f| \in \mathcal{F}([a,b])$ durch

$$f^+(x) = \max\{0, f(x)\}\,,\ f^-(x) = \max\{0, -f(x)\}\,,\ |f|(x) = \max\{f(x), -f(x)\} = |f(x)|$$

für $x \in [a,b]$. Die Funktion $f^+$ ist der **Positivteil** von $f$, $f^-$ ist der **Negativteil** von $f$ und $|f|$ ist der **Absolutbetrag** von $f$.

**Übung 4.23** (Eigenschaften vom Positiv- und Negativteil)**.** *Sei $f \in \mathcal{F}([a,b])$. Zeigen Sie die Gleichungen*

$$f = f^+ - f^-, \quad |f| = f^+ + f^-, \quad f^+ = \frac{|f| + f}{2}, \quad f^- = \frac{|f| - f}{2}.$$

**Satz 4.24** (Monotonie des Riemann-Integrals)**.** *Für zwei Funktionen $f_1, f_2 \in \mathcal{R}([a,b])$ gelten folgende Monotonie-Eigenschaften des Riemann-Integrals:*

*(i) Falls $f_1 \geq 0$ ist, so gilt $\int_a^b f(x)\,\mathrm{d}x \geq 0$.*

*(ii) Falls $f_1 \leq f_2$ ist, so gilt $\int_a^b f_1(x)\,\mathrm{d}x \leq \int_a^b f_2(x)\,\mathrm{d}x$.*

*(iii) Die Funktion $|f_1|$ ist Riemann-integrierbar auf $[a,b]$ und es gilt die **Dreiecksungleichung***

$$\left|\int_a^b f_1(x)\,\mathrm{d}x\right| \leq \int_a^b |f_1(x)|\,\mathrm{d}x.$$

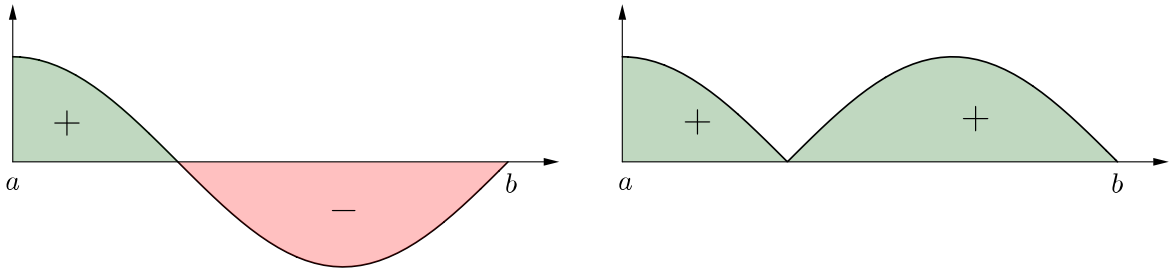
Wir möchten kurz erklären, wieso sich die Ungleichung in Punkt (iii) des obigen Satzes Dreiecksungleichung nennt. Tatsächlich sieht man kein Dreieck, im Gegensatz zur Dreiecksungleichung

$$\left|z_1 + z_2\right| \leq |z_1| + |z_2|$$

für $z_1, z_2 \in \mathbb{C}$, die geometrisch direkt begründet werden kann (wie?). Es gilt auch die verallgemeinerte Dreiecksungleichung

$$\left|\sum_{i=1}^{n} z_i\right| \leq \sum_{i=1}^{n} |z_i|$$

für $z_1, \ldots, z_n \in \mathbb{C}$, wie man direkt aus der Dreiecksungleichung und vollständiger Induktion folgern kann (siehe Übung 3.4). Die Aussage (iii) in Satz 4.24 ist eine „kontinuierliche Version“ der verallgemeinerten Dreiecksungleichung, weswegen wir von der Dreiecksungleichung für das Riemann-Integral sprechen.




Figur 4.3: Wir sehen hier den Graphen einer Funktion $f$ links und der entsprechenden Funkton $|f|$ rechts. Dabei stellt $\int_a^b f(x)\,\mathrm{d}x$ einen Nettoflächeninhalt und $\int_a^b |f(x)|\,\mathrm{d}x$ einen Flächeninhalt dar.

*Beweis.* Für $f_1 \geq 0$ wie in (i) ist die konstante Funktion $u = 0$ eine Treppenfunktion mit $u \leq f_1$ und

$$0 = \int_a^b u(x)\,\mathrm{d}x \leq \sup(\mathcal{U}(f_1)) = \underline{I}(f_1) = \int_a^b f(x)\,\mathrm{d}x$$

folgt.

Falls $f_1 \leq f_2$ wie in (ii) gilt, so ist $f_2 - f_1 \geq 0$ und

$$\int_a^b f_2(x)\,\mathrm{d}x - \int_a^b f_1(x)\,\mathrm{d}x = \int_a^b f_2(x) - f_1(x)\,\mathrm{d}x \geq 0$$

nach Linearität des Riemann-Integrals (Satz 4.19) und Teil (i). Dies zeigt (ii).

Für (iii) wollen wir zuerst zeigen, dass für ein $f \in \mathcal{R}([a,b])$ auch $f^+$ Riemann-integrierbar ist. Dazu bemerken wir zuerst, dass für $s, t \in \mathbb{R}$ die Ungleichung $s \leq t$ impliziert, dass

$$s^+ = \max\{0, s\} \leq t^+ = \max\{0, t\} \text{ und}$$
$$t^+ - s^+ \leq t - s.$$

Dies ergibt sich aus der Unterscheidung der Fälle $s \leq t \leq 0$, $s \leq 0 < t$ und $0 < s \leq t$. (Wieso?). Da $f$ Riemann-integrierbar ist, gibt es nach Proposition 4.12 (iii) zu jedem $\varepsilon > 0$ zwei Treppenfunktion $u, o \in \mathcal{TF}([a,b])$ mit $u \leq f \leq o$ und $\int_a^b (o-u)(x)\,\mathrm{d}x < \varepsilon$. Verknüpfen wir diese mit der Funktion $t \in \mathbb{R} \mapsto t^+ \in \mathbb{R}$, so ergibt sich

$$u^+ \leq f^+ \leq o^+, \quad o^+ - u^+ \leq o - u$$

und daher nach (ii) auch

$$\int_a^b (o^+ - u^+)(x)\,\mathrm{d}x \leq \int_a^b (o-u)(x)\,\mathrm{d}x < \varepsilon.$$

Allerdings sind $u^+, o^+$ wieder Treppenfunktionen. Nach der dritten Charakterisierung in Proposition 4.12 ergibt sich somit, dass $f^+$ Riemann-integrierbar ist, da $\varepsilon > 0$ beliebig war.

Mittels Satz 4.19 erhalten wir, dass $|f| = 2f^+ - f$ auch Riemann-integrierbar ist. Aus $f \leq |f|$ und $-f \leq |f|$ folgt aus (ii) nun

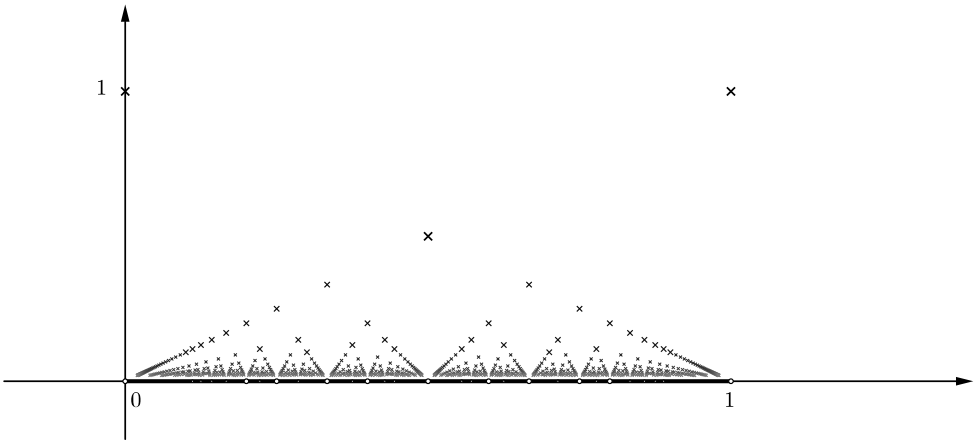
$$\int_a^b f(x)\,\mathrm{d}x \leq \int_a^b |f(x)|\,\mathrm{d}x, \quad \int_a^b -f(x)\,\mathrm{d}x \leq \int_a^b |f(x)|\,\mathrm{d}x,$$

was zur Dreiecksungleichung äquivalent ist. □

**Übung 4.25** (Modifizierte Dirichlet- oder Riemann-Funktion)**.** *Zeigen Sie, dass die Funktion*

$$g : [0,1] \to [0,1],\ x \mapsto \begin{cases} 0 & \textit{falls } x \textit{ irrational} \\ \frac{1}{q} & \textit{falls } x = \frac{p}{q} \textit{ mit } p, q \textit{ teilerfremd} \end{cases}$$

*Riemann-integrierbar ist. Als Hilfestellung stellen wir den Graphen dar, aber überlassen Ihnen die Interpretation des Graphen und die sich daraus ergebenden Überlegungen.*




### 4.3.3 Teilintervalle

Es seien $a < b < c$ drei reelle Zahlen. Dann definiert eine Funktion $f$ auf dem Intervall $[a,c]$ die Funktion $f_1 = f|_{[a,b]}$ auf $[a,b]$ und die Funktion $f_2 = f|_{[b,c]}$ auf $[b,c]$. Dabei gilt

$f_1(b) = f(b) = f_2(b)$. Umgekehrt können wir Funktionen $f_1 \in \mathcal{F}([a,b])$ und $f_2 \in \mathcal{F}([b,c])$ mit $f_1(b) = f_2(b)$ verwenden, um eine Funktion $f \in \mathcal{F}([a,c])$ durch

$$f(x) = \begin{cases} f_1(x) & \text{falls } x \in [a,b] \\ f_2(x) & \text{falls } x \in (b,c] \end{cases}$$

für $x \in [a,c]$ zu definieren. In diesem Sinne entspricht die Funktion $f \in \mathcal{F}([a,c])$ zwei Funktionen $f_1 \in \mathcal{F}([a,b])$, $f_2 \in \mathcal{F}([b,c])$ mit $f_1(b) = f_2(b)$.

**Satz 4.26** (Additionseigenschaft bezüglich Intervallen)**.** *Unter Verwendung obiger Notation gilt, dass $f \in \mathcal{F}([a,c])$ genau dann Riemann-integrierbar ist, wenn $f_1$ und $f_2$ Riemann-integrierbar sind. In diesem Fall ist*

$$\int_a^c f(x)\,\mathrm{d}x = \int_a^b f_1(x)\,\mathrm{d}x + \int_b^c f_2(x)\,\mathrm{d}x.$$

*Beweis.* Wir verifizieren zuerst die behauptete Formel für Treppenfunktionen. Dazu betrachten wir eine Treppenfunktion $t$ auf $[a,c]$ und eine Zerlegung in Konstanzintervalle von $t$

$$\mathfrak{Z} = \{a = x_0 < x_1 < \ldots < x_n = c\}.$$

Dabei dürfen wir wegen Lemma 4.5 ohne Beschränkung der Allgemeinheit annehmen, dass $x_m = b$ für ein $m \in \{1,\ldots,n-1\}$. Für $k \in \{1,\ldots,n\}$ sei $c_k$ der Konstanzwert von $t$ auf $(x_{k-1}, x_k)$. Dann gilt

$$\begin{aligned} \int_a^c t(x)\,\mathrm{d}x &= \sum_{k=1}^n c_k \Delta x_k && (4.4) \\ &= \sum_{k=1}^m c_k \Delta x_k + \sum_{k=m+1}^n c_k \Delta x_k \\ &= \int_a^b t|_{[a,b]}(x)\,\mathrm{d}x + \int_b^c t|_{[b,c]}(x)\,\mathrm{d}x && (4.5) \end{aligned}$$

Sei $f \in \mathcal{F}([a,c])$ eine Funktion und definiere $f_1 = f|_{[a,b]}$, $f_2 = f|_{[b,c]}$. Gegeben $u \in \mathcal{TF}([a,c])$ mit $u \le f$ kann man ebenso $u_1 = u|_{[a,b]}$, $u_2 = u|_{[b,c]}$ definieren. Es gilt $u_1 \le f_1$ und $u_2 \le f_2$. Wegen Gleichung (4.4) erhalten wir, dass

$$\int_a^c u(x)\,\mathrm{d}x = \int_a^b u_1(x)\,\mathrm{d}x + \int_b^c u_2(x)\,\mathrm{d}x, \tag{4.6}$$

was wiederum $\mathcal{U}(f) \subseteq \mathcal{U}(f_1) + \mathcal{U}(f_2)$ zur Folge hat. Umgekehrt kann man, gegeben Treppenfunktionen $u_1, u_2$ mit $u_1 \le f_1$, $u_2 \le f_2$ eine Treppenfunktion $u$ auf $[a,c]$ definieren, die ebenso $u \le f$ und Gleichung (4.6) erfüllt. (Wie?). Dadurch ist

$$\mathcal{U}(f) = \mathcal{U}(f_1) + \mathcal{U}(f_2)$$

und wegen der Additionseigenschaft des Supremums in Proposition 2.63 gilt

$$\underline{I}(f) = \underline{I}(f_1) + \underline{I}(f_2). \tag{4.7}$$

Analog zeigt man, dass

$$\overline{I}(f) = \overline{I}(f_1) + \overline{I}(f_2). \tag{4.8}$$

Ist nun $f$ Riemann-integrierbar, dann ist

$$\underline{I}(f_1) + \underline{I}(f_2) = \underline{I}(f) = \overline{I}(f) = \overline{I}(f_1) + \overline{I}(f_2) \geq \overline{I}(f_1) + \underline{I}(f_2) \geq \underline{I}(f_1) + \underline{I}(f_2).$$

Überall in dieser Kette von Ungleichungen gilt also Gleichheit. Somit ist $\overline{I}(f_1) = \underline{I}(f_1)$ und dadurch auch $\overline{I}(f_2) = \underline{I}(f_2)$. Das heisst, dass $f_1$ und $f_2$ Riemann-integrierbar sind und Gleichung (4.7) wird zur gewünschten Additionseigenschaft für das Riemann-Integral.

Falls $f_1, f_2$ Riemann-integrierbar sind, dann gilt $\overline{I}(f_1) = \underline{I}(f_1)$ und $\overline{I}(f_2) = \underline{I}(f_2)$. Dies impliziert gemeinsam mit den Gleichungen (4.7), (4.8) auch $\underline{I}(f) = \overline{I}(f)$ und die Additionseigenschaft. □

Sei $[a, b]$ ein kompaktes Intervall mit $a < b$. Ist $f$ eine Funktion, die auf einer grösseren Menge als $[a, b]$ definiert ist, so werden wir anstelle von $\int_a^b f|_{[a,b]}\,\mathrm{d}x$ trotzdem meist $\int_a^b f\,\mathrm{d}x$ schreiben, wenn $f|_{[a,b]}$ Riemann-integrierbar ist. Auch definieren wir

$$\int_b^a f\,\mathrm{d}x = -\int_a^b f\,\mathrm{d}x \quad \text{und} \quad \int_a^a f\,\mathrm{d}x = 0. \tag{4.9}$$

Diese Definition vereinfacht die Notation und macht auf Grund der Aussage in folgender Übung Sinn.

**Wichtige Übung 4.27** (Intervalladditivität)**.** *Sei $I = [a_0, b_0]$ für $a_0 < b_0$ ein kompaktes Intervall und sei $f \in \mathcal{R}(I)$. Zeigen Sie die Additionseigenschaft in Satz 4.26 für alle $a, b, c \in I$.*

VO: 01.11.17

**Übung 4.28** (Stetigkeit des partikulären Integrals)**.** *Sei $a < b$ und $f : [a, b] \to \mathbb{R}$ eine Riemann-integrierbare Funktion. Zeigen Sie, dass das sogenannte partikuläre Integral*

$$x \in [a, b] \mapsto \int_a^x f(t)\,\mathrm{d}t$$

*eine stetige reellwertige Funktion auf $[a, b]$ definiert. Ist diese Funktion auch gleichmässig oder Lipschitz-stetig (siehe Übung 3.81 für letzteren Begriff)? Wir visualisieren dies auch in folgendem App.*

## 4.4 Anwendungen

### 4.4.1 Intervallfunktionen

Wir möchten nun spezielle Abbildungen auf der Menge der Teilintervalle eines Intervalles betrachten, wobei wir Ordnungsvertauschungen im Stile von (4.9) zulassen wollen. Genauer untersuchen wir folgenden Begriff.

**Definition 4.29.** Seien $a \leq b$ in $\mathbb{R}$ und sei $\mathcal{I} : (\alpha, \beta) \in [a, b]^2 \mapsto \mathcal{I}(\alpha, \beta) \in \mathbb{R}$ eine Funktion. Wir nennen $\mathcal{I}$ eine **additive Intervallfunktion** auf $[a, b]$, falls

(i) Für alle $\alpha \in [a, b]$ gilt $\mathcal{I}(\alpha, \alpha) = 0$.

(ii) Für alle $\alpha, \beta \in [a, b]$ gilt $\mathcal{I}(\alpha, \beta) = -\mathcal{I}(\beta, \alpha)$.

(iii) Für alle $\alpha, \beta, \gamma \in [a, b]$ mit $\mathcal{I}(\alpha, \beta) + \mathcal{I}(\beta, \gamma) = \mathcal{I}(\alpha, \gamma)$.

Wir wollen hier kurz erklären, woher die Bezeichnung „additive Intervallfunktion“ stammt. Ist $\mathcal{I}$ eine additive Intervallfunktion auf einem kompakten Intervall $[a, b]$, so kann man eine reellwertige Funktion $\mathcal{J}$ auf der Menge der nicht-leeren Teilintervalle von $[a, b]$ durch $\mathcal{J}([\alpha, \beta]) = \mathcal{I}(\alpha, \beta)$ für $[\alpha, \beta] \subseteq [a, b]$ definieren. Diese hat die Eigenschaften

$$\mathcal{J}([\alpha, \alpha]) = 0, \quad \mathcal{J}([\alpha, \beta] \cup [\beta, \gamma]) = \mathcal{J}([\alpha, \beta]) + \mathcal{J}([\beta, \gamma]) \tag{4.10}$$

für alle $\alpha \leq \beta \leq \gamma$ in $[a, b]$ (wieso?). Vor allem letztere Eigenschaft begründet die Bezeichnung „additive Intervallfunktion“.

Hat man umgekehrt eine reellwertige Funktion $\mathcal{J}$ auf der Menge der nicht-leeren Teilintervalle von $[a, b]$ gegeben, die (4.10) genügt, so definiert $\mathcal{I}(\alpha, \beta) = \mathcal{J}([\alpha, \beta])$ für $\alpha \leq \beta$ und $\mathcal{I}(\alpha, \beta) = -\mathcal{J}([\beta, \alpha])$ für $\alpha > \beta$ eine additive Intervallfunktion $\mathcal{I}$ auf $[a, b]$ (wieso?).

Somit haben wir also zwei Arten, wie wir uns additive Intervallfunktionen vorstellen können. Eine grosse Kollektion von Beispielen erhält man mit Satz 4.26 und Übung 4.27, nach welchen die Abbildung

$$\mathcal{I} : (\alpha, \beta) \in [a, b]^2 \mapsto \int_\alpha^\beta f(x)\,\mathrm{d}x \tag{4.11}$$

für jede Riemann-integrierbare Funktion $f : [a, b] \to \mathbb{R}$ eine additive Intervallfunktion ist. Die folgende Proposition charakterisiert derartige additive Intervallfunktionen.

**Proposition 4.30.** *Seien $a < b$ in $\mathbb{R}$, $f : [a, b] \to \mathbb{R}$ eine Riemann-integrierbare Funktion und $\mathcal{I}$ eine additive Intervallfunktion auf $[a, b]$. Angenommen es gilt*

$$(\beta - \alpha) \inf_{x \in [\alpha, \beta]} f(x) \leq \mathcal{I}(\alpha, \beta) \leq (\beta - \alpha) \sup_{x \in [\alpha, \beta]} f(x) \tag{4.12}$$

*für alle $\alpha < \beta$ in $[a, b]$. Dann ist*

$$\mathcal{I}(\alpha, \beta) = \int_\alpha^\beta f(x)\,\mathrm{d}x$$

*für alle* $\alpha, \beta \in [a, b]$.

Wir möchten anmerken, dass jedoch nicht alle additiven Intervallfunktionen von der Form in (4.11) sein müssen.

*Beweis.* Seien $\alpha < \beta$ in $[a, b]$ und sei $\varepsilon \in \left(0, \frac{b-a}{2}\right)$. Auf Grund der Annahme in (4.12) gilt

$$\begin{aligned}\mathcal{I}(\alpha, \beta) &= \mathcal{I}(\alpha, \alpha + \varepsilon) + \mathcal{I}(\alpha + \varepsilon, \beta - \varepsilon) + \mathcal{I}(\beta - \varepsilon, \beta) \\ &\geq \varepsilon \inf_{x \in [\alpha, \alpha+\varepsilon]} f(x) + (\beta - \alpha - 2\varepsilon) \inf_{x \in [\alpha+\varepsilon, \beta-\varepsilon]} f(x) + \varepsilon \inf_{x \in [\beta-\varepsilon, \beta]} f(x) \\ &\geq 2\varepsilon \inf_{x \in [\alpha, \beta]} f(x) + (\beta - \alpha - 2\varepsilon) \inf_{x \in (\alpha, \beta)} f(x).\end{aligned}$$

Da $\varepsilon \in \left(0, \frac{b-a}{2}\right)$ beliebig war, ergibt sich daraus die erste Ungleichung in

$$(\beta - \alpha) \inf_{x \in (\alpha, \beta)} f(x) \leq \mathcal{I}(\alpha, \beta) \leq (\beta - \alpha) \sup_{x \in (\alpha, \beta)} f(x).$$

(Wieso?) Die zweite Ungleichung ergibt sich analog zu obigem aus der zweiten Ungleichung in (4.12).

Sei nun $u \leq f$ eine Treppenfunktion auf $[a, b]$ mit Zerlegung $\mathfrak{Z} = \{a = x_0 < \ldots < x_n = \beta\}$ in Konstanzintervalle von $u$. Seien $c_1, \ldots, c_n$ die Konstanzwerte von $u$ bezüglich $\mathfrak{Z}$. Auf Grund der Annahme $u \leq f$ folgt $c_k \leq \inf_{x \in (x_{k-1}, x_k)} f(x)$ für alle $k \in \{1, \ldots, n\}$. Unter Verwendung der Additivität von $\mathcal{I}$ erhält man damit für die Untersumme

$$\begin{aligned}\int_\alpha^\beta u(x)\,\mathrm{d}x = \sum_{k=1}^n c_k(x_k - x_{k-1}) &\leq \sum_{k=1}^n (x_k - x_{k-1}) \inf_{x \in (x_{k-1}, x_k)} f(x) \\ &\leq \sum_{k=1}^n \mathcal{I}(x_{k-1}, x_k) = \mathcal{I}(\alpha, \beta)\end{aligned}$$

Ebenso ergibt sich $\mathcal{I}(\alpha, \beta) \leq \int_\alpha^\beta o(x)\,\mathrm{d}x$ für jede Treppenfunktion $o$ mit $f \leq o$. Daher gelten für das untere Integral $\underline{I}$ und das obere Integral $\overline{I}$ von $f$ über $[\alpha, \beta]$ die Ungleichungen

$$\underline{I} \leq \mathcal{I}(\alpha, \beta) \leq \overline{I}.$$

Da $f$ aber Riemann-integrierbar ist, gilt $\underline{I} = \overline{I}$ und somit $\mathcal{I}(\alpha, \beta) = \int_\alpha^\beta f(x)\,\mathrm{d}x$. □

Man kann Proposition 4.30 als Wegweiser verwenden, um verschiedene Interpretationen des Riemann-Integrals zu finden. Formal gesehen sind diese Anwendungen jeweils Definitionen.

### 4.4.2 Flächeninhalt

Die einfachste Anwendung von Proposition 4.30 ist die Interpretation von $\int_a^b f(x)\,\mathrm{d}x$ als **Flächeninhalt des Gebietes**

$$\left\{(x, y) \in \mathbb{R}^2 \mid a \leq x \leq b,\ 0 \leq y \leq f(x)\right\}$$

unter dem Graphen einer Riemann-integrierbaren Funktion $f : [a, b] \to \mathbb{R}_{\geq 0}$. Die Argumentation, die zu dieser Definition führt, haben wir bereits in Abschnitt 1.1 besprochen. Formal gesehen erachten wir $\int_a^b f(x)\,\mathrm{d}x$ als Definition des Flächeninhalts des obigen Gebietes.

### 4.4.3 Masse, Momente und Schwerpunkt

Es gibt natürlich auch viele physikalische Beispiele für die Bedeutung des Riemann-Integrals. Sei zum Beispiel $a < b$ und sei $\rho(x)$ die Dichte eines Stabes (in Kilogramm pro Meter, $\frac{kg}{m}$) bei der Koordinate $x \in [a, b]$. Dann ergibt sich aus Proposition 4.30, dass wir $m = \int_a^b \rho(x)\,\mathrm{d}x$ als das **Gesamtmasse** (in $kg$) interpretieren sollten. (Wieso?)

Wir erinnern daran, dass bei einem Hebel das Moment (in $Nm$) einer Krafteinwirkung durch das Produkt der Krafteinwirkung (in Newton $N$) und des Weges (in $m$) definiert ist. Wir stellen uns vor, dass $a = 0 < b$, der obige Stab mit Dichtefunktion $\rho$ im Ursprung fixiert ist, und die Schwerkraft (mit Gravitationskonstante $g$ in $N/kg$) auf den Stab einwirkt. In diesem Fall ergibt sich für $\alpha < \beta$ in $[a, b]$, dass die dem Teilintervall $[\alpha, \beta]$ entsprechende Masse $m(\alpha, \beta)$ die Ungleichung

$$\inf\{\rho(x) \mid x \in [\alpha, \beta]\}\,(\beta - \alpha) \leq m(\alpha, \beta) \leq \sup\{\rho(x) \mid x \in [\alpha, \beta]\}\,(\beta - \alpha)$$

erfüllt, woraus sich für das entsprechende Moment $M(\alpha, \beta)$ die Ungleichung

$$g\alpha \inf\{\rho(x) \mid x \in [\alpha, \beta]\}\,(\beta - \alpha) \leq M(\alpha, \beta) \leq g\beta \sup\{\rho(x) \mid x \in [\alpha, \beta]\}\,(\beta - \alpha)$$

ergibt. Diese Eigenschaft von $M$ unterscheidet sich zwar formal von (4.12) doch lässt sich mit Hilfe der Stetigkeit von $x \in [a, b] \mapsto x$ der Beweis von Proposition 4.30 anpassen. Ebenso ist es physikalisch sinnvoll die Additivität dieser Momentfunktion anzunehmen, dadurch erhalten wir die Definition

$$M(a, b) = \int_a^b \rho(x) g x\,\mathrm{d}x$$

für das **Gesamtmoment** des Stabes.

Der **Schwerpunkt** des Stabes ist definiert als die $x$-Koordinate $x_0$, so dass eine Punktmasse bei $x_0$ mit derselben Masse wie der Stab auch dasselbe Moment besitzt. Also ist

$$x_0 = \frac{M(a, b)}{mg} = \frac{1}{m} \int_a^b \rho(x) x\,\mathrm{d}x.$$

der Schwerpunkt des Stabes.

Die Annahme $a = 0$ ist für diese Diskussion (abgesehen von der Vorstellung dass der Stab am Ursprung gehalten wird) nicht notwendig, falls $a < 0 < b$ erhalten wir physikalisch sinnvolle Integrale von Funktionen, die sowohl positive als auch negative Werte annehmen können.

### 4.4.4 Geleistete Arbeit

Wenn $a < b$ ist und $f : [a, b] \to \mathbb{R}$ eine Riemann-integrierbare Funktion ist, die zu einem Zeitpunkt $t$ den Energieverbrauch $f(t)$ (in Watt $W$) zum Beispiel in Form elektrischer Energie

eines Hauses angibt, so beschreibt $\int_a^b f(t)\,\mathrm{d}t$ die verbrauchte **Energie** oder vom Stromnetz eingespeiste **Arbeit** (in Joule $J = Ws$) zwischen den Zeitpunkten $t = a$ und $t = b$ (in Sekunden $s$). Diese Interpretation ergibt sich wiederum aus Proposition 4.30 und der Definition, dass Arbeit gleich Leistung mal Zeitdauer ist. Hier ist es ebenso physikalisch sinnvoll, Funktionen mit positiven und negativen Werten zuzulassen, wenn zum Beispiel das Hausdach mit einer Solaranlage ausgestattet ist, die bei Schönwetter etwaige Energieüberschüsse des Hauses ins Stromnetz zurückspeist. Das Vorzeichen des Integrals entscheidet in diesem Fall, ob insgesamt innerhalb der Zeitspanne $[a, b]$ das Haus ein Energieverbraucher oder Energielieferant war.

### 4.4.5 Vorteil des Integralbegriffs

Wir haben das Integral abstrakt mittels der Definition 4.6 des Integrals einer Treppenfunktion und der Definition 4.11 des Integrals einer Riemann-integrierbaren Funktion eingeführt. Bei Besprechung dieser Definitionen haben wir uns zwar von einer geometrischen Interpretation des Integrals als (vorzeichenbehafteter) Flächeninhalt leiten lassen, doch war diese Vorstellung formal nicht notwendig für unsere Diskussionen. Wir hoffen, dass der Vorteil dieses abstrakten Zugang nun ersichtlich ist: Das Integral hat je nach Zusammenhang verschiedene (zum Beispiel physikalische) Bedeutungen. Wenn unsere Definition des Integrals „der Flächeninhalt unter der Kurve“ gewesen wäre, dann wäre es nicht klar, was genau der Zusammenahng zwischen einem Flächeninhalt und einer Momentberechnung sein sollte.[1] In diesem Sinne ist unser abstrakter Zugang nicht Selbstzweck, sondern geradezu notwendig auf Grund der vielfältigen Anwendungen des Integralbegriffs.

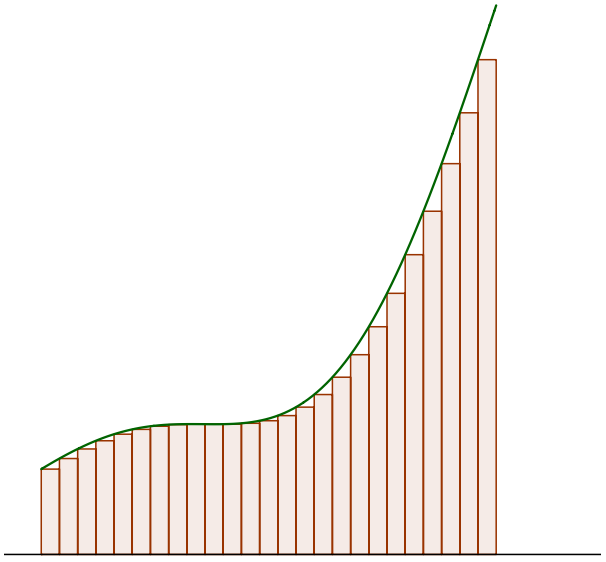
[1] Des Weiteren wäre diese Definition zirkulär gewesen, da wir ohne Definition des Integrals keine Definition des Flächeninhalts unter der Kurve haben.

## 4.5 Integrierbarkeit monotoner Funktionen

Wir betrachten wie zuvor ein kompaktes Intervall $[a, b] \subseteq \mathbb{R}$ für reelle Zahlen $a, b$ mit $a < b$.

**Satz 4.31** (Integrierbarkeit monotoner Funktionen)**.** *Jede monotone Funktion in $\mathcal{F}([a, b])$ ist Riemann-integrierbar.*

*Beweis.* Ohne Beschränkung der Allgemeinheit können wir eine monoton wachsende Funktion $f \in \mathcal{F}([a, b])$ betrachten (ansonsten ersetzt man $f$ mit $-f$ und wendet Satz 4.19 an). Wir möchten die dritte Charakterisierung in Proposition 4.12 anwenden. Das heisst, wir wollen für ein gegebenes $\varepsilon > 0$ zwei Treppenfunktionen $u, o \in \mathcal{TF}([a, b])$ finden, so dass $u \leq f \leq o$ und $\int_a^b (o - u)(x)\,\mathrm{d}x < \varepsilon$ gilt.



Wir konstruieren $u$ und $o$ mittels einer natürlichen Zahl $n \in \mathbb{N}$ (die wir später wählen werden) und der Zerlegung

$$\mathfrak{Z} = \{a = x_0 < x_1 < \ldots < x_n = b\}$$

von $[a, b]$ gegeben durch $x_k = a + \frac{b-a}{n}k$ für $k \in \{0, \ldots, n\}$. Seien $u, o$ gegeben durch

$$u(x) = \begin{cases} f(x_{k-1}) & \text{falls } x \in [x_{k-1}, x_k) \text{ für ein } k \in \{1, \ldots, n\} \\ f(b) & \text{falls } x = b \end{cases}$$

respektive

$$o(x) = \begin{cases} f(a) & \text{falls } x = a \\ f(x_k) & \text{falls } x \in (x_{k-1}, x_k] \text{ für ein } k \in \{1, \ldots, n\} \end{cases}$$

für alle $x \in [a, b]$. Da $f$ monoton wachsend ist, gilt $u \leq f \leq o$. In der Tat ist für $x \in [a, b]$ entweder $x = b$, womit $u(x) = f(x)$, oder es gibt ein $k \in \{1, \ldots, n\}$ mit $x \in [x_{k-1}, x_k)$.

In letzterem Fall erhalten wir $u(x) = f(x_{k-1}) \leq f(x)$ und somit gilt $u \leq f$. Ein analoges Argument liefert $f \leq o$. Des Weiteren gilt

$$\begin{aligned}\int_a^b (o-u)(x)\,\mathrm{d}x &= \sum_{k=1}^n (f(x_k) - f(x_{k-1}))(x_k - x_{k-1}) = \sum_{k=1}^n (f(x_k) - f(x_{k-1}))\frac{b-a}{n} \\ &= \frac{b-a}{n}\sum_{k=1}^n (f(x_k) - f(x_{k-1})) = \frac{b-a}{n}(f(b) - f(a))\end{aligned}$$

nach Vereinfachen der Teleskopsumme. Nach dem Archimedischen Prinzip können wir nun ein $n$ wählen, so dass $\int_a^b (o-u)(x)\,\mathrm{d}x < \varepsilon$ ist. Aus Proposition 4.12 (iii) folgt somit, dass $f$ Riemann-integrierbar ist. □

**Übung 4.32** (Kreisfunktion)**.** *Zeigen Sie, dass die Funktion $x \in [0,1] \mapsto \sqrt{1-x^2} \in \mathbb{R}$ Riemann-integrierbar ist.*

Mit Hilfe der Additionseigenschaft in Satz 4.26 lässt sich die Aussage von Satz 4.31 auf Funktionen erweitern, die nur stückweise monoton sind.

**Definition 4.33** (Stückweise Monotonie)**.** Sei $I = [a,b]$ ein abgeschlossenes, beschränktes Intervall mit $a < b$. Eine Funktion $f \in \mathcal{F}([a,b])$ heisst **stückweise monoton**, falls es eine Zerlegung

$$\mathfrak{Z} = \{a = x_0 < x_1 < \ldots < x_n = b\}$$

von $[a,b]$ gibt, so dass $f|_{(x_{k-1},x_k)}$ monoton ist für alle $k \in \{1,\ldots,n\}$.

Jede monotone Funktion ist stückweise monoton (man braucht dazu nur die Zerlegung $\mathfrak{Z} = \{a < b\}$ zu betrachten). Ein besseres Beispiel einer stückweise monotonen Funktion ist das Polynom $x^2$ auf einem Intervall $I = [a,b]$ für $a < b$. Falls $a \geq 0$ oder $b \leq 0$ ist diese, wie wir schon wissen, monoton. Gilt $a < 0 < b$, so betrachtet man die Zerlegung $\mathfrak{Z} = \{a < 0 < b\}$ und sieht, dass $x^2$ auf den beiden Abschnitten $(a,0)$ und $(0,b)$ monoton ist. Genau gleich sieht man, dass alle Monome stückweise monoton sind.

**Korollar 4.34** (Riemann-Integrierbarkeit von stückweise monotonen Funktionen)**.** *Sei $I = [a,b]$ ein kompaktes Intervall mit $a < b$. Jede stückweise monotone, beschränkte Funktion in $\mathcal{F}([a,b])$ ist Riemann-integrierbar.*

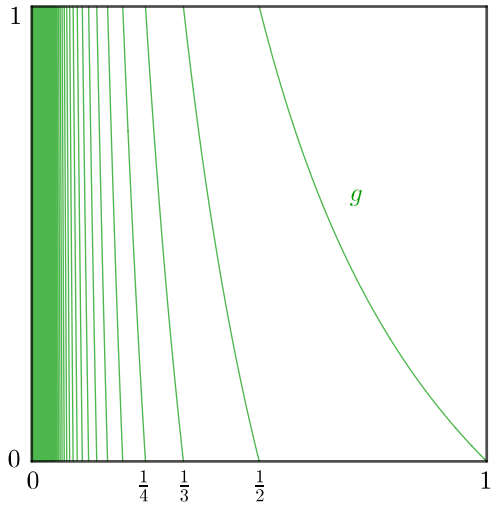
**Übung 4.35** (Riemann-Integrierbarkeit von stückweise monotonen Funktionen)**.** *Beweisen Sie Korollar 4.34 unter Verwendung der Sätze 4.31 und 4.26 und der Übung 4.22.*

Insbesondere sind also alle Monome auf einem abgeschlossenen, beschränkten, nicht-leeren Intervall $I$ Riemann-integrierbar. Mit der Linearität des Riemann-Integrals folgt nun mittels vollständiger Induktion, dass alle Polynome auf $I$ Riemann-integrierbar sind.

**Übung 4.36** (Gauss-Abbildung)**.** *Zeigen Sie, dass die sogenannte Gauss-Abbildung*

$$g : x \in [0,1] \mapsto \begin{cases} \left\{\frac{1}{x}\right\} & \textit{falls } x > 0 \\ 0 & \textit{falls } x = 0 \end{cases}$$

*Riemann-integrierbar ist, wobei $\{\cdot\}$ den gebrochenen Anteil bezeichnet (siehe Abschnitt 2.6.1).*



## 4.6 Integration von Polynomen

Wir betrachten wiederum ein Intervall $[a, b]$ mit Endpunkten $a < b$.

**Satz 4.37** (Riemann-Integrierbarkeit von Polynomen)**.** *Die Einschränkung einer reellen Polynomfunktion auf $[a, b]$ ist Riemann-integrierbar. Für alle Monome $x^d$ mit $d \in \mathbb{N}_0$ gilt*

$$\int_a^b x^d \,\mathrm{d}x = \frac{1}{d+1}\left(b^{d+1} - a^{d+1}\right).$$

*Beweis.* Dass Polynomfunktionen eingeschränkt auf $[a, b]$ Riemann-integrierbar sind, folgt, wie schon diskutiert, aus der Linearität des Riemann-Integrals (Satz 4.19) und der Riemann-Integrierbarkeit von stückweise monotonen Funktionen (Korollar 4.34). Die zweite Aussage behandeln wir hier nur im Spezialfall $0 = a < b$. Der Spezialfall $a < b = 0$ ist analog und die allgemeine Aussage ergibt sich aus diesen beiden Spezialfällen und Satz 4.26 (siehe Übung 4.38).

Da $x \in [0, b] \mapsto x^d \in \mathbb{R}$ monoton wachsend ist, können wir dieselbe Methode wie im Beweis von Satz 4.24 (und daher auch wie in Proposition 1.1) verwenden. Sei also $n \in \mathbb{N}$ und $u, o$ Treppenfunktionen auf $[0, b]$ mit Zerlegung in Konstanzintervalle

$$\mathfrak{Z} = \{0 = x_0 < x_1 < \ldots < x_n\},$$

wobei $x_k = \frac{k}{n}b$ für $k \in \{1, \ldots, n\}$, und Konstanzwert $x_{k-1}^d$ respektive $x_k^d$ auf $(x_{k-1}, x_k)$ für $k \in \{1, \ldots, n\}$ (siehe Beweis von Satz 4.24). Es ergibt sich

$$\sum_{k=0}^{n-1} \left(\frac{k}{n}b\right)^d \frac{b}{n} \leq \int_0^b x^d \,\mathrm{d}x \leq \sum_{k=1}^{n} \left(\frac{k}{n}b\right)^d \frac{b}{n}$$

oder äquivalent

$$\frac{b^{d+1}}{n^{d+1}} \sum_{k=1}^{n-1} k^d \leq \int_0^b x^d \,\mathrm{d}x \leq \frac{b^{d+1}}{n^{d+1}} \sum_{k=1}^{n} k^d \tag{4.13}$$

Nach Proposition 3.31 gilt

$$\sum_{k=1}^{n} k^d = \frac{n^{d+1}}{d+1} + c_d n^d + c_{d-1} n^{d-1} + \ldots + c_0$$

für gewisse Koeffizienten $c_d, \ldots, c_0 \in \mathbb{Q}$. Damit möchten wir die linke und die rechte Summe in (4.13) nach unten respektive nach oben abschätzen. Wir erhalten für die Summe auf der rechten Seite

$$\sum_{k=1}^{n} k^d \leq \frac{n^{d+1}}{d+1} + |c_d| n^d + |c_{d-1}| n^{d-1} + \ldots + |c_0| \leq \frac{n^{d+1}}{d+1} + (|c_d| + |c_{d-1}| + \ldots + |c_0|) n^d.$$

Für die Summe auf der linken Seite von (4.13) erhalten wir analog

$$\begin{aligned}\sum_{k=1}^{n-1} k^d = \sum_{k=1}^{n} k^d - n^d &= \frac{n^{d+1}}{d+1} + (c_d - 1)n^d + c_{d-1}n^{d-1} + \ldots + c_0 \\ &\geq \frac{n^{d+1}}{d+1} - |c_d - 1|n^d - |c_{d-1}|n^{d-1} - \ldots - |c_0| \\ &\geq \frac{n^{d+1}}{d+1} - (|c_d - 1| + |c_{d-1}| + \ldots + |c_0|)n^d\end{aligned}$$

Wir definieren

$$c_- = (|c_d - 1| + |c_{d-1}| + \ldots + |c_0|), \quad c_+ = (|c_d| + |c_{d-1}| + \ldots + |c_0|)$$

und setzen die oben erhaltenen Ungleichungen mit (4.13) zusammen. Wir erhalten

$$\frac{b^{d+1}}{d+1} - \frac{c_- b^{d+1}}{n} \leq \int_0^b x^d \,\mathrm{d}x \leq \frac{b^{d+1}}{d+1} + \frac{c_+ b^{d+1}}{n}. \tag{4.14}$$

Aus dem Archimedischen Prinzip (Satz 2.68) folgt nun, dass $\int_0^b x^d \,\mathrm{d}x = \frac{b^{d+1}}{d+1}$. □

**Übung 4.38** (Allgemeine Grenzen)**.** *Beweisen Sie Satz 4.26 für $a < b = 0$ und dann allgemein.*

**Applet 4.39** (Integral eines Polynoms)**.** *Wir betrachten nochmals das partikuläre Integral, wobei wir diesmal mit einer Polynomfunktion beginnen und dadurch Satz 4.37 anwenden können.*

**Beispiel 4.40.** *Als Anwendung von Satz 4.37 berechnen wir*

$$\begin{aligned}\int_1^2 (x^4 + 5x^2 - x + 1)\,\mathrm{d}x &= \int_1^2 x^4 \,\mathrm{d}x + 5\int_1^2 x^2 \,\mathrm{d}x - \int_1^2 x \,\mathrm{d}x + 1 \\ &= \left[\frac{x^5}{5}\right]_1^2 + 5\left[\frac{x^3}{3}\right]_1^2 - \left[\frac{x^2}{2}\right]_1^2 + 1 \\ &= \frac{2^5 - 1}{5} + 5\frac{2^3 - 1}{3} - \frac{2^2 - 1}{2} + 1 \\ &= \frac{521}{30},\end{aligned}$$

*wobei wir für eine Funktion $f$, deren Definitionsbereich $[a, b]$ enthalten sollte, die Notation $[f(x)]_a^b = f(b) - f(a)$ verwendet haben.*

**Übung 4.41** (Integration der Wurzelfunktion)**.** *Sei $[a, b]$ ein beschränktes, abgeschlossenes Intervall mit $0 \leq a < b$. Zeigen Sie zuerst, dass $x \in [a, b] \mapsto x^{\frac{1}{m}} \in \mathbb{R}$ für $m \in \mathbb{N}$ Riemann-integrierbar ist. In dieser Übung möchten wir des Weiteren das Riemann-Integral von $x \in [0, 1] \mapsto x^{\frac{1}{m}} \in \mathbb{R}$ berechnen. Dazu betrachten wir für $n \in \mathbb{N}$ und $\varepsilon > 0$ die Zerlegung von $[0, 1]$ aus dem Beweis von Satz 4.24 und die dort definierten Treppenfunktionen $u, o$ für das Polynom $x^m$.*

(i) *Finden Sie von $u$ respektive $o$ ausgehend eine Treppenfunktion $o'$ respektive eine Treppenfunktion $u'$ mit $u'(x) \leq x^{\frac{1}{m}} \leq o'(x)$ für $x \in [0,1]$ und*

$$\int_0^1 u(x)\,\mathrm{d}x + \int_0^1 o'(x)\,\mathrm{d}x = 1, \quad \int_0^1 o(x)\,\mathrm{d}x + \int_0^1 u'(x)\,\mathrm{d}x = 1.$$

(ii) *Zeigen Sie, dass*

$$\int_0^1 x^m\,\mathrm{d}x + \int_0^1 x^{\frac{1}{m}}\,\mathrm{d}x = 1$$

*und berechnen Sie damit das Integral $\int_0^1 x^{\frac{1}{m}}\,\mathrm{d}x$.*

## 4.7 Integrierbarkeit stetiger Funktionen

Aus Abschnitt 4.6 wissen wir bereits, dass Polynomfunktionen integrierbar sind. In diesem Abschnitt möchten wir nun unter Verwendung der Beschränktheit und der gleichmässigen Stetigkeit stetiger Funktionen auf kompakten Intervallen (Satz 3.70 respektive Satz 3.78) folgendes allgemeines Resultat beweisen.

**Satz 4.42** (Stetige Funktionen und das Riemann-Integral)**.** *Eine stetige Funktion auf einem kompakten Intervall $[a, b]$ mit $a < b$ ist Riemann-integrierbar.*

*Beweis.* Sei $f \in C([a, b])$ und $\varepsilon > 0$. Nach Satz 3.78 ist $f$ gleichmässig stetig und es gibt ein $\delta > 0$, so dass für alle $x, y \in [a, b]$ gilt

$$|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon.$$

Sei nun $\mathfrak{Z} = \{a = x_0 < x_1 < \ldots < x_n = b\}$ eine Zerlegung von $[a, b]$ mit

$$\max_{k=1,\ldots,n} |x_k - x_{k-1}| < \delta.$$

Wir definieren für jedes $k \in \{1, \ldots, n\}$ die Zahlen

$$\begin{aligned} m_k &= \inf f([x_{k-1}, x_k]) \\ M_k &= \sup f([x_{k-1}, x_k]), \end{aligned}$$

wobei wir Satz 3.70 für die Existenz dieser Infima und Suprema benötigt haben. Wir behaupten nun, dass für alle $k \in \{1, \ldots, n\}$

$$M_k - m_k \leq \varepsilon$$

gilt. In der Tat ist $|x - y| < \delta$ für alle $x, y \in [x_{k-1}, x_k]$, womit nach der gleichmässigen Stetigkeit von $f$ die Ungleichung $|f(x) - f(y)| < \varepsilon$ und insbesondere $f(x) < f(y) + \varepsilon$ erfüllt ist. Da $y \in [x_{k-1}, x_k]$ beliebig war, gilt

$$f(x) \leq \inf \{f(y) + \varepsilon \mid y \in [x_{k-1}, x_k]\} = m_k + \varepsilon$$

Dies beweist aber, dass $m_k + \varepsilon$ eine obere Schranke für $f([x_{k-1}, x_k])$ darstellt, womit wir auf $M_k \leq m_k + \varepsilon$ schliessen.

Wir definieren nun Treppenfunktionen $u, o$ durch

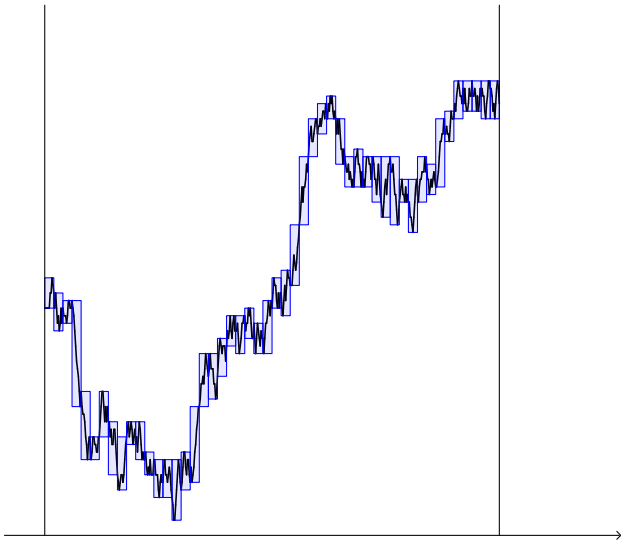
$$u(x) = \begin{cases} m_k & \text{falls } x \in [x_{k-1}, x_k) \text{ für } k \in \{1, \ldots, n\} \\ m_n & \text{falls } x = b \end{cases}$$

$$o(x) = \begin{cases} M_k & \text{falls } x \in [x_{k-1}, x_k) \text{ für } k \in \{1, \ldots, n\} \\ M_n & \text{falls } x = b \end{cases}$$

für $x \in [a,b]$. Nach Definition von $m_k, M_k$ für $k \in \{1,\dots,n\}$ gilt daher $u \leq f \leq o$. Des Weiteren ist

$$\int_a^b (o-u)\,\mathrm{d}x = \sum_{k=1}^n (M_k - m_k)(x_k - x_{k-1}) \leq \varepsilon \sum_{k=1}^n (x_k - x_{k-1}) = \varepsilon(b-a).$$

Da $\varepsilon > 0$ beliebig war (und $b-a$ fix ist), zeigt dies mittels Proposition 4.12, dass $f$ Riemann-integrierbar ist. □



Figur 4.4: Eine stetige Funktion auf einem kompakten Intervall $[a,b]$ mag zwar einen zittrigen Graphen besitzen, ist aber trotzdem Riemann-integrierbar.

**Applet 4.43** (Integrierbarkeit einer „zittrigen“ Funktion)**.** *Wir sehen, dass eine stetige aber zittrige Funktion wie im dargestellten Graphen auch Riemann-integrierbar ist.*

VO: 02.11.17

## 4.8 Weitere Lernmaterialien

### 4.8.1 Verwendung des Kapitels

Im Folgenden werden wir meist nicht direkt auf die Definition des Riemann-Integrals mit Hilfe von Treppenfunktionen zurückgreifen, sondern stattdessen die hier besprochenen Eigenschaften verwenden, um weitere Integrationsgesetze und Integrationsformeln für noch zu findende, weitere Funktionen zu beweisen. Trotzdem ist es wichtig sich an die Definition des Riemann-Integrals und die Vorraussetzungen an Funktion und Integrationsbereich zu erinnern, damit der Unterschied zu etwaigen späteren Verallgemeinerungen klar wird. Das Verständnis der Definition des Riemann-Integrals ist auch wichtig, da wir dieses im zweiten Semester zu einem mehrdimensionalen Integral verallgemeinern wollen und dabei analog vorgehen werden (siehe auch Abschnitt 4.9). Die Berechnung von Riemann-Integralen wird uns später mit Hilfe des Fundamentalsatzes der Differential- und Integralrechnung erheblich einfacher fallen.

Bei einigen Beweisen dieses Kapitels waren Sie vielleicht versucht, den Grenzübergang für $n \to \infty$ oder $\varepsilon \searrow 0$ zu verwenden. Unsere bisherigen Argumente haben diesen Begriff nicht verwendet, aber wir führen Grenzwerte im nächsten Kapitel ein und Sie dürfen daher demnächst die Beweise von Satz 4.31, Satz 4.37 oder Satz 4.42 umformulieren und zum Beispiel in (4.14) den Grenzwert für $n \to \infty$ nehmen.

### 4.8.2 Weitere Übungsaufgaben

**Übung** (Maximum und Minimum). *Charakterisieren Sie die Riemann-integrierbaren Funktionen, für welche sowohl bei den Untersummen als auch bei den Obersummen ein Maximum beziehungsweise ein Minimum in der Definition des Riemann-Integrals angenommen wird.*

**Übung** (Nicht umkehrbar). *Finden Sie eine Funktion $f$ auf einem kompakten Intervall $[a, b]$ für $a < b$, so dass $f$ nicht Riemann-integrierbar ist, aber $|f|$ Riemann-integrierbar ist.*

**Übung** (Verhalten unter Verknüpfung). *Wir möchten in dieser Übung zeigen, dass Verknüpfungen von Riemann-integrierbaren Funktionen im Allgemeinen nicht Riemann-integrierbar sind. Dazu betrachten wir die Riemann-integrierbare Funktion $g : [0, 1] \to [0, 1]$ aus Übung 4.25. Finden Sie eine Riemann-integrierbare Funktion $f : [0, 1] \to \mathbb{R}$, so dass $f \circ g$ die nicht-Riemann-integrierbar ist.*

**Übung** (Definitheit). *Sei $f \in C([a, b])$ eine stetige Funktion auf einem kompakten Intervall $[a, b]$ zu $a < b$, so dass $f \geq 0$ ist (das heisst, $f$ ist nicht-negativ). Zeigen Sie, dass folgende Aussagen äquivalent sind:*

*(i) Es gilt $f(x) = 0$ für alle $x \in [a, b]$.*

*(ii) Es gilt $\int_a^b f(x)\,\mathrm{d}x = 0$.*

**Übung** (Sandwich mit Riemann-integrierbaren Funktionen). *Sei $f \in \mathcal{F}([a, b])$ eine Funktion auf einem kompakten Intervall $[a, b]$ mit $a < b$. Zeigen Sie, dass folgende Aussagen äquivalent sind:*

*(i) Die Funktion $f$ ist Riemann-integrierbar.*

*(ii) Für jedes $\varepsilon > 0$ existieren Riemann-integrierbare Funktionen $f_{\varepsilon,-}, f_{\varepsilon,+} : [a,b] \to \mathbb{R}$ mit $f_{\varepsilon,-} \leq f \leq f_{\varepsilon,+}$ sowie $\int_a^b f_{\varepsilon,+} - f_{\varepsilon,-} \, \mathrm{d}x < \varepsilon$.*

*Wie wir später sehen werden, kann man zu $\varepsilon > 0$ wie oben sogar stetige Funktionen $f_{\varepsilon,-}, f_{\varepsilon,+}$ mit den gewünschten Eigenschaften wählen.*

**Übung** (Funktionen beschränkter Variation)**.** *Sei $I = [a,b]$ ein kompaktes Intervall mit $a < b$. Eine Funktion hat **beschränkte Variation**, falls*

$$\sup \left\{ \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| \mid \mathfrak{Z} = \{a = x_0 < x_1 < \ldots < x_n = b\} \right\} < \infty$$

*In dieser Übung möchten wir zeigen, dass sich jede Funktion $f \in \mathcal{F}([a,b])$ mit beschränkter Variation als Differenz von zwei monotonen Funktionen schreiben lässt und daher auch Riemann-integrierbar ist. Sei also $f \in \mathcal{F}([a,b])$ mit beschränkter Variation und sei*

$$V(f)(x) = \sup \left\{ \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| \mid \mathfrak{Z}_x = \{a = x_0 < x_1 < \ldots < x_n = x\} \right\}.$$

*für $x \in [a,b]$. Zeigen Sie, dass für $x, x'$ mit $a \leq x < x' \leq b$ gilt*

$$|f(x') - f(x)| + V(f)(x) \leq V(f)(x'),$$

*indem Sie von einer beliebigen Zerlegung von $[a,x]$ ausgehen und diese geeignet zu einer Zerlegung von $[a,x']$ erweitern. Schliessen Sie damit, dass die Funktionen $V(f)$ und $V(f) - f$ monoton wachsend sind.*

### 4.8.3 Multiple-Choice Fragen

### 4.8.4 Lernkarten

Sie können wiederum die Lernkarten oder den Graphen für Ihre Wiederholung der Themen des Kapitels verwenden.

## 4.9 Einschub: Mehrdimensionale Integrale*

**Dieser Abschnittes existiert als Hilfestellung für die Physik I-Vorlesung, wo diese Begriffe bereits auftauchten. Der Abschnitt kann natürlich beim Lernen übersprungen werden, da alles sowieso nochmals kommt. Doch könnte eine ungenaue aber knappe Darstellung auch beim Lernen helfen, weswegen dieser Abschnitt kein * bekommt.*

Wir wollen hier unseren vollständigen Aufbau der Analysis kurz unterbrechen und anwendungsbezogen mehrdimensionale Integrale besprechen. Insbesondere werden wir die vorgestellten Methoden informell begründen, aber nicht vollständig erklären oder beweisen können – wir werden dies erst im zweiten Semester nachholen. Der Grund für den Einschub ist einfach zu erklären: Sie werden ein intuitives Verständnis für diese Themen und die wichtigsten Rechenmethoden in den Vorlesungen Physik I und Physik II benötigen.

### 4.9.1 Definition mittels Treppenfunktionen

Wir beginnen unsere Diskussionen damit, die Definition eines mehrdimensionalen Integrals anzudeuten. Für diese Definition sollten wir Funktionen $f$ auf einem $d$-dimensionalen Quader $[a_1, b_1] \times \cdots \times [a_d, b_d]$ betrachten, wobei $d \geq 1$ die Dimension des Quaders angibt und die Zahlen $a_1 < b_1, \ldots, a_d < b_d$ die Koordinaten und Abmessungen des Quaders definieren. Zur Vereinfachung der Notation setzen wir hier vorerst $d = 2$ und betrachten eine reellwertige Funktionen $f : Q = [a_1, b_1] \times [a_2, b_2] \to \mathbb{R}$.

Eine Treppenfunktion $t \in \mathcal{TF}(Q)$ ist in diesem Zusammenhang eine Funktion, so dass man $Q$ in Teilrechtecke zerlegen kann und $t$ auf den einzelnen Teilrechtecken jeweils konstant ist. Genauer sollte die Zerlegung rasterförmig von der Form

$$Q = \bigcup_{j,k} (x_{j-1}, x_j) \times (y_{k-1}, y_k) \cup N$$

sein, wobei

$$\mathfrak{Z}_1 = \{a_1 = x_0 < x_1 < \ldots < x_m = b_1\}$$
$$\mathfrak{Z}_2 = \{a_2 = y_0 < y_1 < \ldots < y_n = b_2\}$$

zwei beliebige Zerlegung der Kanten $[a_1, b_1]$ und $[a_2, b_2]$ des Rechtecks $Q$ sind und die Vereinigung über alle Paare $(j, k)$ läuft mit $j \in \{1, \ldots, m\}$ und $k \in \{1, \ldots, n\}$. Die Menge $N$ besteht hier aus den Rändern der einzelnen Rechtecke und wird im folgenden einfach ignoriert (da dies eine sogenannte Nullmenge darstellt). Falls nun die Treppenfunktion $t$ für jedes Tupel $(j, k)$ auf dem entsprechenden Teilrechteck den Konstanzwert $c_{j,k}$ annimmt, dann definieren wir das Integral der Treppenfunktion durch

$$\int_Q t \,\mathrm{dvol} = \sum_{j,k} c_{j,k}(x_j - x_{j-1})(y_k - y_{k-1}). \tag{4.15}$$

Wir nehmen an, dass $f : Q \to \mathbb{R}$ beschränkt ist. Dies impliziert wiederum, dass es Treppenfunktionen $u, o \in \mathcal{TF}(Q)$ gibt, die $u \leq f \leq o$ erfüllen. Wir bezeichnen $\int u \,\mathrm{dvol}$ beziehungsweise

$\int o \,\mathrm{dvol}$ als Untersumme und Obersumme zu $f$. Das untere Integral $\underline{I}(f)$ ist nun als Supremum der Untersummen und das obere Integral $\overline{I}(f)$ als Infimum der Obersummen definiert. Wenn diese beiden Zahlen übereinstimmen, dann definieren diese das Riemann-Integral

$$\int_Q f \,\mathrm{dvol} = \underline{I}(f) = \overline{I}(f).$$

Wir wollen dies auch im folgenden Applet erklären, wobei wir das zwei-dimensionale Integral als Volumen des Körpers in Figur 4.5 interpretieren.

**Applet 4.44** (Zelt)**.** *Wir sehen, dass wir das Volumen des Zeltes von unten und von oben abschätzen können, wodurch wir immer genauere Annäherungen für das Volumen erhalten können. Das zwei-dimensionale Integral* $\int_{[-1,1]^2}(2 - x^2 - y^2)\,\mathrm{dvol}$ *gibt das Volumen fehlerfrei an.*

Auch dreidimensionale Integrale können konkrete physikalische Bedeutungen besitzen. Falls zum Beispiel $Q = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$ ein drei-dimensionaler Quader mit Abmessungen $b_1 - a_1, b_2 - a_2, b_3 - a_3 > 0$ (in $m$) ist und $\rho(x, y, z)$ die vom Punkt $(x, y, z) \in Q$ abhängige Dichte des Quaders (in $kg/m^3$) angibt, so gibt das drei-dimensionale Integral

$$\int_Q \rho(x, y, z)\,\mathrm{dvol}$$

die Gesamtmasse des Quaders an. Dies ergibt sich durch Verallgemeinerung der Diskussion in Abschnitt 4.4.3.

### 4.9.2 Iterierte Integrale

In der Definition des Begriffes „Integral einer Treppenfunktion“ $t : Q \to \mathbb{R}$ haben wir über alle Paare $(j, k)$ mit $j \in \{1, \ldots, m\}$ und $k \in \{1, \ldots, n\}$ summiert (siehe Definition 4.15). Wollen wir dies genauer mittels der Summennotation aus Abschnitt 3.1 formulieren, so haben wir die zwei äquivalenten Möglichkeiten

$$\begin{aligned}
\int_Q t \,\mathrm{dvol} &= \sum_{j,k} c_{j,k}(x_j - x_{j-1})(y_k - y_{k-1}) \\
&= \sum_{j=1}^{m} \left[ \sum_{k=1}^{n} c_{j,k}(y_k - y_{k-1}) \right] (x_j - x_{j-1}) \\
&= \sum_{k=1}^{n} \left[ \sum_{j=1}^{m} c_{j,k}(x_j - x_{j-1}) \right] (y_k - y_{k-1}).
\end{aligned}$$

Da das mehrdimensionale Integral gewissermassen ein kontinuierliches Analog zu derartigen Doppelsummen darstellt, könnte man erwarten, dass das mehrdimensionale Integral einer

Riemann-integrierbaren Funktion $f : Q \to \mathbb{R}$ analog

$$\begin{aligned}\int_Q f \,\mathrm{dvol} &= \int_{a_1}^{b_1} \left[ \int_{a_2}^{b_2} f(x,y)\,\mathrm{d}y \right] \mathrm{d}x \\ &= \int_{a_2}^{b_2} \left[ \int_{a_1}^{b_1} f(x,y)\,\mathrm{d}x \right] \mathrm{d}y\end{aligned}$$

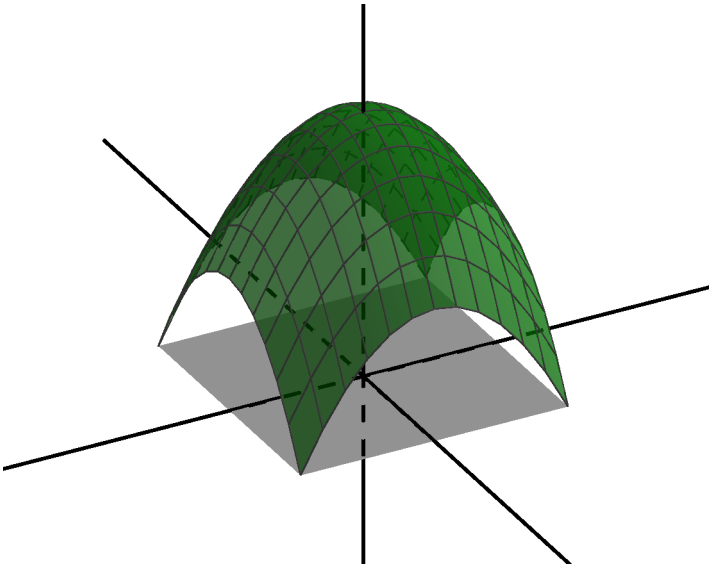
erfüllt, wobei die inneren Integrale (oben das Integral bezüglich $y$) die äussere Integrationsvariable (oben die Variable $x$) als Konstante interpretieren und diese Integrale wiederum eine Funktion bezüglich der äusseren Integrationsvariable (oben $x$) definieren. Dies trifft in der Tat für stetige Funktionen $f$ zu – geeignet interpretiert auch allgemeiner – und wird als der Satz von Fubini bezeichnet. Informell können wir dies in zwei Dimensionen auch durch die Gleichung $\mathrm{dvol} = \mathrm{d}x\,\mathrm{d}y$ ausdrücken.

**Applet 4.45** (Volumen des Zeltes)**.** *Wir können den Satz von Fubini und das Vorgehen der Berechnung des Volumens auch geometrisch veranschaulichen. Dabei bestimmt die $x$-Koordinate einen ebenen Querschnitt durch das Zelt, und die $y$-Koordinate animiert die Berechnung des Flächeninhaltes des Querschnittes. Versuchen Sie mit den Schiebern die Addition der iterierten Summen nachzustellen.*

Der Satz von Fubini ist extrem nützlich, da wir mit diesem Satz die Berechnung von mehrdimensionalen Integrale auf die Berechnung eindimensionaler Integrale zurückführen können (und wir für letztere im Laufe dieses Semester viele Methoden zur Berechnung lernen werden).

**Beispiel 4.46** (Volumen des Zeltes)**.** *Wir definieren das Zelt*

$$Z = \left\{ (x,y,z) \in \mathbb{R}^3 \mid (x,y) \in [-1,1]^2 \text{ und } 0 \le z \le 2 - x^2 - y^2 \right\}.$$



Figur 4.5: Das Zelt $Z$ ist der Bereich unterhalb des Graphen der Funktion $(x,y) \mapsto 2 - x^2 - y^2$.

*Das Volumen des Zeltes ist auf Grund von* $f(x,y) = 2-x^2-y^2 \geq 0$ *für alle* $(x,y) \in [-1,1]^2$ *und obiger Diskussionen durch*

$$\begin{aligned}
\mathrm{vol}(Z) &= \int_{[-1,1]^2} (2-x^2-y^2)\,\mathrm{dvol} \\
&= \int_{-1}^{1} \left[ \int_{-1}^{1} (2-x^2-y^2)\,\mathrm{d}y \right] \mathrm{d}x \\
&= \int_{-1}^{1} \left(\tfrac{10}{3} - 2x^2\right) \mathrm{d}x \\
&= \tfrac{20}{3} - \tfrac{4}{3} = \tfrac{16}{3}
\end{aligned}$$

*gegeben, wobei wir für das innere Integral über* $y \in [-1,1]$ *die Variable* $x$ *als Konstante betrachtet haben und die Rechnung*

$$\begin{aligned}
\int_{-1}^{1} (2-x^2-y^2)\,\mathrm{d}y &= 2\cdot(2-x^2) - \int_{-1}^{1} y^2\,\mathrm{d}y \\
&= 4 - 2x^2 - \left(\tfrac{1}{3}1^3 - \tfrac{1}{3}(-1)^3\right) \\
&= \tfrac{10}{3} - 2x^2
\end{aligned}$$

*verwendet haben.*

### 4.9.3 Schwerpunkt eines Körpers

Wir wollen als weitere Anwendung von mehrdimensionalen Integralen den Schwerpunkt von Körpern $K \subseteq \mathbb{R}^3$ berechnen, wobei $\rho : K \to \mathbb{R}_{\geq 0}$ die vom Punkt abhängige Dichte des Körpers beschreibt. In Analogie zu Abschnitt 4.4.3 sind dann die Gesamtmasse $m$ des Körpers und die Koordinaten $(x_0, y_0, z_0)$ des Schwerpunktes durch die Formeln

$$\begin{aligned}
m &= \int_K \rho(x,y,z)\,\mathrm{dvol} \\
x_0 &= \frac{1}{m}\int_K x\rho(x,y,z)\,\mathrm{dvol} \\
y_0 &= \frac{1}{m}\int_K y\rho(x,y,z)\,\mathrm{dvol} \\
z_0 &= \frac{1}{m}\int_K z\rho(x,y,z)\,\mathrm{dvol}
\end{aligned}$$

gegeben. Wir haben in diesen Definition auch eine Verallgemeinerung des mehrdimensionalen Integrals versteckt, da wir nicht immer annehmen wollen, dass $K = Q$ ein Quader ist. Im Sinne der Anwendung liegt es aber nahe anzunehmen, dass $K$ beschränkt ist. Dadurch existiert ein Quader wie in obiger Diskussion $Q$, der $K$ enthält. Nun setzen wir die Dichtefunktion $\rho$ von $K$ auf ganz $Q$ fort, indem wir $\rho|_{Q\setminus K} = 0$ setzen. Dies macht Sinn, denn wir wollen ja Masse und Schwerpunkt des betrachteten Körpers berechnen und werden dabei davon ausgehen, dass ausserhalb des Körpers Vakuum herrscht. In diesem Sinne ist ein Integral über eine Funktion

$f$ auf $K$ durch

$$\int_K f \,\mathrm{dvol} = \int_Q \mathbb{1}_K f \,\mathrm{dvol}$$

definiert, wobei

$$(\mathbb{1}_K f)(x,y,z) = \begin{cases} f(x,y,z) & \text{für } (x,y,z) \in K \\ 0 & \text{für } (x,y,z) \in Q \setminus K \end{cases}$$

**Beispiel 4.47** (Schwerpunkt des gleichmässig gefüllten Zeltes)**.** *Wir wollen nun diese Formeln ausprobieren und den Schwerpunkt des gleichmässig gefüllten Zeltes (mit Dichte $1kg/m^3$) berechnen. Auf Grund der Symmetrie des Zeltes sind die $x$- und $y$-Koordinaten des gleichmässig gefüllten Zeltes gleich $x_0 = y_0 = 0$. Für die $z$-Koordinate des Zeltes verwenden wir den Quader $Q = [-1,1]^2 \times [0,2]$ und obige Formel, woraus sich*

$$\begin{aligned} z_0 &= \tfrac{1}{m} \int_Z z \,\mathrm{dvol} \\ &= \tfrac{3}{16} \int_{-1}^{1} \int_{-1}^{1} \left[ \int_0^2 \mathbb{1}_Z z \,\mathrm{d}z \right] \mathrm{d}y \,\mathrm{d}x \\ &= \tfrac{3}{16} \int_{-1}^{1} \int_{-1}^{1} \left[ \int_0^{2-x^2-y^2} z \,\mathrm{d}z \right] \mathrm{d}y \,\mathrm{d}x \\ &= \tfrac{3}{16} \int_{-1}^{1} \int_{-1}^{1} \left[ \tfrac{1}{2}(2-x^2-y^2)^2 \right] \mathrm{d}y \,\mathrm{d}x \end{aligned}$$

*ergibt. Wir haben hier die Reihenfolge der Variablen anders gewählt, da in einer anderen Reihenfolge die Betrachtung der Funktion $\mathbb{1}_Z$ erheblich komplizierter wäre. In der Tat hat in dieser Reihenfolge die Funktion $\mathbb{1}_Z$ einfach die Auswirkung, dass das innerste Integral über die Variable $z$ mit den ursprünglichen Integrationsgrenzen $z = 0$ und $z = 2$ (wie in der Definition unseres Quaders $Q$) stattdessen die Integrationsgrenzen $z = 0$ und $z = 2-x^2-y^2$ (was unserer Definition des Zeltes entspricht) verwendet. Um nun $z_0$ tatsächlich zu berechnen, nützen wir nochmals die Symmetrie des Zeltes aus, um die Rechnung ein wenig zu vereinfachen. Dadurch ergibt sich*

$$\begin{aligned} z_0 &= \tfrac{3}{4} \int_0^1 \left[ \int_0^1 \frac{1}{2}(4 + x^4 + y^4 - 4x^2 - 4y^2 + 2x^2y^2) \,\mathrm{d}y \right] \mathrm{d}x \\ &= \tfrac{3}{8} \int_0^1 \left[ 4 + x^4 + \tfrac{1}{5} - 4x^2 - \tfrac{4}{3} + \tfrac{2}{3}x^2 \right] \mathrm{d}x \\ &= \tfrac{3}{8} \left( 4 + \tfrac{1}{5} + \tfrac{1}{5} - \tfrac{4}{3} - \tfrac{4}{3} + \tfrac{2}{9} \right) = \tfrac{11}{15} \end{aligned}$$

**Übung 4.48.** *Wir betrachten nun den Körper*

$$K = \left\{ (x,y,z) \in \mathbb{R}^3 \mid (x,y) \in [0,1]^2 \text{ und } 0 \le z \le x(1-x)y(1-y) \right\}$$

*und die Dichtefunktion $\rho(x, y, z) = xyz$ für $(x, y, z) \in K$.*

*(i) Berechnen Sie das Volumen von $K$.*

*(ii) Berechnen Sie die Masse des Körpers.*

*(iii) Berechnen Sie die Koordinaten des Schwerpunktes. (Auf Grund einer Symmetrie genügt es hierfür zwei dreidimensionale Integrale zu berechnen.)*

Wir erwähnten bereits, dass man den Satz von Fubini für zwei-dimensionale Integrale auf zwei verschiedene Arten anwenden kann. Dies hilft manchmal um die Berechnung des Integrals zu beschleunigen, wie in der nächsten Übungsaufgabe.

**Übung 4.49.** *Berechnen Sie den Flächeninhalt und den Schwerpunkt (bei gleichmässiger Massenverteilung mit Gesamtmasse 1) der Fläche zwischen den Kurven, die durch die Gleichungen $x - y^2 = 0$ und $x - y = 2$ beschrieben wird. Hierzu müssen Sie zuerst eine Skizze des Gebietes erstellen. Versuchen Sie anschliessend die Wahl der Integrationsreihenfolge zu optimieren, so dass Sie möglichst wenige Integrale berechnen müssen (konkret 3 anstatt 6).*

### 4.9.4 Polarkoordinaten

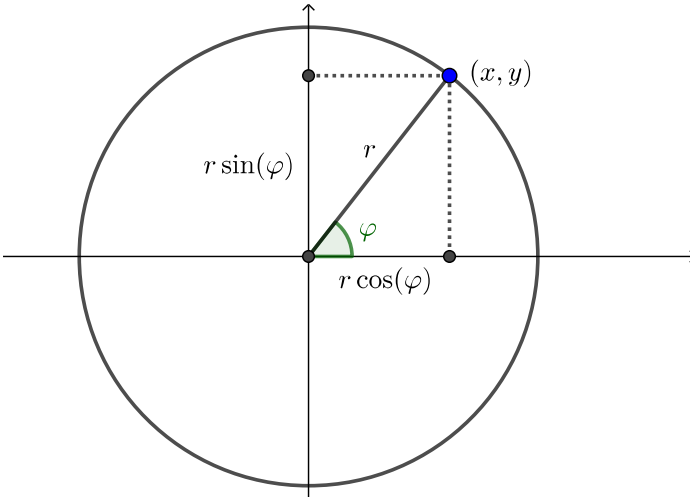
Gelegentlich ist es in gewissen Problemen nützlich, ein Integral in anderen Koordinaten als den Kartesischen Koordinaten $x, y, z$ zu berechnen. Beispielsweise kann eine gegebene Funktion oder ein Integrationsbereich über gewisse Symmetrien verfügen, welche man sich zu Nutzen machen möchte. Wir illustrieren dies hier an den Polarkoordinaten in der Ebene und im nächsten Unterabschnitt an den Kugelkoordinaten im dreidimensionalen Raum.

Jeder Punkt $(x, y) \in \mathbb{R}^2$ lässt sich schreiben als

$$x = r\cos(\varphi),\ \ y = r\sin(\varphi)$$

für den Radius $r = \sqrt{x^2 + y^2}$ und einen Winkel $\varphi \in [0, 2\pi)$.[2] Die Koordinaten $(r, \varphi)$ des Punktes $(x, y)$ werden dabei die **Polarkoordinaten** genannt. Wir bemerken natürlich, dass die Funktionen $\cos : \mathbb{R} \to \mathbb{R}$ und $\sin : \mathbb{R} \to \mathbb{R}$ noch nicht formal definiert wurden; wir werden diesen Mangel später beheben. Für den Moment begnügen wir uns mit folgendem Bild:

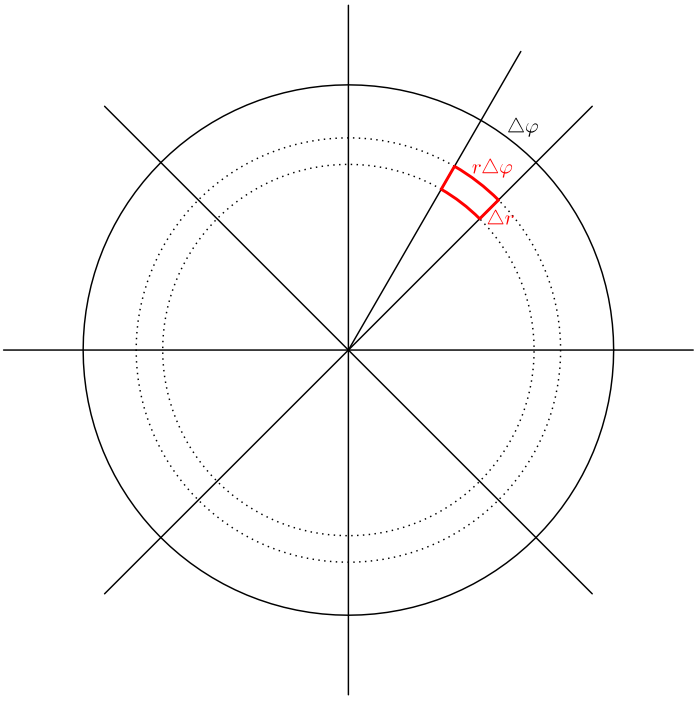
[2] Wie wir später bei der Einführung der Winkelfunktionen sehen werden, stellt die Bogenlänge am Einheitskreis die einzige natürliche Wahl für die Angabe eines Winkels dar.




Gegeben eine Riemann-integrierbare Funktion $f : B_R(0) \to \mathbb{R}$ möchten wir nun das Integral $\int_{B_R(0)} f$ als Integral bezüglich den neuen Koordinaten $(r, \varphi)$ ausdrücken. Dabei können wir aber *nicht* einfach dvol wie in der Diskussion vom Satz von Fubini als $\mathrm{d}\varphi\,\mathrm{d}r$ interpretieren, denn dies würde die vorliegende geometrische Bedeutung der Polarkoordinaten komplett ignorieren. Stattdessen gilt

$$\int_{B_R(0)} f \,\mathrm{dvol} = \int_0^R \int_0^{2\pi} f(r\cos\varphi, r\sin\varphi) r \,\mathrm{d}\varphi\,\mathrm{d}r, \tag{4.16}$$

Der zusätzliche Faktor $r$ beschreibt das Volumen kleiner Quader in den Koordinaten $(r, \varphi)$, wie wir im folgenden Bild erklären möchten.



Figur 4.6: Wir betrachten ein „Rechteck“ in Polarkoordinaten (kurz „Polarrechteck“), das aus jenen Punkten besteht, die Distanz zwischen $r_1$ und $r_2 > r_1$ von Null haben und Winkel zwischen $\varphi_1$ und $\varphi_2 > \varphi_1$ zur $x$-Achse haben. Sind $\triangle r = r_2 - r_1$ und $\triangle\varphi = \varphi_2 - \varphi_1$ klein, so ist dieser Ausschnitt eines Kreisringes fast rechteckig mit „Seitenlängen“ $\triangle r$ und etwa $r\triangle\varphi$. Wir werden diese Idee im 2. Semester zu einem Beweis von (4.16) ausbauen.

**Beispiel 4.50** (Kreisrundes Zelt)**.** *Wir betrachten das adaptierte Zelt*

$$Z_{rund} = \left\{(x,y,z) \in \mathbb{R}^3 \mid (x,y) \in B_1(0) \text{ und } 0 \le z \le 2 - x^2 - y^2\right\}$$

*mit kreisförmiger Basis und berechnen das Volumen. Es gilt*

$$\mathrm{vol}(Z_{rund}) = \int_{B_1(0)} (2 - x^2 - y^2)\,\mathrm{d}x\,\mathrm{d}y = \int_0^{2\pi}\int_0^1 (2 - r^2) r\,\mathrm{d}r\,\mathrm{d}\varphi$$

*unter Verwendung der Formel* (4.16)*. Nun berechnet man*

$$\begin{aligned}\int_0^{2\pi}\int_0^1 (2 - r^2) r\,\mathrm{d}r\,\mathrm{d}\varphi &= 2\pi \int_0^1 2r - r^3\,\mathrm{d}r \\ &= 2\pi[r^2 - \tfrac{1}{4}r^4]_0^1 = \tfrac{3\pi}{2}.\end{aligned}$$

**Beispiel 4.51** (Trägheitsmoment der Kreisscheibe)**.** *Wir betrachten zu einem Radius $R > 0$ die Kreisscheibe $B_R(0) \subseteq \mathbb{R}^2$, welche wir nun um die Null rotieren lassen möchten. Sei $\omega \in \mathbb{R}$ die dazugehörige Winkelgeschwindigkeit (mit Einheit $s^{-1}$). Betrachtet man nun einen Punkt $p \in B_R(0)$ und ein sehr kleines „Polarrechteck“ $U$ um diesen Punkt wie in Figur 4.6, so rotieren Punkte in $U$ etwa mit Geschwindigkeit $\|p\|\omega$. Die kinetische Energie für die Bewegung von $U$ ist also in etwa gegeben durch $\frac{1}{2}\|p\|^2\omega^2\triangle m$, wobei $\triangle m$ die Masse von $U$ bezeichnet. Summiert man dies über alle Polarrechtecke, so erhält man eine intuitive Begründung für die folgende Formel für die kinetische Energie der Rotation (kurz Rotationsenergie)*

$$E_{rot} = \tfrac{1}{2}\omega^2 \int_{B_R(0)} (x^2 + y^2)\rho(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

*Dabei ist $\rho$ die Massenverteilung auf $B_R(0)$. Die Grösse $J = \int_{B_R(0)} (x^2 + y^2)\rho(x,y)\,\mathrm{d}x\,\mathrm{d}y$ (mit Einheit $\mathrm{kg\,m^2}$) verhält sich also wie die Masse für die geradlinige Bewegung und ist in diesem Sinne intrinsisch. Sie wird das Trägheitsmoment von $B_R(0)$ um Null genannt. Wir wollen dieses nun berechnen, wobei wir annehmen wollen, dass $\rho$ konstant ist und $B_R(0)$ Masse $m$ hat. Wir haben also $\rho = \frac{m}{\pi R^2}$, und damit ist das Trägheitsmoment durch*

$$J = \int_{B_R(0)} (x^2 + y^2)\rho\,\mathrm{d}x\,\mathrm{d}y = \rho \int_0^{2\pi}\int_0^R r^3\,\mathrm{d}r\,\mathrm{d}\varphi = \rho 2\pi \tfrac{R^4}{4} = \tfrac{1}{2}mR^2$$

*gegeben.*

*Im Vergleich dazu wäre das Trägheitsmoment für einen Kreisring mit Masse $m$ am Kreis mit Radius $R$ gleich $mR^2$: Denn bei vernachlässigbarer Dicke des Kreisrings hat jeder Teil der*

*Masse Geschwindigkeit $\omega R$, womit die kinetische Energie der Rotation durch*

$$E_{rot} = \frac{1}{2}m(\omega R)^2 = \frac{1}{2}\omega^2 \underbrace{mR^2}_{J}$$

*gegeben ist.*

**Applet 4.52** (Trägheitsmomente)**.** *Wir sehen verschiedene Körper, welche an einer Rampe frei runter rollen. Dabei kommt es je nach Trägheitsmoment des Körpers zu unterschiedlichen Geschwindigkeiten, da die potentiellen Energie in Rotationsenergie und kinetische Energie umgewandelt wird und $E_{rot} = \frac{1}{2}\omega^2 J$ ist. Zum Vergleich wird auch noch ein nicht rotierender Würfel dargestellt, der ohne Reibung die Rampe runter rutscht. Wir werden die Trägheitsmomente der anderen dargestellten Körper unten berechnen.*

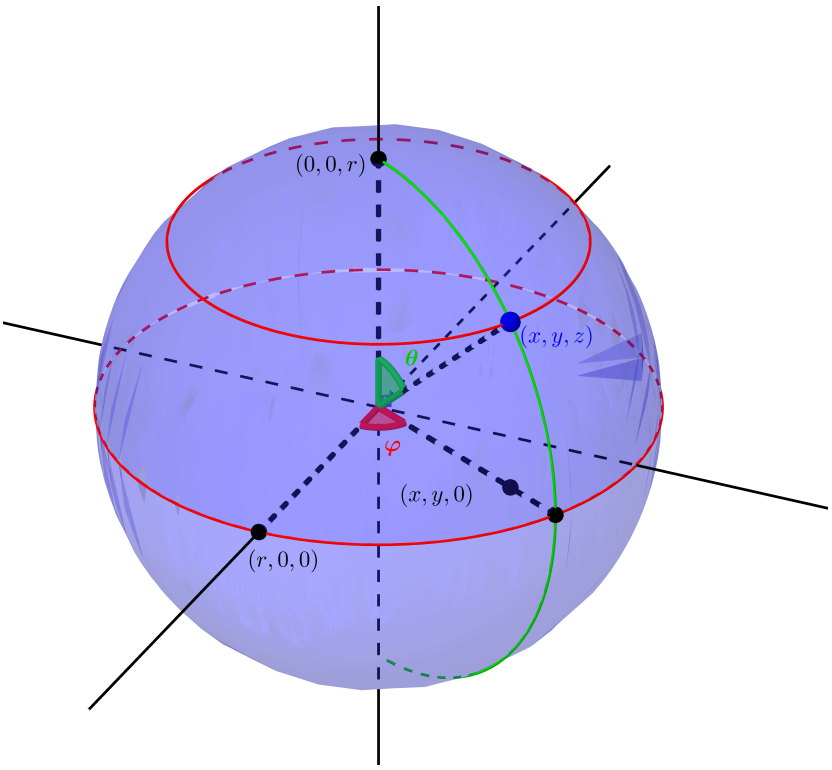
*Was passiert, wenn wir am Ende der Rampe alle Objekte (mit Hilfe einer Stange durch die Rotationsachsen) stoppen ohne die Rotation zu stören und dann nochmals gleichzeitig weiterrollen lassen? Es ist klar, dass der Würfel dann einfach liegen bleibt, da wir beim Stoppen seine kinetische Energie auf Null gesetzt haben und er keine Rotationsenergie hat. Was passiert mit den anderen Körpern?*

### 4.9.5 Kugelkoordinaten

Ähnlich zum zweidimensionalen Fall gibt es im dreidimensionalen Raum sphärische Koordinaten. Jeder Punkt $(x, y, z) \in \mathbb{R}^3$ lässt sich schreiben als

$$\begin{aligned} x &= r\sin(\theta)\cos(\varphi), \\ y &= r\sin(\theta)\sin(\varphi), \\ z &= r\cos(\theta) \end{aligned}$$

für den Radius $r = \sqrt{x^2 + y^2 + z^2}$ und Winkel $\varphi \in [0, 2\pi)$, $\theta \in [0, \pi)$ wie im folgenden Bild.




Figur 4.7: Der Winkel $\varphi$ definiert (abgesehen von der Einheit) den Längengrad des Punktes $(x,y,z)$ und $\theta$ entspricht dem Breitengrad. Formal konstruiert sich der Winkel $\varphi$ als den Winkel, den man mit Polarkoordinaten erhält, wenn man den Punkt $(x,y,z)$ auf die $xy$-Ebene projiziert. Weiter ist $\theta$ der Winkel zwischen $(x,y,z)$ und der positiven $z$-Achse.

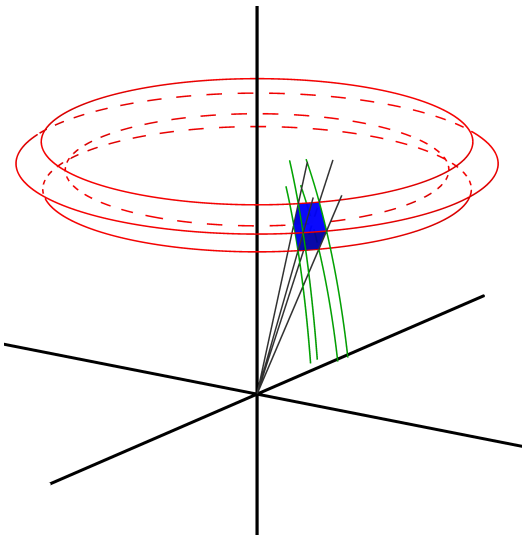
Für eine Riemann-integrierbare Funktion

$$f : B_R(0) = \left\{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2+z^2 \leq R^2\right\} \to \mathbb{R}$$

gilt dann

$$\int_{B_R(0)} f \,\mathrm{dvol} = \int_0^R \int_0^\pi \int_0^{2\pi} f(r\sin\theta\cos\phi, r\sin\theta\sin\varphi, r\cos\theta) r^2 \sin\theta \,\mathrm{d}\varphi\,\mathrm{d}\theta\,\mathrm{d}r. \tag{4.17}$$

Wie im vorherigen Abschnitt beschreibt der Faktor $r^2\sin\theta$ das Verhältnis des Volumens eines sehr kleinen Quaders bezüglich den neuen Kugelkoordinaten $(r,\theta,\varphi)$ im Vergleich zu dem Produkt der Differenzen der einzelnen (Kugel-)Koordinaten.



Figur 4.8: Wie schon beim Bild 4.6 möchten wir hier erklären, wie der Faktor $r^2 \sin\theta$ in der Integrationsformel 4.17 zustande kommt. Betrachtet man den durch Radien $r_1 < r_2$ und Winkel $\theta_1 < \theta_2$, $\varphi_1 < \varphi_2$ gegebenen „Kugelquader", so ist dessen Volumen in etwa das Produkt seiner Seitenlängen, falls $\triangle r = r_2 - r_1$, $\triangle\theta = \theta_2 - \theta_1$, $\triangle\varphi = \varphi_2 - \varphi_1$ klein sind. Die Seitenlänge in radialer Richtung ist $\triangle r$ (unabhängig von beiden Winkeln) und die Seitenlänge in Richtung von $\theta$ ist etwa $r\triangle\theta$ (unabhängig von $\varphi$). Die Seitenlänge in Richtung $\varphi$ ist gegeben durch $r\sin(\theta)\triangle\varphi$, da für festes $r$ und $\theta$ und sich verändernden $\varphi$ eine Bewegung auf einem Kreis mit Radius $r\sin\theta$ beschrieben wird. Daraus ergibt sich, das das gesuchte Volumen des Kugelquaders in etwa durch $r^2\sin\theta\triangle\varphi\triangle\theta\triangle r$ gegeben ist, was die Formel (4.17) geometrisch erklärt.

**Beispiel 4.53** (Volumen des Balles mit Radius $R$)**.** *Wir berechnen das Volumen des Balles* $B_R(0) = \{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2+z^2 \leq R^2\}$. *Es gilt*

$$\begin{aligned}
\operatorname{vol}(B_R(0)) &= \int_{B_R(0)} 1 \,\mathrm{dvol} = \int_0^R \int_0^\pi \int_0^{2\pi} r^2 \sin\theta \,\mathrm{d}\varphi \,\mathrm{d}\theta \,\mathrm{d}r \\
&= 2\pi \left( \int_0^\pi \sin(\theta) \,\mathrm{d}\theta \right) \cdot \left( \int_0^R r^2 \,\mathrm{d}r \right) \\
&= 2\pi \cdot [-\cos(\theta)]_0^\pi \cdot [\tfrac{1}{3}r^3]_0^R = \tfrac{4}{3}\pi R^3,
\end{aligned}$$

*wobei wir die Integrationsregel* $\int_a^b \sin(\theta)\,\mathrm{d}\theta = [-\cos(\theta)]_a^b$ *verwendet haben.*

**Beispiel 4.54** (Trägheitsmoment des Balles)**.** *Wir wollen das Trägheitsmoment des Balles mit Radius $R$ berechnen, wobei wir annehmen wollen, dass die Masse $m$ gleichmässig mit Dichte* $\rho = \frac{3m}{4\pi R^3}$ *im Ball verteilt ist. Wir gehen hier ähnlich wie in Beispiel 4.51 vor und wollen annehmen, dass der Ball $B_R(0)$ mit Mittelpunkt 0 gegeben ist und wir diesen um die $z$-Achse*

*rotieren lassen wollen. Daraus ergibt sich*

$$\begin{aligned}
J &= \int_{B_R(0)} (x^2+y^2)\rho \,\mathrm{dvol} \\
&= \rho \int_0^R \int_0^\pi \int_0^{2\pi} (r\sin\theta)^2 r^2 \sin\theta \,\mathrm{d}\varphi \,\mathrm{d}\theta \,\mathrm{d}r \\
&= \rho \cdot 2\pi \cdot \left( \int_0^\pi (1-\cos^2\theta)\sin(\theta)\,\mathrm{d}\theta \right) \cdot \left( \int_0^R r^4 \,\mathrm{d}r \right) \\
&= \tfrac{3m}{4\pi R^3} \cdot 2\pi \cdot [-\cos\theta + \tfrac{1}{3}\cos^3\theta]_0^\pi \cdot [\tfrac{1}{5}r^5]_0^R \\
&= \tfrac{2mR^2}{5}.
\end{aligned}$$

**Beispiel 4.55** (Trägheitsmoment einer Kugelschale)**.** *Wir wollen nun annehmen, dass die Masse $m$ mit gleichmässiger Dichte $\rho$ (in $kg/m^2$) an der Oberfläche des Balles mit Radius $R$ verteilt ist, und wiederum das Trägheitsmoment berechnen. Da die Oberfläche zwei-dimensional ist, liegt es nahe zu erwarten, dass wir auch ein zwei-dimensionales Integral berechnen müssen. Wir werden auch dies im zweiten Semester genauer definieren und dessen Eigenschaften vollständig erklären, doch begnügen wir uns hier mit folgenden beiden Rechnungen.*

*Die Oberfläche der Kugel ist gegeben durch*

$$\begin{aligned}
A &= \int_0^\pi \int_0^{2\pi} R^2 \sin\theta \,\mathrm{d}\varphi \,\mathrm{d}\theta \\
&= 2\pi R^2 \cdot \int_0^\pi \sin(\theta)\,\mathrm{d}\theta \\
&= 4\pi R^2,
\end{aligned}$$

*wobei wir uns die Kugeloberfläche als Vereinigung von kleinen „Sphärenrechtecken“ (ähnlich wie in Figur 4.8) vorgestellt haben und dabei das Ihnen wahrscheinlich bekannte Ergebnis erhalten haben.*

*Durch diesen Erfolg bestätigt berechnen wir die Dichte $\rho = \frac{m}{4\pi R^2}$ und das Trägheitsmoment*

$$\begin{aligned}
J &= \int_0^\pi \int_0^{2\pi} (x^2+y^2) R^2 \sin\theta \rho \,\mathrm{d}\varphi \,\mathrm{d}\theta \\
&= \rho R^2 \int_0^\pi \int_0^{2\pi} (R\sin\theta)^2 \sin\theta \,\mathrm{d}\varphi \,\mathrm{d}\theta \\
&= \tfrac{m}{4\pi R^2} \cdot 2\pi R^4 \cdot \int_0^\pi \sin^3\theta \,\mathrm{d}\theta \\
&= \tfrac{mR^2}{2} \cdot \tfrac{4}{3} = \tfrac{2mR^2}{3}.
\end{aligned}$$

### 4.9.6 Zusammenfassung

Wir hoffen, dass Sie in dieser Diskussion folgende Punkte erkennen konnten:

- Der Satz von Fubini erlaubt uns mehrdimensionale Integrale mit Hilfe von eindimensionalen Integralen zu berechnen. Dies liefert zusätzliche Motivation weitere Methoden zur

Berechnung von eindimensionalen Integralen zu finden, da wir diese Methoden auch für die Berechnung von mehrdimensionalen Riemann-Integralen benötigen werden.

- Die geometrische Anschauung ist sehr hilfreich – fast schon notwendig – um die betrachteten Integralausdrücke zu finden. Vor allem bei Polar- und Kugelkoordinaten ist es wichtig den zusätzlichen „geometrischen Faktor“ in die Integrale einzubauen.

- Das Wort „etwa“ ist in diesem Abschnitt unüblich oft aufgetreten, da eine genauere Begründung oder sogar ein Beweis der Aussagen für uns erst im nächsten Semester möglich ist. In der Tat hatten wir bei der Besprechung sehr viel Vertrauen in die Welt, da wir des Öfteren kleine Fehler erlaubten, aber dann über alle Teilquader summierten, ohne uns Gedanken zu machen, ob denn die kleinen Fehler auch in der Summe (über sehr viele kleine Teilquader) noch klein bleiben.

- Wir hoffen, dass Sie dies auch als Motivation sehen, unsere Theorie weiterhin schrittweise und ausführlich aufzubauen. Damit wir eben nicht wie oben „Integral-Alchemie“ betreiben sondern auch die mehrdimensionale Integralrechnung und die dafür nötigen Theorien vollständig verstehen. Zum Beispiel werden wir dann auch die geometrischen Faktoren der Polar- und Kugelkoordinaten vollständig erklären und für beliebige „glatte Koordinatensysteme“ berechnen können.

# Kapitel 5

# Folgen und Grenzwerte

Wir beginnen in diesem Kapitel „Bewegungen“ zu untersuchen, indem wir Grenzwerte von Folgen und allgemeineren Funktionen diskutieren. Dies wird uns insbesondere erlauben, die Exponentialabbildung zu definieren. Wir werden auch den Zusammenhang zwischen dem Riemann-Integral und Riemann-Summen als eine Art von Grenzwertprozess untersuchen.

Wir haben in den letzten beiden Kapiteln wiederholt Supremum und Infimum als wichtigste Hilfsmittel für den Aufbau der Theorie verwendet. Allerdings haben diese die fundamentale Einschränkung, nur für die reellen Zahlen sinnvoll zu sein. Wir werden hier bedeutende alternative Hilfsmittel, nämlich den Begriff der Cauchy-Folgen und die Existenz von konvergenten Teilfolgen einführen. Wie wir zum Teil in diesem Kapitel aber in grösserem Ausmass im zweiten Semester sehen werden sind diese Hilfsmittel auch im $\mathbb{R}^d$ für $d \geq 2$ sinnvoll und nützlich.

## 5.1 Konvergenz von Folgen

**Definition 5.1** (Folge)**.** Sei $X$ eine Menge. Eine **Folge** in $X$ ist eine Abbildung $a : \mathbb{N} \to X$. Das Bild $a(n)$ von $n \in \mathbb{N}$ schreibt man auch als $a_n$ und bezeichnet es als das $n$-te **Folgenglied** von $a$. Anstatt $a : \mathbb{N} \to X$ schreibt man oft $(a_n)_{n\in\mathbb{N}}$, $(a_n)_{n=1}^{\infty}$ oder kurz $(a_n)_n$. Die Menge der Folgen in $X$ wird auch als $X^{\mathbb{N}}$ bezeichnet. Eine Folge $(a_n)_n$ heisst **konstant**, falls $a_n = a_m$ für alle $m, n \in \mathbb{N}$, und **schliesslich konstant**, falls ein $N \in \mathbb{N}$ existiert mit $a_n = a_m$ für alle $m, n \in \mathbb{N}$ mit $m, n \geq N$.

Wir werden vorerst nur Folgen in $\mathbb{C}$ betrachten. Diese bilden, wie wir in den Abschnitten 3.4 und 3.5.1 gesehen haben, einen Vektorraum ($\mathbb{C}^{\mathbb{N}} = F_{\mathbb{C}}(\mathbb{N})$) mit den Verknüpfungen

$$(a_n)_n + (b_n)_n = (a_n + b_n)_n, \quad \alpha \cdot (a_n)_n = (\alpha a_n)_n$$

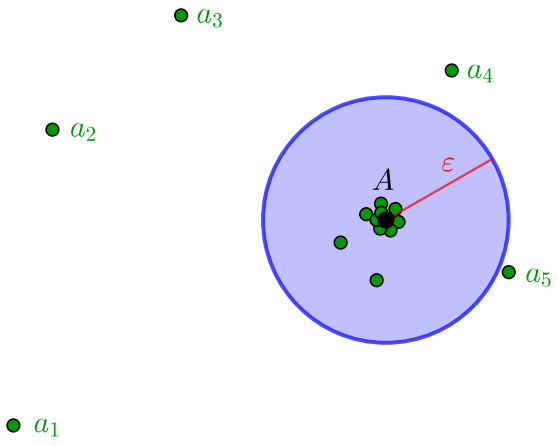
für $\alpha \in \mathbb{C}$ und Folgen $(a_n)_n, (b_n)_n \in \mathbb{C}^{\mathbb{N}}$.

Für eine schliesslich konstante Folge $(a_n)_n$ ist $A \in \mathbb{C}$ mit $a_n = A$ für alle hinreichend grossen $n \in \mathbb{N}$ eine besondere Zahl, die wir mit der schliesslich konstanten Folge assoziieren können. Wir wollen diese Assoziation verallgemeinern, wobei wir eine beliebig kleine Fehlerschranke $\varepsilon > 0$ erlauben und wiederum ein $A \in \mathbb{C}$ suchen, so dass für alle hinreichend

grossen $n \in \mathbb{N}$ das Folgenglied $a_n$ – bis auf einen Fehler kleiner als $\varepsilon$ – gleich $A$ sein soll. Wir formalisieren diesen Begriff wie folgt.

**Definition 5.2** (Grenzwert)**.** Wir sagen, dass eine Folge $(a_n)_n$ in $\mathbb{C}$ gegen eine Zahl $A \in \mathbb{C}$ **strebt**, gegen die Zahl $A$ **konvergiert** oder **Grenzwert** $A$ hat, falls es für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass $|a_n - A| < \varepsilon$ für alle $n \geq N$. In diesem Fall nennen wir die Folge **konvergent**. Falls es keinen Grenzwert in $\mathbb{C}$ gibt, nennen wir die Folge **divergent**.

Nochmals anders formuliert ist eine Folge $(a_n)_n$ nach $A$ konvergent, falls hinreichend späte Folgenglieder der Zahl $A$ beliebig nahe kommen müssen – siehe folgendes Bild:




In Prädikatenlogik ist Konvergenz gegen $A$ durch

$$\forall \varepsilon > 0 \; \exists N \in \mathbb{N} \; \forall n \geq N : |a_n - A| < \varepsilon$$

gegeben. Wir bemerken noch, dass eine Folge $(a_n)_n$ in $\mathbb{C}$ genau dann gegen $A \in \mathbb{C}$ konvergiert, wenn die Folge $(|a_n - A|)_n$ gegen Null konvergiert.

**Lemma 5.3** (Eindeutigkeit des Grenzwerts)**.** *Der Grenzwert einer konvergenten Folge $(a_n)_n$ ist eindeutig bestimmt. Wir bezeichnen ihn mit $\lim_{n\to\infty} a_n$.*

*Beweis.* Die Folge $(a_n)_n$ konvergiere gegen $A_1 \in \mathbb{C}$ und $A_2 \in \mathbb{C}$. Wir nehmen an, dass $A_1 \neq A_2$ und sei $\varepsilon = |A_1 - A_2| > 0$. Dann können wir $N_1, N_2 \in \mathbb{N}$ finden, so dass für alle $n \geq N_1$ gilt $|a_n - A_1| < \frac{\varepsilon}{2}$ und für alle $n \geq N_2$ gilt $|a_n - A_2| < \frac{\varepsilon}{2}$. Für $N = \max\{N_1, N_2\}$ gelten also beide Aussagen und es folgt

$$\varepsilon = |A_1 - A_2| = |A_1 - a_N + a_N - A_2| \leq |a_N - A_1| + |a_N - A_2| < \varepsilon.$$

Dies ist ein Widerspruch. Das heisst, $A_1 = A_2$ und das Lemma folgt. □

Wir empfehlen Ihnen, sich ein Bild zu obigem Beweis zu zeichnen. Falls $A = \lim_{n\to\infty} a_n$ für eine Folge $(a_n)_n$, dann schreibt man oft auch $a_n \to A$ für $n \to \infty$.

**Beispiel 5.4** (Konvergente und divergente Folgen)**.**

- *Eine konstante Folge* $(a_n)_n$ *mit* $a_n = A \in \mathbb{C}$ *für alle* $n \in \mathbb{N}$ *konvergiert gegen* $A$*. Genauso konvergieren schliesslich konstante Folgen gegen den Wert, den sie schliesslich annehmen.*

- *Die Folge* $(\frac{1}{n})_n$ *konvergiert gegen Null, das heisst* $\lim_{n\to\infty} \frac{1}{n} = 0$*. Denn für alle* $\varepsilon > 0$ *existiert nach dem Archimedischen Prinzip (Satz 2.68) ein* $N \in \mathbb{N}$ *mit* $\frac{1}{N} < \varepsilon$ *und für jedes* $n \in \mathbb{N}$ *mit* $n \geq N$ *gilt nun* $0 \leq \frac{1}{n} \leq \frac{1}{N} < \varepsilon$ *(und damit* $|\frac{1}{n} - 0| = \frac{1}{n} < \varepsilon$*).*

- *Die Folge* $(a_n)_n$ *gegeben durch* $a_n = (-1)^n$ *für* $n \in \mathbb{N}$ *ist divergent, da die Folgenglieder* $1, -1, 1, -1, 1, -1, \ldots$ *zwischen* $1$ *und* $-1$ *hin und her wechseln und sich insbesondere keiner bestimmten Zahl nähern.*

  *Formal argumentiert: Für jede Zahl* $A \in \mathbb{C}$ *ist entweder* $A \neq 1$ *und somit* $\varepsilon = |A-1| > 0$ *oder* $A = 1$*. Im ersten Fall gibt es für jedes* $N \in \mathbb{N}$ *ein gerades* $n \geq N$ *mit* $a_n = (-1)^n = 1$ *und damit* $|a_n - A| = |A-1| = \varepsilon$ *(anstatt* $|a_n - A| < \varepsilon$*). Im zweiten Fall gibt es für jedes* $N \in \mathbb{N}$ *ein ungerades* $n \geq N$ *mit* $a_n = (-1)^n = -1$ *und* $|A - a_n| = |1 - a_n| = 2$ *(anstatt* $|a_n - A| < 2$*).*

Oft werden wir anstelle von Folgen in $\mathbb{C}$ auch Folgen in $\mathbb{R}$ betrachten. Ist eine solche Folge konvergent, so liegt ihr Grenzwert in $\mathbb{R}$, wie wir in folgender Übung zeigen.

**Wichtige Übung 5.5** (Reelle Grenzwerte)**.** *Sei* $(a_n)_n$ *eine konvergente Folge reeller Zahlen. Zeigen Sie, dass der Grenzwert* $\lim_{n\to\infty} a_n$ *reell ist.*

Eine Folge $(a_n)_n$ heisst **beschränkt**, falls es ein $M > 0$ gibt, so dass $|a_n| \leq M$ für alle $n \in \mathbb{N}$. Wie in Übung 3.39 kann man zeigen, dass die Menge der beschränkten Folgen in $\mathbb{C}^{\mathbb{N}}$ einen Unterraum bildet.

**Lemma 5.6** (Beschränktheit)**.** *Jede konvergente Folge ist beschränkt.*

*Beweis.* Sei $(a_n)_n$ eine konvergente Folge und $A = \lim_{n\to\infty} a_n$. Dann existiert ein $N \in \mathbb{N}$, so dass $|a_n - A| < 1$ für alle $n \geq N$. Daraus folgt

$$|a_n| = |a_n - A + A| \leq |a_n - A| + |A| < 1 + |A|$$

für alle $n \geq N$ und

$$|a_n| \leq \max\{|a_1|, |a_2|, \ldots, |a_{N-1}|, 1 + |A|\}$$

für alle $n \in \mathbb{N}$. □

Wie schon bei der Stetigkeit von Funktionen möchten wir auch hier nicht jedesmal „von Hand“ mit $\varepsilon > 0$ und $N \geq 1$ Grenzwerte berechnen müssen. Dazu ist folgende Proposition hilfreich.

**Proposition 5.7** (Additive und multiplikative Eigenschaften des Grenzwerts)**.** *Seien* $(a_n)_n$*,* $(b_n)_n$ *zwei konvergente Folgen.*

*(i) Die Folge $(a_n)_n + (b_n)_n$ ist konvergent und es gilt*

$$\lim_{n\to\infty}(a_n + b_n) = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n.$$

*(ii) Die Folge $(a_n b_n)_n$ ist konvergent und es gilt*

$$\lim_{n\to\infty}(a_n b_n) = (\lim_{n\to\infty} a_n)(\lim_{n\to\infty} b_n).$$

*Insbesondere ist für $\alpha \in \mathbb{R}$ die Folge $\alpha(a_n)_n$ konvergent und*

$$\lim_{n\to\infty}(\alpha a_n) = \alpha \lim_{n\to\infty} a_n.$$

*(iii) Angenommen $a_n \neq 0$ für alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} a_n \neq 0$. Dann ist die Folge $(\frac{1}{a_n})_n$ konvergent und es gilt*

$$\lim_{n\to\infty} \frac{1}{a_n} = \frac{1}{\lim_{n\to\infty} a_n}.$$

*Nach (i) und (ii) bildet die Menge der konvergenten Folgen in $\mathbb{C}^{\mathbb{N}}$ einen Unterraum und der Grenzwert stellt eine lineare Abbildung von diesem Unterraum nach $\mathbb{C}$ dar.*

*Beweis.* Für (i) und (ii) verweisen wir auf Übung 5.8. Also angenommen $a_n \neq 0$ für alle $n \in \mathbb{N}$ und $A = \lim_{n\to\infty} a_n \neq 0$. Dann gilt

$$\left|\frac{1}{a_n} - \frac{1}{A}\right| = \frac{|A - a_n|}{|a_n A|}$$

Wir sehen also, dass wir erzwingen können, dass $\left|\frac{1}{a_n} - \frac{1}{A}\right|$ klein ist, wenn $|A - a_n|$ klein ist. Dazu müssen wir allerdings verhindern, dass $a_n$ zu klein wird. Für den formalen Beweis sei $\varepsilon > 0$. Nach Definition von $A = \lim_{n\to\infty} a_n$ existiert ein $N \in \mathbb{N}$, so dass

$$|a_n - A| < \min\left\{\frac{|A|}{2}, \frac{\varepsilon|A|^2}{2}\right\}$$

für alle $n \geq N$. Für $n \geq N$ gilt dann nach der umgekehrten Dreiecksungleichung

$$|a_n| = |a_n - A + A| \geq |A| - |a_n - A| > |A| - \frac{|A|}{2} = \frac{|A|}{2}.$$

Also wird $a_n$ nicht zu klein und

$$\left|\frac{1}{a_n} - \frac{1}{A}\right| = \frac{|A - a_n|}{|a_n||A|} < \frac{|a_n - A|}{|A|^2/2} < \frac{\varepsilon|A|^2/2}{|A|^2/2} = \varepsilon,$$

was zu zeigen war. □

**Übung 5.8.** *Modifizieren Sie die Argumente für Proposition 3.51, um einen Beweis von Proposition 5.7 (i) und (ii) zu erhalten.*

**Übung 5.9** (Rationale Funktionen als Folgen)**.**

*(i)) Berechnen Sie folgende Grenzwerte, wenn sie existieren:*

$$\lim_{n\to\infty} \frac{7n^4+15}{3n^4+n^3+n-1}, \quad \lim_{n\to\infty} \frac{n^2+5}{n^3+n+1}, \quad \lim_{n\to\infty} \frac{n^5-10}{n^2+1}.$$

*(ii)) Formulieren und beweisen Sie allgemeine Versionen von den Beispielen in a).*

*Verwenden Sie hier und auch sonst kein früher erlerntes Kochrezept, das Sie nicht begründen können.*

Eine konvergente Folge mit Grenzwert Null wird auch eine **Nullfolge** genannt.

**Übung 5.10** (Nullfolgen und Divergenz)**.** *Sei $(a_n)_n$ eine komplex-wertige Folge mit $a_n \neq 0$ für alle $n$, so dass $(a_n^{-1})_n$ gegen $0$ konvergiert. Zeigen Sie, dass $(a_n)_n$ divergiert.*

**Lemma 5.11** (Indexverschiebung)**.** *Für eine konvergente Folge $(a_n)_n$ und $\ell \in \mathbb{N}_0$ ist die Folge $(a_{n+\ell})_n$ konvergent und es gilt*

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} a_{n+\ell}.$$

**Übung 5.12.** *Beweisen Sie Lemma 5.11.*

**Beispiel 5.13** (Geometrische Folgen)**.** *Sei $q \in \mathbb{C}$. Die Folge $n \in \mathbb{N} \mapsto q^n \in \mathbb{R}$ bezeichnen wir als **geometrische Folge** zum Skalierungsfaktor $q$. Wir untersuchen nun diese geometrische Folge auf Konvergenz.*

*(i) Für $q = 1$ ist $q^n = 1$ für alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} q^n = 1$.*

*(ii) Für $q = -1$ wissen wir bereits, dass die Folge $n \in \mathbb{N} \mapsto (-1)^n$ divergiert (also keinen Grenzwert hat).*

*(iii) Allgemeiner gilt, dass für $q \in \mathbb{C}$ mit $|q| = 1$ und $q \neq 1$ die Folge $(q^n)_n$ divergiert.*

*(iv) Für $|q| > 1$ ist $(q^n)_n$ unbeschränkt und daher divergent.*

*(v) Für $q \in \mathbb{C}$ mit $|q| < 1$ gilt $\lim_{n\to\infty} q^n = 0$.*

*Wir müssen noch (iii)-(v) beweisen. Für (iii) argumentieren wir indirekt. Angenommen $q \in \mathbb{C} \setminus \{1\}$ erfüllt $|q| = 1$ und $\lim_{n\to\infty} q^n = A$. Dann gilt*

$$A = \lim_{n\to\infty} q^n = \lim_{n\to\infty} q^{-1}q^{n+1} = q^{-1} \lim_{n\to\infty} q^{n+1} = q^{-1}A$$

*nach Proposition 5.7 und Lemma 5.11. Dies impliziert $(q^{-1}-1)A = 0$ und wegen $q \neq 1$, dass $A = 0$. Da aber $|q^n - A| = |q^n| = |q|^n = 1$ gilt, kann $A = 0$ nicht der Grenzwert der Folge sein.*

*Für (iv) sei nun $q \in \mathbb{C}$ mit $|q| > 1$. Wir zeigen, dass die Folge $(q^n)_n$ unbeschränkt ist, womit (iv) aus Lemma 5.6 folgt. Sei $M > 0$ und $x = |q| - 1 > 0$. Nach dem Archimedischen*

*Prinzip (Satz 2.68) existiert ein $N \in \mathbb{N}$ mit $1+Nx > M$. Nun ergibt die Bernoulli-Ungleichung (Lemma 3.5) $M < 1 + Nx \leq (1+x)^N = |q|^N$, womit die Behauptung gezeigt ist.*

*Für (v) sei $q \in \mathbb{C}$ mit $|q| < 1$ und sei $\varepsilon > 0$. Nach (iv) existiert ein $N \in \mathbb{N}$, so dass $|q|^{-N} > \frac{1}{\varepsilon}$. Somit gilt für alle $n \in \mathbb{N}$ mit $n \geq N$*

$$\left|q^n - 0\right| = |q|^n \leq |q|^N < \varepsilon.$$

**Übung 5.14.** *Sei $q \in \mathbb{C}$ mit $|q| < 1$. Zeigen Sie, dass $\lim_{n\to\infty} nq^n = 0$.*

**Übung 5.15** (Cesàro-Mittel)**.** *Sei $(a_n)_n$ eine konvergente Folge in $\mathbb{C}$. Zeigen Sie, dass die Folge der **Cesàro-Mittel** (auch **arithmetische Mittel** oder **Cauchy-Mittel** genannt) $(b_n)_n$ gegeben durch*

$$b_n = \frac{1}{n}\sum_{k=1}^{n} a_k$$

*für $n \in \mathbb{N}$ konvergiert und denselben Grenzwert wie $(a_n)_n$ hat.*

*Überzeugen Sie sich auch davon, dass die umgekehrte Implikation nicht gilt, das heisst, dass die Konvergenz der Cesàro-Mittel nicht Konvergenz der Folge impliziert.*

**Applet 5.16** (Einige Folgen)**.** *Wir betrachten verschiedene Folgen und können mittels Verkleinern der $x$-Achse die Konvergenz- und Divergenzeigenschaften der Folgen beobachten.*

VO: 06.11.17

### 5.1.1 Zusammenhang zur Stetigkeit

Stetigkeit lässt sich auch mit Hilfe von Folgen charakterisieren, wie wir in folgender Proposition diskutieren wollen.

**Proposition 5.17** (Folgenstetigkeit bei einem Punkt)**.** *Sei $D \subseteq \mathbb{C}$ eine Teilmenge, $f : D \to \mathbb{C}$ eine Funktion und $z_0 \in D$. Die Funktion $f$ ist genau dann stetig bei $z_0$, wenn für jede Folge $(a_n)_n$ in $D$ mit $\lim_{n\to\infty} a_n = z_0$ auch $\lim_{n\to\infty} f(a_n) = f(z_0)$ gilt.*

Sehr intuitiv ausgedrückt ist eine Funktion also stetig, wenn sie konvergente Folgen auf konvergente Folgen (mit dem richtigen Grenzwert) abbildet. Dies bezeichnet man auch als **Folgenstetigkeit**. Diese Eigenschaft kann für Grenzwertberechnungen von Folgen verwendet werden, indem man eine gegebene Folge als Bild einer Folge (deren Grenzwert idealerweise schon bekannt ist) unter einer stetigen Funktion darstellt.

*Beweis.* Angenommen $f$ ist bei $z_0$ stetig und $(a_n)_n$ ist eine Folge in $D$ mit $\lim_{n\to\infty} a_n = z_0$. Dann existiert für $\varepsilon > 0$ ein $\delta > 0$ mit

$$|z - z_0| < \delta \implies |f(z) - f(z_0)| < \varepsilon.$$

Des Weiteren existiert ein $N \in \mathbb{N}$ mit

$$n \geq N \implies |a_n - z_0| < \delta,$$

was gemeinsam

$$n \geq N \implies |f(a_n) - f(z_0)| < \varepsilon$$

ergibt. Die Folge $(f(a_n))_n$ konvergiert also gegen $f(z_0)$.

Für die Umkehrung nehmen wir an, dass $f$ nicht stetig ist in $z_0$. Dann existiert ein $\varepsilon > 0$, so dass für alle $\delta > 0$ ein $z \in D$ existiert mit

$$|z - z_0| < \delta \wedge |f(z) - f(z_0)| \geq \varepsilon$$

Wir verwenden dies für $n \in \mathbb{N}$ und $\delta = \frac{1}{n} > 0$ und finden also ein $a_n \in D$ mit

$$|a_n - z_0| < \frac{1}{n} \tag{5.1}$$

und

$$|f(a_n) - f(z_0)| \geq \varepsilon. \tag{5.2}$$

Aus Ungleichung (5.1) schliessen wir, dass die Folge $(a_n)_n$ gegen $z_0$ konvergiert und aus Ungleichung (5.2) folgt, dass $(f(a_n))_n$ nicht gegen $f(z_0)$ konvergiert. □

Nach Proposition 5.17 verfügen wir nun über zwei Varianten, wie wir Stetigkeit (oder Nicht-Stetigkeit) einer Funktion nachweisen können: mit der Definition (dem $\varepsilon, \delta$-Spiel) oder mit Konvergenz von Folgen. Letztere Variante kann unter anderem sehr nützlich sein, wenn man Nicht-Stetigkeit einer Funktion zeigen will, da man demnach bloss eine spezielle Folge konstruieren muss, die auf eine nicht-konvergente Folge abgebildet wird.

**Übung 5.18.** *Sei $D \subseteq \mathbb{C}$ eine Teilmenge und $f : D \to \mathbb{C}$ eine stetige Funktion. Angenommen $(a_n)_n$ ist eine Folge in $D$, so dass $(f(a_n))_n$ konvergiert. Muss auch $(a_n)_n$ konvergieren?*

### 5.1.2 Teilfolgen

Oft möchte man anstelle einer Folge $(a_n)_n$ nur einen „Teil“ der Folge betrachten, wobei wir im Gegensatz zur Indexverschiebung in Lemma 5.11 manchmal auch unendlich viele Folgenglieder wegstreichen wollen. Dies ist insbesondere dann der Fall, wenn die Folge nicht konvergiert.

**Definition 5.19** (Teilfolge)**.** Wenn $(a_n)_n$ eine Folge ist und $(n_k)_k : k \in \mathbb{N} \to n_k \in \mathbb{N}$ eine streng monoton wachsende Folge ist, dann wird $(a_{n_k})_k$ eine **Teilfolge** von $(a_n)_n$ genannt.

Liegt eine Teilfolge einer konvergenten Folge vor, so konvergiert diese gegen denselben Grenzwert wie die Folge.

**Lemma 5.20** (Konvergenz von Teilfolgen)**.** *Sei $(a_n)_n$ eine konvergente Folge. Jede Teilfolge $(a_{n_k})_k$ von $(a_n)_n$ konvergiert und hat denselben Grenzwert $\lim_{k\to\infty} a_{n_k} = \lim_{n\to\infty} a_n$.*

**Wichtige Übung 5.21.** *Beweisen Sie Lemma 5.20.*

Eine Folge kann konvergente Teilfolgen besitzen, ohne selbst zu konvergieren. Beispielsweise hat die Folge $n \in \mathbb{N} \mapsto (-1)^n \in \mathbb{R}$ die konvergente (konstante) Teilfolge $n \in \mathbb{N} \mapsto (-1)^{2n} \in \mathbb{R}$, konvergiert aber nicht, wie wir schon gesehen haben. In der Tat haben wir mit Lemma 5.20 jetzt ein kürzeres Argument. Falls die Folge $(a_n)_n = ((-1)^n)$ gegen $A \in \mathbb{R}$ konvergieren würde, so müssten die beiden konstanten Folgen $(a_{2n})_n$, $(a_{2n+1})_n$ auch gegen $A$ konvergieren. Dies ist natürlich nicht möglich, da die eine gegen 1 und die andere gegen $-1$ konvergiert.

In gewissen Situationen lässt sich aus dem Konvergenzverhalten von Teilfolgen trotzdem etwas über das Konvergenzverhalten der gesamten Folge sagen.

**Übung 5.22** (Teilfolgen von Teilfolgen und Konvergenz)**.** *Sei $(a_n)_n$ eine Folge und sei $A \in \mathbb{C}$. Zeigen Sie, dass die Folge $(a_n)_n$ genau dann gegen $A$ konvergiert, wenn jede Teilfolge von $(a_n)_n$ eine Teilfolge besitzt, die gegen $A$ konvergiert.*

**Proposition 5.23** (Häufungspunkte einer Folge)**.** *Sei $(a_n)_n$ eine Folge in $\mathbb{C}$. Eine Zahl $A \in \mathbb{C}$ heisst **Häufungspunkt** von $(a_n)_n$, falls die folgenden äquivalenten Bedingungen erfüllt sind.*

*(a) Es gibt eine Teilfolge $a_{n_k}$, so dass $\lim_{k\to\infty} a_{n_k} = A$.*

*(b) Für alle $\varepsilon > 0$ und $N \in \mathbb{N}$ gibt es ein $n \geq N$ mit $|a_n - A| < \varepsilon$.*

*Beweis.* Angenommen (a) gilt. Sei also $(a_{n_k})_k$ eine konvergente Teilfolge von $(a_n)_n$ mit Grenzwert $A$ und sei $\varepsilon > 0$. Dann existiert ein $K \in \mathbb{N}$ mit $|a_{n_k} - A| < \varepsilon$ für alle $k \geq K$. Sei nun $k \geq K$ mit $n_k \geq N$. Dann erfüllt $n = n_k$ die Bedingung $|a_n - A| < \varepsilon$ wie gewollt und (b) ist erfüllt.

Angenommen (b) gilt. Wir möchten rekursiv eine Teilfolge $(a_{n_k})_k$ finden mit

$$|a_{n_k} - A| < \frac{1}{k}$$

für alle $k \in \mathbb{N}$. Diese konvergiert dann gegen $A$, da für $\varepsilon > 0$ die Ungleichung $|a_{n_\ell} - A| < \varepsilon$ für alle $\ell > \frac{1}{\varepsilon}$ erfüllt ist.

Sei $\varepsilon = 1$ und $N = 1$. Dann gibt es ein $n_1 \geq N = 1$ mit $|a_{n_1} - A| < 1$. Nun nehmen wir an, dass $n_1 < n_2 < \ldots < n_k$ bereits konstruiert sind mit

$$|a_{n_\ell} - A| < \frac{1}{\ell}$$

für $\ell = 1, \ldots, k$. Wir setzen $\varepsilon = \frac{1}{k+1}$ und $N = n_k + 1$. Dann existiert nach Voraussetzung ein $n_{k+1} \geq N > n_k$ mit

$$|a_{n_{k+1}} - A| < \frac{1}{k+1}.$$

Dies beendet den Induktionsschritt und wir erhalten durch Rekursion die gewünschte Teilfolge $(a_{n_k})_k$ mit Grenzwert $A$. □

*Bemerkung.* Obiger Beweis ist formal nicht ganz unproblematisch. Denn zum Unterschied von der Rekursion, welche wir am Ende von Abschnitt 2.2.1 besprochen haben, müssen wir hier eigentlich eine Wahl für $n_{k+1}$ treffen. Da diese Wahl nicht nur einmal notwendig ist,

sondern abzählbar oft, so haben wir in obiger Formulierung eigentlich eine (schwache) Version des Auswahlaxioms verwendet. Wir werden uns diese Freiheit hier und auch in ähnlichen Situationen erlauben, ohne dieses Auswahlaxiom genauer zu besprechen. Mit ein Grund dafür ist, dass ein Grossteil der modernen Mathematik an diesem Auswahlaxiom der Axiomatischen Mengenlehre gebunden ist. Es ist aber auch möglich (wenn auch anstrengend) die Verwendung dieses Auswahlaxioms in obigem Beweis zu vermeiden indem man bei jeder Wahl sicherstellt, dass man das minimale $n_{k+1} \in \mathbb{N}$ mit allen gewünschten Eigenschaften verwendet.

### 5.1.3 Cauchy-Folgen

Wir führen eine weitere Eigenschaft von Folgen ein.

**Definition 5.24** (Cauchy-Folge)**.** Eine Folge $(a_n)_n$ in $\mathbb{C}$ ist eine Cauchy-Folge, falls es für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass

$$|a_m - a_n| < \varepsilon$$

für alle $m, n \geq N$.

Intuitiv ausgedrückt ist eine Cauchy-Folge also eine Folge, deren Folgenglieder für grosse Indizes immer näher beieinanderliegen. In dieser Formulierung macht die Aussage folgender Übung Sinn.

**Wichtige Übung 5.25** (Konvergente Folgen sind Cauchy-Folgen)**.** *Zeigen Sie, dass eine konvergente Folge in $\mathbb{C}$ eine Cauchy-Folge ist.*

Ein Ziel dieses Kapitels ist die Umkehrung zu zeigen, also dass jede Cauchy-Folge selbst bereits einen Grenzwert besitzt. Dies ist nützlich, da wir für den Beweis der Konvergenz nach Definition eigentlich den Grenzwert bereits kennen müssen. Wollen wir hingegen zeigen, dass eine Folge eine Cauchy-Folge ist, so müssen wir nur die Folgenglieder der Folge betrachten. Damit werden wir im nächsten Kapitel viele neue Funktionen definieren können.[1]

Um Konvergenz einer Cauchy-Folge zu zeigen, kann man folgendes nützliches Kriterium verwenden.

**Wichtige Übung 5.26** (Konvergente Teilfolgen von Cauchy-Folgen)**.** *Zeigen Sie, dass eine Cauchy-Folge genau dann konvergiert, wenn sie eine konvergente Teilfolge besitzt.*

### 5.1.4 Reduktion auf reelle Folgen

Anstelle von Folgen in $\mathbb{C}$ ist es meistens ausreichend, Folgen in $\mathbb{R}$ zu betrachten.

**Lemma 5.27** (Reduktion)**.** *Eine komplexwertige Folge $(a_n)_n$ ist genau dann konvergent (mit Grenzwert $a \in \mathbb{C}$), wenn die beiden reellwertigen Folgen $(\mathrm{Re}(a_n))_n$ und $(\mathrm{Im}(a_n))_n$ konvergent sind (mit Grenzwerten $\mathrm{Re}(a)$ respektive $\mathrm{Im}(a)$).*

[1] Falls wir den Grenzwert für die Konstruktion dieser Funktionen bereits kennen müssten, so könnten wir ja damit keine *neuen* Funktionen definieren.

*Beweis.* Angenommen $(a_n)_n$ konvergiert gegen $a \in \mathbb{C}$. Für $\varepsilon > 0$ existiert dann ein $N \in \mathbb{N}$, so dass $|a_n - a| < \varepsilon$ für $n \in \mathbb{N}$. Da für alle $b \in \mathbb{C}$ gilt $|\operatorname{Re}(b)| \leq |b|$ und $|\operatorname{Im}(b)| \leq |b|$, folgt

$$|\operatorname{Re}(a_n) - \operatorname{Re}(a)| = |\operatorname{Re}(a_n - a)| \leq |a_n - a| < \varepsilon$$
$$|\operatorname{Im}(a_n) - \operatorname{Im}(a)| = |\operatorname{Im}(a_n - a)| \leq |a_n - a| < \varepsilon$$

für alle $n \geq N$. Dies zeigt wie gewünscht, dass die Folgen $(\operatorname{Re}(a_n))_n$ und $((\operatorname{Im}(a_n))_n$ gegen $\operatorname{Re}(a)$ respektive $\operatorname{Im}(a)$ konvergieren.

Die Umkehrung folgt aus $a_n = \operatorname{Re}(a_n) + \operatorname{Im}(a_n)\mathrm{i}$ für alle $n \in \mathbb{N}$ und Proposition 5.7(i). □

**Übung 5.28.** *Sei $(a_n)_n$ eine konvergente Folge in $\mathbb{C}$. Zeigen Sie, dass $(|a_n|)_n$ konvergiert und geben Sie den Grenzwert an. Impliziert umgekehrt die Konvergenz von $(|a_n|)_n$ die Konvergenz von $(a_n)_n$?*

**Übung 5.29.** *Sei $(a_n)_n$ eine Folge in $\mathbb{C}$. Dann ist $(a_n)_n$ genau dann eine Cauchy-Folge, wenn $(\operatorname{Re}(a_n))_n$ und $(\operatorname{Im}(a_n))_n$ Cauchy-Folgen sind.*

## 5.2 Reelle Folgen

Wie wir im letzten Abschnitt gesehen haben, lässt sich die Konvergenz von komplexwertigen Folgen auf die Konvergenz von reellwertigen Folgen zurückführen. Unter anderem deswegen möchten wir nun reelle Folgen (also Folgen in $\mathbb{R}$) genauer betrachten und beginnen damit, das Verhalten von Ungleichungen bei Konvergenz zu untersuchen.

**Proposition 5.30** (Grenzwerte und Ungleichungen)**.** *Seien $(a_n)_n, (b_n)_n$ reelle Folgen mit Grenzwerten $a = \lim_{n\to\infty} a_n, b = \lim_{n\to\infty} b_n$.*

*(i) Falls $a_n \leq b_n$ für alle $n \in \mathbb{N}$, dann gilt auch $a \leq b$.*

*(ii) Falls $a < b$, dann existiert ein $N \in \mathbb{N}$, so dass $a_n < b_n$ für alle $n \geq N$.*

Wir möchten noch kurz anmerken, dass die Annahme der strikten Ungleichung in (i) nicht $a < b$ impliziert. In der Tat gilt für die beiden reellen Folgen $(a_n)_n, (b_n)_n$ mit $a_n = 0 < b_n = \frac{1}{n}$, dass die Grenzwerte $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = 0$ übereinstimmen.

*Beweis.* Wir beginnen mit der zweiten Behauptung. Angenommen $a < b$, dann ist $\varepsilon = \frac{b-a}{2} > 0$ und es existiert ein $N \in \mathbb{N}$ (definiert als ein Maximum), so dass

$$n \geq N \implies a - \varepsilon < a_n < a + \varepsilon$$
$$n \geq N \implies b - \varepsilon < b_n < b + \varepsilon$$

Da aber $a + \varepsilon = b - \varepsilon$ nach Wahl von $\varepsilon$, ergibt sich $a_n < a + \varepsilon = b - \varepsilon < b_n$ für alle $n \geq N$.

Dies beweist auch die erste Behauptung. Denn falls $a > b$ wäre, dann gäbe es nach obigem Folgenglieder $a_n, b_n$ mit $a_n > b_n$. □

Sie sollten sich ein Bild zu obigem Beweis überlegen, da das Bild es deutlich einfacher machen sollte, sich diesen Beweis (oder auch ähnliche Beweise) zu merken. Wir können Ungleichungen zwischen verschiedenen Folgen auch dazu verwenden, Konvergenz zu zeigen.

**Lemma 5.31** (Sandwich)**.** *Es seien $(a_n)_n, (b_n)_n, (c_n)_n$ drei reelle Folgen, so dass $a_n \leq b_n \leq c_n$ für alle $n \in \mathbb{N}$ gilt. Angenommen $(a_n)_n$ und $(c_n)_n$ sind konvergent und $A = \lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n$. Dann ist auch die Folge $(b_n)_n$ konvergent und $A = \lim_{n\to\infty} b_n$.*

**Wichtige Übung 5.32.** *Beweisen Sie Lemma 5.31.*

Lemma 5.31 ist sehr nützlich um Konvergenz spezifischer Folgen zu zeigen, wie wir in folgendem Beispiel illustrieren wollen.

**Beispiel 5.33** (Anwendungen von Lemma 5.31)**.**

*(i) Wir wissen bereits, dass $\lim_{n\to\infty} \frac{1}{n} = 0 = \lim_{n\to\infty} \frac{-1}{n}$. Für jedes $n \in \mathbb{N}$ gilt jedoch $\frac{-1}{n} \leq \frac{(-1)^n}{n} \leq \frac{1}{n}$ und somit haben wir auch $\lim_{n\to\infty} \frac{(-1)^n}{n} = 0$.*

*(ii) Für eine beliebige Zahl $a > 0$ definieren wir die Folge $(a_n)_n$ durch $a_n = \sqrt[n]{a} = a^{\frac{1}{n}}$. Wir behaupten, dass*

$$\lim_{n\to\infty} \sqrt[n]{a} = 1.$$

*Angenommen $a > 1$. Wir definieren $b_n = a_n - 1 \geq 0$ für $n \in \mathbb{N}$ und wollen zeigen, dass die Folge $(b_n)_n$ gegen Null konvergiert. Es gilt*

$$a = (1 + b_n)^n \geq 1 + nb_n \geq nb_n$$

*nach der Bernoulli-Ungleichung (Lemma 3.5). Daher ist $0 \leq b_n \leq \frac{a}{n}$ für alle $n \in \mathbb{N}$, womit $\lim_{n\to\infty} b_n = 0$ nach Lemma 5.31 und damit $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n + 1 = 1$ nach Proposition 5.7.*

*Der Fall $a \in (0, 1)$ folgt auch aus obigem, denn nach Proposition 5.7 ist*

$$\lim_{n\to\infty} \sqrt[n]{a} = \lim_{n\to\infty} \frac{1}{\sqrt[n]{a^{-1}}} = 1.$$

*(iii) Wir behaupten des Weiteren, dass*

$$\lim_{n\to\infty} \sqrt[n]{n} = 1$$

*und argumentieren dafür ähnlich, verwenden aber statt der Bernoulli-Ungleichung den Binomialsatz (Satz 3.28). Definiere $b_n = \sqrt[n]{n} - 1 \geq 0$, so dass*

$$n = (1 + b_n)^n = \sum_{k=0}^{n} \binom{n}{k} b_n^k \geq \binom{n}{2} b_n^2 = \frac{n(n-1)}{2} b_n^2$$

*für alle $n \geq 2$. Wir dividieren durch $\frac{n(n-1)}{2}$ und erhalten*

$$0 \leq b_n \leq \sqrt{\frac{2}{n-1}}$$

*für alle $n \geq 2$. Da die Wurzelfunktion nach dem Umkehrsatz (Satz 3.65) stetig ist, konvergiert nach Satz 5.17 die Folge $\left(\sqrt{\frac{2}{n}}\right)_n$ gegen Null. Verwenden wir nun die Indexverschiebung (Lemma 5.11) und das Sandwich-Lemma (Lemma 5.31) so erhalten wir $\lim_{n\to\infty} b_n = 0$ und unsere Behauptung folgt.*

### 5.2.1 Monotone Folgen

Für monotone Folgen gibt es folgendes hinreichendes und notwendiges Kriterium, das den Nachweis von Konvergenz deutlich vereinfacht. Dies bringt das erste Mal das Vollständigkeitsaxiom in Beziehung zur Konvergenz einer Folge.

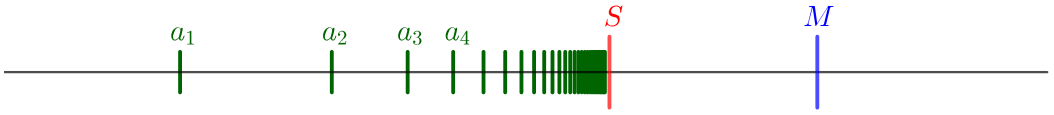
**Satz 5.34** (Konvergenz von monotonen Folgen)**.** *Eine monotone reelle Folge $(a_n)_n$ konvergiert genau dann, wenn sie beschränkt ist. Falls die Folge $(a_n)_n$ monoton wachsend ist, gilt*

$$\lim_{n\to\infty} a_n = \sup\{a_n \mid n \in \mathbb{N}\}.$$

*Falls die Folge $(a_n)_n$ monoton fallend ist, gilt*

$$\lim_{n\to\infty} a_n = \inf\{a_n \mid n \in \mathbb{N}\}.$$

Wir illustrieren den Satz an einem Bild.




Figur 5.1: Die (durch $M > 0$) beschränkte, monoton wachsende Folge $(a_n)_n$ im Bild hat keine andere Wahl als gegen das Supremum $S$ der Folgenglieder zu konvergieren. Wenn Sie sich an die charakterisierenden Eigenschaften des Supremums erinnern (was wir Ihnen empfehlen), dann können Sie den Beweis vielleicht schon selbst durchführen.

*Beweis.* Falls $(a_n)_n$ konvergent ist, ist $(a_n)_n$ beschränkt nach Lemma 5.6. Sei also $(a_n)_n$ eine beschränkte, reelle Folge. Ohne Beschränkung der Allgemeinheit können wir annehmen, dass $(a_n)_n$ monoton wachsend ist (sonst ersetzen wir einfach $(a_n)_n$ durch $(-a_n)_n$). Sei $a = \sup\{a_n \mid n \in \mathbb{N}\}$. Dann existiert nach der Charakterisierung des Supremums in Satz 2.59 für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$ mit $a_N > a - \varepsilon$. Für $n \geq N$ folgt damit aus der Monotonie von $(a_n)_n$, dass

$$a - \varepsilon < a_N \leq a_n \leq a < a + \varepsilon,$$

was zu zeigen war. □

**Übung 5.35** (Eine rekursiv definierte konvergente Folge)**.** *Sei $(a_n)_n$ die durch*

$$a_1 = 1,\ a_{n+1} = \frac{2}{3}\left(a_n + \frac{1}{a_n}\right)$$

*für $n \geq 1$ rekursiv definierte Folge. Zeigen Sie, dass $(a_n)_n$ konvergiert und bestimmen Sie den Grenzwert.*

**Übung 5.36** (Gegengerichtete Folgen)**.** *Sei $(a_n)_n$ eine monoton wachsende, reelle Folge und $(b_n)_n$ eine monoton fallende, reelle Folge mit $a_n \leq b_n$ für alle $n \in \mathbb{N}$. Zeigen Sie, dass beide Folgen konvergieren und dass $\lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n$ gilt.*

VO: 08.11.17

### 5.2.2 Limes superior und Limes inferior

Wir bringen nochmals das Vollständigkeitsaxiom (in der Form der Existenz des Supremums – Satz 2.59) in die Diskussion ein, aber betrachten im Gegensatz zum letzten Teilabschnitt eine allgemeine beschränkte reelle Folge $(a_n)_n$. Wir bemerken, dass für die Definition

der Konvergenz der Folge $(a_n)_n$ die sogenannten **Endabschnitte** $(a_k)_{k=n}^{\infty}$ der Folge (formal definiert durch die Einschränkung von $k \mapsto a_k$ auf $\{k \in \mathbb{N} \mid k \geq n\}$) eine entscheidende Rolle spielen. Deswegen definieren wir für jedes $n \in \mathbb{N}$ das Supremum

$$S_n = \sup \{a_k \mid k \geq n\}$$

über den Endabschnitt $\{a_k \mid k \geq n\}$ der Folge. Wir schreiben diese Definition auch in der Kurzform

$$S_n = \sup_{k \geq n} a_k.$$

Da $\{a_k \mid k \geq n+1\} \subseteq \{a_k \mid k \geq n\}$ ist, folgt (siehe die Bemerkung direkt nach Satz 4.19), dass $S_{n+1} \leq S_n$. Die Folge $(S_n)_n$ ist also monoton fallend. Da $(a_n)_n$ nach Annahme beschränkt ist, existiert ein $M > 0$ mit $|a_k| \leq M$ für alle $k \in \mathbb{N}$. Dies impliziert für alle $n \in \mathbb{N}$, dass $S_n \leq M$, da $M$ eine obere Schranke von $\{a_k \mid k \geq n\}$ ist, und dass $S_n \geq -M$, da $S_n \geq a_n \geq -M$. Unter dem Strich ist $(S_n)_n$ also eine monoton fallende, beschränkte Folge und muss daher nach Satz 5.34 gegen das Infimum der Folge konvergieren.

**Definition 5.37** (Limes superior)**.** Für eine beschränkte reelle Folge $(a_n)_n$ ist der **Limes superior** definiert durch

$$\overline{\lim_{n\to\infty}}\, a_n = \limsup_{n\to\infty} a_n = \lim_{n\to\infty}\Big(\sup_{k\geq n} a_k\Big) = \inf_{n\geq 1}\Big(\sup_{k\geq n} a_k\Big).$$

Es ergibt sich beispielsweise für die Folge $(a_n)_n$ gegeben durch $a_n = (-1)^n$ für $n \in \mathbb{N}$ unmittelbar $S_n = 1$ für alle $n \in \mathbb{N}$ und somit $\limsup_{n\to\infty}(-1)^n = 1$.

Sei $(a_n)_n$ eine beschränkte, reelle Folge. Analog zu obiger Diskussion definiert $I_n = \inf_{k\geq n} a_k$ für $n \in \mathbb{N}$ eine monoton wachsende, beschränkte Folge $(I_n)_n$, die nach Satz 5.34 gegen $\sup_{n\in\mathbb{N}} I_n$ konvergiert.

**Definition 5.38** (Limes inferior)**.** Für eine beschränkte, reelle Folge $(a_n)_n$ ist der **Limes inferior** definiert durch

$$\underline{\lim_{n\to\infty}}\, a_n = \liminf_{n\to\infty} a_n = \lim_{n\to\infty}\Big(\inf_{k\geq n} a_k\Big) = \sup_{n\geq 1}\Big(\inf_{k\geq n} a_k\Big).$$

Wie zuvor erhält man beispielsweise $\liminf_{n\to\infty}(-1)^n = -1$. Wir wollen die Begriffe Limes superior und Limes inferior an einem konkreten Zahlenbeispiel erproben.

**Beispiel 5.39.** *Sei $(a_n)_n$ die reelle Folge definiert durch $a_n = (-1)^n + \frac{1}{n}$ für $n \in \mathbb{N}$. Wir stellen $a_n$, $S_n = \sup_{k\geq n} a_k$, $I_n = \inf_{k\geq n} a_k$ in folgender Tabelle dar.*

| $n$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | ... |
|---|---|---|---|---|---|---|---|---|---|
| $a_n$ | $-1+\frac{1}{1}$ | $1+\frac{1}{2}$ | $-1+\frac{1}{3}$ | $1+\frac{1}{4}$ | $-1+\frac{1}{5}$ | $1+\frac{1}{6}$ | $-1+\frac{1}{7}$ | $1+\frac{1}{8}$ | ... |
| $S_n$ | $1+\frac{1}{2}$ | $1+\frac{1}{2}$ | $1+\frac{1}{4}$ | $1+\frac{1}{4}$ | $1+\frac{1}{6}$ | $1+\frac{1}{6}$ | $1+\frac{1}{8}$ | $1+\frac{1}{8}$ | ... |
| $I_n$ | $-1$ | $-1$ | $-1$ | $-1$ | $-1$ | $-1$ | $-1$ | $-1$ | ... |

*Man beachte dabei, dass* $S_n = a_n$ *ist, wenn* $n$ *gerade ist und sonst* $S_n = a_{n+1}$. *Daher ist* $\limsup_{n\to\infty}((-1)^n + \frac{1}{n}) = \lim_{n\to\infty}((-1)^{2n} + \frac{1}{2n}) = 1$. *Weiters ist wegen* $\lim_{n\to\infty} a_{2n+1} = -1$ *das Infimum* $I_n = -1$ *für alle* $n \in \mathbb{N}$ *(wieso?), womit* $\liminf_{n\to\infty}((-1)^n + \frac{1}{n}) = -1$.

**Satz 5.40** (Eigenschaften des Limes superior)**.** *Für eine reelle, beschränkte Folge* $(a_n)_n$ *erfüllt der Limes superior* $\limsup_{n\to\infty} a_n$ *die folgenden Eigenschaften:*

- *Für alle* $\varepsilon > 0$ *gibt es nur endlich viele Folgenglieder* $a_n$ *mit* $a_n > \limsup_{n\to\infty} a_n + \varepsilon$.
- *Für alle* $\varepsilon > 0$ *gibt es unendlich viele Folgenglieder* $a_n$ *mit* $a_n \geq \limsup_{n\to\infty} a_n - \varepsilon$.

Intuitiv bleibt die Folge also nicht lange weit über dem Limes superior und nähert sich immer wieder dem Limes superior. Wir empfehlen Ihnen zur Veranschaulichung die Aussagen des Satzes im Beispiel 5.39 direkt zu überprüfen und im Bild dazu zu veranschaulichen.

*Beweis.* Zur ersten Aussage: Sei $\varepsilon > 0$ und $S = \limsup_{n\to\infty} a_n$. Da $S_n = \sup_{k\geq n} a_k$ gegen $S$ konvergiert für $n \to \infty$, gibt es ein $N \in \mathbb{N}$, so dass $S_N < S + \varepsilon$. Damit ist $a_k \leq S_N < S + \varepsilon$ für alle $k \geq N$, was die erste Eigenschaft von $S = \limsup_{n\to\infty} a_n$ beweist.

Für die zweite Aussage sei $\varepsilon > 0$ und wieder $S = \limsup_{n\to\infty} a_n$. Sei $N \in \mathbb{N}$. Dann gilt $S_N \geq S = \inf_{n\geq 1} S_n$. Nach Definition von $S_N = \sup_{k\geq n} a_k$ und Satz 2.59 existiert ein $k \geq N$ mit $a_k \geq S_N - \varepsilon > S - \varepsilon$. Da $N \in \mathbb{N}$ beliebig war, beweist dies die zweite Eigenschaft des Limes superior. □

**Übung 5.41** (Charakterisierung des Limes Superior)**.** *Zeigen Sie, dass die beiden Eigenschaften in Satz 5.40 den Limes superior einer beschränkten, reellen Folge eindeutig charakterisieren.*

Der Limes Inferior hat zu dem Limes Superior analoge Eigenschaften (siehe Übung 5.42).

**Übung 5.42** (Limes superior und Limes inferior)**.** *Sei* $(a_n)_n$ *eine beschränkte reelle Folge.*

*(i) Beweisen Sie*

$$\limsup_{n\to\infty}(-a_n) = -\liminf_{n\to\infty} a_n.$$

*(ii) Verwenden Sie diese Gleichheit, um eine analoge Version von Satz 5.40 für den Limes inferior zu formulieren und zu beweisen.*

*(iii) Zeigen Sie* $\liminf_{n\to\infty} a_n \leq \limsup_{n\to\infty} a_n$.

*(iv) Zeigen Sie, dass der Limes superior als Abbildung definiert auf dem Unterraum der beschränkten Folgen in* $\mathbb{R}^{\mathbb{N}}$ *mit Zielraum* $\mathbb{R}$ *nicht linear ist. Genauer: Verifizieren Sie, dass sich der Limes superior weder unter Addition noch unter skalarer Multiplikation geeignet verhält.*

**Korollar 5.43** (Charakterisierung der Konvergenz)**.** *Für eine reelle, beschränkte Folge* $(a_n)_n$ *gilt* $\liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n$ *genau dann wenn* $(a_n)_n$ *konvergent ist.*

*Beweis.* Angenommen $A = \liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n$ und $\varepsilon > 0$. Dann existiert nach Satz 5.40 ein $N$ so dass $a_n \leq A + \varepsilon/2$ für alle $n \geq N$ (auf Grund der Eigenschaften des Limes Superior) und $a_n \geq A - \varepsilon/2$ für alle $n \geq N$ (auf Grund der Eigenschaften des Limes Inferior). Zusammen erhalten wir $|a_n - A| < \varepsilon$ für alle $n \geq N$, was zu beweisen war.

Wir nehmen nun an, dass $A = \lim_{n\to\infty} a_n$ existiert. Sei $\varepsilon > 0$. Dann existiert ein $N$ so dass $A - \varepsilon < a_n < A + \varepsilon$ für alle $n \geq N$. Aus $I_n = \inf\{a_k \mid k \geq n\}$ und $S_n = \sup\{a_k \mid k \geq n\}$ folgt nun

$$A - \varepsilon \leq I_N \leq I_n \leq S_n \leq S_N \leq A + \varepsilon$$

für alle $n \geq N$, und daher

$$A - \varepsilon \leq \liminf_{n\to\infty} a_n \leq \limsup_{n\to\infty} a_n \leq A + \varepsilon.$$

Da dies für alle $\varepsilon > 0$ gilt, folgt daraus $A = \liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n$. □

### 5.2.3 Konvergente Teilfolgen

Wir betrachten wiederum eine beschränkte, reelle Folge $(a_n)_n$ und wollen die Konvergenz von Teilfolgen von $(a_n)_n$ untersuchen.

**Satz 5.44** (Konvergenz von Teilfolgen)**.** *Für jede konvergente Teilfolge* $(a_{n_k})_k$ *einer beschränkten, reellen Folge* $(a_n)_n$ *gilt*

$$\lim_{k\to\infty} a_{n_k} \in [\liminf_{n\to\infty} a_n, \limsup_{n\to\infty} a_n].$$

*Des Weiteren existiert eine konvergente Teilfolge* $(a_{n_k})_k$ *mit* $\lim_{k\to\infty} a_{n_k} = \limsup_{n\to\infty} a_n$ *und eine konvergente Teilfolge* $(a_{m_k})_k$ *mit* $\lim_{k\to\infty} a_{m_k} = \liminf_{n\to\infty} a_n$.

Betrachtet man die Folge $(a_n)_n$ mit $a_n = (-1)^n$ für $n \in \mathbb{N}$, so erfüllt beispielsweise die Teilfolge $(a_{2k})_k$, dass $\lim_{k\to\infty} a_{2k} = \limsup_{n\to\infty} a_n = 1$, und die Teilfolge $(a_{2k+1})_k$, dass $\lim_{k\to\infty} a_{2k+1} = \liminf_{n\to\infty} a_n = -1$, wie wir schon gesehen haben.

*Beweis von Satz 5.44.* Sei $(a_{n_k})_k$ eine konvergente Teilfolge von $(a_n)_n$, $I = \liminf_{n\to\infty} a_n$, $S = \limsup_{n\to\infty} a_n$ und $\varepsilon > 0$. Dann gibt es nach der ersten Eigenschaft in Satz 5.40 ein $N \in \mathbb{N}$, so dass

$$a_n \leq S + \varepsilon \tag{5.3}$$

für alle $n \geq N$. Wenn nötig können wir $N$ noch grösser wählen, so dass ebenso gilt

$$a_n \geq I - \varepsilon \tag{5.4}$$

für alle $n \geq N$ (siehe auch Übung 5.42). Insbesondere gelten (5.3) und (5.4) auch für $a_{n_k}$ und genügend grosse $k$ (zum Beispiel $k \geq N$, da dann $n_k \geq k \geq N$). Für den Limes der Folge $(a_{n_k})_k$ ergibt sich daraus

$$I - \varepsilon \leq \lim_{k\to\infty} a_{n_k} \leq S + \varepsilon.$$

(siehe Proposition 5.30 und Lemma 5.11). Da $\varepsilon > 0$ beliebig war und $\lim_{k\to\infty} a_{n_k}$ nicht von $\varepsilon$ abhängt, ergibt sich daraus $I \leq \lim_{k\to\infty} a_{n_k} \leq S$, wie im Satz behauptet wurde.

Wir wollen nun eine konvergente Teilfolge von $(a_n)_n$ mit Grenzwert $\limsup_{n\to\infty} a_n$ finden. In anderen Worten wollen wir also zeigen, dass $\limsup_{n\to\infty} a_n$ ein Häufungspunkt der Folge $(a_n)_n$ ist. Wir bedienen uns dabei der zweiten äquivalenten Bedingung in Proposition 5.23. Sei also $\varepsilon > 0$ und $N \in \mathbb{N}$. Dann existiert ein $M \geq N$ mit $S \leq S_M < S + \varepsilon$, da

$$S = \lim_{n\to\infty} S_n = \inf \{S_n \mid n \in \mathbb{N}\}.$$

Auf Grund der Definition des Supremums existiert damit ein $n \geq M \geq N$ mit

$$S - \varepsilon \leq S_M - \varepsilon < a_n < S_M < S + \varepsilon.$$

Der Beweis der Existenz einer Teilfolge mit Grenzwert $\liminf_{n\to\infty} a_n$ ist analog. □

Wir wollen noch einen zweiten Beweis der Existenz von konvergenten Teilfolgen unter Verwendung des Intervallschachtelungsprinzips andeuten. Sei $(a_n)_n$ eine reelle, beschränkte Folge und $c_1, d_1 \in \mathbb{R}$, so dass $c_1 \leq a_n \leq d_1$ für alle $n \in \mathbb{N}$. Wir definieren $n_1 = 1$. Als nächsten Schritt teilen wir das Intervall $[c_1, d_1]$ in zwei Hälften

$$[c_1, d_1] = \big[c_1, \frac{c_1 + d_1}{2}\big] \cup \big[\frac{c_1 + d_1}{2}, d_1\big]$$

auf und erkennen, dass zumindest eine der beiden Hälften unendlich viele Folgenglieder enthalten muss. Wir definieren $c_2 = c_1$, $d_2 = \frac{c_1+d_1}{2}$, falls $a_n \in [c_1, \frac{c_1+d_1}{2}]$ für unendlich viele $n \in \mathbb{N}$ und ansonsten $c_2 = \frac{c_1+d_1}{2}$, $d_2 = d_1$. Insbesondere existiert ein $n_2 > n_1$ mit $a_{n_2} \in [c_2, d_2]$. Wiederum teilen wir $[c_2, d_2]$ in zwei Hälften auf und wählen eine Hälfte $[c_3, d_3]$ aus, die unendlich viele Folgenglieder $a_n$ enthält. Auch wählen wir $n_3 > n_2$ mit $a_{n_3} \in [c_3, d_3]$.

Iterieren wir dieses Argument, so erhalten wir zwei weitere Folgen $(c_k)_k, (d_k)_k$ und eine Teilfolge $(a_{n_k})_k$ von $(a_n)_n$, so dass

$$c_k \leq a_{n_k} \leq d_k, \quad d_k - c_k = \frac{d_1 - c_1}{2^{k-1}}, \quad [c_{k+1}, d_{k+1}] \subseteq [c_k, d_k]$$

für alle $k \in \mathbb{N}$. Nach Satz 2.77 (wo ebenso wie in Satz 5.44 die Existenz von Supremum und Infimum verwendet wurde) gibt es ein $A \in \bigcap_{k\in\mathbb{N}} [c_k, d_k]$. Aus $c_k \leq A \leq d_k$ und $c_k \leq a_{n_k} \leq d_k$ folgt

$$A - \frac{d_1 - c_1}{2^{k-1}} \leq a_{n_k} \leq A + \frac{d_1 - c_1}{2^{k-1}}$$

für alle $k \in \mathbb{N}$. Nach Lemma 5.31 und Beispiel 5.13 gilt nun $\lim_{k\to\infty} a_{n_k} = A$. Dieses Argument beweist die Existenz von konvergenten Teilfolgen, ohne diese aber in Zusammenhang mit dem Limes Superior und dem Limes Inferior wie in Satz 5.44 zu bringen.

Für eine beschränkte, reelle Folge $(a_n)_n$ liegen nach Satz 5.44 alle Häufungspunkte zwischen $\liminf_{n\to\infty} a_n$ und $\limsup_{n\to\infty} a_n$ und diese beiden Punkte sind Häufungspunkte von $(a_n)_n$. Lässt man die Annahme der Beschränktheit fallen, so muss eine Folge nicht unbedingt Häufungspunkte besitzen. Ein Beispiel einer Folge ohne Häufungspunkte ist $(n)_n$.

Wir möchten auch anmerken, dass die Menge der Häufungspunkte einer Folge $(a_n)_n$ in $\mathbb{R}$ nicht gleich den Häufungspunkten der Teilmenge $\{a_n \mid n \in \mathbb{N}\}$ (siehe Definition 2.73) sein muss. Beispielsweise hat die Folge $(a_n)_n$ gegeben durch $a_n = (-1)^n$ für $n \in \mathbb{N}$ die beiden Häufungspunkte $-1$ und $1$ und die Menge $\{-1, 1\}$ hat aber keine Häufungspunkte.

**Übung 5.45** (Häufungspunkte)**.**

*(i) Finden Sie eine Folge $(a_n)_n$, so dass jede reelle Zahl $A$ ein Häufungspunkt der Folge $(a_n)_n$ ist.*

*(ii) Sei $(a_n)_n$ eine Folge in $\mathbb{R}$. Zeigen Sie, dass ein Häufungspunkt der Menge $\{a_n \mid n \in \mathbb{N}\}$ auch ein Häufungspunkt der Folge $(a_n)_n$ ist.*

*(iii) Sei $(a_n)_n$ eine Folge in $\mathbb{R}$ und nehmen Sie an, dass $n \in \mathbb{N} \mapsto a_n$ injektiv ist. Zeigen Sie, dass ein Häufungspunkt der Folge $(a_n)_n$ in diesem Fall auch ein Häufungspunkt der Menge $\{a_n \mid n \in \mathbb{N}\}$ ist.*

**Übung 5.46** (Mischung von drei Folgen)**.** *Seien $(a_n)_n, (b_n)_n, (c_n)_n$ konvergente reelle Folgen und seien $A = \lim_{n\to\infty} a_n$, $B = \lim_{n\to\infty} b_n$, $C = \lim_{n\to\infty} c_n$. Betrachten Sie die reelle Folge $(x_n)_n$ definiert durch*

$$x_n = \begin{cases} a_n & \textit{falls } 3|n \\ b_n & \textit{falls } 3|n-1 \\ c_n & \textit{falls } 3|n-2 \end{cases}$$

*für $n \in \mathbb{N}$ und berechnen Sie $\limsup_{n\to\infty} x_n$, $\liminf_{n\to\infty} x_n$ und die Menge der Häufungspunkte von $(x_n)_n$.*

Zu jeder natürlichen Zahl $K$ lässt sich eine Folge konstruieren, die $K$ Häufungspunkte hat. Die Folge $(a_n)_n$ definiert für $n \in \mathbb{N}$ durch $a_n = j$ falls $K|n-j$ und $j \in \{1, \ldots, K\}$ ist ein Beispiel dafür. Es lassen sich auch Folgen mit abzählbar vielen Häufungspunkten konstruieren. In folgender Übung zeigt sich, dass auch Folgen existieren, deren Häufungspunkte ein ganzes abgeschlossenes Intervall bilden.

**Übung 5.47** (Folgen mit vielen Häufungspunkten)**.**

*(i) Angenommen $(a_n)_n$ ist eine beschränkte reelle Folge mit $\lim_{n\to\infty}(a_{n+1} - a_n) = 0$. Sei $I = \liminf_{n\to\infty} a_n$ und $S = \limsup_{n\to\infty} a_n$. Dann ist die Menge der Häufungspunkte gegeben durch $[I, S]$.*

*(ii) Konstruieren Sie ein Beispiel einer Folge wie in a) mit* $[I, S] = [0, 1]$

*Hinweis für a): Für jedes* $L \in (I, S)$, $\varepsilon > 0$ *und* $N \in \mathbb{N}$ *mit* $|a_{k+1} - a_k| < \varepsilon$ *für alle* $k \geq N$ *gibt es ein* $n \geq N$ *mit* $a_n > S - \varepsilon > L - \varepsilon$. *Zeigen Sie, dass es auch ein* $n \geq N$ *gibt mit* $a_n \in (L - \varepsilon, L + \varepsilon)$.

*Hinweis für b): Setzen Sie* $a_1 = 0$, $a_2 = 1$, $a_3 = 1$, $a_4 = \frac{1}{2}$, $a_5 = 0$, $a_6 = 0$, $a_7 = \frac{1}{3}$, $a_8 = \frac{2}{3}$, $a_9 = 1$, $a_{10} = 1$, $a_{11} = \frac{3}{4}$, *... geeignet zu einer Folge fort.*

Wir wollen noch einen wichtigen topologischen Begriff einführen, welchen wir im zweiten Semester genauer untersuchen wollen. Eine Teilmenge $K \subseteq \mathbb{R}$ heisst **folgenkompakt** falls jede Folge in $K$ eine konvergente Teilfolge mit Grenzwert in $K$ besitzt. Wir bemerken, dass für $a < b$ in $\mathbb{R}$ das kompakte Intervall $[a, b]$ tatsächlich folgenkompakt ist. Denn nach Satz 5.44 hat jede Folge in $[a, b]$ eine konvergente Teilfolge, für welche wegen Proposition 5.30 der Grenzwert wieder in $[a, b]$ liegt.

### 5.2.4 Reelle Cauchy-Folgen

Wir haben Cauchy-Folgen bereits in Abschnitt 5.1.3 eingeführt und bemerkt, dass wir, um Konvergenz einer solchen Folge zu zeigen, nur eine konvergente Teilfolge finden müssen. In der Zwischenzeit haben wir Methoden gelernt, um genau dies zu erreichen und können deswegen nun folgenden Satz beweisen.

**Satz 5.48** (Cauchy-Kriterium für Folgen)**.** *Eine reelle Folge ist genau dann konvergent, wenn sie eine Cauchy-Folge ist.*

Der Begriff der Cauchy-Folge gemeinsam mit Satz 5.48 haben gegenüber der Definition der Konvergenz den entscheidenden Vorteil, dass wir den Grenzwert nicht kennen müssen, um zu zeigen, dass eine Folge eine Cauchy-Folge ist (und damit nach Satz 5.48 einen Grenzwert besitzt). Dies ist zum Beispiel dann wichtig, wenn man eine Zahl (die man nicht konkret kennt) als Grenzwert einer Folge definieren will (die man kennt). Des Weiteren hat Satz 5.48 gegenüber Satz 5.34 den Vorteil, dass er nicht nur für spezielle Folgen anwendbar ist.

*Beweis.* Angenommen $(a_n)_n$ ist eine reelle Folge mit $a_n \to A \in \mathbb{R}$ für $n \to \infty$. Sei $\varepsilon > 0$. Dann existiert ein $N \in \mathbb{N}$, so dass für alle $n \geq N$ gilt $|a_n - A| < \frac{\varepsilon}{2}$. Für $m, n \geq N$ gilt somit auch

$$|a_m - a_n| \leq |a_m - A| + |A - a_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Dies beweist, dass $(a_n)_n$ eine Cauchy-Folge ist.

Sei nun umgekehrt $(a_n)_n$ eine Cauchy-Folge. Für $\varepsilon = 1$ existiert dann ein $N \in \mathbb{N}$, so dass $|a_m - a_n| < 1$ für $m, n \geq N$. Insbesondere gilt also

$$|a_n| \leq |a_n - a_N| + |a_N| < 1 + |a_N|$$

für alle $n \geq N$. Daher ist $(a_n)_n$ eine beschränkte Folge (wieso? – siehe auch den Beweis von Lemma 5.6). Des Weiteren existiert nach Annahme für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$, so dass

$|a_m - a_n| < \varepsilon$ für alle $m, n \geq N$. Wir setzen $m = N$ und erhalten

$$a_m - \varepsilon < a_n < a_m + \varepsilon.$$

Wir betrachten nun Limes Inferior und Limes Superior der Folge (welche ja nach Satz 5.44 Grenzwerte konvergenter Teilfolgen sind) und erhalten

$$a_m - \varepsilon \leq \liminf_{n\to\infty} a_n \leq \limsup_{n\to\infty} a_n \leq a_m + \varepsilon.$$

Insbesondere gilt $|\liminf_{n\to\infty} a_n - \limsup_{n\to\infty} a_n| \leq 2\varepsilon$. Da aber $\varepsilon > 0$ beliebig war, erhalten wir Gleichheit von Limes Superior und Limes Inferior und daher Konvergenz der Folge nach Korollar 5.43. □

**Beispiel 5.49** (Falsches Kriterium)**.** *Sei $(a_n)_n$ eine reelle Folge. Wir behaupten, dass die Bedingung*

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \geq N : |a_n - a_{n+1}| < \varepsilon$$

*nicht zur Konvergenz der Folge äquivalent ist.*

*Wir setzen $1, 2, 2, 2 + \frac{1}{2}, 3, 3 + \frac{1}{3}, 3 + \frac{2}{3}, 4, 4, 4 + \frac{1}{4}, 4 + \frac{2}{4}, 4 + \frac{3}{4}, 5, 5, \ldots$ zu einer Folge fort, in dem wir alle rationalen Zahlen zwischen $\ell$ und $\ell + 1$ mit Nenner $\ell$ für $\ell \in \mathbb{N}$ aufsteigend auflisten. Falls $\ell \leq a_n \leq a_{n+1} \leq \ell + 1$ für $n \in \mathbb{N}$, dann haben $a_n$ und $a_{n+1}$ höchstens Abstand $\frac{1}{\ell}$. Dies zeigt, dass $\lim_{n\to\infty} |a_{n+1} - a_n| = 0$, aber $(a_n)_n$ ist trotzdem nicht beschränkt und insbesondere nicht konvergent.*

Die Aussage von Satz 5.48 ist fundamental für die Analysis und ist das, was man in einem allgemeineren Kontext unter Vollständigkeit der reellen Zahlen versteht (siehe auch Kapitel 9). In anderen Worten ist Satz 5.48 zum Vollständigkeitsaxiom äquivalent im Sinne der folgenden Übung.

**Übung 5.50** (Vollständigkeit der reellen Zahlen)**.** *Wir möchten hier erklären, wie aus dem Archimedischen Prinzip und der Aussage von Satz 5.48 das Vollständigkeitsaxiom folgt. Genauer sei $\mathbb{R}$ ein geordneter Körper, der $\mathbb{Q}$ als dichte Teilmenge enthält (im Sinne von Korollar 2.70) und in dem alle Cauchy-Folgen konvergent sind. Zeigen Sie, dass $\mathbb{R}$ das Vollständigkeitsaxiom erfüllt.*

### 5.2.5 Uneigentliche Grenzwerte

Folgende Erweiterung des Konvergenzbegriffs ist für uns auch wichtig.

**Definition 5.51** (Uneigentliche Grenzwerte)**.** Eine reelle Folge $(a_n)_n$ **divergiert gegen** $\infty$, wenn für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert, so dass $a_n \geq \varepsilon^{-1}$ gilt für alle $n \geq N$. In diesem Fall schreiben wir $\lim_{n\to\infty} a_n = \infty$. Genauso sagen wir, dass $(a_n)_n$ **gegen** $-\infty$ **divergiert**, falls für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert, so dass $a_n \leq -\varepsilon^{-1}$ für alle $n \geq N$. In letzterem Fall schreiben wir $\lim_{n\to\infty} a_n = -\infty$. In beiden Fällen spricht man auch von **uneigentlicher Konvergenz** und **uneigentlichen Grenzwerten**.

Man beachte, dass eine divergente Folge (oder auch eine unbeschränkte Folge) nicht gegen $\infty$ oder $-\infty$ divergieren muss. Zum Beispiel ist die Folge $(a_n)_n$ definiert durch

$$a_n = \begin{cases} 1 & \text{falls } n \text{ gerade} \\ n & \text{falls } n \text{ ungerade} \end{cases}$$

divergent, aber sie divergiert nicht gegen $\infty$. Es gilt jedoch folgende Aussage.

**Übung 5.52** (Teilfolgen von unbeschränkten Folgen)**.** *Sei $(a_n)_n$ eine unbeschränkte, reelle Folge. Zeigen Sie, dass eine Teilfolge existiert, die gegen $\infty$ oder $-\infty$ divergiert.*

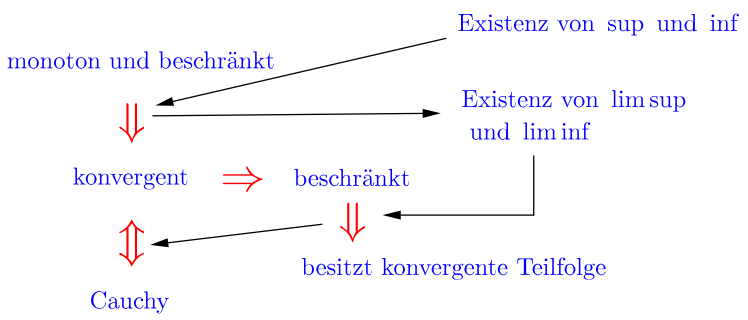
Wir können auch uneigentliche Grenzwerte verwenden, um den Limes superior und den Limes inferior für unbeschränkte Folgen zu definieren. In der Tat, falls die reellwertige Folge $(a_n)_n$ nicht von oben beschränkt ist, dann gilt $\sup_{k\geq n} a_k = \infty$ für alle $n \in \mathbb{N}$ (siehe Abschnitt 2.5.3) und wir setzen $\limsup_{n\to\infty} a_n = \infty$. Falls $(a_n)_n$ zwar von oben, aber nicht von unten beschränkt ist, dann setzen wir $\limsup_{n\to\infty} a_n = \lim_{n\to\infty}\big(\sup_{k\geq n} a_k\big)$ (wobei aber $\sup_{k\geq n} a_k$ möglicherweise gegen $-\infty$ divergiert). Diese Erweiterungen der Definitionen gelten analog für den Limes inferior.

**Übung 5.53** (Sandwich für uneigentlich Grenzwerte)**.** *Beweisen Sie folgende uneigentliche Sandwich-Lemmata für zwei reelle Folgen $(a_n)_n$ und $(b_n)_n$ mit $a_n \leq b_n$ für alle $n \in \mathbb{N}$:*

$$\lim_{n\to\infty} a_n = \infty \implies \lim_{n\to\infty} b_n = \infty$$
$$\lim_{n\to\infty} b_n = -\infty \implies \lim_{n\to\infty} a_n = -\infty.$$

### 5.2.6 Ein Diagramm für die Zusammenhänge der Begriffe und Sätze

Wir fassen obiges Wissen über reelle Folgen in folgendem Diagramm zusammen und empfehlen Ihnen sich zu überlegen, was genau die einzelnen Pfeile bedeuten und welche Sätze sie andeuten.




Figur 5.2: Einige wichtige Eigenschaften von reellen Folgen und wichtige Sätze, die diese Eigenschaften in Verbindung bringen.

## 5.3 Die Exponentialfunktion

Wir werden jetzt Grenzwerte von Folgen und insbesondere Satz 5.34 anwenden, um die Exponentialfunktion zu definieren und einige ihrer Eigenschaften zu beweisen. Die **Exponentialfunktion** $\exp : \mathbb{R} \to \mathbb{R}_{>0}$ ist definiert durch

$$\exp(x) = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n > 0 \tag{5.5}$$

für alle $x \in \mathbb{R}$. Des Weiteren ist die **Eulersche Zahl** definiert als

$$\mathrm{e} = \exp(1) = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n \in [2, 3].$$

Wir wollen zeigen, dass (5.5) Sinn ergibt (also der Grenzwert tatsächlich existiert) und dass dadurch die Abbildung $\exp : \mathbb{R} \to \mathbb{R}_{>0}$ definiert wird. Dies führt uns dann auch zum natürlichen Logarithmus und zu allgemeinen Potenzfunktionen.

**Proposition 5.54** (Reelle Exponentialfunktion)**.** *Für alle $x \in \mathbb{R}$ existiert der Grenzwert in* (5.5) *und dies definiert die streng monotone, bijektive, stetige Abbildung* $\exp : \mathbb{R} \to \mathbb{R}_{>0}$*, die die Additionsformel*

$$\exp(x + y) = \exp(x)\exp(y) \tag{5.6}$$

*für alle $x, y \in \mathbb{R}$ erfüllt.*

Der Beweis der Proposition erfolgt in den Unterabschnitten 5.3.2–5.3.7.

VO: 09.11.17

### 5.3.1 Eine Interpretation

Die Definition (5.5) hat für $x \in (0, 1)$ folgende ökonomische Interpretation. Angenommen $x$ steht für den jährlichen Zinssatz in der Bank 1. Bank 2 verrechnet die Zinsen halbjährlich und gibt $\frac{x}{2}$ Zinsen in einem halben Jahr, ..., die Bank $n$ verrechnet die Zinsen $n$-mal im Jahr und gibt in einem $n$-tel Jahr genau $\frac{x}{n}$ Zinsen. Bei welcher Bank sollte man sein Geld deponieren? Auf Grund des Zinseszinses sollte man wahrscheinlich Kunde der Bank mit dem grössten $n$ werden. Also drängt sich die Vermutung auf, dass $a_n = (1 + \frac{x}{n})^n$ eine monoton wachsende Folge ist. Aber kann man seinen jährlichen Gewinn grenzenlos steigern, in dem man immer weiter sucht und bei einer Bank mit noch grösserem $n$ um ein Konto anfragt? Dies klingt vielleicht ein bisschen zu optimistisch. Es drängt sich also die Vermutung auf, dass $(a_n)_n$ eine beschränkte monoton wachsende Folge ist.

### 5.3.2 Konvergenz der Folge

Sei $x \in \mathbb{R}$ fest gewählt. Falls $x \geq 0$ ist, dann ist

$$\frac{x}{(n+1)(n+x)} \leq \frac{x+n}{(n+1)(n+x)} \leq 1$$

und damit

$$a_{n,x} = -\frac{x}{(n+1)(n+x)} \geq -1$$

für alle $n \in \mathbb{N}$. Ansonsten ist $x < 0$ und es gelten obige Ungleichungen zumindest für alle $n \in \mathbb{N}$ mit $n > -x$. Für diese $n \in \mathbb{N}$ können wir die Bernoulli-Ungleichung in Lemma 3.5 verwenden und erhalten

$$\begin{aligned}
\frac{\left(1+\frac{x}{n+1}\right)^{n+1}}{\left(1+\frac{x}{n}\right)^{n}} &= \left(1+\frac{x}{n}\right)\left(\frac{1+\frac{x}{n+1}}{1+\frac{x}{n}}\right)^{n+1} = \frac{n+x}{n}\left(\frac{(n+1+x)n}{(n+1)(n+x)}\right)^{n+1} \\
&= \frac{n+x}{n}\left(\frac{n^2+nx+n}{(n+1)(n+x)}\right)^{n+1} = \frac{n+x}{n}\left(\frac{(n+1)(n+x)-x}{(n+1)(n+x)}\right)^{n+1} \\
&= \frac{n+x}{n}\left(1-\frac{x}{(n+1)(n+x)}\right)^{n+1} = \frac{n+x}{n}\left(1+a_{n,x}\right)^{n+1} \\
&\geq \frac{n+x}{n}\left(1+(n+1)a_{n,x}\right) = \frac{n+x}{n}\left(1-\frac{x}{n+x}\right) = 1.
\end{aligned}$$

Für $x \geq 0$ beweist dies die Monotonie der Folge $\left(1+\frac{x}{n}\right)^n$. Für $x < 0$ beweist dies die „schlussendliche“ Monotonie. Genauer formuliert existiert ein $N_x \in \mathbb{N}$ so dass für alle $n \geq N_x$ sowohl $1+\frac{x}{n} > 0$ also auch $\left(1+\frac{x}{n+1}\right)^{n+1} \geq \left(1+\frac{x}{n}\right)^n$ gilt. Da monoton wachsende, beschränkte Folgen konvergieren (Satz 5.34) und da die ersten paar Glieder der Folge nicht über Konvergenz entscheiden (Lemma 5.11) reicht es für die Konvergenz somit, Beschränktheit zu zeigen.

Für $x \leq 0$ gilt $\left(1+\frac{x}{n}\right)^n \leq 1$. Daher gilt

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = \sup\left\{\left(1+\tfrac{x}{n}\right)^n \mid n \geq N_x\right\} \in (0,1],$$

wobei $N_x$ wie oben gewählt wurde.

Für $x \geq 0$ verwenden wir

$$\left(1+\frac{x}{n}\right)^n\left(1-\frac{x}{n}\right)^n = \left(1-\frac{x^2}{n^2}\right)^n \leq 1,$$

woraus für alle $n \geq N_{-x}$ die Abschätzung

$$\left(1+\frac{x}{n}\right)^n \leq \left(1-\frac{x}{n}\right)^{-n} = a_n$$

folgt. Da aber die Folge $a_n$ auf Grund von obigem und Proposition 5.7(iii) konvergent und damit beschränkt ist, folgt nun die Beschränktheit der Folge $\left(\left(1+\frac{x}{n}\right)^n\right)_n$.

Wir wollen ein zweites Argument für die Beschränktheit der Folge für ein $x \geq 0$ skizzieren. Hierfür betrachten wir für ein $n \in \mathbb{N}$ die Umformung

$$\begin{aligned}\left(1+\frac{x}{n}\right)^n &= \sum_{k=0}^{n} \binom{n}{k} \left(\frac{x}{n}\right)^k = \sum_{k=0}^{n} \frac{1}{k!}\left(\prod_{j=n-k+1}^{n} j\right)\frac{1}{n^k}x^k = \sum_{k=0}^{n} \frac{1}{k!}x^k \frac{1}{n^k}\prod_{\ell=0}^{k-1}(n-\ell) \\ &= \sum_{k=0}^{n} \frac{1}{k!}x^k \prod_{\ell=0}^{k-1} \frac{n-\ell}{n} = \sum_{k=0}^{n} \frac{1}{k!}x^k \prod_{\ell=0}^{k-1}\left(1-\frac{\ell}{n}\right).\end{aligned}$$

unter Verwendung des Binomialsatz (Satz 3.28). Damit erhalten wir für $x \in (0,1]$, dass

$$\left(1+\frac{x}{n}\right)^n = \sum_{k=0}^{n} \frac{1}{k!}\left(\prod_{\ell=0}^{k-1}\left(1-\frac{\ell}{n}\right)\right)x^k \leq \sum_{k=0}^{n} \frac{x^k}{k!} \leq 1 + \sum_{k=1}^{n} \frac{1}{2^{k-1}} = 1 + \frac{1-\frac{1}{2^n}}{1-\frac{1}{2}} \leq 3,$$

wobei wir $k! \geq 2^{k-1}$ für $k \in \mathbb{N}$ und die geometrische Summenformel (Proposition 3.8) verwendet haben.

**Übung 5.55** (Alternative obere Schranke)**.** *Verallgemeinern Sie obige Abschätzung für beliebige $x \geq 0$.*

*Hinweis: Für $x \in [0,1]$ und $\ell, n \in \mathbb{N}$ können Sie die Abschätzung*

$$\left(1+\frac{\ell x}{n}\right)^n \leq \left(1+\frac{x}{n}\right)^{\ell n} \leq 3^\ell$$

*beweisen und verwenden.*

Auf Grund von Satz 5.34 ergibt sich daher, dass

$$\exp(x) = \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n \in \mathbb{R}_{>0}$$

für alle $x \in \mathbb{R}$ existiert. Insbesondere für $x = 1$ erhalten wir $\mathrm{e} = \exp(1) \in [2,3]$ auf Grund obiger Abschätzungen.

Für ein beliebiges $x \in \mathbb{R}$ ist $\frac{x}{n} \geq -1$ für alle hinreichend grossen $n \in \mathbb{N}$ und damit $1 + x \leq 1 + n\frac{x}{n} \leq (1+\frac{x}{n})^n$ nach der Bernoulli-Ungleichung (Lemma 3.5). Daraus folgt

$$1 + x \leq \exp(x) \tag{5.7}$$

für alle $x \in \mathbb{R}$.

**Übung 5.56** (Rosinen im Brot)**.** *Angenommen wir schneiden ein Brot, das $n = 10$ Rosinen enthält, in $n$ Stücke. Wir nehmen nun ein Stück. Wie gross ist die Wahrscheinlichkeit, dass dieses keine Rosine enthält? Wie verhält sich diese Wahrscheinlichkeit für $n \to \infty$.*

**Übung 5.57** (Quadratisches Wachstum)**.** *Zeigen Sie, dass für $x \geq 0$ gilt $1 + x + \frac{x^2}{2} \leq \exp(x)$.*

### 5.3.3 Inversionsformel

Wir behaupten nun, dass

$$\exp(-x) = \exp(x)^{-1} \tag{5.8}$$

für alle $x \in \mathbb{R}$. Es gilt

$$\exp(x)\exp(-x) = \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n \lim_{n\to\infty}\left(1-\frac{x}{n}\right)^n = \lim_{n\to\infty}\left(1-\frac{x^2}{n^2}\right)^n$$

Wir betrachten also die Folge $(b_n)_n$ definiert durch $b_n = \left(1-\frac{x^2}{n^2}\right)^n$. Nach der Bernoulli-Ungleichung gilt für $n \in \mathbb{N}$ mit $n \geq |x|$ (und damit $-\frac{x^2}{n^2} \geq -1$)

$$1-\frac{x^2}{n} = 1 + n\left(-\tfrac{x^2}{n^2}\right) \leq \left(1-\tfrac{x^2}{n^2}\right)^n = b_n \leq 1,$$

was gemeinsam mit dem Sandwich-Lemma (Lemma 5.31) $\lim_{n\to\infty}\left(1-\frac{x^2}{n^2}\right)^n = 1$ zur Folge hat und Gleichung (5.8) zeigt.

### 5.3.4 Additionsformel

Seien $x, y \in \mathbb{R}$. Für $x = 0$ oder $y = 0$ ist die Additionsformel (5.6) gültig (wieso?). Wir wollen den verbleibenden Fall ($x \neq 0$ und $y \neq 0$) durch ein ähnliches Argument wie oben beweisen. Deswegen berechnen wir zuerst für $n \in \mathbb{N}$ das Produkt

$$\begin{aligned}\left(1-\frac{x}{n}\right)\left(1-\frac{y}{n}\right)\left(1+\frac{x+y}{n}\right) &= \left(1-\frac{x+y}{n}+\frac{xy}{n^2}\right)\left(1+\frac{x+y}{n}\right)\\ &= 1-\frac{(x+y)^2}{n^2}+\frac{xy}{n^2}\left(1+\frac{x+y}{n}\right) = 1+\frac{c_n}{n^2},\end{aligned}$$

wobei die konvergente reelle Folge $(c_n)_n$ durch

$$\begin{aligned}c_n = -(x+y)^2 + xy\left(1+\frac{x+y}{n}\right) &= -(x^2+y^2) - 2xy + xy + xy\frac{x+y}{n}\\ &= -(x^2+y^2) - xy + xy\frac{x+y}{n}\end{aligned}$$

gegeben ist. Damit erhalten wir

$$\begin{aligned}\exp(x+y) &= \lim_{n\to\infty}\left(1+\frac{x+y}{n}\right)^n\\ &= \lim_{n\to\infty}\frac{\left(1+\frac{c_n}{n^2}\right)^n}{\left(1-\frac{x}{n}\right)^n\left(1-\frac{y}{n}\right)^n} = \exp(x)\exp(y)\lim_{n\to\infty}\left(1+\frac{c_n}{n^2}\right)^n\end{aligned}$$

wegen (5.8). Wir zeigen nun, dass $\lim_{n\to\infty}\left(1+\frac{c_n}{n^2}\right)^n$ gleich 1 ist. Da $xy\frac{x+y}{n} \to 0$ für $n \to \infty$, erhalten wir $c_n \to -(x^2+y^2) - xy$ für $n \to \infty$. Weiters ist $-(x^2+y^2) - xy < 0$ (unter Verwendung von $\max\{|x|,|y|\} < \sqrt{x^2+y^2}$ wegen $x \neq 0$ und $y \neq 0$), womit wir $c_n < 0$ und

$\frac{c_n}{n^2} \geq -1$ für hinreichend grosse $n$ erhalten. Aus der Bernoulli-Ungleichung folgt nun

$$1 + \frac{c_n}{n} = 1 + n\frac{c_n}{n^2} \leq \left(1 + \frac{c_n}{n^2}\right)^n \leq 1$$

und daher gilt $\lim_{n\to\infty} \left(1 + \frac{c_n}{n^2}\right)^n = 1$. Dies beweist die Additionsformel (5.6).

### 5.3.5 Stetigkeit

Wir zeigen zuerst die Stetigkeit von $\exp : \mathbb{R} \to \mathbb{R}_{>0}$ bei $0 \in \mathbb{R}$. Sei also $\varepsilon > 0$ und wähle $\delta = \min\left\{\varepsilon, 1 - \frac{1}{1+\varepsilon}\right\}$ (womit $\delta < 1$ und auch $\frac{1}{1-\delta} \leq 1 + \varepsilon$ nach einer kurzen Rechnung). Für $x \in (-\delta, 0]$ wenden wir (5.7) an und erhalten

$$1 - \varepsilon \leq 1 - \delta < 1 + x \leq \exp(x) \leq 1$$

(also insbesondere $|\exp(x) - \exp(0)| < \varepsilon$). Für $x \in [0, \delta)$ wenden wir obiges Argument für $-x$ an und erhalten $1 - \delta \leq \exp(-x) \leq 1$ oder äquivalenterweise $1 \leq \exp(x) \leq \frac{1}{1-\delta} < 1 + \varepsilon$ nach Wahl von $\delta$ (und dadurch wiederum $|\exp(x) - \exp(0)| < \varepsilon$).

Um Stetigkeit bei jedem $x_0 \in \mathbb{R}$ zu zeigen, verwenden wir die Additionseigenschaft. Denn es gilt für alle $x \in \mathbb{R}$

$$\exp(x) = \exp(x - x_0 + x_0) = \exp(x - x_0)\exp(x_0),$$

wodurch wir $\exp(x)$ als Verknüpfung der Abbildungen

$$\begin{aligned} h &: x \in \mathbb{R} \mapsto x - x_0 \in \mathbb{R} \\ g &: y \in \mathbb{R} \mapsto \exp(y) \in \mathbb{R} \\ f &: a \in \mathbb{R} \mapsto a\exp(x_0) \end{aligned}$$

schreiben können, wobei $h$ bei $x_0$, $g$ bei $0 = h(x_0)$, beziehungsweise $f$ bei $1 = g(0)$ stetig sind. Es folgt die Stetigkeit von exp bei $x_0$ aus Proposition 3.53.

### 5.3.6 Strenge Monotonie

Für $x > 0$ gilt $\exp(0) = 1 < 1 + x \leq \exp(x)$ wegen (5.7). Falls $x < y$, dann folgt aus $\exp(x) > 0$ und $\exp(y - x) > 1$ mit der Additionsformel (5.6), dass

$$\exp(y) = \exp(x)\exp(y - x) > \exp(x).$$

Daher ist $\exp : \mathbb{R} \to \mathbb{R}_{>0}$ streng monoton wachsend und insbesondere injektiv.

### 5.3.7 Surjektivität

Wir verwenden (5.7) und den Zwischenwertsatz (Satz 3.59) um $\exp(\mathbb{R}) = \mathbb{R}_{>0}$ zu zeigen. Sei also $y > 0$. Dann gilt $y < \exp(y)$ nach (5.7). Weiters ist $\frac{1}{y} < \exp(\frac{1}{y})$ auf Grund desselben Arguments und damit $\exp(-\frac{1}{y}) < y < \exp(y)$. Da exp auf ganz $\mathbb{R}$ stetig ist, ergibt sich aus

dem Zwischenwertsatz (Satz 3.59), dass es ein $x \in \mathbb{R}$ (zwischen den Punkten $-\frac{1}{y}$ und $y$) mit $\exp(x) = y$ gibt. Dies beendet den Beweis von Proposition 5.54.

### 5.3.8 Der Logarithmus und Potenzen

Zusammenfassend haben wir also gezeigt, dass $\exp : \mathbb{R} \to \mathbb{R}_{>0}$ eine bijektive, streng monoton wachsende stetige Funktion darstellt, so dass die Additionsformel $\exp(x + y) = \exp(x)\exp(y)$ für alle $x, y \in \mathbb{R}$ gilt.

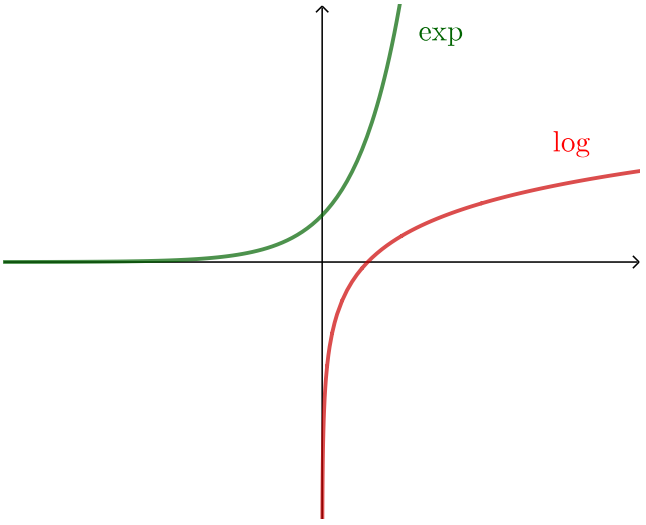
Die Umkehrfunktion der bijektiven Abbildung $\exp : \mathbb{R} \to \mathbb{R}_{>0}$ nennen wir den (natürlichen) **Logarithmus** $\log : \mathbb{R}_{>0} \to \mathbb{R}$. Aus dem Umkehrsatz (Satz 3.65) folgt nun folgendes Korollar.

**Korollar 5.58** (Natürlicher Logarithmus)**.** *Der natürliche Logarithmus* $\log : \mathbb{R}_{>0} \to \mathbb{R}$ *ist eine streng monoton wachsende, stetige und bijektive Funktion. Des Weiteren gilt*

$$\log(ab) = \log a + \log b$$

*für alle* $a, b \in \mathbb{R}_{>0}$.

Wir bemerken, dass die letzte Aussage in obigem Korollar aus der Additionsformel der Exponentialabbildung folgt wenn wir $x = \log a$ und $y = \log b$ setzen.




Figur 5.3: Die Graphen der Exponentialfunktion und des Logarithmus log.

Der Logarithmus und die Exponentialabbildung können wir verwenden, um allgemeinere Potenzen zu definieren. Für eine positive Basis $x > 0$ und beliebige Exponenten $a \in \mathbb{R}$ setzen wir

$$x^a := \exp(a \log(x)).$$

Insbesondere gilt $\mathrm{e}^x = \exp(x \log(\mathrm{e})) = \exp(x)$ für alle $x \in \mathbb{R}$.

**Übung 5.59** (Rechenregel für Potenzen)**.** *Zeigen Sie, dass für* $a \in \mathbb{Q}$ *diese Definition mit der Definition von rationalen Potenzen aus Beispiel 3.66 übereinstimmt. Verifizieren Sie des*

*Weiteren die Rechenregeln*

$$\log(x^a) = a\log(x), \quad x^a x^b = x^{a+b}, \quad (x^a)^b = x^{ab}$$

*für* $x, y > 0$ *und* $a, b \in \mathbb{R}$.

**Übung 5.60** (Obere Schranke für den Logarithmus)**.** *Sei* $\alpha > 0$ *eine positive Zahl. Zeigen Sie, dass eine Konstante* $C_\alpha > 0$ *existiert mit* $\log(x) \leq C_\alpha x^\alpha$ *für alle* $x > 0$.

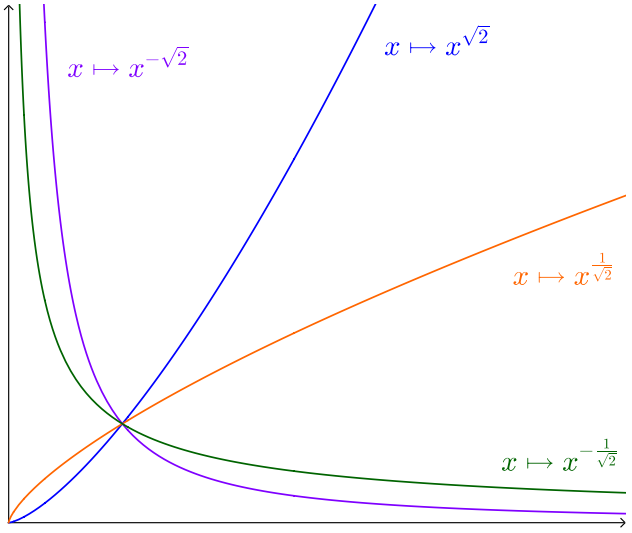
**Übung 5.61** (Eine kontinuierliche Bernoulli-Ungleichung)**.** *Zeigen Sie, dass für alle* $x \geq -1$ *und* $p \geq 1$ *gilt*

$$(1+x)^p \geq 1 + px.$$

*Hinweis: Analysieren Sie das Argument für die Monotonie aus Abschnitt 5.3.2 genauer, um zu zeigen, dass für alle* $m \leq n$ *und* $t > -m$

$$\left(1 + \frac{t}{n}\right)^n \geq \left(1 + \frac{t}{m}\right)^m$$

*gilt. Betrachten Sie nun* $x = \frac{t}{n}$*, um die gewünschte Ungleichung für rationale* $p$ *zu zeigen. Verwenden Sie dann Stetigkeit und Dichtheit von* $\mathbb{Q}$.




Figur 5.4: Die Graphen von $x \in \mathbb{R}_{>0} \mapsto x^a \in \mathbb{R}_{>0}$ für verschiedene (hier irrationale) $a \in \mathbb{R}$.

*Bemerkung.* Sie dürfen nun auch den Logarithmus $\log_a : \mathbb{R}_{>0} \to \mathbb{R}$ zu einer Basis $a > 1$ definieren. In der Tat können Sie $\log_a(x) = \frac{\log x}{\log a}$ für alle $x \in \mathbb{R}_{>0}$ setzen und nun überprüfen, dass $a^{\log_a x} = x$ für alle $x \in \mathbb{R}_{>0}$ gilt. Wir werden diese Definition aber nicht benötigen, auch nicht für $a = 10$, und $\log(x) = \ln(x)$ wird immer den natürlichen Logarithmus von $x \in \mathbb{R}_{>0}$ zur Basis $a = \mathrm{e}$ bezeichnen.

**Applet 5.62** (Rechenschieber)**.** *Falls Sie dies noch nicht gesehen haben, empfehlen wir Ihnen mit dem Rechenschieber einige Produkte und Quotienten zu berechnen. Erinnern Sie sich an*

*die Eigenschaften des Logarithmus um zu erkennen, wie man diese Berechnungen durchführt. Vor der Einführung von elektronischen Taschenrechnern waren diese mechanischen Hilfsmittel sehr verbreitet.*

## 5.4 Grenzwerte von Funktionen

Wir betrachten jetzt wieder allgemeine Funktionen $f : D \to \mathbb{R}$ auf einer allgemeinen Teilmenge $D \subseteq \mathbb{R}$ und wollen (eigentliche und uneigentliche) Grenzwerte für den Fall definieren, wenn $x \in D$ gegen ein $x_0 \in \mathbb{R}$ strebt (oder auch wenn $x \in D$ gegen $+\infty$ oder gegen $-\infty$ divergiert). Da es mit allen Kombinationen viele verschiedene Möglichkeiten gibt, wollen wir einige konkrete Fälle definieren und dann ein allgemeines Schema für solche Definitionen betrachten.

### 5.4.1 Grenzwerte und punktierte Umgebungen

Sei $D \subseteq \mathbb{R}$ eine Teilmenge und $x_0 \in \mathbb{R}$. Wir erinnern daran, dass $x_0$ ein **Häufungspunkt** ist falls

$$D \cap (x_0 - \delta, x_0 + \delta) \setminus \{x_0\} \neq \emptyset \tag{5.9}$$

für alle $\delta > 0$. Wir bemerken, dass $x_0$ genau dann ein Häufungspunkt ist, wenn es eine Folge in $D \setminus \{x_0\}$ gibt, die gegen $x_0$ strebt.

Für eine Funktion $f : D \to \mathbb{R}$ ist $A = \lim_{x \to x_0} f(x)$ **der Grenzwert von** $f(x)$ **für** $x \to x_0$, oder auch **der Grenzwert bei** $x_0$, falls

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x \in D \cap (x_0 - \delta, x_0 + \delta) \setminus \{x_0\} : |f(x) - A| < \varepsilon.$$

Informell ausgedrückt bedeutet dies, dass die Funktionswerte von $f$ beliebig nahe bei $A$ liegen wenn $x \in D \setminus \{x_0\}$ nahe an $x_0$ heranrückt. Der Grenzwert von $f(x)$ für $x \to x_0$ muss natürlich nicht existieren; wenn er existiert, ist er aber eindeutig bestimmt (diese Eigenschaft ist der Grund, wieso wir (5.9) angenommen haben, siehe Übung 5.63).

Der Grenzwert erfüllt, analog zu Proposition 5.7, die gewohnten Eigenschaften. Er ist

- linear (das heisst, falls $\lim_{x \to x_0} f(x)$ und $\lim_{x \to x_0} g(x)$ existieren, so existiert auch der Grenzwert $\lim_{x \to x_0} f(x) + g(x) = \lim_{x \to x_0} f(x) + \lim_{x \to x_0} g(x)$ und analog für skalare Multiplikation),
- multiplikativ (das heisst, falls $\lim_{x \to x_0} f(x)$ und $\lim_{x \to x_0} g(x)$ existieren, so existiert auch $\lim_{x \to x_0} f(x)g(x) = (\lim_{x \to x_0} f(x))(\lim_{x \to x_0} g(x))$),
- monoton ($f \leq g$ impliziert $\lim_{x \to x_0} f(x) \leq \lim_{x \to x_0} g(x)$, falls die Grenzwerte existieren)
- und erfüllt ein Sandwich-Lemma.

**Übung 5.63** (Erste Eigenschaften). *(i) Beweisen Sie, dass der Grenzwert* $\lim_{x \to x_0} f(x)$ *eindeutig bestimmt ist, falls er existiert.*

*(ii) Beweisen Sie die oben aufgelisteten Eigenschaften des Grenzwerts von Funktionen auf $D$ für $x \to x_0$.*

*(iii) Formulieren und beweisen Sie ein Sandwich-Lemma für den Grenzwert von Funktionen auf $D$ für $x \to x_0$.*

**Lemma 5.64** (Grenzwerte und Stetigkeit)**.** *Sei $D \subseteq \mathbb{R}$ eine Teilmenge, $x_0 \in D$ ein Häufungspunkt von $D$ und $f$ eine reellwertige Funktion auf $D$. Dann ist $f$ genau dann stetig bei $x_0$, wenn $\lim_{x\to x_0} f(x) = f(x_0)$.*

*Beweis.* Falls $f$ bei $x_0$ stetig ist, dann existiert zu jedem $\varepsilon > 0$ ein $\delta > 0$, so dass für alle $x \in D$ die Implikation $|x-x_0| < \delta \implies |f(x)-f(x_0)| < \varepsilon$ gilt. Vergleicht man dies mit der Definition von $\lim_{x\to x_0} f(x)$, erhält man $\lim_{x\to x_0} f(x) = f(x_0)$.

Falls umgekehrt $\lim_{x\to x_0} f(x) = f(x_0)$ gilt, so müssen wir wiederum nur die Definition der Stetigkeit (und die Gleichheit des Grenzwerts mit $f(x_0)$) verwenden, um Stetigkeit von $f$ bei $x_0$ zu erhalten. □

Obiges Lemma hat auch eine Interpretation für den Fall $x_0 \notin D$, denn in diesem Fall wäre der Grenzwert $\lim_{x\to x_0} f(x)$ (falls dieser existiert) ein guter Kandidat für eine Fortsetzung der Funktion auf die Menge $D \cup \{x_0\}$, da diese Fortsetzung dann bei $x_0$ stetig wird.

Man nennt einen Häufungspunkt $x_0 \in D$ eine **hebbare Unstetigkeitsstelle** von $f$, falls $\lim_{x\to x_0} f(x)$ existiert, aber nicht gleich $f(x_0)$ ist (siehe auch Figur 5.5). In diesem Fall kann man eine neue Funktion $f_{\text{neu}} : D \to \mathbb{R}$ durch

$$f_{\text{neu}}(x) = \begin{cases} f(x) & \text{falls } x \in D \setminus \{x_0\} \\ \lim_{x\to x_0} f(x) & \text{falls } x = x_0 \end{cases}$$

für $x \in D$ definieren, die bei $x_0$ stetig ist.

Genauso wie in dem Beweis von Proposition 5.17 sieht man nun, dass für eine Folge $(x_n)_n$ in $D \setminus \{x_0\}$ mit $x_n \to x_0$ für $n \to \infty$ die Gleichheit $\lim_{n\to\infty} f(x_n) = \lim_{x\to x_0} f(x)$ gilt, falls letzter Grenzwert existiert.

**Lemma 5.65** (Grenzwerte mittels Folgen)**.** *Sei $D \subseteq \mathbb{R}$ eine Teilmenge, $f : D \to \mathbb{R}$ eine Funktion und $x_0 \in D$ ein Häufungspunkt von $D$. Dann gilt $A = \lim_{x\to x_0} f(x)$ genau dann, wenn für jede Folge $(a_n)_n$ in $D \setminus \{x_0\}$ mit $\lim_{n\to\infty} a_n = x_0$ auch $\lim_{n\to\infty} f(a_n) = A$ gilt.*

*Beweis.* Angenommen $A = \lim_{x\to x_0} f(x)$ und $(a_n)_n$ ist eine Folge in $D\setminus\{x_0\}$ mit $\lim_{n\to\infty} a_n = x_0$. Dann existiert für $\varepsilon > 0$ ein $\delta > 0$ mit

$$0 < |x - x_0| < \delta \implies |f(x) - A| < \varepsilon$$

für alle $x \in D$. Des Weiteren existiert ein $N \in \mathbb{N}$ mit

$$n \geq N \implies 0 < |a_n - x_0| < \delta,$$

was gemeinsam

$$n \geq N \implies |f(a_n) - A| < \varepsilon$$

ergibt. Die Folge $(f(a_n))_n$ konvergiert also gegen $A$.

Für die Umkehrung nehmen wir an, dass $A = \lim_{x\to x_0} f(x)$ nicht erfüllt ist (also entweder der Grenzwert nicht existiert oder nicht gleich $A$ ist). Dann existiert ein $\varepsilon > 0$, so dass für alle

$\delta > 0$ ein $x \in D$ existiert mit

$$0 < |x - x_0| < \delta \wedge |f(x) - A| \geq \varepsilon$$

Wir verwenden dies für $n \in \mathbb{N}$ und $\delta = \frac{1}{n} > 0$ und finden also ein $a_n \in D$ mit

$$0 < |a_n - x_0| < \frac{1}{n} \tag{5.10}$$

und

$$|f(a_n) - A| \geq \varepsilon. \tag{5.11}$$

Aus Ungleichung (5.10) schliessen wir, dass die Folge $(a_n)_n$ Werte in $D \setminus \{x_0\}$ annimmt und gegen $x_0$ konvergiert. Aus Ungleichung (5.11) folgt, dass $(f(a_n))_n$ nicht gegen $A$ konvergiert. □

**Proposition 5.66** (Grenzwerte und Verknüpfung mit stetigen Funktionen)**.** *Seien $D, E \subseteq \mathbb{R}$, $x_0$ ein Häufungspunkt von $D$, $f : D \to E$ eine Funktion, $y_0 = \lim_{x \to x_0} f(x) \in E$, und $g : E \to \mathbb{R}$ eine bei $y_0$ stetige Funktion. Dann gilt $\lim_{x \to x_0} g(f(x)) = g(y_0)$.*

*Beweis.* Sei $(a_n)_n$ eine Folge in $D$ mit $\lim_{n \to \infty} a_n = x_0$. Nach Lemma 5.65 gilt $\lim_{n \to \infty} f(a_n) = \lim_{x \to x_0} f(x) = y_0$. Die Stetigkeit von $g$ bei $y_0$ impliziert nun gemeinsam mit Proposition 5.17, dass $\lim_{n \to \infty} g(f(a_n)) = g(y_0)$. Da $(a_n)_n$ eine beliebige Folge in $D$ mit $\lim_{n \to \infty} a_n = x_0$ war, folgt wiederum aus Lemma 5.65, dass $\lim_{x \to x_0} g(f(x)) = g(y_0)$. □

Diese Eigenschaften von Grenzwerten können bereits für die Berechnung von vielen Grenzwerten verwendet werden.

Weiters können wir uneigentliche Grenzwerte definieren. Wir sagen zum Beispiel, dass $f(x)$ gegen $+\infty$ für $x \to x_0$ divergiert und schreiben $\lim_{x \to x_0} f(x) = +\infty$, falls

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x \in D \cap (x_0 - \delta, x_0 + \delta) \setminus \{x_0\} : f(x) > \varepsilon^{-1}.$$

Wir nennen die Menge $\dot{U}_\delta(x_0) = (x_0 - \delta, x_0 + \delta) \setminus \{x_0\}$ die **punktierte $\delta$-Umgebung** um $x_0$, und bemerken, dass diese Mengen implizit in der Definition des Grenzwerts aufgetreten sind.

### 5.4.2 Links- und rechtsseitige Grenzwerte

Angenommen $D \subseteq \mathbb{R}$ ist eine Teilmenge und $x_0 \in \mathbb{R}$ hat die Eigenschaft $D \cap (x_0, x_0 + \delta) \neq \emptyset$ für alle $\delta > 0$. Intuitiv hat der Punkt $x_0$ also die Eigenschaft, dass ihm $D$ von rechts beliebig nahe kommt, was also eine stärkere Forderung ist als (5.9). Einen solchen Punkt $x_0$ wollen wir einen **rechtsseitigen Häufungspunkt** von $D$ nennen. Für eine Funktion $f : D \to \mathbb{R}$ ist $A = \lim_{x \searrow x_0} f(x)$ (alternativ $\lim_{x \to x_0^+} f(x)$ oder auch $\lim_{x \to x_0,\, x > x_0} f(x)$) der **rechtsseitige**

**Grenzwert von** $f(x)$ **bei** $x_0$, falls

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in D \cap (x_0, x_0 + \delta) : |f(x) - A| < \varepsilon.$$

Wir schreiben $\lim_{x \searrow x_0} f(x) = +\infty$, falls

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in D \cap (x_0, x_0 + \delta) : f(x) > \frac{1}{\varepsilon}$$

und $\lim_{x \searrow x_0} f(x) = -\infty$, falls

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in D \cap (x_0, x_0 + \delta) : f(x) < -\frac{1}{\varepsilon}.$$

Falls $x_0$ die Eigenschaft $D \cap (x_0 - \delta, x_0) \neq \emptyset$ für alle $\delta > 0$ hat ($x_0$ ist ein **linksseitiger Häufungspunkt**), können wir ebenso den **linksseitigen Grenzwert** $\lim_{x \nearrow x_0} f(x)$ (alternativ $\lim_{x \to x_0^-} f(x)$ oder auch $\lim_{x \to x_0,\ x < x_0} f(x)$) definieren.

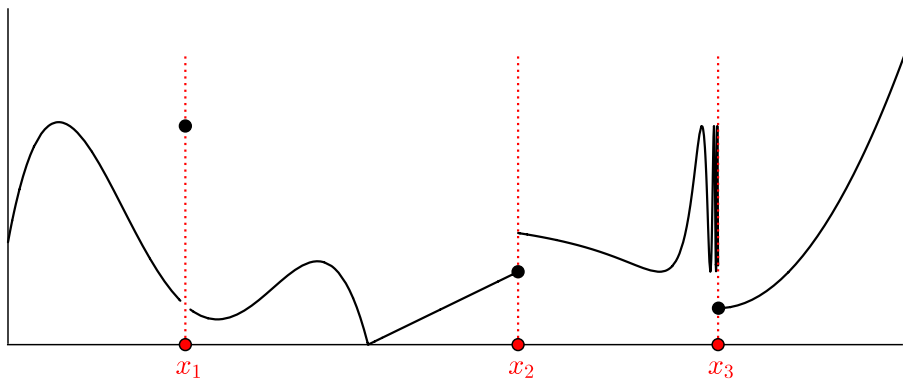
Falls $x_0$ ein links- und rechtsseitiger Häufungspunkt ist, dann existiert der Grenzwert $\lim_{x \to x_0} f(x)$ genau dann, wenn die links- und rechtseitigen Grenzwerte von $f(x)$ bei $x_0$ existiert und $\lim_{x \nearrow x_0} f(x) = \lim_{x \searrow x_0} f(x)$ erfüllt ist.

Beispiele von links- und rechtsseitigen Grenzwerten sind

$$\lim_{x \searrow 0} \frac{1}{x} = +\infty, \quad \lim_{x \nearrow 0} \frac{1}{x} = -\infty, \quad \lim_{x \searrow 0} \log(x) = -\infty, \quad \lim_{x \searrow 0} \sqrt{x} = \lim_{x \to 0} \sqrt{x} = 0.$$

### 5.4.3 Einseitige Stetigkeit und Sprungstellen

Sei $D \subseteq \mathbb{R}$ eine Teilmenge, $x_0 \in D$ ein rechtseitiger Häufungspunkt von $D$ und $f : D \to \mathbb{R}$ eine Funktion. Falls $\lim_{x \searrow x_0} f(x)$ existiert und gleich $f(x_0)$ ist, dann sagen wir, dass $f$ **rechtsseitig stetig** bei $x_0$ ist. Ist $x_0 \in D$ ein linksseitiger Häufungspunkt von $D$, dann sagen wir analog, dass $f$ **linksseitig stetig** bei $x_0$ ist, falls $\lim_{x \nearrow x_0} f(x)$ existiert und gleich $f(x_0)$ ist.




Figur 5.5: Der Graph einer Funktion, die eine hebbare Unstetigkeitsstelle bei $x_1$ hat, bei $x_2$ linksseitig stetig (aber nicht rechtseitig stetig) ist und bei $x_3$ rechtsseitig stetig ist.

Sei nun $D \subseteq \mathbb{R}$ und $x_0 \in D$ ein links- und rechtsseitiger Häufungspunkt von $D$ (insbesondere ein Häufungspunkt von $D$). Für eine Funktion $f : D \to \mathbb{R}$ heisst $x_0$ eine **Sprungstelle**,

falls die einseitigen Grenzwerte $\lim_{x\nearrow x_0} f(x)$ und $\lim_{x\searrow x_0} f(x)$ existieren, aber verschieden sind.

**Applet 5.67** (Grenzwerte einer Funktion). *Wir sehen eine Funktion mit Definitionsbereich* $D = \{0\} \cup (\frac{1}{2}, 2) \cup (2, 5) \cup (5, 8) \cup (8, 11]$, *und betrachten verschiedene Bewegungen im Definitionsbereich und Grenzwerte für diese Funktion.*

### 5.4.4 Die Bewegung nach Unendlich

Angenommen $D \subseteq \mathbb{R}$ ist eine nicht von oben beschränkte Teilmenge (das heisst, für alle $\delta > 0$ gilt $(\frac{1}{\delta}, \infty) \cap D \neq \emptyset$) und $f : D \to \mathbb{R}$ ist eine Funktion. Wir sagen, dass $f$ gegen $A \in \mathbb{R}$ strebt für $x \to \infty$, und schreiben $\lim_{x\to\infty} f(x) = A$, falls

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in (\tfrac{1}{\delta}, \infty) \cap D : |f(x) - A| < \varepsilon.$$

**Übung 5.68** (Beispiele für uneigentliche Grenzwerte). *Definieren Sie für $D$ wie oben und eine Funktion $f : D \to \mathbb{R}$ die uneigentlichen Grenzwerte $\lim_{x\to\infty} f(x) = \infty$, $\lim_{x\to\infty} f(x) = -\infty$ und finden Sie je eine Funktion $f$ auf $(0, \infty)$ mit $\lim_{x\to\infty} f(x) = 1$, $\lim_{x\to\infty} f(x) = +\infty$ und $\lim_{x\to\infty} f(x) = -\infty$.*

VO: 13.11.17

### 5.4.5 Umgebungsfilter und Konvergenz entlang eines Filters*

**Alle natürlich auftretenden Konvergenzbegriffe (siehe Applet) sind auch Prüfungsstoff, doch der allgemeine Rahmen der Filterkonvergenz wird nicht bei der Prüfung vorkommen.*

Wir haben oben insgesamt 6 Konvergenzen ausführlich definiert, wobei es aber insgesamt 15 Kombinationen (wieso?) für reellwertige Funktionen $D \subseteq \mathbb{R}$ zu definieren gäbe. Betrachten wir Teilmengen $D \subseteq \mathbb{C}$, reellwertige Funktionen $f : D \to \mathbb{R}$ und Punkte $z_0 \in \mathbb{C}$, so gibt es nochmals drei Möglichkeiten (reelle oder uneigentliche Grenzwerte). Für komplexwertige Funktionen gibt es noch eine weitere Definition für jede der möglichen Bewegungen. Es wäre wohl eher langweilig, all diese Definitionen einzeln auszuformulieren und ihre (jeweils sehr analogen) Eigenschaften zu besprechen (siehe aber folgendes Applet).

**Applet 5.69** (40 Definitionen). *Wir fassen alle (und einige weitere) Definitionen für Konvergenz in diesem Applet zusammen. Versuchen Sie die Ähnlichkeiten und Unterschiede der verschiedenen Definitionen zu finden.*

Stattdessen analysieren wir kurz die Struktur der Definition. In allen Fällen der obigen Definitionen verschiedener Typen von Konvergenz werden gewisse Teilmengen des Definitionsbereichs (zum Beispiel punktierte oder einseitige Umgebungen) gemeinsam mit $\varepsilon$-Umgebungen des vermeintlichen Grenzwerts $A$ verwendet, um Konvergenz zu definieren.

Wir erinnern daran, dass für $\varepsilon > 0$ die $\varepsilon$-Umgebung von $A \in \mathbb{R}$ das offene Intervall

$$U_\varepsilon(A) = (A - \varepsilon, A + \varepsilon)$$

ist. Für das Symbol $+\infty$ definieren wir die $\varepsilon$-Umgebung durch $U_\varepsilon(\infty) = (\frac{1}{\varepsilon}, \infty)$ und für $-\infty$ analog $U_\varepsilon(-\infty) = (-\infty, -\frac{1}{\varepsilon})$. Für $A \in \mathbb{C}$ ist die $\varepsilon$-Umgebung durch den offenen Ball $U_\varepsilon(A) = \{z \in \mathbb{C} \mid |z - A| < \varepsilon\}$ gegeben.

Eine Menge ist eine **Umgebung von** $A$ (wobei $A \in \mathbb{R}$, $A = \pm\infty$ oder $A \in \mathbb{C}$), falls sie die $\varepsilon$-Umgebung $U_\varepsilon(A)$ für ein $\varepsilon > 0$ enthält. Dies motiviert die folgende Definition.

**Definition 5.70** (Filter)**.** Für eine gegebene Menge $X$ ist eine nicht-leere Familie $\mathcal{F}$ von Teilmengen von $X$ ein **Filter** auf $X$, falls folgende drei Eigenschaften erfüllt sind:

- Die leere Menge ist kein Element von $\mathcal{F}$.
- Für $U_1, U_2 \in \mathcal{F}$ gilt auch $U_1 \cap U_2 \in \mathcal{F}$.
- Falls $U \in \mathcal{F}$ und $V \subseteq X$ eine Teilmenge, die $U$ enthält, dann ist auch $V \in \mathcal{F}$.

Die in $\mathcal{F}$ enthaltenen Mengen werden auch **Filtermengen** genannt.

Intuitiv ausgedrückt enthält jede Filtermenge alle wesentlichen Punkte für die durch den Filter beschriebene Bewegung. Zum Beispiel enthalten die Filtermengen $U$ im Filter auf $\mathbb{N}$ gegeben durch

$$\mathcal{F} = \{U \subseteq \mathbb{N} \mid \exists N \in \mathbb{N}\ \forall n \geq N : n \in U\}$$

immer einen ganzen „Endabschnitt von $\mathbb{N}$“ und damit alle Punkte, die relevant sind für die Bewegung in $\mathbb{N}$ gegen unendlich (Folgenkonvergenz). Wir definieren also die „Bewegung“ indirekt mittels dem Begriff „Filter“, welcher beschreibt in welchen Mengen man schlussendlich liegen muss.

**Beispiel 5.71** (Umgebungsfilter)**.** *Folgende Mengenfamilien sind weitere Beispiele für Filter.*

- *Die Familie der Umgebungen von $A \in \overline{\mathbb{R}}$ bildet den Filter*

$$\mathcal{U}_A = \{V \subseteq \mathbb{R} \mid \exists \varepsilon > 0 : U_\varepsilon(A) \subseteq V\}$$

  *auf $X = \mathbb{R}$, den wir den **Umgebungsfilter** von $A$ nennen. In der Tat ist die Familie $\mathcal{U}_A$ nicht-leer, da sie alle $\varepsilon$-Umgebungen von $A$ enthält. Da jede Umgebung von $A$ den Punkt $A$ enthält, ist die leere Menge kein Element des Umgebungsfilters. Des Weiteren gibt es für $V_1, V_2 \in \mathcal{U}_A$ per Definition $\varepsilon_1, \varepsilon_2 > 0$ mit $U_{\varepsilon_i}(A) \subseteq V_i$ für $i \in \{1, 2\}$ und somit ist $U_{\min\{\varepsilon_1,\varepsilon_2\}}(A) = U_{\varepsilon_1}(A) \cap U_{\varepsilon_2}(A) \subseteq V_1 \cap V_2$. Also ist $V_1 \cap V_2 \in \mathcal{U}_A$. Ist $U \in \mathcal{U}_A$ und $V \supseteq U$, dann enthält $V$ alle $\varepsilon$-Umgebungen, die in $U$ enthalten sind, und ist somit auch in $\mathcal{U}_A$.*

- *Für $A \in \mathbb{C}$ erhalten wir ebenso einen Umgebungsfilter $\mathcal{U}_A$ von $A$ auf $X = \mathbb{C}$.*

- *Sei $D \subseteq \mathbb{R}$ eine Teilmenge und $x_0 \in \mathbb{R}$ ein Häufungspunkt von $D$. Die Familie von Teilmengen von $D$*

$$\dot{\mathcal{U}}_{x_0} = \{V \subseteq D \mid \exists \varepsilon > 0 : D \cap U_\varepsilon(x_0) \setminus \{x_0\} \subseteq V\}$$

*ist ein Filter auf $D$, den wir den Filter der punktierten Umgebungen von $x_0$ auf $D$ nennen. Die Annahme, dass $x_0$ ein Häufungspunkt von $D$ ist, ist hier zwingend notwendig, damit der erste Punkt in Definition 5.70 erfüllt ist.*

- *Für eine Teilmenge $D \subseteq \mathbb{R}$ und einen rechtsseitigen Häufungspunkt $x_0 \in D$ ist*

$$\mathcal{U}_{x_0}^{+} = \{V \subseteq D \mid \exists \varepsilon > 0 : D \cap (x_0, x_0 + \varepsilon) \subseteq V\}$$

  *der Filter der rechtsseitigen Umgebungen von $x_0$ auf $D$.*

- *Für eine Teilmenge $D \subseteq \mathbb{R}$ und einen linksseitigen Häufungspunkt $x_0 \in D$ ist der Filter $\mathcal{U}_{x_0}^{-}$ der linksseitigen Umgebungen von $x_0$ auf $D$ analog definiert.*

**Definition 5.72** (Filterkonvergenz)**.** Seien nun $D \subseteq \mathbb{R}$ eine beliebige Teilmenge, $f : D \to \mathbb{R}$ eine Funktion, $\mathcal{F}$ ein Filter auf $D$ und $A \in \overline{\mathbb{R}}$ gegeben. Dann sagen wir, dass $f$ **entlang** $\mathcal{F}$ **gegen** $A$ konvergiert und schreiben $\lim_{\mathcal{F}} f(x) = A$, falls

$$\forall V \in \mathcal{U}_A \; \exists F \in \mathcal{F} \; \forall x \in F : f(x) \in V.$$

Für komplexwertige Funktionen verwendet man auf die gleiche Weise den Umgebungsfilter von möglichen Grenzwerten $A \in \mathbb{C}$ für die Definition der Konvergenz entlang des Filters. Alle oben besprochenen Konvergenzen lassen sich nun mit diesem Begriff von Konvergenz entlang von Filtern auffassen. Es gilt zum Beispiel

$$\lim_{\dot{\mathcal{U}}_{x_0}} f(x) = \lim_{x \to x_0} f(x), \quad \lim_{\mathcal{U}_{x_0}^{+}} f(x) = \lim_{x \searrow x_0} f(x), \quad \lim_{\mathcal{U}_{x_0}^{-}} f(x) = \lim_{x \nearrow x_0} f(x)$$

jeweils unter geeigneten Annahmen an $D \subseteq \mathbb{R}$, $x_0 \in \mathbb{R}$ und $f : D \to \mathbb{R}$. Sie sollten nun zum Beispiel die nicht besprochenen Definitionen von $\lim_{x \nearrow x_0} f(x) = A \in \mathbb{C}$ für eine komplexwertige Funktion oder $\lim_{x \to -\infty} f(x) = +\infty$ für eine reellwertige Funktion als Spezialfälle obiger Definition sehen. Der Grenzwert entlang eines Filters ist nach folgender Übung eindeutig, womit man sinnvoll von dem Grenzwert sprechen kann.

**Übung 5.73** (Eindeutigkeit)**.** *Seien $D$ eine nicht-leere Menge, $f : D \to \mathbb{R}$ eine Funktion und $\mathcal{F}$ ein Filter auf $D$. Zeigen Sie, dass der Grenzwert $A \in \overline{\mathbb{R}}$ von $f$ entlang des Filters $\mathcal{F}$ eindeutig durch $f$ bestimmt ist, falls er existiert.*

Um die Nützlichkeit der vereinheitlichten Herangehensweise via Filter noch stärker hervorzuheben, wollen wir das Analogon von Lemma 5.31 und Übung 5.63(iii) für alle diskutierten Grenzwerte von Funktionen in folgendem Lemma beweisen.

**Lemma 5.74** (Sandwich für Filterkonvergenz)**.** *Sei $D$ eine nicht-leere Menge und $\mathcal{F}$ ein Filter auf $D$. Angenommen $f, g_1, g_2 : D \to \mathbb{R}$ sind Funktionen mit $g_1 \leq f \leq g_2$ und wir haben $\lim_{\mathcal{F}} g_1(x) = \lim_{\mathcal{F}} g_2(x) = A \in \bar{\mathbb{R}}$. Dann existiert auch der Grenzwert $\lim_{\mathcal{F}} f$ und ist durch $A$ gegeben.*

*Beweis.* Sei $V \in \mathcal{U}_A$ und $\varepsilon > 0$ mit $U_\varepsilon(A) \subseteq V$. Da auch $U_\varepsilon(A) \in \mathcal{U}_A$, existieren per Annahme $F_1, F_2 \in \mathcal{F}$ , so dass für $i \in \{1, 2\}$ und für alle $x \in F_i$ gilt $g_i(x) \in U_\varepsilon(A)$. Ohne Beschränkung

der Allgemeinheit können wir annehmen, dass $F_1 = F_2 = F$, in dem wir $F_1$ respektive $F_2$ durch $F_1 \cap F_2 \in \mathcal{F}$ ersetzen. Des Weiteren liegt für $x \in F$ der Punkt $f(x)$ zwischen $g_1(x)$ und $g_2(x)$, die beide in $U_\varepsilon(A)$ liegen. Da aber $U_\varepsilon(A)$ ein Intervall ist, muss auch $f(x)$ in diesem liegen. Somit gilt für alle $x \in F$, dass $f(x) \in U_\varepsilon(A) \subseteq V$ und, da $V \in \mathcal{U}_A$ beliebig war, gilt somit $\lim_{\mathcal{F}} f(x) = A$. □

**Übung 5.75** (Weitere Filter)**.** *In dieser Übung möchten wir zwei Beispiele von Filtern auf einer allgemeinen, nicht-leeren Menge $X$ diskutieren.*

*(i) Sei $B \subseteq X$ eine nicht-leere Teilmenge. Zeigen Sie, dass $\mathcal{F}_B = \{V \subseteq X \mid B \subseteq V\}$ ein Filter auf $X$ ist. Verifizieren Sie des Weiteren, dass der Filter $\mathcal{F}_B$ genau dann in keinem grösseren Filter enthalten ist, wenn die Teilmenge $B$ aus einem Punkt besteht.*

*(ii) Angenommen $X$ ist unendlich. Zeigen Sie, dass $\mathcal{F} = \{V \subseteq X \mid X \setminus V$ ist endlich$\}$ einen Filter auf $X$ bildet.*

*(iii) Beschreiben Sie den „Konvergenzbegriff“ für reellwertige Funktionen $f$ auf $X$, der durch diese beiden Filter beschrieben wird. Ist Ihnen einer der beiden für $X = \mathbb{N}$ bereits bekannt?*

### 5.4.6 Einige Rechenbeispiele

Wir werden bei Rechnungen wie den folgenden oft davon ausgehen, dass ein Formelausdruck eine Funktion mit dem maximalen für den Formelausdruck sinnvollen Definitionsbereich definiert.

**Beispiel 5.76.** *Wir wollen hier $\lim_{x\to 0} x^x = \lim_{x\searrow 0} x^x$ berechnen, und müssen für dies zwei weitere Grenzwerte berechnen.*

- *Wir behaupten zuerst*

$$\lim_{y\to\infty} y\exp(-y) = 0. \tag{5.12}$$

*In der Tat gilt $\exp(y) \geq (1+\frac{y}{2})^2$ für $y \geq 0$ auf Grund der Monotonie der Folge $(1+\frac{y}{n})^n$, die in Abschnitt 5.3 für die Definition der Exponentialabbildung verwendet wurde. Daraus ergibt sich $0 \leq y\exp(-y) \leq \frac{y}{(1+\frac{y}{2})^2} \leq \frac{4}{y}$, was wegen dem Sandwich-Lemma (Lemma 5.74) eben (5.12) impliziert.*

- *Als nächstes wollen wir*

$$\lim_{x\to 0} x\log x = 0 \tag{5.13}$$

*zeigen. Sei also $\varepsilon > 0$. Dann gibt es wegen (5.12) ein $\delta > 0$ so dass $|y\exp(-y)| < \varepsilon$ für alle $y > \frac{1}{\delta}$. Sei nun $x \in (0, \exp(-\frac{1}{\delta}))$ und $y = -\log x$, dann ist $y > \frac{1}{\delta}$ auf Grund der strengen Monotonie der Logarithmus-Abbildung und damit $|x\log x| = |\exp(-y)y| < \varepsilon$, was zu zeigen war.*

- *Auf Grund von Proposition 5.66 und da die Exponentialabbildung stetig ist, ergibt sich aus* (5.13) *nun*

$$\lim_{x\to 0} x^x = \lim_{x\to 0} \exp(x \log x) = \exp(0) = 1.$$

*Hieraus ergibt sich auch ein weiterer Beweis für Beispiel 5.33 (iii). (Wieso?)*

**Übung 5.77.** *Berechnen Sie die folgenden Grenzwerte (falls sie existieren)*

$$\lim_{x\to 2} \frac{x^3 - x^2 - x - 2}{x - 2}, \quad \lim_{x\to\infty} \frac{3e^{2x} + e^x + 1}{2e^{2x} - 1}, \quad \lim_{x\to\infty} \frac{e^x}{x^a}, \quad \lim_{x\to\infty} \frac{\log(x)}{x^a}$$

*für ein* $a \in \mathbb{R}$. *Beschreiben Sie weiters, wie man die Grenzwerte*

$$\lim_{x\to\infty} p(x), \quad \lim_{x\to\infty} \frac{p(x)}{q(x)}$$

*für zwei Polynome* $p, q \in \mathbb{R}[x]$ *mit* $q \neq 0$ *berechnet.*

## 5.5 Riemann-Summen

Riemann gab 1854 eine formale Definition des Integrals mit Hilfe sogenannter Riemann-Summen und eines Grenzübergangs, welchen wir hier besprechen wollen. Wie zuvor sei $f$ eine reellwertige Funktion auf einem kompakten Intervall $[a, b]$ mit $a < b$.

**Definition 5.78** (Riemann-Summen)**.** Für eine Zerlegung $\mathfrak{Z} = \{a = x_0 < x_1 < \ldots < x_n = b\}$ von $[a, b]$ definieren wir die **Maschenweite der Zerlegung** $\mathfrak{Z}$ als $|\mathfrak{Z}| = \max_{k=1,\ldots,n}(x_k - x_{k-1})$. Weiters bezeichnen wir $\mathbf{z} = (z_1, \ldots, z_n) \in [a, b]^n$ als eine **erlaubte Wahl von Zwischenpunkten** der Zerlegung $\mathfrak{Z}$ falls $z_k \in [x_{k-1}, x_k]$ für $k \in \{1, \ldots, n\}$. Für eine reellwertige Funktion $f$ auf $[a, b]$, eine Zerlegung $\mathfrak{Z}$ von $[a, b]$ und eine erlaubte Wahl von Zwischenpunkten $\mathbf{z}$ definieren wir die **Riemann-Summe** durch

$$R(f, \mathfrak{Z}, \mathbf{z}) = \sum_{k=1}^{n} f(z_k)(x_k - x_{k-1}).$$

Das heisst, wir betrachten beliebige Punkte $z_k \in [x_{k-1}, x_k]$, die Funktionswerte $f(z_k)$ an diesen Punkten und hoffen, dass diese halbwegs repräsentativ für die Funktionswerte von $f|_{[x_{k-1},x_k]}$ sind. Diese Hoffnung mag zwar nicht in allen Teilintervallen immer zutreffen, trotzdem ist die Riemann-Summe eine Approximation des Riemann-Integrals in folgendem Sinne.

**Satz 5.79** (Riemann-Integral über Riemann-Summen)**.** *Sei $f$ eine Riemann-integrierbare reellwertige Funktion auf $[a, b]$. Dann ist $\int_a^b f(x)\,\mathrm{d}x$ der Grenzwert der Riemann-Summen $R(f, \mathfrak{Z}, z)$, wenn die Maschenweite $|\mathfrak{Z}|$ der Zerlegung gegen Null geht. Formal geschrieben gilt also*

$$\forall \varepsilon > 0,\ \exists \delta > 0\ \forall \mathfrak{Z}\ \forall \mathbf{z} : |\mathfrak{Z}| < \delta \implies \left| R(f, \mathfrak{Z}, \mathbf{z}) - \int_a^b f(x)\,\mathrm{d}x \right| < \varepsilon,$$

*wobei $\mathfrak{Z}$ über die Zerlegungen von $[a, b]$ läuft und $\mathbf{z}$ über die erlaubten Wahlen von Zwischenpunkten der Zerlegung $\mathfrak{Z}$ (wie in Definition 5.78) läuft.*

*Bemerkung.* (a) Die Konvergenz der Riemann-Summen wie in obigem Satz ist sogar eine Charakterisierung der Riemann-Integrierbarkeit. Es gibt auch noch weitere, äquivalente Bedingungen, aber wir begnügen uns mit der Aussage in Satz 5.79.
(b) Die Konvergenz in Satz 5.79 lässt sich auch mittels dem Filter $\mathcal{F}$ auf der Menge

$$X = \{(\mathfrak{Z}, \mathbf{z}) \mid \mathfrak{Z} \text{ ist eine Zerlegung und } \mathbf{z} \text{ ist ein Vektor von erlaubten Zwischenpunkten}\}$$

formulieren, wobei $F \in \mathcal{F}$ genau dann, wenn es ein $\delta > 0$ gibt so dass $F_\delta \subseteq F$, wobei $F_\delta$ die Menge der Zerlegungen mit erlaubten Zwischenpunkten $(\mathfrak{Z}, \mathbf{z})$ ist mit $|\mathfrak{Z}| < \delta$.

Die Idee des Beweises ist einfach formuliert, denn eigentlich sollten die Untersummen untere Schranken und die Obersummen obere Schranken für die Riemann-Summen darstellen. Doch gibt es hierbei mehrere technische Probleme, die mit den Endpunkten (welche bei den Unter- und Obersummen ignoriert werden und bei den Riemann-Summen erlaubt sind) zu tun

haben. Wir müssen also im korrekten Beweis die quantitativen Auswirkung dieser Endpunkte untersuchen und abschätzen.

*Beweis von Satz 5.79.* Sei $\varepsilon > 0$. Dann existieren nach Proposition 4.12 Treppenfunktionen $u, o \in TF([a,b])$ mit $u \leq f \leq o, \int_a^b (o-u)\,\mathrm{d}x < \varepsilon$ und $\int_a^b u\,\mathrm{d}x \leq \int_a^b f\,\mathrm{d}x \leq \int_a^b o\,\mathrm{d}x$ . Es sei

$$\mathfrak{Z}_0 = \{a = y_0 < y_1 < \ldots < y_N = b\}$$

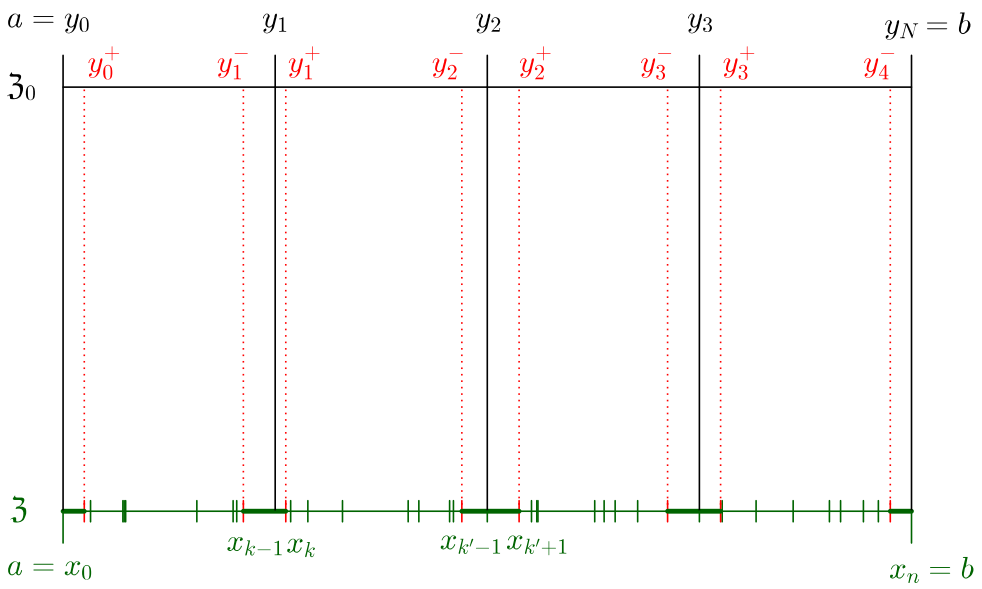
eine Zerlegung von $[a,b]$ in gemeinsame Konstanzintervalle für $u$ und $o$ und seien

$$u|_{(x_{\ell-1},x_\ell)} = c_\ell,$$
$$o|_{(x_{\ell-1},x_\ell)} = d_\ell$$

die Konstanzwerte von $u$ und $o$ für $\ell \in \{1,\ldots,N\}$. Da $f, u, o$ beschränkt sind, existiert auch ein $M > 0$ mit $|f(x)|, |u(x)|, |o(x)| \leq M$ für alle $x \in [a,b]$.

Wir wollen $\delta > 0$ wie im Satz erst später definieren, doch versprechen wir bereits, $\delta$ nur abhängig von $\mathfrak{Z}_0$ und $M$ zu definieren.

Sei nun $\mathfrak{Z} = \{a = x_0 < x_1 < \ldots < x_n = b\}$ eine Zerlegung von $[a,b]$ mit Maschenweite kleiner als ein $\delta > 0$ und $\mathbf{z} \in [a,b]^n$ eine erlaubte Wahl von Zwischenpunkten (so dass $z_k \in [x_{k-1}, x_k]$ für $k \in \{1,\ldots,n\}$). Wir werden folgendes Bild im Beweis verwenden.




Figur 5.6: Wir haben die Zerlegung $\mathfrak{Z}_0$ (in Abhängigkeit von $\varepsilon > 0$) fix gewählt und werden sie auch nicht ändern, während wir $\delta > 0$ verkleinern und damit $\mathfrak{Z}$ komplizierter machen. In diesem Sinne ist es am besten sich vorzustellen, dass $\mathfrak{Z}_0$ eine Zerlegung in wenige (hier im Bild $N = 4$) Punkte ist. Viele der Zerlegungsintervalle $(x_{k-1}, x_k)$ von $\mathfrak{Z}$ sind dann in einem Zerlegungsintervall $(y_{\ell-1}, y_\ell)$ von $\mathfrak{Z}_0$ enthalten, die Ausnahmen (höchstens $2N$) sind markiert.

Für $\ell \in \{1,\ldots,N\}$ definieren wir $y_{\ell-1}^+$ als das kleinste $x_k$ in $\mathfrak{Z}$ mit $x_k > y_{\ell-1}$. Wegen $|\mathfrak{Z}| < \delta$ gilt daher $y_{\ell-1}^+ - y_{\ell-1} < \delta$ (siehe Figur 5.6). Analog definieren wir die Zahl $y_\ell^-$ für

$\ell \in \{1, \ldots, N\}$ als das grösste $x_k$ in $\mathfrak{Z}$ mit $x_k < y_\ell$, womit wiederum $y_\ell - y_\ell^- < \delta$. Zur Vereinfachung des Arguments nehmen wir an, dass $\delta \leq \frac{1}{3} \min_{\ell=1,\ldots,N}(y_\ell - y_{\ell-1})$. Aus dieser Annahme folgt

$$y_\ell^- - y_{\ell-1}^+ > y_\ell - \delta - y_{\ell-1} - \delta = y_\ell - y_{\ell-1} - 2\delta > 0,$$

was die Abfolge der Punkte

$$y_0 < y_0^+ < y_1^- < y_1 < y_1^+ < \ldots < y_n^- < y_n$$

wie in Figur 5.6 bestätigt. Wir verkleinern nun die Intervalle $[y_{\ell-1}, y_\ell]$ zu den Intervallen $[y_{\ell-1}^+, y_\ell^-]$, definieren

$$U' = \sum_{\ell=1}^{N} c_\ell (y_\ell^- - y_{\ell-1}^+)$$

und erhalten wegen $|u(x)| \leq M$ für alle $x \in [a, b]$

$$\Big| \int_a^b u(x)\,\mathrm{d}x - U' \Big| = \Big| \sum_{\ell=1}^{N} c_\ell (y_\ell - y_\ell^- + y_{\ell-1}^+ - y_{\ell-1}) \Big| \leq 2NM\delta.$$

Ebenso gilt für $O' = \sum_{\ell=1}^{N} d_\ell (y_\ell^- - y_{\ell-1}^+)$, dass $\Big| \int_a^b o(x)\,\mathrm{d}x - O' \Big| \leq 2NM\delta$.

Für $\ell \in \{1, \ldots, N\}$ wollen wir den Teil der Riemann-Summe über die Teilintervalle der Zerlegung $\mathfrak{Z}$, die in $(y_{\ell-1}, y_\ell)$ enthalten sind, als $R_\ell$ bezeichnen. Genauer formuliert ist

$$R_\ell = \sum_{k:[x_{k-1},x_k] \subseteq (y_{\ell-1},y_\ell)} f(z_k)(x_k - x_{k-1}).$$

Summieren wir über $\ell \in \{1, \ldots, N\}$, so erhalten wir $R' = \sum_{\ell=1}^{N} R_\ell$, wobei sich $R'$ von $R(f, \mathfrak{Z}, \mathbf{z})$ um höchstens $2N$ Summanden unterscheidet. (Falls $\mathfrak{Z}$ feiner ist als $\mathfrak{Z}_0$, dann sind es in der Tat $2N$ Summanden, sonst weniger.) Daher ist

$$|R' - R(f, \mathfrak{Z}, \mathbf{z})| \leq 2NM\delta.$$

Für $\ell \in \{1, \ldots, N\}$ und $k \in \{1, \ldots, n\}$ mit $[x_{k-1}, x_k] \subseteq (y_{\ell-1}, y_\ell)$ folgt aus $u \leq f \leq o$ und $z_k \in [x_{k-1}, x_k] \subseteq (y_{\ell-1}, y_\ell)$, dass

$$c_\ell = u(z_k) \leq f(z_k) \leq o(z_k) = d_\ell$$

(da $u$ und $o$ auf $(y_{\ell-1}, y_\ell)$ konstant sind). Wir multiplizieren diese Ungleichung mit $(x_k - x_{k-1})$ und summieren über jene $k \in \{1, \ldots, n\}$ mit $[x_{k-1}, x_k] \subseteq (y_{\ell-1}, y_\ell)$. Es ergibt sich dank zwei Teleskopsummen

$$c_\ell (y_\ell^- - y_{\ell-1}^+) \leq R_\ell \leq d_\ell (y_\ell^- - y_{\ell-1}^+).$$

Summiert man jetzt über $\ell \in \{1, \dots, N\}$, so erhält man

$$U' \leq R' \leq O'. \tag{5.14}$$

Zusammenfassend haben wir also gezeigt, dass man durch Entfernen von höchstens $2N$ Teilintervallen die Untersumme $\int_a^b u(x)\,\mathrm{d}x$, die Obersumme $\int_a^b o(x)\,\mathrm{d}x$ und die Riemann-Summe $R(f, \mathfrak{Z}, \mathbf{z})$ um jeweils höchstens $2NM\delta$ verändern und dann (5.14) erreichen kann. Kombinieren wir dies mit

$$\int_a^b f(x)\,\mathrm{d}x - \varepsilon \leq \int_a^b u(x)\,\mathrm{d}x \leq \int_a^b f(x)\,\mathrm{d}x \leq \int_a^b o(x)\,\mathrm{d}x \leq \int_a^b f(x)\,\mathrm{d}x + \varepsilon,$$

so erhalten wir

$$R(f, \mathfrak{Z}, \mathbf{z}) \leq R' + 2MN\delta \leq O' + 2MN\delta \leq \int_a^b o(x)\,\mathrm{d}x + 4MN\delta \leq \int_a^b f(x)\,\mathrm{d}x + 4MN\delta + \varepsilon$$

und

$$R(f, \mathfrak{Z}, \mathbf{z}) \geq R' - 2MN\delta \geq U' - 2MN\delta \geq \int_a^b u(x)\,\mathrm{d}x - 4MN\delta \geq \int_a^b f(x)\,\mathrm{d}x - 4MN\delta - \varepsilon.$$

Wir setzen $\delta = \min\left\{\frac{\varepsilon}{MN}, \frac{1}{3}(y_1 - y_0), \dots, \frac{1}{3}(y_N - y_{N-1})\right\}$ und erhalten

$$\left|R(f, \mathfrak{Z}, \mathbf{z}) - \int_a^b f(x)\,\mathrm{d}x\right| \leq 5\varepsilon.$$

Da $\varepsilon > 0$, die Zerlegung $\mathfrak{Z}$ mit Maschenweite kleiner $\delta$ und die erlaubt Wahl der Zwischenpunkte $\mathbf{z}$ beliebig waren, folgt der Satz. □

**Applet 5.80** (Riemann-Summen für die Parabel)**.** *Wir sehen Riemann-Summen für die Parabel aus Abschnitt 1.1, wobei die Zwischenpunkte zufällig gewählt werden.*

**Übung 5.81** (Charakterisierung)**.** *Seien $a < b$ reellen Zahlen und $f : [a, b] \to \mathbb{R}$. Zeigen Sie die Umkehrung zu Satz 5.79, also dass die Konvergenz der Riemann-Summen für $f$ zu einer Zahl $I$ (wie in Satz 5.79) die Riemann-Integrierbarkeit von $f$ und die Gleichung $\int_a^b f(x)\,\mathrm{d}x = I$ impliziert.*

## 5.6 Landau Notation

Wir führen nun zwei geläufige Notationen ein, die das asymptotische Verhalten einer Funktion mit dem asymptotischen Verhalten einer anderen Funktion vergleichen – also ein relatives asymptotisches Verhalten beschreiben.

Sei $D \subseteq \mathbb{R}$ eine Teilmenge und $x_0 \in \overline{\mathbb{R}}$ ein Häufungspunkt (also mit $\dot{U}_\delta(x_0) \cap D \neq \emptyset$ für alle $\delta > 0$). Seien $f, g : D \to \mathbb{R}$ Funktionen. Wir schreiben

$$f(x) = O(g(x)) \text{ für } x \to x_0,$$

falls ein $\delta > 0$ und eine Konstante $M > 0$ existieren, so dass $|f(x)| \leq M|g(x)|$ für alle $x \in D \cap \dot{U}_\delta(x_0)$. In anderen Worten, $f$ ist „**Gross-O**“ von $g$ für $x \to x_0$, falls in einer punktierten Umgebung von $x_0$ die Funktion $f$ durch eine Konstante mal $|g|$ beschränkt werden kann. Obwohl dies für obige Defintion nicht notwendig ist, werden wir eigentlich immer vorraussetzen, dass $g(x) \neq 0$ für alle $x \in D$ oder zumindest für alle $x \in \dot{U}_{\delta_0}(x_0)$ für ein $\delta_0 > 0$. In diesem Fall ist $f$ genau dann Gross-O von $g$, wenn $\frac{f}{g}$ in einer $\delta$-Umgebung von $x_0$ beschränkt ist. Nochmals in anderen Worten ist $f = O(g)$ gleichbedeutend damit, dass $f$ nicht viel grösser als $g$ ist wenn $x$ in der Nähe von $x_0$ liegt. Zum Beispiel gilt

- $\frac{x}{x+1} = O(1)$ für $x \to \infty$,
- für jedes $x_0 \in \mathbb{R}$ gilt $x^2 = O(x)$ für $x \to x_0$, und insbesondere auch
- $x^2 = O(x)$ für $x \to 0$, aber
- $x^2$ ist nicht gleich $O(x)$ für $x \to \infty$ da $\frac{x^2}{x} = x$ in keiner Umgebung von $\infty$ beschränkt ist.

Der Vorteil der Notation ist, dass wir den Namen (oben $M$) für die obere Schranke nicht einführen. Falls uns diese Konstante nicht besonders interessiert, dann können wir uns dadurch bei Rechnungen von einer Zeile zur nächsten auf das Wesentlich konzentrieren. Man spricht in diesem Zusammenhang auch von der **impliziten Konstante**, falls diese nach einigen Rechenschritten doch erwähnt werden muss.

Wenn $f$ nicht nur durch $g$ beschränkt ist, sondern asymptotisch gegenüber $g$ vernachlässigbar ist, dann sagen wir, dass $f$ „**Klein-o**“ von $g$ ist für $x \to x_0$. Genauer formuliert, wir schreiben

$$f(x) = o(g(x)) \text{ für } x \to x_0,$$

falls für jedes $\varepsilon > 0$ ein $\delta > 0$ existiert mit $|f(x)| \leq \varepsilon|g(x)|$ für alle $x \in D \cap \dot{U}_\delta(x_0)$. Wie zuvor wollen wir meist $g(x) \neq 0$ auf $D$ annehmen. In diesem Fall gilt $f(x) = o(g(x))$ für $x \to x_0$ genau dann, wenn

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = 0.$$

Man beachte, dass, falls die eigentlichen Grenzwerte $\lim_{x\to x_0} f(x)$ und $\lim_{x\to x_0} g(x)$ existieren und nicht Null sind, sicherlich $f(x) = O(g(x))$ für $x \to x_0$ erfüllt ist, aber die stärkere Aussage $f(x) = o(g(x))$ für $x \to x_0$ falsch ist. Beide Notationen sind also vor allem dann interessant, wenn die Grenzwerte entweder null oder unendlich sind. Zum Beispiel gilt

- $x = o(x^2)$ für $x \to \infty$ (aber nicht umgekehrt) und
- $x^2 = o(x)$ für $x \to 0$ (aber nicht umgekehrt).

Diese Notationen machen analog Sinn für andere Bewegungen wie zum Beispiel $x \searrow x_0$, allgemeine Filter, und können insbesondere auch für Folgen verwendet werden. VO: 15.11.17

**Übung 5.82** (Klein-o Asymptotiken)**.** *Zeigen Sie, dass die Asymptotiken*

$$x^p = o(x) \text{ für } x \to 0, \qquad x = o(x^p) \text{ für } x \to \infty$$
$$x^a = o(e^x) \text{ für } x \to \infty, \qquad \log(x) = o(x^b) \text{ für } x \to \infty$$

*für jedes $p > 1$, $a \in \mathbb{R}$ und $b > 0$ zutreffen, wobei sie Übung 5.60 verwenden dürfen.*

**Übung 5.83** (Rechnen mit der Landau Notation)**.** *Seien $D, x_0$ wie oben und $f, f_1, f_2, g$ reellwertige Funktionen auf $D$. Zeigen Sie, dass falls $f(x) = o(g(x))$, $f_1(x) = o(g(x))$ und $f_2(x) = o(g(x))$ für $x \to x_0$, dann auch*

$$f_1(x) + f_2(x) = o(g(x)) \quad \text{für } x \to x_0,$$
$$\alpha f(x) = o(g(x)) \quad \text{für } x \to x_0$$

*für $\alpha \in \mathbb{R}$ und analog für Gross-O.*

Die Landau Notation wird in vielen Situationen auch als Platzhalter verwendet, um beispielsweise auszudrücken, dass ein Term in einer Summe schneller anwächst oder abfällt als die anderen. In einem Ausdruck der Form

$$f(x) + o(g(x)) \quad \text{für } x \to x_0$$

steht der Term $o(g(x))$ für eine implizite Funktion $h : D \to \mathbb{R}$ mit der Eigenschaft

$$h(x) = o(g(x)) \quad \text{für } x \to x_0,$$

also soll $\tilde{f}(x) = f(x) + o(f(x))$ die Asymptotik $\tilde{f}(x) - f(x) = h(x) = o(g(x))$ für $x \to x_0$ erfüllen. Dies gilt analog ebenso für die Gross-O Notation.

Beispielsweise schreibt man (nach Polynomdivision mit Rest)

$$\begin{aligned}\frac{x^3 - 7x^2 + 6x + 2}{x^2 + x - 34} &= x - 8 + o(1) \quad \text{für } x \to \infty \\ &= x + O(1) \quad \text{für } x \to \infty \\ &= x + o(x) \quad \text{für } x \to \infty,\end{aligned}$$

und erinnert sich auf der rechten Seite somit nur an jene Terme, die den Hauptteil der Bewegung $x \to \infty$ ausmacht. Es mag vielleicht überraschen, dass im obigen Beispiel alle drei Formeln zutreffen oder nützlich sein könnten. Doch folgen diese Behauptungen direkt aus der Polynomdivision und je nach Zusammenhang will man vielleicht die etwas genauere Aussage mit Fehler $o(1)$ oder die gröbere Aussage mit Hilfe des Fehlers $o(x)$ verwenden.

Derartige asymptotische Aussagen helfen in komplizierteren Argumenten und Berechnungen den Fokus auf die wesentlichen Teile einer Berechnungen zu lenken. Doch liegen in diesem Verstecken von gewissen Ausdrücken auch Risiken für Fehler, zum Beispiel wenn wir die Fehlerterme unbeschränkt oft addieren wollen oder die Fehlerterme von weiteren Parametern abhängen und diese Abhängigkeit auf Grund der Notation vergessen wird. Wir werden die Landau Notation sporadisch aber doch immer wieder einmal einsetzen um das Wesentliche an einer Aussage zu betonen.

## 5.7 Normen und Konvergenz auf Vektorräumen

In Kapitel 2 haben wir bereits gesehen, wie man Distanzen auf $\mathbb{R}$ oder $\mathbb{C}$ messen kann. In Analogie dazu möchten wir hier verschiedene Varianten von Normen definieren, welche die Rolle des Absolutbetrags übernehmen und Abstände in Vektorräumen messen werden. Da wir im zweiten Semester auf das Konzept einer Norm nochmals zu sprechen kommen werden, werden wir uns hier vor allem auf die Vektorräume $\mathbb{R}^d$ oder $\mathbb{C}^d$ konzentrieren. Wir fixieren für den ganze Abschnitt eine Dimension $d \in \mathbb{N}$ (wobei wir eigentlich an den Fall $d \geq 2$ interessiert sind). Wir schreiben Vektoren in $\mathbb{R}^d$ oder $\mathbb{C}^d$ in der Form

$$\mathbf{v} = (v_1, \ldots, v_d)^t = \begin{pmatrix} v_1 \\ \vdots \\ v_d \end{pmatrix},$$

wobei $t$ die „Transposition“ des platzsparenden Zeilenvektors zu einem Spaltenvektor bezeichnet.

**Definition 5.84** (Normen)**.** Sei $V$ ein Vektorraum über $\mathbb{K} = \mathbb{R}$ (oder $\mathbb{K} = \mathbb{C}$). Eine **Norm** auf $V$ ist eine Abbildung $\|\cdot\| : v \in V \mapsto \|v\| \in \mathbb{R}_{\geq 0}$, die folgende drei Eigenschaften erfüllt.

- (Definitheit) Für alle $v \in V$ gilt $\|v\| = 0 \iff v = 0$.
- (Homogenität) Für alle $v \in V$ und alle $\alpha \in \mathbb{K}$ gilt $\|\alpha v\| = |\alpha|\|v\|$.
- (Dreiecksungleichung) Für alle $v_1, v_2 \in V$ gilt $\|v_1 + v_2\| \leq \|v_1\| + \|v_2\|$.

Man nennt $V$ gemeinsam mit der Norm $\|\cdot\|$ auch einen **normierten Vektorraum**.

**Beispiel 5.85** (Maximumsnorm und Einsnorm)**.** *Sei $d \in \mathbb{N}$. Zu $j \in \{1, \ldots, d\}$ bezeichnen wir mit $\pi_j$ die **Projektion***

$$\pi_j : \mathbf{v} = (v_1, \ldots, v_d)^t \in \mathbb{C}^d \mapsto v_j \in \mathbb{C}$$

*auf die $j$-te Komponente. Die **Maximumsnorm** oder **Unendlichnorm** $\|\cdot\|_\infty$ ist definiert durch*

$$\|v\|_\infty = \max_{j=1,\ldots,d} |\pi_j(v)|$$

*für $v \in \mathbb{C}^d$ und die **1-Norm** ist definiert durch*

$$\|\mathbf{v}\|_1 = \sum_{j=1}^{d} |\pi_j(\mathbf{v})|$$

*für $\mathbf{v} \in \mathbb{C}^d$. Die Maximumsnorm und die 1-Norm auf $\mathbb{R}^d$ sind durch die gleichen Formeln definiert (oder äquivalent dazu durch Einschränkung auf $\mathbb{R}^d$). Wir überlassen Ihnen die Überprüfung der Eigenschaften in Definition 5.84.*

### 5.7.1 Die euklidsche Norm

Sei $d \in \mathbb{N}$. Wir möchten nun eine für die sogenannte „Euklidische Geometrie“ natürliche Norm auf $V = \mathbb{C}^d$ definieren und besprechen. Das **Euklidische innere Produkt** (oder **Skalarprodukt**) von

$$\begin{aligned} \mathbf{v} &= (v_1, \dots, v_d)^t \quad \text{und} \\ \mathbf{w} &= (w_1, \dots, w_d)^t \end{aligned}$$

ist definiert durch

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{k=1}^{d} v_k \overline{w_k}.$$

Dieses erfüllt folgende Eigenschaften:

- (Sesquilinearität) Für alle $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}, \mathbf{w}_1, \mathbf{w}_2, \mathbf{w} \in V$ und $\alpha_1, \alpha_2 \in \mathbb{C}$ gilt

$$\begin{aligned} \langle \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2, \mathbf{w} \rangle &= \alpha_1 \langle \mathbf{v}_1, \mathbf{w} \rangle + \alpha_2 \langle \mathbf{v}_2, \mathbf{w} \rangle \\ \langle \mathbf{v}, \alpha_1 \mathbf{w}_1 + \alpha_2 \mathbf{w}_2 \rangle &= \overline{\alpha}_1 \langle \mathbf{v}, \mathbf{w}_1 \rangle + \overline{\alpha}_2 \langle \mathbf{v}, \mathbf{w}_2 \rangle . \end{aligned}$$

- (Symmetrie) Für alle $\mathbf{v}, \mathbf{w} \in V$ gilt $\langle \mathbf{v}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{v} \rangle}$.
- (Definitheit) Für $\mathbf{v} \in V$ gilt $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ und $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ genau dann, wenn $\mathbf{v} = 0$ ist.

Wir bemerken, dass das Wort „sesqui“ für eineinhalb steht: das innere Produkt ist linear im ersten Argument und „halblinear“ im zweiten Argument. Das **reelle innere Produkt** auf $\mathbb{R}^d$ ist durch dieselbe Formel definiert und erfüllt an Stelle der Sesquilinearität die **Bilinearität**, also die Linearität in beiden Argumenten (bei festgehaltenem anderem Argument).

Den Beweis der Sesquilinearität und der Symmetrie überlassen wir als Übung. Wir beweisen Definitheit. Sei also $\mathbf{v} = (v_1, \dots, v_d)^t \in V = \mathbb{C}^d$. Dann gilt

$$\langle \mathbf{v}, \mathbf{v} \rangle = \sum_{k=1}^{d} |v_k|^2 \geq 0.$$

Wenn $\mathbf{v} = 0$ ist, dann ist auch $\langle \mathbf{v}, \mathbf{v} \rangle = 0$. Wenn $\langle \mathbf{v}, \mathbf{v} \rangle = \sum_{k=1}^{n} |v_k|^2 = 0$ ist, dann muss jeder Summand verschwinden. Also gilt $v_k = 0$ für alle $k \in \{1, \dots, d\}$ und damit $\mathbf{v} = 0$.

Unter Verwendung der obigen Eigenschaften des Euklidschen inneren Produkts lässt sich nun eine Norm definieren. Die **Euklidsche Norm** auf $V = \mathbb{C}^d$ ist gegeben durch

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = \sqrt{\sum_{k=1}^{d} |v_k|^2}$$

für alle $\mathbf{v} = (v_1, \dots, v_d)^t \in \mathbb{C}^d$. Sie wird auch die **2-Norm** genannt und dementsprechend als $\| \cdot \|_2$ geschrieben.

Wir möchten im Folgenden zeigen, dass die Euklidsche Norm in der Tat eine Norm ist. Definitheit und Homogenität des Euklidschen Norm folgen direkt aus den Eigenschaften des Euklidschen inneren Produkts (wieso?). Um die Dreiecksungleichung zu beweisen, benötigen wir folgende fundamentale Abschätzung.

**Proposition 5.86** (Cauchy-Schwarz Ungleichung)**.** *Sei $d \in \mathbb{N}$ und $V = \mathbb{C}^d$. Dann gilt für alle $\mathbf{v}, \mathbf{w} \in V$ die Ungleichung*

$$|\langle \mathbf{v}, \mathbf{w} \rangle| \leq \|\mathbf{v}\|\|\mathbf{w}\|. \tag{5.15}$$

*Des Weiteren gilt Gleichheit in* (5.15) *genau dann, wenn $\mathbf{v}, \mathbf{w}$ linear abhängig sind (das heisst, wenn ein $\alpha \in \mathbb{C}$ existiert mit $\alpha\mathbf{v} = \mathbf{w}$ oder $\mathbf{v} = \alpha\mathbf{w}$).*

Das innere Produkt zweier Vektoren lässt sich also durch die „Normen“ der beiden Vektoren auf eine konkrete Art und Weise kontrollieren. Wir merken an, dass der folgende Beweis nur die „Axiome“ des inneren Produktes Sesquilinearität, Symmetrie und Definitheit und nicht die konkrete Formel in der Definition des Euklidschen inneren Produktes verwendet.

*Beweis.* Falls $\mathbf{v} = 0$ oder $\mathbf{w} = 0$ ist, so steht auf beiden Seiten von (5.15) Null und die Vektoren $\mathbf{v}, \mathbf{w}$ sind linear abhängig. Wir nehmen also an, dass $\mathbf{v} \neq 0$ und $\mathbf{w} \neq 0$. Dann gilt für $\alpha = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{w}\|^2}$

$$\begin{aligned}
\|\mathbf{v} - \alpha\mathbf{w}\|^2 &= \langle \mathbf{v} - \alpha\mathbf{w}, \mathbf{v} - \alpha\mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{v} - \alpha\mathbf{w} \rangle - \alpha \langle \mathbf{w}, \mathbf{v} - \alpha\mathbf{w} \rangle \\
&= \langle \mathbf{v}, \mathbf{v} \rangle - \overline{\alpha} \langle \mathbf{v}, \mathbf{w} \rangle - \alpha \langle \mathbf{w}, \mathbf{v} \rangle + |\alpha|^2\|\mathbf{w}\|^2 = \|\mathbf{v}\|^2 - \overline{\alpha} \langle \mathbf{v}, \mathbf{w} \rangle - \alpha\overline{\langle \mathbf{v}, \mathbf{w} \rangle} + |\alpha|^2\|\mathbf{w}\|^2 \\
&= \|\mathbf{v}\|^2 - 2\frac{|\langle \mathbf{v}, \mathbf{w} \rangle|^2}{\|\mathbf{w}\|^2} + \frac{|\langle \mathbf{v}, \mathbf{w} \rangle|^2}{\|\mathbf{w}\|^4}\|\mathbf{w}\|^2 = \|\mathbf{v}\|^2 - \frac{|\langle \mathbf{v}, \mathbf{w} \rangle|^2}{\|\mathbf{w}\|^2}.
\end{aligned}$$

Der Ausdruck $\|\mathbf{v} - \alpha\mathbf{w}\|^2$ ist aber nicht-negativ und es folgt

$$\|\mathbf{v}\|^2 - \frac{|\langle \mathbf{v}, \mathbf{w} \rangle|^2}{\|\mathbf{w}\|^2} \geq 0.$$

Somit folgt $\|\mathbf{v}\|^2\|\mathbf{w}\|^2 \geq |\langle \mathbf{v}, \mathbf{w} \rangle|^2$, was die gewünschte Ungleichung (5.15) impliziert. Gleichheit gilt genau dann, wenn $\|\mathbf{v} - \alpha\mathbf{w}\| = 0$ und somit $\mathbf{v} = \alpha\mathbf{w}$ ist. □

Alternativ lässt sich die Cauchy-Schwarz-Ungleichung auch wie folgt beweisen.

**Übung 5.87** (Cauchy-Ungleichung mit einem $\varepsilon$)**.** *Sei $\varepsilon > 0$. Zeigen Sie, dass alle $\mathbf{v}, \mathbf{w} \in \mathbb{R}^d$ die Abschätzung*

$$|\langle \mathbf{v}, \mathbf{w} \rangle| \leq \tfrac{\varepsilon^2}{2}\|\mathbf{v}\|^2 + \tfrac{1}{2\varepsilon^2}\|\mathbf{w}\|^2$$

*erfüllen und schliessen Sie daraus auf die Cauchy-Schwarz Ungleichung* (5.15)*.*

**Korollar 5.88** (Euklidische Norm)**.** *Sei $d \in \mathbb{N}$. Die Euklidische Norm definiert eine Norm auf $\mathbb{C}^d$.*

*Beweis.* Es verbleibt die Dreiecksungleichung zu beweisen. Seien $\mathbf{v}, \mathbf{w} \in \mathbb{C}^d$. Wir schätzen direkt ab unter Verwendung der Cauchy-Schwarz-Ungleichung

$$\begin{aligned}\|\mathbf{v}+\mathbf{w}\|^2 &= \langle \mathbf{v}+\mathbf{w}, \mathbf{v}+\mathbf{w}\rangle = \langle \mathbf{v}, \mathbf{v}+\mathbf{w}\rangle + \langle \mathbf{w}, \mathbf{v}+\mathbf{w}\rangle = \|\mathbf{v}\|^2 + \langle \mathbf{v}, \mathbf{w}\rangle + \langle \mathbf{w}, \mathbf{v}\rangle + \|\mathbf{w}\|^2 \\ &= \|\mathbf{v}\|^2 + \langle \mathbf{v}, \mathbf{w}\rangle + \overline{\langle \mathbf{v}, \mathbf{w}\rangle} + \|\mathbf{w}\|^2 = \|\mathbf{v}\|^2 + 2\,\mathrm{Re}(\langle \mathbf{v}, \mathbf{w}\rangle) + \|\mathbf{w}\|^2 \\ &\leq \|\mathbf{v}\|^2 + 2|\,\langle \mathbf{v}, \mathbf{w}\rangle\,| + \|\mathbf{w}\|^2 \leq \|\mathbf{v}\|^2 + 2\|\mathbf{v}\|\|\mathbf{w}\| + \|\mathbf{w}\|^2 = (\|\mathbf{v}\| + \|\mathbf{w}\|)^2,\end{aligned}$$

womit die Aussage nach Ziehen der Wurzel folgt. □

Durch Einschränkung auf $\mathbb{R}^d \subseteq \mathbb{C}^d$ erhalten wir auch das **Euklidische innere Produkt** und die **Euklidische Norm** auf $\mathbb{R}^d$. Alle oben bewiesenen Aussagen gelten analog für $\mathbb{R}^d$.

### 5.7.2 Normäquivalenz

Sei $d \in \mathbb{N}$. Auf $\mathbb{R}^d$ haben wir soweit 3 verschiedene Normen kennengelernt: die 1-Norm $\|\cdot\|_1$ (siehe Beispiel 5.85), die obige 2-Norm $\|\cdot\|_2$ (die Euklidsche Norm) und die Maximumsnorm (siehe Beispiel 5.85). Um diese zu vergleichen ist folgender Begriff nützlich.

**Definition 5.89** (Normäquivalenz)**.** Sei $V$ ein Vektorraum über $\mathbb{K} = \mathbb{R}$ (oder $\mathbb{K} = \mathbb{C}$) und seien $\|\cdot\|$ und $\|\cdot\|'$ zwei Normen auf $V$. Wir nennen $\|\cdot\|$ und $\|\cdot\|'$ **äquivalent**, falls Konstanten $c, C > 0$ existieren mit

$$c\|v\| \leq \|v\|' \leq C\|v\|$$

für alle $v \in V$.

In der Tat sind die 1-Norm, die euklidische Norm und die Maximumsnorm äquivalent. Denn es gelten die Ungleichungen

$$\begin{aligned}&\|\mathbf{v}\|_\infty \leq \|\mathbf{v}\|_1 \leq d\|\mathbf{v}\|_\infty \text{ und} \\ &\|\mathbf{v}\|_\infty \leq \|\mathbf{v}\|_2 \leq \sqrt{d}\|\mathbf{v}\|_\infty\end{aligned}$$

für alle $\mathbf{v} \in \mathbb{R}^d$. In der Tat behauptet die erste Ungleichung bloss, dass der maximale Absolutbetrag kleiner gleich der Summe der Absolutbeträge, und die Summe der Absolutbeträge kleiner gleich $d$ mal dem maximalen Absolutbetrages ist. Der erste Teil der zweiten Ungleichung ergibt sich analog und der zweite Teil folgt aus der Cauchy-Schwarz Ungleichung indem man den Vektor

$$\mathbf{w} = (\mathrm{sgn}(v_1), \ldots, \mathrm{sgn}(v_d))^t$$

definiert und damit

$$\|\mathbf{v}\|_1 = \langle \mathbf{v}, \mathbf{w}\rangle \leq \|\mathbf{v}\|_2\|\mathbf{w}\|_2 \leq \sqrt{d}\|\mathbf{v}\|_2$$

beweist.

Daher sind also die 1-, die 2- und die Maximumsnorm auf $\mathbb{R}^d$ äquivalent. (Tatsächlich werden wir im zweiten Semester zeigen, dass alle Normen auf $\mathbb{R}^d$ äquivalent sind.)

**Übung 5.90** (Zu Normäquivalenz)**.** *Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ und sei $V$ ein Vektorraum über dem Körper $\mathbb{K}$. Zeigen Sie, dass die Normäquivalenz (wie der Name sagt) eine Äquivalenzrelation auf der Menge der Normen auf $V$ definiert.*

### 5.7.3 Folgenkonvergenz

Wir möchten nun damit beginnen, Folgen und vor allem den Konvergenzbegriff für diese zu untersuchen. In Analogie zu Definition 5.2 führen wir folgenden Konvergenzbegriff ein.

**Definition 5.91** (Konvergenz von Folgen in normierten Vektorräumen)**.** Sei $V$ ein reeller oder komplexer Vektorraum und sei $\|\cdot\|$ eine Norm auf $V$. Ein **Grenzwert** einer Folge $(v_n)_n$ in $V$ ist ein $v \in V$ mit der Eigenschaft, dass es für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass $\|v_n - v\| < \varepsilon$ für alle $n \geq N$. Weiter heisst eine Folge $(v_n)_n$ in $V$ **konvergent**, falls ein **Grenzwert** existiert. Um die Abhängigkeit von der Norm hervorzuheben, sagt man auch, dass $(v_n)_n$ **bezüglich der Norm** $\|\cdot\|$ gegen $v$ konvergiert.

Genauso wie in Lemma 5.3 kann man nun zeigen, dass ein Grenzwert einer Folge eindeutig bestimmt ist, wenn er existiert. Man kann also wieder von *dem* Grenzwert sprechen.

Wir untersuchen nun die Abhängigkeit des Konvergenzbegriffes von der Wahl der Norm.

**Lemma 5.92** (Normäquivalenz und Konvergenz)**.** *Sei $V$ ein reeller oder komplexer Vektorraum und seien $\|\cdot\|$, $\|\cdot\|'$ zwei äquivalente Normen auf $V$. Sei $(v_n)_n$ eine Folge in $V$. Dann ist die Folge $(v_n)_n$ genau dann bezüglich der Norm $\|\cdot\|$ konvergent, wenn sie bezüglich der Norm $\|\cdot\|'$ konvergent ist. Weiter sind die Grenzwerte in diesem Fall gleich.*

*Beweis.* Da die Normen $\|\cdot\|$, $\|\cdot\|'$ äquivalent sind, gibt es positive Zahlen $c, C$ mit

$$c\|v\| \leq \|v\|' \leq C\|v\|$$

für alle $v \in V$. Wir nehmen nun an, dass $(v_n)_n$ bezüglich $\|\cdot\|$ gegen $v \in V$ konvergiert. Sei $\varepsilon > 0$ beliebig. Dann gibt es ein $N \in \mathbb{N}$ mit $\|v_n - v\| < \varepsilon/C$ für alle $n \geq N$. Insbesondere gilt

$$\|v_n - v\|' \leq C\|v_n - v\| < \varepsilon$$

für alle $n \geq N$. Also konvergiert $(v_n)_n$ auch bezüglich der Norm $\|\cdot\|'$ gegen $v$.

Für die Umkehrung können wir im obigem Argument die beiden Normen vertauschen. □

### 5.7.4 Konvergenz in endlich dimensionalen Vektorräumen

Lemma 5.92 und die Diskussion in Abschnitt 5.7.2 zeigen, dass in $\mathbb{R}^d$ Konvergenzbegriffe bezüglich den Normen $\|\cdot\|_1$, $\|\cdot\|_2$ und $\|\cdot\|_\infty$ äquivalent sind. Wie schon erwähnt, ist dies in der Tat für jede Norm auf $\mathbb{R}^d$ der Fall, da alle Normen auf $\mathbb{R}^d$ äquivalent sind (was wir im zweiten Semester beweisen werden). Insbesondere werden wir bei Konvergenz einer Folge

in $\mathbb{R}^d$ jeweils nicht dazusagen, ob wir die Konvergenz bezüglich $\|\cdot\|_1$, bezüglich $\|\cdot\|_2$ oder bezüglich $\|\cdot\|_\infty$ betrachten.

**Lemma 5.93** (Koordinatenweise Konvergenz)**.** *Sei $d \in \mathbb{N}$. Eine Folge im $\mathbb{R}^d$ konvergiert genau dann, wenn sie koordinatenweise konvergiert. Genauer konvergiert eine Folge $(\mathbf{v}_n)_n$ in $\mathbb{R}^d$ genau dann gegen $\mathbf{v}$, wenn für alle $j = 1, \ldots, d$ die Folge der Komponenten $(\pi_j(\mathbf{v}_n))_n$ gegen $\pi_j(\mathbf{v})$ konvergiert.*

*Beweis.* Sei $(\mathbf{v}_n)_n$ eine Folge in $\mathbb{R}^d$. Angenommen $(\mathbf{v}_n)_n$ konvergiert gegen $\mathbf{v} \in V$ und sei $\varepsilon > 0$. Es gibt somit ein $N \in \mathbb{N}$ mit $\|\mathbf{v}_n - \mathbf{v}\|_\infty < \varepsilon$ für alle $n \geq N$. Insbesondere gilt für $j = 1, \ldots, d$

$$|\pi_j(\mathbf{v}_n) - \pi_j(\mathbf{v})| = |\pi_j(\mathbf{v}_n - \mathbf{v})| \leq \|\mathbf{v}_n - \mathbf{v}\|_\infty < \varepsilon$$

für alle $n \geq N$, womit $(\pi_j(\mathbf{v}_n))_n$ gegen $\pi_j(\mathbf{v})$ konvergiert, da $\varepsilon > 0$ beliebig war.

Wir nehmen nun umgekehrt an, dass $(\pi_j(\mathbf{v}_n))_n$ gegen $\pi_j(\mathbf{v})$ konvergiert für jedes $j = 1, \ldots, d$. Sei $\varepsilon > 0$. Dann gibt es zu $j \in \{1, \ldots, d\}$ ein $N_j \in \mathbb{N}$ mit $|\pi_j(\mathbf{v}_n) - \pi_j(\mathbf{v})| < \varepsilon$ für alle $n \geq N_j$. Sei $N = \max\{N_1, \ldots, N_d\}$. Dann gilt für alle $n \geq N$

$$\|\mathbf{v}_n - \mathbf{v}\|_\infty = \max_{j=1,\ldots,d} |\pi_j(\mathbf{v}_n) - \pi_j(\mathbf{v})| < \varepsilon.$$

Da $\varepsilon > 0$ beliebig ist, folgt die Konvergenz von $(\mathbf{v}_n)$ gegen $\mathbf{v}$, was das Lemma beweist. □

Sei $X$ eine beliebige nicht-leere Menge und $V$ ein normierter reeller oder komplexer Vektorraum mit Norm $\|\cdot\|$. Wir sagen, dass eine Funktion $f : X \to V$ **beschränkt** ist, falls es eine reelle Zahl $M > 0$ gibt so dass $\|f(x)\| \leq M$ für alle $x \in X$. Diese Definition macht insbesondere für $V = \mathbb{R}^d$ Sinn, wobei wir wiederum nicht angeben müssen, welche der äquivalenten Normen $\|\cdot\|_1$, $\|\cdot\|_2$ oder $\|\cdot\|_\infty$ wir betrachten. Falls $X = \mathbb{N}$ erhalten wir den Begriff der Beschränktheit für Folgen.

**Theorem 5.94** (Heine-Borel)**.** *Sei $d \in \mathbb{N}$. Jede beschränkte Folge in $\mathbb{R}^d$ besitzt eine konvergente Teilfolge.*

Der Begriff der Folgenkompaktheit lässt sich nun ohne Probleme auf Teilmengen von $\mathbb{R}^d$ erweitern. Eine Teilmenge $K \subseteq \mathbb{R}^d$ heisst **folgenkompakt**, falls jede Folge in $K$ eine konvergente Teilfolge mit Grenzwert in $K$ besitzt. Die Aussage von Theorem 5.94 zeigt dann unter Verwendung dieses Begriffes, dass zum Beispiel Teilmengen von $\mathbb{R}^d$ der Form $[-M, M]^d$ für $M > 0$ folgenkompakt sind.

*Beweis.* Wir beweisen die Aussage mittels Induktion nach $d$. Für $d = 1$ gilt die Aussage des Satzes bereits wegen Satz 5.44.

Sei nun $(\mathbf{v}_n)_n$ eine beschränkte Folge in $\mathbb{R}^{d+1}$ und sei $M > 0$ mit $\|\mathbf{v}_n\|_\infty \leq M$ für alle $n \in \mathbb{N}$. Wir fassen die ersten $d$ Kompontenten von $\mathbf{v}_n \in \mathbb{R}^{d+1}$ zu einem Vektor $\mathbf{w}_n \in \mathbb{R}^d$ zusammen so dass $\mathbf{v}_n = (\mathbf{w}_n, \pi_{d+1}(\mathbf{v}_n))^t$. Es gilt $\|\mathbf{w}_n\| \leq M$ und $|\pi_{d+1}(\mathbf{v}_n)| \leq M$ für alle $n \in \mathbb{N}$. Also ist $(\mathbf{w}_n)$ eine beschränkte Folge in $\mathbb{R}^d$ und nach Induktionsvorraussetzung existiert eine Teilfolge

$n_k$ so dass $(\mathbf{w}_{n_k})_k$ in $\mathbb{R}^d$ konvergiert. Des Weiteren ist ebenso $\pi_{d+1}(\mathbf{v}_{n_k})$ eine beschränkte Folge in $\mathbb{R}$ und es existiert nach Satz 5.44 eine weitere Teilfolge $(\mathbf{v}_{n_{k_\ell}})_\ell$, so dass der Grenzwert $\lim_{\ell\to\infty} \pi_1(\mathbf{v}_{n_{k_\ell}})$ existiert.

Zusammenfassend erhalten wir mit Lemma 5.93 also, dass jede Komponente von der Folge $(\mathbf{v}_{n_{k_\ell}})_\ell$ konvergiert und daher auch die Folge in $\mathbb{R}^{d+1}$ konvergiert. Dies beendet den Induktionsschritt. □

In Analogie zu Definition 5.24 nennen wir eine Folge $(\mathbf{v}_n)_n$ in einem normierten Vektorraum $(V, \|\cdot\|)$ eine **Cauchy-Folge**, falls für alle $\varepsilon > 0$ eine $N \in \mathbb{N}$ existiert, so dass

$$\|\mathbf{v}_m - \mathbf{v}_n\| < \varepsilon$$

für alle $m, n \geq N$ gilt. Sind zwei Normen $\|\cdot\|$, $\|\cdot\|'$ äquivalent, so ist eine Folge genau dann eine Cauchy-Folge bezüglich $\|\cdot\|$, wenn sie bezüglich $\|\cdot\|'$ eine Cauchy-Folge ist. Dies kann man wie in Lemma 5.92 verifizieren. Insbesondere können wir schlicht von Cauchy-Folgen im $\mathbb{R}^d$ sprechen und werden die Norm nicht angeben.

Wie in Übung 5.50 erklärt, ist die Aussage von Satz 5.48 („alle Cauchy-Folgen in $\mathbb{R}$ sind konvergent“) äquivalent zum Vollständigkeitsaxiom. Wir möchten diese Aussage nun auf höherdimensionale Vektorräume erweitern.

**Satz 5.95** (Vollständigkeit endlich-dimensionaler Vektorräume)**.** *Sei $d \in \mathbb{N}$. Jede Cauchy-Folge im $\mathbb{R}^d$ ist konvergent.*

Die wesentliche Idee des Beweises ist zu zeigen, dass eine Cauchy-Folge beschränkt sein muss und damit nach dem Satz von Heine-Borel eine konvergente Teilfolge haben muss. Da aber eine Cauchy-Folge genau dann konvergiert, wenn sie eine konvergente Teilfolge hat, sind alle Cauchy-Folgen konvergent.

*Beweis.* Sei $(\mathbf{v}_n)_n$ eine Cauchy-Folge im $\mathbb{R}^d$. Dann gibt es ein $N \in \mathbb{N}$, so dass $\|\mathbf{v}_m - \mathbf{v}_n\|_2 < 1$ für alle $m, n \geq N$. Sei $M = \max\{\|\mathbf{v}_1\|_2, \ldots, \|\mathbf{v}_{N-1}\|_2, \|\mathbf{v}_N\|_2 + 1\}$. Dann gilt $\|\mathbf{v}_n\|_2 \leq M$ für alle $n \in \mathbb{N}$, womit $(\mathbf{v}_n)_n$ beschränkt ist. In der Tat, gilt dies per Definition für $n < N$ und folgt für $n \geq N$ aus der Abschätzung

$$\|\mathbf{v}_n\|_2 \leq \|\mathbf{v}_N\|_2 + \|\mathbf{v}_n - \mathbf{v}_N\|_2 < \|\mathbf{v}_N\|_2 + 1.$$

Nach Theorem 5.94 besitzt $(\mathbf{v}_n)_n$ eine konvergente Teilfolge $(\mathbf{v}_{n_k})_k$ mit Grenzwert $\mathbf{v}$. Sei nun $\varepsilon > 0$ und $K \in \mathbb{N}$ mit $\|\mathbf{v}_{n_k} - \mathbf{v}\|_2 < \varepsilon$ für alle $k \geq K$. Des Weiteren sei $N_0 \in \mathbb{N}$ mit $\|\mathbf{v}_m - \mathbf{v}_n\|_2 < \varepsilon$ für alle $m, n \geq N_0$. Wir wählen nun ein $k \geq K$ mit $n_k \geq N_0$. Für jedes $n \geq N_0$ gilt dann

$$\|\mathbf{v}_n - \mathbf{v}\|_2 \leq \|\mathbf{v}_n - \mathbf{v}_{n_k}\|_2 + \|\mathbf{v}_{n_k} - \mathbf{v}\|_2 < 2\varepsilon.$$

Also konvergiert auch $(\mathbf{v}_n)_n$ gegen $\mathbf{v}$, womit der Satz folgt. □

### 5.7.5 Stetigkeit

**Definition 5.96** (Stetigkeit)**.** Seien $(V, \|\cdot\|_V)$, $(W, \|\cdot\|_W)$ normierte Vektorräume, $D \subseteq V$ eine Teilmenge und sei $f : D \to W$ eine Abbildung. Wir nennen $f$ **stetig bei einem Punkt** $x_0 \in D$, falls für alle $\varepsilon > 0$ ein $\delta > 0$ mit

$$\|x - x_0\|_V < \delta \implies \|f(x) - f(x_0)\|_W < \varepsilon$$

für alle $x \in D$ existiert. Die Funktion $f$ ist **stetig**, falls sie bei jedem Punkt in $D$ stetig ist.

Da wir später in grösserer Allgemeinheit auf den Stetigkeitsbegriff zurückkehren werden, wollen wir uns hier ausschliesslich mit stetigen Funktionen auf einem Intervall $[a, b]$ mit Werten in $\mathbb{R}^d$ befassen. Solche Funktionen werden wir später auch Wege nennen. Ist $f : [a, b] \to \mathbb{R}^d$ stetig so können wir die Komponenten $f_1, \ldots, f_d : [a, b] \to \mathbb{R}$ von $f$ definieren durch $f_j = \pi_j \circ f$ für $j = 1, \ldots, d$, wobei $\pi_j$ wieder die Projektion auf die $j$-te Koordinate bezeichnet. Die Komponenten von $f$ beschreiben $f$ vollständig, da für alle $x \in [a, b]$ gilt

$$f(x) = (f_1(x), \ldots, f_d(x))^t.$$

Auch die Stetigkeit von $f$ lässt sich mit ihren Komponenten charakterisieren.

**Lemma 5.97** (Komponentenweise Stetigkeit)**.** *Sei $d \in \mathbb{N}$, sei $[a, b]$ ein Intervall zu $a < b$ in $\mathbb{R}$ und sei $f : [a, b] \to \mathbb{R}^d$ eine Funktion. Dann sind folgende Aussagen äquivalent:*

*(i) Die Funktion $f$ ist stetig.*

*(ii) Die Komponenten von $f$ sind stetig.*

**Wichtige Übung 5.98.** *Beweisen Sie Lemma 5.97.*

### 5.7.6 Vektorwertige Integrale

Unsere ursprüngliche Definition des Riemann-Integrals in Kapitel 4 verwendete die Ungleichung $\leq$ in $\mathbb{R}$ in zentraler Weise und kann deswegen nicht auf diese Weise für vektorwertige Funktionen verallgemeinert werden. Riemann-Summen lassen sich hingegen leicht verallgemeinern. Für $\mathbf{f} : [a, b] \to \mathbb{R}^d$, $\mathfrak{Z} = \{a = x_0 < x_1 < \ldots < x_n = b\}$ eine Zerlegung von $[a, b]$ und $\mathbf{z} \in [a, b]^n$ eine zulässige Wahl von Zwischenpunkten gemäss Definition 5.78 setzen wir wie zuvor

$$R(\mathbf{f}, \mathfrak{Z}, \mathbf{z}) = \sum_{k=1}^{n} \mathbf{f}(z_k)(x_k - x_{k-1}).$$

Des Weiteren sagen wir, dass $\mathbf{f} : [a, b] \to \mathbb{R}^d$ Riemann-integrierbar ist, falls

$$\mathbf{f}(x) = (f_1(x), \ldots, f_d(x))^t$$

für alle $x \in [a,b]$, und die Komponentenfunktionen $f_j : [a,b] \to \mathbb{R}$ für $j = 1, \dots, d$ Riemann-integrierbar sind (man vergleiche dies zu Lemma 5.97). Das Riemann-Integral wird dann komponentenweise definiert durch

$$\int_a^b \mathbf{f}(x)\,\mathrm{d}x = \left( \int_a^b f_1(x)\,\mathrm{d}x,\ \dots\ , \int_a^b f_d(x)\,\mathrm{d}x \right)^t.$$

Satz 5.79 gilt nun analog für Riemann-integrierbare Funktionen von $[a,b]$ nach $\mathbb{R}^d$. In der Tat gilt $\| \int_a^b \mathbf{f}(x)\,\mathrm{d}x - R(\mathbf{f}, \mathfrak{Z}, z)\|_\infty < \varepsilon$ genau dann, wenn für alle $j \in \{1, \dots, d\}$ die Ungleichung $\left| \int_a^b f_j(x)\,\mathrm{d}x - R(f_j, \mathfrak{Z}, z) \right| < \varepsilon$ erfüllt ist, was wir gemäss Satz 5.79 für genügend kleine Maschenweiten von $\mathfrak{Z}$ erzielen können.

Des Weiteren gilt auch die Dreiecksungleichung (vergleiche zu Satz 4.24) für vektorwertige Integrale: Für eine stetige Funktion $\mathbf{f} : [a,b] \to \mathbb{R}^d$ ist $\|\mathbf{f}\|_2 : [a,b] \to \mathbb{R}$ Riemann-integrierbar (da stetig) und es gilt

$$\left\| \int_a^b \mathbf{f}(x)\,\mathrm{d}x \right\|_2 \leq \int_a^b \|\mathbf{f}(x)\|_2\,\mathrm{d}x. \tag{5.16}$$

Die analoge Ungleichung gilt auch für andere Normen.

**Wichtige Übung 5.99** (Dreiecksungleichung für vektorwertige Integrale)**.** *Beweisen Sie Ungleichung 5.16.*

Obige Diskussion enthält mit $d = 2$ auch den Fall von komplexwertigen Funktionen

$$f : [a,b] \to \mathbb{C},$$

welche also genau dann Riemann-integrierbar sind, wenn $\mathrm{Re}(f)$ und $\mathrm{Im}(f)$ Riemann-integrierbar sind. Das Riemann-Integral ist in diesem Fall gegeben durch

VO: 16.11.17

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^b \mathrm{Re}(f(x))\,\mathrm{d}x + \mathrm{i} \int_a^b \mathrm{Im}(f(x))\,\mathrm{d}x.$$

## 5.8 Weitere Lernmaterialien

### 5.8.1 Verwendung des Kapitels

Dieses Kapitel stellt die Grundlagen für Konvergenzbetrachtungen bereit, wobei wir auch einige elementare Grenzwerte bereits berechnen konnten. Die Berechnung dieser Grenzwerte erforderte mit unserem derzeitigen Wissen noch sehr viel Geschick, was mit Hilfe der Regel von de l'Hôpital später erheblich einfacher werden wird. Das heisst, dass die wichtigsten Resultate dieses Kapitels nicht durch die konkreten Beispiele oder auch die ersten speziellen Berechnungsmethoden gegeben sind, sondern vielmehr durch die folgenden Sätze:

- Zusammenhang zwischen Folgenkonvergenz und Stetigkeit in Proposition 5.17, da dies für die Berechnung von Grenzwerten auch später oft notwendig sein wird.
- Konvergenzverhalten für monotone Folgen in Satz 5.34.
- Definition und Eigenschaften von Limes Superior in Satz 5.40 (und analog für Limes Inferior).
- Die Existenz von konvergenten Teilfolgen in Satz 5.44 und der mehrdimensionalen Verallgemeinerung in Theorem 5.94.
- Der Begriff der Cauchy-Folge und das Cauchy-Kriterium in Satz 5.48.
- Sandwich-Lemma für Folgen und Funktionen.

Wir bemerken noch, dass Limes Superior und Limes Inferior nützliche allgemeine Werkzeuge sind, die mitunter auch im Beweis der Konvergenz einer Folge auftreten können. Denn $\limsup_{n\to\infty} a_n \in \overline{\mathbb{R}}$ und $\liminf_{n\to\infty} a_n \in \overline{\mathbb{R}}$ können für jede reellwertige Folge $(a_n)_n$ betrachtet werden, und das Erfüllen der Gleichung $\limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n \in \mathbb{R}$ ist nach Korollar 5.43 zur Konvergenz der Folge äquivalent. Dies ist vergleichbar damit, dass (wie zum Beispiel im Beweis von Korollar 3.72 über das Maximum von stetigen Funktionen) das Supremum im Beweis für die Existenz eines Maximums wichtig sein kann.

Des Weiteren konnten wir die reelle Exponentialfunktion mit einem Grenzwert definieren, welche wir gemeinsam mit der Logarithmusfunktion und allgemeinen Potenzen ab nun mit den gewohnten Eigenschaften verwenden dürfen. Die Definition der Exponentialfunktion hat auch zu der Ungleichung

$$\left(1+\frac{x}{n}\right)^n \leq \exp(x)$$

für alle $x \geq 0$ und $n \in \mathbb{N}$ (siehe Abschnitt 5.3.2) geführt, welche für einige Grenzwertberechnungen nützlich war.

Auch haben wir gesehen, dass das Riemann-Integral sich als ein Grenzwert der sogenannten Riemann-Summen auffassen lässt, was auch zu einer Definition eines vektorwertigen Riemann-Integrals geführt hat.

Schlussendlich haben wir gesehen, dass sich die Begriffe der Stetigkeit und der Konvergenz auf Vektorräume übertragen lassen. Letzteres werden wir im zweiten Semester zu einer vollständigeren Theorie weiter ausbauen.

### 5.8.2 Übungen

**Übung** (Stetige Fortsetzung). *Seien $f_1, f_2 : \mathbb{R} \to \mathbb{R}$ stetige Funktionen mit der Eigenschaft, dass $f_1|_{\mathbb{Q}} = f_2|_{\mathbb{Q}}$. Zeigen Sie, dass $f_1 = f_2$ gilt.*

**Übung** (Eine Umkehrung des Zwischenwertsatzes). *Sei $I = [a, b] \subseteq \mathbb{R}$ ein Intervall zu $a < b$ und sei $f : [a, b] \to \mathbb{R}$ eine Funktion, die folgende Eigenschaften erfüllt:*

*(i) Für alle $y \in \mathbb{R}$ ist das Urbild $f^{-1}(\{y\})$ abgeschlossen.*

*(ii) $f$ erfüllt den Zwischenwertsatz, das heisst, für alle $x_1 < x_2$ in $I$ und für alle $c \in \mathbb{R}$ zwischen $f(x_1)$ und $f(x_2)$ gibt es ein $x \in [x_1, x_2]$ mit $f(x) = c$.*

*Zeigen Sie, dass $f$ stetig ist.*

**Übung.** *Zeigen Sie die Ungleichung $\mathrm{e}^{1-n} \leq \frac{n!}{n^n}$ für alle $n \in \mathbb{N}$. In der Tat werden wir später eine explizite Form der Asymptotik von $\frac{n!}{n^n}$ (das Gesetz von Stirling) sehen, welche diese Ungleichung verschärft.*

**Übung** (Häufungspunkte). *Sei $A \subseteq \mathbb{R}$ und $x_0 \in \mathbb{R}$. Zeigen Sie, dass folgende drei Aussagen äquivalent sind.*

*(i) $x_0$ ist ein Häufungspunkt der Menge $A$.*

*(ii) $x_0$ ist ein Häufungspunkt einer injektiven Folge $(a_n)_n$ mit Folgengliedern $a_n \in A$ für alle $n \in \mathbb{N}$.*

*(iii) $x_0$ ist der Grenzwert einer injektiven Folge $(a_n)_n$ mit Folgengliedern $a_n \in A$ für alle $n \in \mathbb{N}$.*

**Übung** (Abgeschlossene Menge der Häufungspunkte). *Zeigen Sie, dass die Menge der Häufungspunkte einer reellwertigen Folge (oder einer Teilmenge $A \subseteq \mathbb{R}$) eine abgeschlossene Teilmenge von $\mathbb{R}$ bildet.*

**Übung** (Landau Notation). *Begründen Sie, inwiefern die Gleichungen zu $k, \ell \in \mathbb{N}$*

$$o(x^k) + o(x^\ell) = o(x^{\max\{k,\ell\}}),\ o(x^k)o(x^\ell) = o(x^{k+\ell})$$

*für $x \to \infty$ Sinn ergeben. Verwenden Sie dies, um die Asymptotik für $x \to \infty$ von*

$$\frac{3x^4 - 5x + 2}{5x^2 + 2x - 13} + \frac{x^5 - 3x^3 + 7x + 17}{3x^2 + 2x - 18}$$

*sowie*

$$\frac{3x^4 - 5x + 2}{5x^4 + 5x^2 + 2x - 13} + \frac{x^5 - 3x^3 + 7x + 17}{3x^5 + 2x - 18}$$

*zu beschreiben.*

**Übung** (Gross- und Klein-Omega). *Seien zwei Funktionen $f, g : D \to \mathbb{R}$ auf einer Teilmenge $D \subseteq \mathbb{R}$ gegeben und sei $x_0 \in \overline{\mathbb{R}}$ ein Häufungspunkt von $D$. Definieren Sie in Analogie zur Definition von Gross-O und Klein-o die Beziehungen*

$$f(x) = \Omega(g(x)) \text{ für } x \to x_0$$

*und*

$$f(x) = \omega(g(x)) \text{ für } x \to x_0,$$

*welche zum Ausdruck bringen, dass $g$ in der Nähe von $x_0$ durch ein positives Vielfaches von $|f|$ beschränkt ist respektive dass $\frac{g(x)}{f(x)}$ gegen Null geht für $x \to x_0$.*

Wir wollen in der nächsten Übung den Zusammenhang zwischen „unseren axiomatisch eingeführten reellen Zahlen“ und den „reellen Zahlen als Steigung von quasi-linearen Abbildungen“ von Abschnitt 2.7.5 besprechen.

**Übung** (Steigungen von quasi-linearen Abbildungen). *(i) Sei $f : \mathbb{Z} \to \mathbb{Z}$ eine quasi-lineare Abbildung wie in Abschnitt 2.7.5. Zeigen Sie, dass*

$$\lim_{n\to\infty} \frac{f(n)}{n}$$

*in $\mathbb{R}$ existiert.*

*(ii) Sei nun $\mathcal{Q}$ die additive Gruppe der quasi-linearen Abbildungen. Zeigen Sie, dass die Abbildung*

$$\Psi : f \in \mathcal{Q} \mapsto \lim_{n\to\infty} \frac{f(n)}{n} \in \mathbb{R}$$

*ein Homomorphismus ist und $\Psi(\mathcal{K}) = \{0\}$.*

*(iii) Konstruieren Sie zu jedem $a \in \mathbb{R}$ eine quasi-lineare Abbildung $f \in \mathcal{Q}$ so dass $\Psi(f) = a$ ist.*

*(iv) Sei $f \in \mathcal{Q}$ quasi-linear so dass $\Psi(f) = 0$. Zeigen Sie, dass $f \in \mathcal{K}$ nur endlich viele Werte annimmt.*

*Zusammen sehen wir also in der Tat, dass $\mathcal{Q}/\mathcal{K}$ als abelsche Gruppe isomorph zu $\mathbb{R}$ ist. Mit etwas mehr Arbeit lässt sich beweisen, dass in der Tat ein Körperisomorphismus vorliegt.*

**Übung** (Folgenstetigkeit). *Seien $D \subseteq \mathbb{R}$ nicht-leer, $d \in \mathbb{N}$, und $f : D \to \mathbb{R}^d$ eine Funktion. Zeigen Sie, dass $f$ genau dann stetig ist, wenn $f$ folgenstetig ist, dass heisst, wenn für jede Folge $(x_n)_n$ in $D$ mit Grenzwert $x \in D$ die Folge $(f(x_n))_n$ in $\mathbb{R}^d$ nach $f(x)$ konvergiert.*

**Übung** (Cauchy-Folgen). *Zeigen Sie direkt, dass eine Folge im $\mathbb{R}^d$ genau dann eine Cauchy-Folge ist, wenn für jedes $j \in \{1, \ldots, d\}$ die reelle Folge der $j$-ten Komponenten eine Cauchy-Folge ist.*

**Übung** (Bilder von Cauchy-Folgen)**.** *Sei $D \subseteq \mathbb{R}$ eine Teilmenge und $f : D \to \mathbb{R}$ eine gleichmässig stetige Funktion. Zeigen Sie, dass $f$ Cauchy-Folgen auf Cauchy-Folgen abbildet (für jede Cauchy-Folge $(x_n)_n$ in $[a, b]$ ist auch $(f(x_n))_n$ eine Cauchy-Folge). Gilt dies auch für Funktionen, die stetig, aber nicht gleichmässig stetig sind?*

Unter Verwendung von Folgen und Satz 5.44 lassen sich viele Aussagen aus Kapitel 3 anders beweisen, was wir in den folgenden Übung illustrieren möchten.

**Übung** (Beschränktheit mit Hilfe von Folgen)**.** *Sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion auf einem kompakten Intervall $[a, b]$ zu $a < b$. Wir wollen zeigen, dass $f$ beschränkt ist.*

*(i) Gehen Sie per Widerspruch vor und finden Sie eine Folge $(x_n)_n$ in $[a, b]$ mit $|f(x_n)| > n$ für alle $n \in \mathbb{N}$.*

*(ii) Wenden Sie nun Satz 5.44 an.*

**Übung** (Gleichmässige Stetigkeit mit Hilfe von Folgen)**.** *Sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion auf einem kompakten Intervall $[a, b]$ zu $a < b$. Gehen Sie nach einem ähnlichen Prinzip vor wie in obiger Übung, um zu zeigen, dass $f$ gleichmässig stetig ist.*

### 5.8.3 Lernkarten

Sie können wiederum die Lernkarten oder den Graphen für Ihre Wiederholung der Themen des Kapitels verwenden.

# Kapitel 6

# Reihen, Funktionenfolgen und Potenzreihen

Wir werden in diesem Kapitel sogenannte Reihen also „Summen von allen Gliedern einer Folge“ betrachten, was uns auch zu den Definitionen vieler weiterer Ihnen bekannten Funktionen führen wird.

## 6.1 Reihen

**Definition 6.1** (Reihen)**.** Sei $(a_k)_k$ eine Folge reeller oder komplexer Zahlen. Wir wollen die **(unendliche) Reihe** $\sum_{k=1}^{\infty} a_k$ betrachten, wobei $a_k$ für $k \in \mathbb{N}$ das **$k$-te Glied** oder der **$k$-te Summand** der Reihe genannt wird. Für $n \in \mathbb{N}$ ist die $n$-te **Partialsumme** der Reihe $\sum_{k=1}^{\infty} a_k$ durch $s_n = \sum_{k=1}^{n} a_k$ gegeben. Wir nennen die Reihe $\sum_{k=1}^{\infty} a_k$ **konvergent**, falls der Grenzwert

$$\sum_{k=1}^{\infty} a_k = \lim_{n\to\infty} \sum_{k=1}^{n} a_k = \lim_{n\to\infty} s_n$$

in $\mathbb{C}$ existiert, wobei wir diesen dann als **Wert der Reihe** bezeichnen. Ansonsten nennen wir die Reihe **divergent**.

Eine kleine Warnung: Mit „Sei $\sum_{k=1}^{\infty} a_k$ eine Reihe ...“ meinen wir trotz der Notation nicht wirklich, dass $\sum_{k=1}^{\infty} a_k$ effektiv eine Zahl darstellt. Vor allem bevor wir wissen, ob die Reihe konvergent ist, ist $\sum_{k=1}^{\infty} a_k$ vielmehr als formales Objekt zu verstehen (gewissermassen als die Folge der Partialsummen), dessen Konvergenzeigenschaften wir untersuchen wollen.[1] Der erste Summand der Reihe muss nicht immer dem Index $k = 1$ zugeordnet sein und obige Definition ist in solchen Fällen entsprechend anzupassen.

Eine einfache aber auch sehr wichtige Eigenschaft konvergenter Reihen ist in folgender Proposition enthalten.

[1] Manche Autoren verwenden in diesem Zusammenhang auch $\sum_n a_n$ für die Reihe als formales Objekt, welches mit der Folge der Partialsummen identifiziert werden kann, und $\sum_{n=1}^{\infty} a_n$ für den Wert der Reihe.

**Proposition 6.2** (Nullfolgen)**.** *Falls die Reihe* $\sum_{k=1}^{\infty} a_k$ *konvergiert, dann ist die Folge* $(a_n)_n$ *eine* ***Nullfolge****, das heisst* $\lim_{n\to\infty} a_n = 0$.

*Beweis.* Nach Annahme haben die Partialsummen $s_n = \sum_{k=1}^{n} a_k$ für $n \in \mathbb{N}$ einen Grenzwert $\lim_{n\to\infty} s_n = S = \sum_{k=1}^{\infty} a_k$ und damit gilt ebenso

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} (s_n - s_{n-1}) = S - S = 0.$$

□

**Beispiel 6.3** (Geometrische Reihe)**.** *Die* ***geometrische Reihe*** $\sum_{n=0}^{\infty} q^n$ *zu* $q \in \mathbb{R}$ *(oder* $\mathbb{C}$*) konvergiert genau dann, wenn* $|q| < 1$ *ist. In diesem Fall ist*

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}.$$

*In der Tat impliziert Konvergenz der Reihe mittels Proposition 6.2, dass* $|q| < 1$*. Umgekehrt gilt für* $|q| < 1$ *auf Grund der geometrischen Summenformel in Proposition 3.8 und der Konvergenz der geometrischen Folge in Beispiel 5.13, dass*

$$\sum_{k=0}^{n} q^k = \frac{1-q^{n+1}}{1-q} \to \frac{1}{1-q}$$

*für* $n \to \infty$.

**Beispiel 6.4** (Harmonische Reihe)**.** *Die Umkehrung von Proposition 6.2 gilt nicht. Beispielsweise ist die* ***harmonische Reihe*** $\sum_{k=1}^{\infty} \frac{1}{k}$ *divergent.*

*Wir beweisen die Divergenz mit einer konkreten Abschätzung. Sei* $n = 2^\ell$*, dann erfüllt die Partialsumme der harmonischen Reihe für* $n$ *die Abschätzung*

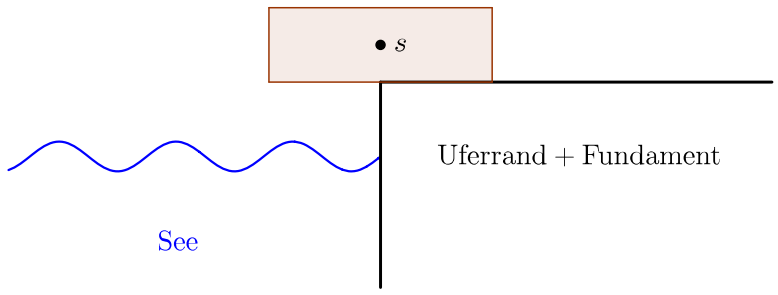
$$\begin{aligned}\sum_{k=1}^{2^\ell} \frac{1}{k} &= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \frac{1}{9} + \cdots + \frac{1}{2^{\ell-1}+1} + \cdots + \frac{1}{2^\ell} \\ &\geq 1 + \frac{1}{2} + \underbrace{\frac{1}{4} + \frac{1}{4}}_{=\frac{1}{2}} + \underbrace{\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}}_{=\frac{1}{2}} + \frac{1}{16} + \cdots + \underbrace{\frac{1}{2^\ell} + \cdots + \frac{1}{2^\ell}}_{=\frac{1}{2}} = 1 + \frac{\ell}{2}.\end{aligned}$$

*Da* $\ell \in \mathbb{N}$ *beliebig war, erkennen wir, dass die Partialsummen nicht beschränkt sind, und daher ist die harmonische Reihe divergent.*

Wir präsentieren noch eine kleine Anwendung der Divergenz der harmonischen Reihe aus dem Alltag.

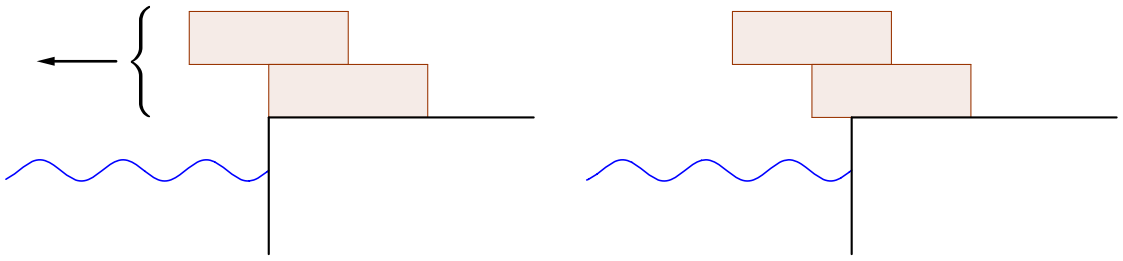
**Beispiel 6.5** (Harmonischer Springturm)**.** *Wir wollen am Rande des Zürichsees einen Springturm bauen, der aus einzelnen quaderförmigen Bausteinen (von gleicher Form und gleichem Material) besteht und möglichst weit in den See hineinragen soll. Wie weit können wir kommen, ohne die Bausteine aneinander oder an das ebene, äusserst stabile, am Uferrand liegende*

*Fundament zu befestigen? Wir wollen $n$ Bausteine von 2 Metern Länge verwenden und rechnen von oben weg jeweils aus, wie weit die Bausteine zueinander verschoben sein dürfen, ohne dass der Turm einstürzt. Für $n = 1$ sehen wir in Figur 6.1, dass der Baustein $1m$ in den See ragen kann.*




Figur 6.1: Der Schwerpunkt $s$ des Bausteins muss oberhalb des Ufers liegen, denn sonst kippt der Stein in den See. Wir wollen im Folgenden immer den Grenzfall, wo der Schwerpunkt genau über dem Uferrand liegt, ebenso als stabil erklären.

*Wir schieben jetzt einen Baustein von unten ein und wollen beide Bausteine soweit wie möglich in den See schieben.*



Figur 6.2: Bei beiden Steinen gibt es jetzt auch zwei Punkte, bei denen der Turm kippen und zumindest teilweise in den See stürzen könnte. Der obere Kipppunkt ist kein Problem, da wir beide Steine gemeinsam verschoben haben und somit den Schwerpunkt des oberen Steins genau am Rand des unteren liegt.

*Um zu bestimmen, wie weit man beide Steine in Richtung See schieben darf, berechnen wir den Schwerpunkt der beiden Steine gemeinsam. Hierfür verwenden wir ein geeignetes Koordinatensystem; nämlich messen wir nach rechts vom Uferrand (also der linken Kante des unteren Steins) aus – siehe dazu das linke Bild in Figur 6.2. Der Schwerpunkt des oberen Steins hat in diesem Koordinatensystem die Koordinate 0, der untere die Koordinate 1 und damit beide zusammen die Koordinate $\frac{1}{2} = \frac{0+1}{2}$. Also können wir beide $\frac{1}{2}m$ in Richtung See verschieben und kommen somit total $1m + \frac{1}{2}m = \frac{3}{2}m$ in den See hinein.*

*Wir heben jetzt diese beiden an und fügen einen weiteren Stein so hinzu, dass die linke Kante genau unter dem Schwerpunkt der ersten zwei und damit am Uferrand zu liegen kommt. Wir müssen also wieder den gemeinsamen Schwerpunkt dieser drei Steine bestimmen. Die oberen beiden haben den gemeinsamen Schwerpunkt 0, der untere hat die Koordinate 1 und somit haben alle drei zusammen den Schwerpunkt $\frac{1}{3} = \frac{2\cdot 0+1\cdot 1}{3}$. Wir verschieben also alle drei um $\frac{1}{3}m$ in Richtung See, was eine totale Verschiebung von $1m + \frac{1}{2}m + \frac{1}{3}m$ ergibt, und wiederholen den Vorgang so oft wie wir wollen. Da aber die Partialsummen der harmonischen Reihe*

*unbeschränkt sind, können wir damit beliebig weit in den See hineinbauen. Eine interaktive Darstellung dieses Vorgehens findet man unter diesem Link und ein Video unter diesem.*

*Für die Praxis ist diese Methode kaum zu empfehlen, zum einen haben wir das Gewicht des Turmspringers ignoriert, und weiters haben wir nicht beschrieben, wie hoch der Turm wirklich wird, wenn wir auch nur* $10m$ *in den See hineinreichen wollen (da in diesem Fall* $n = 12367$ *Bausteine notwendig sind).*

Die folgenden drei Lemmata sind einfache Konsequenzen der Definition der Konvergenz von Reihen.

**Lemma 6.6** (Linearität)**.** *Seien* $\sum_{k=1}^{\infty} a_k$, $\sum_{k=1}^{\infty} b_k$ *konvergente Reihen und* $\alpha \in \mathbb{C}$. *Dann sind die Reihen* $\sum_{k=1}^{\infty}(a_k + b_k)$, $\sum_{k=1}^{\infty}(\alpha a_k)$ *konvergent und es gilt*

$$\sum_{k=1}^{\infty}(a_k + b_k) = \sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} b_k, \quad \sum_{k=1}^{\infty}(\alpha a_k) = \alpha \sum_{k=1}^{\infty} a_k.$$

*Also bilden konvergente Reihen einen Vektorraum über* $\mathbb{C}$ *und der Wert der Reihe stellt eine lineare Abbildung auf diesem Vektorraum nach* $\mathbb{C}$ *dar.*

**Übung 6.7.** *Beweisen Sie Lemma 6.6.*

**Lemma 6.8** (Indexverschiebung für Reihen)**.** *Sei* $\sum_{k=1}^{\infty} a_k$ *eine Reihe. Für jedes* $N \in \mathbb{N}$ *ist die Reihe* $\sum_{k=N}^{\infty} a_k = \sum_{\ell=1}^{\infty} a_{\ell+N-1}$ *genau dann konvergent, wenn die Reihe* $\sum_{k=1}^{\infty} a_k$ *konvergent ist. In diesem Fall gilt*

$$\sum_{k=1}^{\infty} a_k = \sum_{k=1}^{N-1} a_k + \sum_{k=N}^{\infty} a_k.$$

Insbesondere zeigt Lemma 6.8, dass das Konvergenzverhalten einer Reihe sich nicht ändert, wenn endlich viele Glieder der Reihe weggelassen, hinzugefügt oder geändert werden. Wir werden diese zentrale Eigenschaft oft und deswegen mitunter auch implizit verwenden.

*Beweis.* Für $n \geq N$ gilt

$$\sum_{k=1}^{n} a_k = \sum_{k=1}^{N-1} a_k + \sum_{k=N}^{n} a_k.$$

Insbesondere konvergieren die Partialsummen von $\sum_{k=N}^{\infty} a_k$ genau dann, wenn die Partialsummen von $\sum_{k=1}^{\infty} a_k$ konvergieren und das Lemma folgt. □

**Lemma 6.9** (Zusammenfassen von *benachbarten* Gliedern)**.** *Sei* $\sum_{n=1}^{\infty} a_n$ *eine konvergente Reihe und* $(n_k)_k$ *eine streng monoton wachsende Folge natürlicher Zahlen. Definiere* $A_1 = a_1 + \cdots + a_{n_1}$ *und* $A_k = a_{n_{k-1}+1} + \cdots + a_{n_k}$ *für* $k \geq 2$. *Dann gilt*

$$\sum_{k=1}^{\infty} A_k = \sum_{n=1}^{\infty} a_n.$$

*Beweis.* Die $K$-te Partialsumme von $\sum_{k=1}^{\infty} A_k$ ist

$$\sum_{k=1}^{K} A_k = (a_1 + \ldots + a_{n_1}) + (a_{n_1+1} + \ldots + a_{n_2}) + \ldots + (a_{n_{K-1}+1} + \ldots + a_{n_K}) = \sum_{n=1}^{n_K} a_n.$$

Somit bilden die Partialsummen von $\sum_{k=1}^{\infty} A_k$ eine Teilfolge der konvergenten Folge der Partialsummen von $\sum_{n=1}^{\infty} a_n$. □

**Beispiel 6.10.** *Die Umkehrung von Lemma 6.9 gilt im Allgemeinen nicht. Beispielsweise ist die Reihe $\sum_{n=1}^{\infty}(-1)^n$ nach Proposition 6.2 divergent, aber die Reihe $\sum_{k=1}^{\infty}((-1)^{2k-1} + (-1)^{2k})$, die aus zusammengefügten Gliedern von $\sum_{n=1}^{\infty}(-1)^n$ besteht, ist konvergent, da jedes Glied Null ist.*

### 6.1.1 Reihen mit nicht-negativen Gliedern

Für Reihen mit nicht-negativen Gliedern gilt folgende fundamentale Eigenschaft.

**Proposition 6.11** (Monotone Partialsummen)**.** *Für eine Reihe $\sum_{k=1}^{\infty} a_k$ mit nicht-negativen Gliedern $a_k \geq 0$ für alle $k \in \mathbb{N}$ bilden die Partialsummen $s_n = \sum_{k=1}^{n} a_k$ eine monoton wachsende Folge. Falls diese Folge der Partialsummen beschränkt ist, dann konvergiert die Reihe $\sum_{k=1}^{\infty} a_k$. Ansonsten gilt*

$$\sum_{k=1}^{\infty} a_k = \lim_{n\to\infty} s_n = \infty.$$

Insbesondere können wir für die harmonische Reihe in Beispiel 6.4

$$\sum_{n=1}^{\infty} \frac{1}{n} = \infty$$

schreiben.

*Beweis.* Aus $a_{n+1} \geq 0$ folgt $s_{n+1} = s_n + a_{n+1} \geq s_n$ für alle $n \in \mathbb{N}$. Falls die Partialsummen $\{s_n \mid n \in \mathbb{N}\}$ zusätzlich noch beschränkt sind, dann sind diese (und damit auch die Reihe) konvergent nach Satz 5.34. □

**Korollar 6.12** (Vergleichssatz)**.** *Seien $\sum_{k=1}^{\infty} a_k$, $\sum_{k=1}^{\infty} b_k$ zwei Reihen mit der Eigenschaft $0 \leq a_k \leq b_k$ für alle $k \in \mathbb{N}$. Dann gilt $\sum_{k=1}^{\infty} a_k \leq \sum_{k=1}^{\infty} b_k$ und insbesondere gelten die Implikationen*

$$\sum_{k=1}^{\infty} b_k \text{ konvergent} \implies \sum_{k=1}^{\infty} a_k \text{ konvergent}$$
$$\sum_{k=1}^{\infty} a_k \text{ divergent} \implies \sum_{k=1}^{\infty} b_k \text{ divergent}.$$

*Diese beiden Implikationen treffen auch dann zu, wenn $0 \leq a_n \leq b_n$ nur für alle hinreichend grossen $n \in \mathbb{N}$ gilt.*

Man nennt unter den Annahmen des Korollars die Reihe $\sum_{k=1}^{\infty} b_k$ eine Majorante der Reihe $\sum_{k=1}^{\infty} a_k$, und letztere auch eine Minorante der Reihe $\sum_{k=1}^{\infty} b_k$. Daher spricht man auch von dem **Majoranten-** und dem **Minorantenkriterium**.

*Beweis.* Aus $a_k \leq b_k$ für alle $k \in \mathbb{N}$ folgt $\sum_{k=1}^{n} a_k \leq \sum_{k=1}^{n} b_k$ für alle $n \in \mathbb{N}$. Somit gilt nach Monotonie der Folge der Partialsummen

$$\sum_{k=1}^{\infty} a_k = \sup\left\{\sum_{k=1}^{n} a_k \mid n \in \mathbb{N}\right\} \leq \sup\left\{\sum_{k=1}^{n} b_k \mid n \in \mathbb{N}\right\} = \sum_{k=1}^{\infty} b_k.$$

Die letzte Aussage der Proposition ist nun eine Konsequenz von Lemma 6.8 (wieso?). □

**Übung 6.13.** *Zeigen Sie, dass die Annahme in Korollar 6.12, dass die Reihen $\sum_{k=1}^{\infty} a_k$, $\sum_{k=1}^{\infty} b_k$ nicht-negative Glieder haben, notwendig ist.*

**Beispiel 6.14** (Reihe der Kehrwerte der Quadratzahlen)**.** *Die Reihe $\sum_{k=1}^{\infty} \frac{1}{k^2}$ ist konvergent. Tatsächlich gilt $a_k = \frac{1}{k^2} \leq \frac{1}{k(k-1)} = b_k$ für $k \geq 2$ und die Reihe $\sum_{k=1}^{\infty} b_k$ ist konvergent, da deren $n$-te Partialsumme unter Auflösen einer Teleskopsumme (siehe Abschnitt 3.1.1) durch*

$$\sum_{k=2}^{n} \frac{1}{k(k-1)} = \sum_{k=2}^{n} \left(\frac{1}{k-1} - \frac{1}{k}\right) = 1 - \frac{1}{n}$$

*gegeben ist.*

**Beispiel 6.15** (Reihenkonvergenz und Folgenasymptotik)**.** *Wir betrachten die Reihe*

$$\sum_{n=1}^{\infty} \frac{2n-10}{n^3 - 10n + 100}$$

*und wollen zeigen, dass diese konvergiert. Für dies bemerken wir, dass $a_n = \frac{2n-1}{n^3-10n+100}$ im Wesentlichen sich wie $\frac{1}{n^2}$ verhalten sollte. Genauer formuliert gilt $a_n = O(\frac{1}{n^2})$ für $n \to \infty$ (siehe Abschnitt 5.6) da*

$$\lim_{n\to\infty} \frac{a_n}{1/n^2} = \lim_{n\to\infty} \frac{n^2(2n-10)}{n^3 - 10n + 100} = 2.$$

*Daher gibt es ein $M > 0$ mit $a_n \leq \frac{M}{n^2}$ für alle $n \in \mathbb{N}$ (wieso?) und ein $N \in \mathbb{N}$ mit $0 \leq a_n$ für alle $n \geq N$. Verwenden wir nun Korollar 6.12 und Beispiel 6.14 ergibt sich die Konvergenz von $\sum_{n=1}^{\infty} \frac{2n-10}{n^3-10n+100}$.*

**Proposition 6.16** (Verdichtung)**.** *Eine Reihe $\sum_{k=1}^{\infty} a_k$ mit nicht-negativen, monoton abnehmenden Gliedern $a_1 \geq a_2 \geq \ldots \geq 0$ ist genau dann konvergent, wenn $\sum_{k=1}^{\infty} 2^k a_{2^k}$ konvergent ist.*

*Beweis.* Es gelten auf Grund der angenommenen Monotonie von $(a_n)_n$ die Ungleichungen

$$\begin{aligned} a_2 &\leq a_2 \leq a_1 \\ 2a_4 &\leq a_3 + a_4 \leq 2a_2 \\ 2^2 a_8 &\leq a_5 + a_6 + a_7 + a_8 \leq 2^2 a_4 \end{aligned}$$

und allgemeiner

$$2^n a_{2^{n+1}} \leq a_{2^n+1} + \ldots + a_{2^{n+1}} \leq 2^n a_{2^n}$$

für $n \in \mathbb{N}$. Für die Summen ergibt sich daher die Ungleichung

$$\sum_{k=1}^{n+1} 2^{k-1} a_{2^k} \leq \sum_{\ell=2}^{2^{n+1}} a_\ell \leq \sum_{k=0}^{n} 2^k a_{2^k}$$

und wir erhalten die Proposition durch den Grenzübergang $n \to \infty$ und Korollar 6.12. □

**Beispiel 6.17** ($p$-Test)**.** *Die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^p}$ für $p \in \mathbb{R}$ konvergiert genau dann, wenn $p > 1$. Für $p \leq 0$ ist $\frac{1}{n^p} \geq 1$ für alle $n \in \mathbb{N}$ und die Reihe nach Proposition 6.2 somit divergent. Für $p \leq 1$ gilt $\frac{1}{n} \leq \frac{1}{n^p}$ für alle $n \in \mathbb{N}$ und die Reihe divergiert nach Korollar 6.12, da die harmonische Reihe divergiert. Wir wenden nun Proposition 6.16 an. Für $p \geq 0$ ist $\left(\frac{1}{n^p}\right)_n$ eine monoton abnehmende Folge und wir erhalten aus Proposition 6.16, dass $\sum_{n=1}^{\infty} \frac{1}{n^p}$ genau dann konvergiert, wenn*

$$\sum_{k=1}^{\infty} 2^k \frac{1}{(2^k)^p} = \sum_{k=1}^{\infty} (2^{1-p})^k$$

*konvergiert. Diese geometrische Reihe konvergiert aber nach Beispiel 6.3 genau dann, wenn $p > 1$ ist.*

**Übung 6.18.**

*(i) Zeigen Sie für $p \in \mathbb{R}$, dass die Reihe $\sum_{n=2}^{\infty} \frac{1}{n \log(n)^p}$ genau dann konvergiert, wenn $p > 1$ ist.*

*(ii) Ist die Reihe $\sum_{n=3}^{\infty} \frac{1}{n \log(n) \log(\log(n))}$ konvergent oder divergent?*

**Übung 6.19** (q-äre Darstellungen)**.** *Sei $q \in \mathbb{N}$, $q > 1$. In Übung 3.7 haben wir gezeigt, dass jede ganze Zahl eine Ziffernentwicklung zur Basis $q$ besitzt. In dieser Übung wollen wir die analoge Aussage für reelle Zahlen formulieren und beweisen. Sei $x \in \mathbb{R}$. Wegen Übung 3.7 wollen wir sogar annehmen, dass $x \in [0, 1)$. Zeigen Sie, dass eine Folge von Ziffern $(\alpha_k)_k$ mit $\alpha_k \in \{1, \ldots, q-1\}$ für alle $k \in \mathbb{N}$ existiert, so dass die Reihe $\sum_{k=1}^{\infty} \alpha_k q^k$ konvergiert und $x$ darstellt im Sinne von*

$$\sum_{k=1}^{\infty} \alpha_k q^{-k} = x.$$

*Sind die Ziffern $(\alpha_k)_k$ zu $x \in \mathbb{R}$ wie oben eindeutig bestimmt?*

### 6.1.2 Bedingte Konvergenz

Wir sagen, dass eine Reihe $\sum_{n=1}^{\infty} a_n$ mit komplexen Summanden **absolut konvergiert**, falls die Reihe $\sum_{n=1}^{\infty} |a_n|$ konvergiert. Die Reihe $\sum_{n=1}^{\infty} a_n$ ist **bedingt konvergent**, falls sie konvergiert, aber nicht absolut konvergiert.

Wie wir im nächsten Abschnitt besprechen werden, haben absolut konvergente Reihen, im Gegensatz zu bedingt konvergenten Folgen, sehr robuste Eigenschaften. Inwiefern letztere über gewisse (widersprüchlich erscheinende) Eigenheiten verfügen, wollen wir in diesem Teilabschnitt erklären.

**Beispiel 6.20** (Alternierende harmonische Reihe)**.** *Wir wollen zuerst zeigen, dass die* ***alternierende harmonische Reihe***

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} + \dots$$

*bedingt konvergiert. Nach Beispiel 6.4 divergiert die harmonische Reihe*

$$\sum_{n=1}^{\infty} \frac{1}{n} = \sum_{n=1}^{\infty} \frac{|(-1)^{n+1}|}{n}$$

*(nach unendlich). Wir müssen also nur noch Konvergenz der alternierenden Reihe beweisen. Wir betrachten zuerst zu $n \in \mathbb{N}$ die $2n$-te Partialsumme*

$$\begin{aligned} s_{2n} &= \left(1 - \frac{1}{2}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \dots + \left(\frac{1}{2n-1} - \frac{1}{2n}\right) = \sum_{k=1}^{n} \left(\frac{1}{2k-1} - \frac{1}{2k}\right) \\ &= \sum_{k=1}^{n} \frac{1}{2k(2k-1)} \leq \sum_{k=1}^{n} \frac{1}{k^2} \leq \sum_{k=1}^{\infty} \frac{1}{k^2} \end{aligned}$$

*Die Folge $(s_{2n})_n$ ist somit monoton wachsend und beschränkt (wegen Beispiel 6.14) und konvergiert damit. Wegen $s_{2n-1} = s_{2n} - \frac{1}{2n}$ konvergiert aber ebenso die Folge $(s_{2n-1})_n$ und gegen den gleichen Limes.*[2] *Also konvergiert die Reihe $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$ (wieso?). Des Weiteren folgt $\sum_{n=1} \frac{(-1)^{n+1}}{n} \geq \frac{1}{2}$, da nach obigem Argument die Partialsumme $s_{2n}$ als Summe von positiven Summanden geschrieben werden kann, wovon der erste Term gleich $\frac{1}{2}$ ist.*

Folgender Satz mag zuerst überraschend sein und zeigt, dass man mit bedingter Konvergenz vorsichtig umgehen muss, da diese sehr zerbrechliche Eigenschaften besitzt.

**Satz 6.21** (Riemannscher Umordnungssatz)**.** *Sei $\sum_{n=1}^{\infty} a_n$ eine bedingt konvergente Reihe mit reellen Gliedern. Dann gibt es zu jedem $A \in \mathbb{R}$ eine bijektive Funktion (eine Umordnung) $\varphi : \mathbb{N} \to \mathbb{N}$, so dass die Reihe $\sum_{n=1}^{\infty} a_{\varphi(n)}$ bedingt konvergiert und $\sum_{n=1}^{\infty} a_{\varphi(n)} = A$ ist. Weiters gibt es eine Umordnung der Reihe, die divergiert.*

[2] Wir werden den Wert dieser Reihe erst später berechnen können.

**Beispiel 6.22** (Umordnen der alternierenden harmonischen Reihe)**.** *Wir ordnen die alternierende harmonische Reihe um und erhalten*

$$\left(1-\frac{1}{2}\right)-\frac{1}{4}+\left(\frac{1}{3}-\frac{1}{6}\right)-\frac{1}{8}+\left(\frac{1}{5}-\frac{1}{10}\right)-\frac{1}{12}+\dots$$

*Wir bemerken, dass die so erhaltene Umordnung der alternierenden harmonischen Reihe konvergiert und den halben Wert der alternierenden harmonischen Reihe annimmt. Hierbei sind die Klammern als Hilfestellung gedacht, denn rechnet man die Klammern aus, so erhält man die Reihe in der jeder Summand genau die Hälfte der Summanden der alternierenden harmonischen Reihe ausmacht. Lässt man hingegen die Klammern weg, so erhält man die umgeordnete Reihe der alternierenden harmonischen Reihe. Des Weiteren weiss man für diese Reihe, dass die Teilfolge $(s_{3n})_n$ der Partialsummen konvergiert. Da aber $s_{3n+1}-s_{3n}$ und $s_{3n+2}-s_{3n}$ Nullfolgen sind, können wir daraus schliessen, dass die umgeordnete Reihe konvergiert und den halben Wert der ursprünglichen Reihe hat.*

Da dieser Satz eher negativer Natur ist, begnügen wir uns mit einer Beweisskizze und verweisen auf [Wal04, Satz 5.17]. Sei also $\sum_{n=1}^{\infty} a_n$ eine bedingt konvergente Reihe (wobei es helfen könnte, an $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$ zu denken). Dann gilt $a_n \to 0$ für $n \to \infty$ und $\sum_{n=1}^{\infty} |a_n| = \infty$ nach Annahme. Wir teilen die natürlichen Zahlen $\mathbb{N}$ in die zwei Mengen

$$P = \{n \in \mathbb{N} \mid a_n \geq 0\}\,, \quad N = \{n \in \mathbb{N} \mid a_n < 0\}$$

auf. Dann müssen $P$ und $N$ beide unendliche Kardinalität haben, denn wenn zum Beispiel $N$ endlich wäre, dann würden sich $\sum_{n=1}^{\infty} a_n$ und $\sum_{n=1}^{\infty} |a_n|$ nur um endlich viele Terme unterscheiden. (Für $a_n = \frac{(-1)^{n+1}}{n}$ wäre $P = 2\mathbb{N} - 1$ und $N = 2\mathbb{N}$.) Wir zählen die Elemente in $P$ und $Q$ so auf, dass

$$P = \{p_1 < p_2 < p_3 < \dots\}$$

und

$$N = \{n_1 < n_2 < n_3 < \dots\}\,.$$

Weiters ist $\sum_{k=1}^{\infty} a_{p_k} = +\infty$ und $\sum_{k=1}^{\infty}(-a_{n_k}) = +\infty$. Denn falls beide Summen endlich wären, dann wäre $\sum_{n=1}^{\infty} |a_n| < \infty$. Wäre zum Beispiel $\sum_{k=1}^{\infty} a_{p_k} = \infty$, aber $\sum_{k=1}^{\infty}(-a_{n_k}) < \infty$, dann wäre auch $\sum_{n=1}^{\infty} a_n = \infty$. (Für $a_n = \frac{(-1)^{n+1}}{n}$ sind diese beiden Reihen $\sum_{k=1}^{\infty} \frac{1}{2k-1} = \infty$ und $\sum_{k=1}^{\infty} \frac{1}{2k} = \infty$.)

Für ein gegebenes $A \in \mathbb{R}$ konstruieren wir die bijektive Abbildung $\varphi : \mathbb{N} \to \mathbb{N}$ gemeinsam mit der Reihe $\sum_{k=1}^{\infty} a_{\varphi(k)}$ auf folgende Weise.

Wir beginnen die Reihe mit den ersten nicht-negativen Gliedern

$$a_{p_1} + \dots + a_{p_{k_1}}$$

und wählen $k_1 \geq 1$ minimal, so dass die obige Summe grösser als $A$ ist (was wegen$\sum_{k=1}^{\infty} a_{p_k} =$

$+\infty$ möglich ist). Anschliessend addieren wir die ersten negativen Glieder und wählen $\ell_1 \geq 1$, so dass die Summe

$$a_{p_1} + \ldots + a_{p_{k_1}} + a_{n_1} + \ldots + a_{n_{\ell_1}}$$

kleiner als $A$ ist. Als nächstes addieren wir, beginnend mit $p_{k_1+1}$, nicht-negative Terme, um die Summe grösser als $A$ werden zu lassen. Wir führen dies fort und weil $\sum_{k=1}^{\infty} a_{p_k} = +\infty$ und $\sum_{k=1}^{\infty}(-a_{n_k}) = +\infty$ können wir immer wieder nach endlich vielen Summanden von der einen Seite von $A$ zu der anderen Seite von $A$ wechseln. Da noch dazu $a_n \to 0$ für $n \to \infty$ gilt, werden die einzelnen Schritte über $A$ hinweg immer kleiner und wir können auf diese Art $A$ als Grenzwert der umgeordneten Folge realisieren.

**Übung 6.23.** *Füllen Sie die unterlassenen Schritte am Ende der obigen Beweisskizze ein, um einen vollständigen Beweis von Satz 6.21 zu erhalten.*

Falls Sie jetzt denken, dass der Riemann'sche Umordnungssatz (Satz 6.21) einen Widerspruch in der Mathematik darstellt, dann täuschen Sie sich. Denn wir haben eine klare Definition für den Wert einer Reihe in Definition 6.1 ausformuliert. Wir haben auch besprochen, wie sich das Zusammenfassen von benachbarten (!) Gliedern einer Reihe für diese Definition auswirkt (siehe Lemma 6.9) – hierbei geht man von der ursprünglichen Folge der Partialsummen zu einer Teilfolge der Partialsummen über und weder das Konvergenzverhalten noch der Wert ändern sich hier. Was hingegen passiert mit der Folge der Partialsummen, wenn sie die Summanden mittels einer beliebigen Bijektion $\varphi : \mathbb{N} \to \mathbb{N}$ permutieren? Dies ist unmöglich zu beantworten, denn es gibt im Allgemeinen überhaupt keinen Zusammenhang zwischen der Folge $\sum_{k=1}^{n} a_k$ der Partialsummen der ursprünglichen Reihe $\sum_{k=1}^{\infty} a_k$ und der Folge $\sum_{k=1}^{n} a_{\varphi(k)}$ der Partialsummen der permutierten Reihe $\sum_{k=1}^{\infty} a_{\varphi(k)}$ (da der Riemann'sche Umordnungssatz ja zeigt, dass wir das Konvergenzverhalten auf diese Weise komplett ändern können).

Wir sind immer mittels formaler Definitionen und Beweise vorgegangen und versuchen natürlich auch ein intuitives Verständnis für die so entstehenden Theorien zu entwickeln, doch wenn es (wie zum Beispiel hier) zu einer Diskrepanz zwischen unseren Sätzen und unserer Anschauung kommt, dann müssen wir daran arbeiten unsere Anschauung den gegebenen Fakten (also Definitionen und Sätzen) anzupassen. Für den Fall einer bedingt konvergenten Reihe müssen wir uns daran erinnern, dass der Wert der Reihe nicht als die Summe aller Summanden definiert wurde – wie sollen wir denn unendlich viele Additionen gleichzeitig durchführen? Stattdessen wurde der Wert der Reihe als der Grenzwert der Partialsummen definiert und für diese Definition müssen wir die Reihenfolge der Summanden kennen. Ändert sich die Reihenfolge, dann könnte sich dies auf die Definition auswirken (was bei bedingt konvergenten Reihen auf Grund von Satz 6.21 in der Tat der Fall ist).

Bedingt konvergente Reihen sind für uns am Rande interessant, da wir zum Beispiel zeigen werden, dass die alternierende harmonische Reihe den Wert $\log 2$ hat. Doch die weitaus meisten Reihen, die wir betrachten werden, werden absolut konvergent sein und robusteres Verhalten zeigen.

**Applet 6.24** (Drei Reihen). *Wir stellen in diesem Applet die harmonische Reihe, die alternierende harmonische Reihe und die Reihe mit den reziproken Quadratzahlen gegenüber und sehen drei unterschiedliche Verhaltensweisen. Erklären Sie diese Unterschiede. Wie nennen wir diese Verhaltensweisen?*

### 6.1.3 Alternierende Reihen

Das in Beispiel 6.20 aufgetretene Phänomen tritt auch in folgendem Resultat auf (welches wegen der sehr einfachen Abschätzung auch für absolut konvergente Reihen von Interesse sein wird). Für eine Folge $(a_n)_n$ positiver Zahlen bezeichnen wir die Reihe $\sum_{n=1}^{\infty}(-1)^{n+1}a_n$ als eine **alternierende** Reihe.

**Proposition 6.25** (Leibniz-Kriterium). *Gegeben sei eine monoton fallende Folge $(a_n)_n$ positiver Zahlen, die gegen Null konvergiert. Dann konvergiert die zugehörige alternierende Reihe $\sum_{k=1}^{\infty}(-1)^{k+1}a_k$ und es gilt, dass*

$$\left|\sum_{k=1}^{\ell}(-1)^{k+1}a_k - \sum_{k=1}^{\infty}(-1)^{k+1}a_k\right| \leq a_{\ell+1}. \tag{6.1}$$

*für alle $\ell \in \mathbb{N}$. Weiters ist*

$$\sum_{k=1}^{2n}(-1)^{k+1}a_k \leq \sum_{k=1}^{\infty}(-1)^{k+1}a_k \leq \sum_{k=1}^{2n-1}(-1)^{k+1}a_k$$

*für alle $n \in \mathbb{N}$.*

Abschätzungen des Typs (6.1) werden meist auch als Fehlerabschätzungen oder Fehlerschranken bezeichnet. Intuitiv beschreibt man damit, wie gross der Fehler höchstens ist, wenn man anstatt des Wertes der Reihe nur die Summe bis zu einem gewissen Glied (als Approximation gewissermassen) betrachtet.

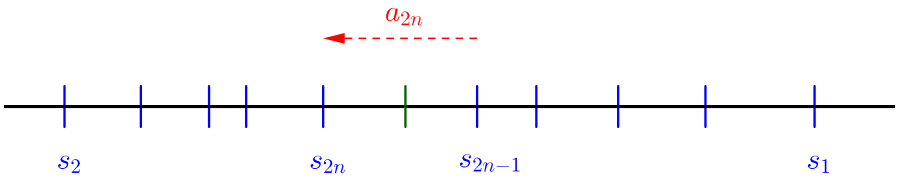
*Beweis.* Ähnlich wie in Beispiel 6.20 spielen wir die Folge der Partialsummen zu geraden und ungeraden Indices gegeneinander aus. Für $n \in \mathbb{N}$ sei $s_n = \sum_{k=1}^{n}(-1)^{k+1}a_k$. Es gilt

$$s_{2n+1} = s_{2n-1} - a_{2n} + a_{2n+1} \leq s_{2n-1}, \quad s_{2n+2} = s_{2n} + a_{2n+1} - a_{2n+2} \geq s_{2n}$$

für alle $n \in \mathbb{N}$. Insbesondere ist die Folge $(s_{2n-1})_n$ monoton fallend und die Folge $(s_{2n})_n$ ist monoton wachsend. Wegen $s_{2n} = s_{2n-1} - a_{2n}$ und der Monotonieeigenschaften gilt

$$s_2 \leq s_{2n} \leq s_{2n-1} \leq s_1$$

für alle $n \in \mathbb{N}$. Somit ist $(s_{2n})_n$ von oben beschränkt und damit konvergent. Analog ist auch $(s_{2n-1})_n$ von unten beschränkt und konvergent. Wir fassen die erhaltenen Erkenntnisse in folgendem Bild zusammen.




Da aber $(a_n)_n$ gegen Null konvergiert, haben die Folgen $(s_{2n})_n$ und $(s_{2n-1})_n$ wegen $s_{2n} = s_{2n-1} - a_{2n}$ für alle $n \in \mathbb{N}$ den gleichen Grenzwert. Insbesondere konvergiert die Reihe $\sum_{k=1}^{\infty}(-1)^{k+1}a_k = S \in \mathbb{R}$ (wieso?).

Wir zeigen nun die Fehlerabschätzung und die behauptete Ungleichung. Für $n \in \mathbb{N}$ gilt auf Grund der besprochenen Monotonieeigenschaften, dass

$$s_{2n} \leq \sup\{s_{2m} \mid m \in \mathbb{N}\} = S = \inf\{s_{2m-1} \mid m \in \mathbb{N}\} \leq s_{2n+1}.$$

Für $\ell = 2n$ ist aber $s_{2n+1} - s_{2n} = a_{\ell+1}$ und wir erhalten (6.1). Für $\ell = 2n-1$ ungerade, gilt ebenso $s_{2n} = s_{2n-1} - a_{2n} \leq S \leq s_{2n-1}$ woraus sich (6.1) ergibt. Dies beweist die Fehlerabschätzung sowohl für einen geraden als auch für einen ungeraden Index und damit die Proposition. □

VO: 20.11.17

### 6.1.4 Das Cauchy-Kriterium

Der nächste Satz übernimmt den Grossteil unserer Vorarbeiten über reellwertige und komplexwertige Folgen und gibt uns für den weiteren Aufbau der Theorie ein sehr wichtiges und genaues Kriterium für die Konvergenz von Reihen.

**Satz 6.26** (Cauchy-Kriterium)**.** *Die Reihe $\sum_{k=1}^{\infty} a_k$ konvergiert genau dann, wenn es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass für $n \geq m \geq N$*

$$\left|\sum_{k=m}^{n} a_k\right| < \varepsilon$$

*erfüllt ist.*

*Beweis.* Dies folgt aus dem Cauchy-Kriterium für Folgen (Satz 5.48) angewendet auf die Folge der Partialsummen $s_n = \sum_{k=1}^{n} a_k$, da für $n \geq m$

$$s_n - s_{m-1} = \sum_{k=1}^{n} a_k - \sum_{k=1}^{m-1} a_k = \sum_{k=m}^{n} a_k. \qquad \square$$

**Beispiel 6.27** (Harmonische Reihe)**.** *Um die Divergenz der harmonischen Reihe zu sehen, können wir auch das Cauchy-Kriterium verwenden. Wir setzen dazu $\varepsilon = \frac{1}{2}$. Für ein beliebiges*

*$N \in \mathbb{N}$ gilt dann*

$$\sum_{k=N}^{2N} \frac{1}{k} = \frac{1}{N} + \frac{1}{N+1} + \ldots + \frac{1}{2N} \geq \frac{N+1}{2N} > \frac{1}{2},$$

*was wegen dem Cauchy-Kriterium für Reihen (Satz 6.26) die Divergenz impliziert.*

## 6.2 Absolute Konvergenz

In diesem Abschnitt wollen wir uns vor allem mit absolut konvergenten Reihen auseinandersetzen und einige Konvergenzkriterien beweisen. Auch möchten wir zeigen, dass absolut konvergente Reihen im Gegensatz zu bedingt konvergenten Reihen stabilere Eigenschaften haben.

**Proposition 6.28** (Absolute Konvergenz)**.** *Eine absolut konvergente Reihe $\sum_{n=1}^{\infty} a_n$ ist auch konvergent und es gilt die verallgemeinerte Dreiecksungleichung*

$$\left|\sum_{n=1}^{\infty} a_n\right| \leq \sum_{n=1}^{\infty} |a_n|.$$

*Beweis.* Der erste Teil folgt unmittelbar aus zweimaliger Anwendung des Cauchy-Kriteriums für Reihen (Satz 6.26). Da die Reihe $\sum_{n=1}^{\infty} |a_n|$ konvergiert, gibt es für $\varepsilon > 0$ nach dem Cauchy-Kriterium ein $N \in \mathbb{N}$, so dass für $n \geq m \geq N$ die Abschätzung

$$\sum_{k=m}^{n} |a_k| < \varepsilon$$

gilt. Daraus folgt

$$\left|\sum_{k=m}^{n} a_k\right| \leq \sum_{k=m}^{n} |a_k| < \varepsilon$$

mit der Dreiecksungleichung. Da $\varepsilon > 0$ beliebig war, beweist dies nach dem Cauchy-Kriterium die Konvergenz der Reihe $\sum_{n=1}^{\infty} a_n$. Der zweite Teil folgt nun aus der Ungleichung

$$\left|\sum_{k=1}^{n} a_k\right| \leq \sum_{k=1}^{n} |a_k|$$

für alle $n \in \mathbb{N}$ und dem Grenzübergang für $n \to \infty$. □

### 6.2.1 Hinreichende Kriterien für absolute Konvergenz

Falls sich die Glieder einer Reihe im Absolutbetrag durch die Glieder einer konvergenten Reihe abschätzen lassen, so ist die Reihe konvergent, wie wir in folgendem Korollar des Vergleichssatzes (Korollar 6.12) zeigen.

**Korollar 6.29** (Majorantenkriterium von Weierstrass)**.** *Sei $(a_n)_n$ eine komplexe und $(b_n)_n$ eine reelle Folge mit $|a_n| \leq b_n$ für alle hinreichend grossen $n \in \mathbb{N}$. Falls $\sum_{n=1}^{\infty} b_n$ konvergiert, dann ist $\sum_{n=1}^{\infty} a_n$ absolut konvergent und daher auch konvergent.*

*Beweis.* Nach dem Vergleichssatz (Korollar 6.12) ist $\sum_{n=1}^{\infty} a_n$ absolut konvergent und die Konvergenz folgt nun aus Proposition 6.28. □

Wir möchten nun zwei Korollare des Majorantenkriteriums diskutieren.

**Korollar 6.30** (Cauchy-Wurzelkriterium)**.** *Sei $(a_n)_n$ eine Folge komplexer Zahlen und*

$$\alpha = \limsup_{n\to\infty} \sqrt[n]{|a_n|} \in \mathbb{R} \cup \{\infty\}\,.$$

*Dann gilt*

$$\alpha < 1 \implies \sum_{n=1}^{\infty} a_n \text{ ist absolut konvergent,}$$
$$\alpha > 1 \implies \sum_{n=1}^{\infty} a_n \text{ ist divergent und } (a_n)_n \text{ ist keine Nullfolge.}$$

Sehr schwammig ausgedrückt lässt sich eine Folge $(a_n)_n$ für $\alpha < 1$ wie in Korollar 6.30 bis auf endlich viele Glieder und einen kleinen Fehler $\varepsilon > 0$ von oben durch die geometrische Folge $(\alpha + \varepsilon)^n$ abschätzen. Somit kann man Korollar 6.29 anwenden. Nun aber genauer.

*Beweis.* Angenommen $\alpha < 1$. Dann gibt es ein $n \in \mathbb{N}$ mit

$$\sup_{k\geq n} \sqrt[k]{|a_k|} < q = \frac{1+\alpha}{2} < 1$$

und somit $|a_k| < q^k$ für alle $k \geq n$. Die Reihe $\sum_{k=1}^{\infty} a_k$ konvergiert somit absolut wegen dem Majorantenkriterium (Korollar 6.29) und der geometrischen Reihe in Beispiel 6.3.

Falls $\alpha > 1$ gilt, gibt es nach Satz 5.44 eine Teilfolge $(a_{n_k})_k$ mit $\sqrt[n_k]{|a_{n_k}|} > 1$ für alle $k$. Daraus folgt aber $|a_{n_k}| > 1$. Insbesondere ist $(a_n)_n$ keine Nullfolge und $\sum_{n=1}^{\infty} a_n$ divergiert nach Proposition 6.2. □

**Beispiel 6.31** (Der Fall $\alpha = 1$ in Korollar 6.30)**.** *Sei $(a_n)_n$ eine Folge komplexer Zahlen und $\alpha = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$ wie im Wurzelkriterium (Korollar 6.30). Falls $\alpha = 1$, dann kann anhand des Wurzelkriteriums keine Entscheidung über Konvergenz oder Divergenz der Reihe $\sum_{n=1}^{\infty} a_n$ getroffen werden.*

- *Ist $a_n = \frac{1}{n}$ für alle $n \in \mathbb{N}$, dann gilt $\sqrt[n]{|a_n|} = \frac{1}{\sqrt[n]{n}} \to 1$ für $n \to \infty$ wegen Beispiel 5.33 und wegen Beispiel 6.4 divergiert die Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$ gegen Unendlich.*
- *Ist $a_n = \frac{1}{n^2}$ für alle $n \in \mathbb{N}$, dann gilt $\sqrt[n]{|a_n|} = \frac{1}{\sqrt[n]{n^2}} \to 1$ für $n \to \infty$ und $\sum_{n=1}^{\infty} \frac{1}{n^2}$ konvergiert nach Beispiel 6.14.*

**Korollar 6.32** (D'Alemberts Quotientenkriterium)**.** *Sei $(a_n)_n$ eine Folge komplexer Zahlen mit $a_n \neq 0$ für alle $n \in \mathbb{N}$, so dass*

$$\alpha = \lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|}$$

*existiert. Dann gilt*

$$\alpha < 1 \implies \sum_{n=1}^{\infty} a_n \text{ ist absolut konvergent.}$$
$$\alpha > 1 \implies \sum_{n=1}^{\infty} a_n \text{ ist divergent und } (a_n)_n \text{ konvergiert nicht gegen Null.}$$

**Wichtige Übung 6.33.** *Beweisen Sie Korollar 6.32.*

**Übung 6.34.** *Wieso kann man im Quotientenkriterium (Korollar 6.32) nicht auch den Limes superior anstelle des Limes verwenden?*

### 6.2.2 Umordnen von Reihen

Im Gegensatz zur bedingten Konvergenz in Teilabschnitt 6.1.2 ist absolute Konvergenz sehr robust. Der erste dieser Robustheitssätze ist folgender positive Umordnungssatz.

**Satz 6.35** (Umordnen absolut konvergenter Reihen)**.** *Sei $\sum_{n=1}^{\infty} a_n$ eine absolut konvergente Reihe mit komplexen Gliedern. Sei $\varphi : \mathbb{N} \to \mathbb{N}$ eine Bijektion. Dann ist $\sum_{n=1}^{\infty} a_{\varphi(n)}$ ebenso absolut konvergent und es gilt*

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} a_{\varphi(n)}. \tag{6.2}$$

*Beweis.* Sei $\varphi : \mathbb{N} \to \mathbb{N}$ eine Bijektion und $\varepsilon > 0$. Nach dem Cauchy-Kriterium (Satz 6.26) gibt es ein $N \in \mathbb{N}$, so dass $\sum_{k=m}^{n} |a_k| < \varepsilon$ für alle natürliche Zahlen $n \geq m \geq N$. Daher gilt auch nach der Dreiecksungleichung in Proposition 6.28, dass

$$\left| \sum_{k=m}^{\infty} a_k \right| \leq \sum_{k=m}^{\infty} |a_k| \leq \varepsilon$$

für alle $m \geq N$. (Es hilft vielleicht für das Folgende diese Abschätzung als „Die Summanden $a_1, \ldots, a_N$ der ursprünglichen Reihe sind wichtig, aber die restlichen Summanden sind weniger wichtig.“ zu interpretieren.)

Wir definieren $M = \max\left\{\varphi^{-1}(k) \mid k \leq N\right\}$ und wählen ein $n \geq M$. Dann gilt

$$\begin{aligned}\left| \sum_{\ell=1}^{n} a_{\varphi(\ell)} - \sum_{k=1}^{\infty} a_k \right| &= \left| \sum_{\ell=1}^{n} a_{\varphi(\ell)} - \sum_{k=1}^{N} a_k - \sum_{k=N+1}^{\infty} a_k \right| \\ &\leq \left| \sum_{\substack{\ell \leq n, \\ \varphi(\ell) > N}} a_{\varphi(\ell)} \right| + \underbrace{\left| \sum_{k=N+1}^{\infty} a_k \right|}_{\leq \varepsilon} \leq \sum_{\substack{\ell \leq n, \\ \varphi(\ell) > N}} |a_{\varphi(\ell)}| + \varepsilon < \varepsilon + \varepsilon,\end{aligned}$$

wobei wir verwendet haben, dass $\varphi$ eine Bijektion ist. Insbesondere treten damit und wegen $n \geq M$ alle $k \in \{1, \ldots, N\}$ genau einmal als $\varphi(\ell)$ für $\ell \in \{1, \ldots, n\}$ auf und die Differenz $\sum_{\ell=1}^{n} a_{\varphi(\ell)} - \sum_{k=1}^{N} a_k$ enthält nach Wegstreichen dieser Terme nur mehr eine Summe

über gewisse $a_k$ mit $k \geq N$ (welche wir im Sinne obiger Interpretation als „weniger wichtig“ betrachten und formal eben insgesamt durch ein $\varepsilon$ abschätzen können). Da $\varepsilon > 0$ beliebig war, zeigt dies die Gleichung (6.2).

Wenden wir dasselbe Argument wie oben auf die Reihe $\sum_{n=1}^{\infty} |a_{\varphi(n)}|$ an, ergibt sich auch die absolute Konvergenz von $\sum_{n=1}^{\infty} a_{\varphi(n)}$. □

VO: 22.11.17

### 6.2.3 Produkte

Wir zeigen nun, dass wir absolut konvergente Reihen gliedweise ausmultiplizieren können.

**Satz 6.36** (Produktsatz)**.** *Seien $\sum_{n=1}^{\infty} a_n$ und $\sum_{n=1}^{\infty} b_n$ zwei absolut konvergente Reihen und $\varphi : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ eine bijektive Abbildung. Dann ist*

$$\sum_{n=1}^{\infty} a_{\varphi(n)_1} b_{\varphi(n)_2}$$

*eine absolut konvergente Reihe, wobei $\varphi(n) = (\varphi(n)_1, \varphi(n)_2)$ für alle $n \in \mathbb{N}$. Weiters gilt*

$$\sum_{n=1}^{\infty} a_{\varphi(n)_1} b_{\varphi(n)_2} = \left(\sum_{n=1}^{\infty} a_n\right)\left(\sum_{n=1}^{\infty} b_n\right). \tag{6.3}$$

Informell ausgedrückt kann man schreiben

$$\left(\sum_{m=1}^{\infty} a_m\right)\left(\sum_{n=1}^{\infty} b_n\right) = \sum_{m=1}^{\infty}\left(\sum_{n=1}^{\infty} b_n\right) a_m = \sum_{(m,n)\in\mathbb{N}^2} a_m b_n.$$

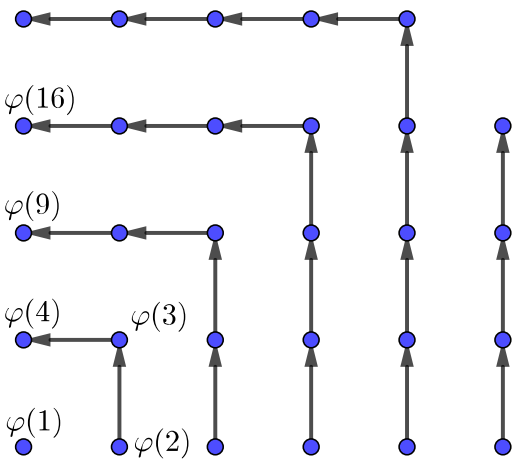
Die beiden (internen) Indices $(m, n)$ würde man nun gerne anders ausdrücken, damit aus der Doppelsumme auf der rechten Seite (die wir eigentlich nicht definiert haben) eine einfache Summe wird. Wählt man eine Bijektion $\varphi : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$, so durchläuft $\varphi(k)$ alle $(m, n)$ und somit wird aus der Doppelsumme eine einfache Summe $\sum_{k=1}^{\infty} a_{\varphi(k)_1} b_{\varphi(k)_2}$. Satz 6.36 besagt nun, dass diese Reihe effektiv konvergiert und gleich dem gewünschten Produkt ist.

*Beweis.* Wir wählen zuerst die Bijektion $\varphi : \mathbb{N} \to \mathbb{N}^2$ so dass

$$\{\varphi(1), \varphi(2), \ldots, \varphi(n^2)\} = \{1, 2, \ldots, n\} \times \{1, 2, \ldots, n\}$$

für alle $n \in \mathbb{N}$. Zum Beispiel könnte $\varphi$ wie im folgenden Bild definiert sein.[3]

[3] Wir verwenden hier das Bild um uns eine sonst eher langweilige formale Definition der Abbildung $\varphi$ zu ersparen. Sie sollten sich aber davon überzeugen, dass man diese Definition durchaus formal machen kann oder einfach formal mittels Induktion nach $n$ die Existenz einer Bijektion mit der gewünschten Eigenschaft zeigen kann. Wir verwenden hier also das Bild als eine Abkürzung für einen formalen Beweis und nicht als einen Ersatz.




Für jedes $n \in \mathbb{N}$ gilt dann für die Partialsumme bis $n^2$ der gliedweise multiplizierten Reihe der Absolutbeträge das (endliche verallgemeinerte) Distributivgesetz

$$\sum_{k=1}^{n^2} |a_{\varphi(k)_1}||b_{\varphi(k)_2}| = \left(\sum_{\ell=1}^{n} |a_\ell|\right)\left(\sum_{m=1}^{n} |b_m|\right).$$

Insbesondere folgt also

$$\sum_{k=1}^{n^2} |a_{\varphi(k)_1}||b_{\varphi(k)_2}| \leq \left(\sum_{\ell=1}^{\infty} |a_\ell|\right)\left(\sum_{m=1}^{\infty} |b_m|\right)$$

für alle $n \in \mathbb{N}$. Da aber für eine Reihe mit nichtnegativen Termen die Folge die Partialsummen monoton wachsend sind, folgt daraus dass die Reihe $\sum_{k=1}^{\infty} |a_{\varphi(k)_1}||b_{\varphi(k)_2}|$ konvergiert und damit die Reihe $\sum_{k=1}^{\infty} a_{\varphi(k)_1} b_{\varphi(k)_2}$ absolut konvergent ist.

Obiges Distributivgesetz gilt auch in der Form

$$\sum_{k=1}^{n^2} a_{\varphi(k)_1} b_{\varphi(k)_2} = \left(\sum_{\ell=1}^{n} a_\ell\right)\left(\sum_{m=1}^{n} b_m\right).$$

für alle $n \in \mathbb{N}$. Mit Hilfe des Grenzwertübergangs $n \to \infty$ erhalten wir daraus

$$\sum_{k=1}^{\infty} a_{\varphi(k)_1} b_{\varphi(k)_2} = \left(\sum_{\ell=1}^{\infty} a_\ell\right)\left(\sum_{m=1}^{\infty} b_m\right).$$

Betrachten wir eine beliebige Bijektion $\psi : \mathbb{N} \to \mathbb{N}^2$, so ist $\varphi^{-1} \circ \psi : \mathbb{N} \to \mathbb{N}$ eine Bijektion und die Formel

$$\sum_{k=1}^{\infty} a_{\psi(k)_1} b_{\psi(k)_2} = \left(\sum_{\ell=1}^{\infty} a_\ell\right)\left(\sum_{m=1}^{\infty} b_m\right).$$

folgt aus obigem und dem Umordnungssatz (Satz 6.35). □

Wie wir in Abschnitt 3.2 gesehen haben, lassen sich Polynome mittels der Regel

$$\left(\sum_{n=0}^{N} a_n x^n\right)\left(\sum_{n=0}^{N} b_n x^n\right) = \sum_{n=0}^{2N}\left(\sum_{k=0}^{n} a_{n-k} b_k\right) x^k$$

multiplizieren, wobei wir $a_{N+1} = \cdots = a_{2N} = b_{N+1} = \cdots = b_{2N} = 0$ setzen. Setzt man $x = 1$, erhält man insbesondere

$$\left(\sum_{n=0}^{N} a_n\right)\left(\sum_{n=0}^{N} b_n\right) = \sum_{n=0}^{2N}\left(\sum_{k=0}^{n} a_{n-k} b_k\right).$$

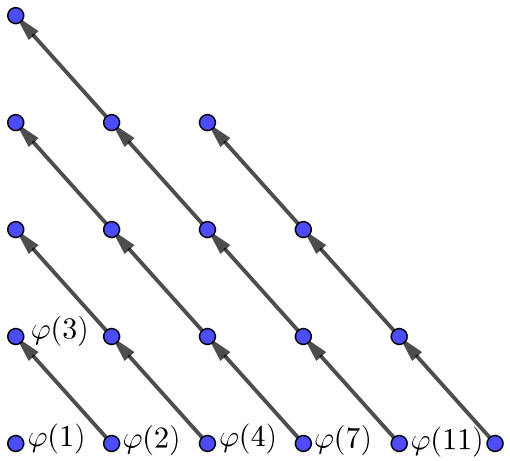
Diese Identität trifft, wie sich herausstellt, analog für Reihen zu, was wir im folgenden Korollar des Produktsatzes (Satz 6.36) festhalten wollen.

**Korollar 6.37** (Cauchy-Produkt)**.** *Falls $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ absolut konvergente Reihen mit komplexen Gliedern sind, dann gilt*

$$\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_{n-k} b_k\right) = \left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right),$$

*wobei die Reihe $\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_{n-k} b_k\right)$ absolut konvergent ist.*

*Beweis.* Dies folgt, indem wir die Abzählung $\varphi$ von $\mathbb{N}_0 \times \mathbb{N}_0$ aus dem Bild unten auf den Produktsatz (Satz 6.36) anwenden und dann Glieder zusammenfassen (Lemma 6.9).




Die absolute Konvergenz folgt ebenso aus Satz 6.36 und

$$\sum_{n=0}^{\infty}\left|\sum_{k=0}^{n} a_{n-k} b_k\right| \leq \sum_{n=0}^{\infty}\sum_{k=0}^{n} |a_{n-k} b_k| < \infty.$$

□

**Beispiel 6.38.** *Sei $q \in \mathbb{C}$ mit $|q| < 1$. Dann konvergiert $\sum_{n=0}^{\infty} q^n$ absolut. Wenden wir das Cauchy-Produkt auf diese Reihe und sich selbst an, so erhalten wir*

$$\begin{aligned}\frac{1}{(1-q)^2} &= \left(\sum_{n=0}^{\infty} q^n\right)^2 \\ &= \sum_{n=0}^{\infty}\sum_{k=0}^{n} q^{n-k}q^k = \sum_{n=0}^{\infty}(n+1)q^n.\end{aligned}$$

*Auf diese Weise erhalten wir auch eine Summenformel für*

$$\sum_{n=1}^{\infty} nq^n = q\sum_{n=1}^{\infty} nq^{n-1} = q\sum_{k=0}^{\infty}(k+1)q^k = \frac{q}{(1-q)^2},$$

*wobei wir die Indexverschiebung $k = n-1$ durchgeführt haben.*

**Übung 6.39.** *Formal lässt sich auch für bedingt konvergente Reihen $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ das Cauchy-Produkt $\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_{n-k}b_k\right)$ bilden. Es muss jedoch nicht mehr konvergent sein: Zeigen Sie, dass das Cauchy-Produkts der bedingt konvergenten Reihe*

$$\sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{\sqrt{n+1}}$$

*mit sich selbst divergiert.*

## 6.3 Konvergenz von Funktionenfolgen

### 6.3.1 Punktweise Konvergenz

**Definition 6.40** (Funktionenfolgen und punktweise Konvergenz)**.** Eine reellwertige (oder komplexwertige) **Funktionenfolge** auf einer Menge $X$ ist eine Folge $(f_n)_n$ von Funktionen $f_n : X \to \mathbb{R}$ (oder $f_n : X \to \mathbb{C}$). Wir sagen, dass eine Funktionenfolge $(f_n)_n$ **punktweise** gegen eine Funktion $f : X \to \mathbb{R}$ (oder $f : X \to \mathbb{C}$) **konvergiert**, falls $f_n(x) \to f(x)$ für $n \to \infty$ und alle $x \in X$. Wir bezeichnen die Funktion $f$ als den **punktweisen Grenzwert** (oder auch **Grenzfunktion** oder **Limes**) der Funktionenfolge $(f_n)_n$. In Prädikatenlogik ist punktweise Konvergenz durch

$$\forall x \in X \ \forall \varepsilon > 0 \ \exists N \in \mathbb{N} \ \forall n \in \mathbb{N} : \big(n \geq N \implies |f_n(x) - f(x)| < \varepsilon\big)$$

gegeben.

**Übung 6.41** (Eindeutigkeit)**.** *Sei $(f_n)_n$ eine Funktionenfolge auf einer Menge $X$. Zeigen Sie, dass der Grenzwert $f$ einer Funktionenfolge eindeutig bestimmt ist, falls er existiert.*

Wir haben in Abschnitt 5.3 bereits ein Beispiel einer punktweise konvergenten Funktionenfolge gesehen, da wir die reelle Exponentialabbildung exp durch

$$\exp(x) = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n$$

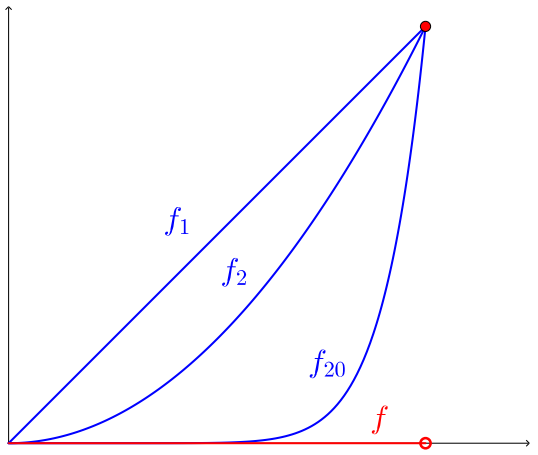
für $x \in \mathbb{R}$ definiert haben.

Wir betrachten einige weitere Beispiele, die die Nachteile der punktweisen Konvergenz aufzeigen werden.

**Beispiel 6.42** (Punktweise konvergent)**.** *Sei $X = [0,1]$ und $f_n : x \in [0,1] \to x^n \in \mathbb{R}$. Dann konvergieren die stetigen Funktionen $f_n$ punktweise gegen die Funktion $f : [0,1] \to \mathbb{R}$ gegeben durch*

$$f(x) = \mathbb{1}_{\{1\}}(x) = \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} x^n = \begin{cases} 0 & \text{für } x < 1 \\ 1 & \text{für } x = 1 \end{cases}$$

*für $x \in [0,1]$, die nicht mehr stetig ist.*




**Beispiel 6.43** (Punktweise konvergent)**.** *Sei wiederum* $X = [0,1]$ *und definiere* $f_n : [0,1] \to \mathbb{R}$ *durch*

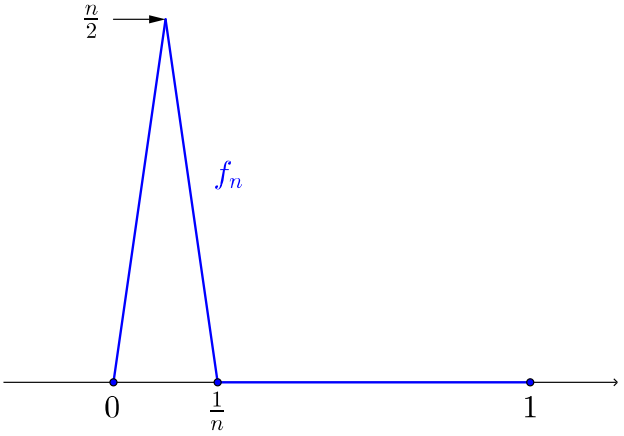
$$f_n(x) = \begin{cases} n^2 x & \text{für } x \in \left[0, \frac{1}{2n}\right] \\ n^2\left(\frac{1}{n} - x\right) & \text{für } x \in \left[\frac{1}{2n}, \frac{1}{n}\right] \\ 0 & \text{für } x \in \left[\frac{1}{n}, 1\right] \end{cases}$$

*für* $x \in [0,1]$ *und* $n \in \mathbb{N}$*. Dann ist* $f_n$ *stetig (und somit auch Riemann-integrierbar) und konvergiert punktweise gegen die stetige Funktion* $f : x \in [0,1] \mapsto \lim_{n\to\infty} f_n(x) = 0$*.*

*Es gilt jedoch für alle* $n \in \mathbb{N}$

$$\int_0^1 f_n(x)\,\mathrm{d}x = \frac{1}{4} \neq 0 = \int_0^1 f(x)\,\mathrm{d}x.$$

*Also ist der Grenzwert der Integrale nicht gleich dem Integral der Limesfunktion, obwohl alle Funktionen stetig sind und die Limesfunktion stetig ist.*




**Beispiel 6.44** (Punktweise konvergent)**.** *Sei wieder* $X = [0,1]$ *und* $\mathbb{Q} \cap [0,1] = \{q_1, q_2, \ldots\}$ *eine Abzählung. Dann ist für jedes* $n \in \mathbb{N}$ *die charakteristische Funktion*

$$f_n = \mathbb{1}_{\{q_1,\ldots,q_n\}}$$

*der ersten $n$ rationalen Zahlen $\{q_1, \ldots, q_n\}$ in $[0,1]$ Riemann-integrierbar mit $\int_0^1 f_n(x)\,\mathrm{d}x = 0$. Die Limesfunktion der Folge $(f_n)_n$ ist aber die charakteristische Funktion $\mathbb{1}_{\mathbb{Q}\cap[0,1]}$, die nach Beispiel 4.17 nicht Riemann-integrierbar ist.*

Zusammenfassend hat also der Begriff der punktweisen Konvergenz weder für die Stetigkeit noch für das Riemann-Integral besonders gute Eigenschaften. Wir wenden uns deswegen einem neuen Konvergenzbegriff zu.

### 6.3.2 Gleichmässige Konvergenz

**Definition 6.45** (Gleichmässige Konvergenz)**.** Sei $(f_n)_n$ eine komplexwertige Funktionenfolge auf einer Menge $X$ und $f$ eine weitere komplexwertige Funktion auf $X$. Wir sagen, $f_n$ **strebt gleichmässig** gegen $f$ für $n \to \infty$, oder dass $f$ der **gleichmässige Grenzwert** der Funktionenfolge $(f_n)_n$ ist, falls es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass für alle $n \geq N$ und alle $x \in X$ die Abschätzung

$$|f_n(x) - f(x)| < \varepsilon$$

gilt. In Prädikatenlogik ist gleichmässige Konvergenz durch

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \in \mathbb{N} : \big(n \geq N \implies (\forall x \in X : |f_n(x) - f(x)| < \varepsilon)\big)$$

gegeben.

Im Gegensatz zur punktweisen Konvergenz von Funktionenfolgen nimmt man bei der gleichmässigen Konvergenz also an, dass ein $N \in \mathbb{N}$ existiert, so dass $|f_n(x) - f(x)| < \varepsilon$ für alle $x \in X$ und alle natürlichen $n \geq N$ gilt, wobei die Zahl $N$ nicht vom Punkt $x \in X$ abhängt. Gleichmässigkeit bezieht sich meistens (wie hier und zum Beispiel auch bei der gleichmässigen Stetigkeit) auf die Unabhängigkeit einer gewissen Zahl (hier $N$ und bei gleichmässiger Stetigkeit $\delta$) von der Wahl eines Punktes (hier $x$).

**Übung 6.46** (Gleichmässige Konvergenz)**.** *Sei $(f_n)_n$ eine komplexwertige Funktionenfolge auf einer Menge $X$ und $f$ eine weitere komplexwertige Funktion auf $X$.*

*(i) Zeigen Sie, dass $f_n$ genau dann gleichmässig gegen $f$ konvergiert, wenn*

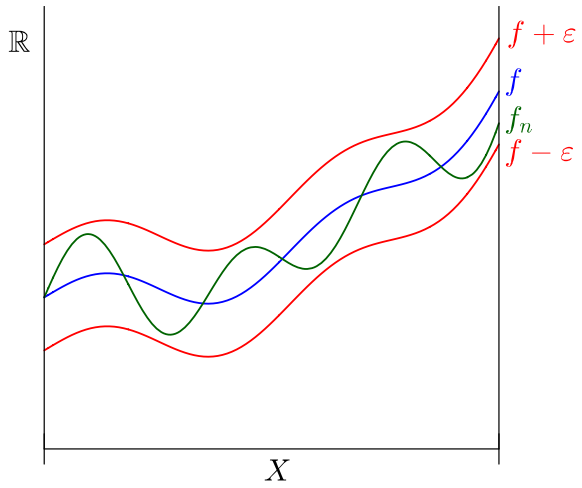
$$\sup_{x \in X} |f_n(x) - f(x)| \to 0$$

*für $n \to \infty$.*

*(ii) Zeigen Sie, dass $f_n \to f$ gleichmässig für $n \to \infty$ auch $f_n \to f$ punktweise für $n \to \infty$ impliziert.*

*(iii) Zeigen Sie, dass die punktweise konvergenten Funktionenfolgen aus den Beispielen 6.42, 6.43 und 6.44 nicht gleichmässig konvergieren. Insbesondere ist punktweise Konvergenz eine schwächere Forderung als gleichmässige Konvergenz.*

Für reellwertige Funktionen $f_n, f$, $\varepsilon > 0$ und $x$ im Definitionsbereich ist die Abschätzung $|f_n(x) - f(x)| < \varepsilon$ zu $f(x) - \varepsilon \leq f_n(x) \leq f(x) + \varepsilon$ äquivalent. Dadurch lässt sich gleichmässige Konvergenz auch durch den Graphen einer Funktionenfolge und deren Limesfunktion beschreiben, wie wir in folgender Figur demonstrieren wollen.




Figur 6.3: Die Funktionenfolge $f_n$ konvergiert also gleichmässig gegen $f$, wenn für jedes $\varepsilon > 0$ die Graphen von $f_n$ für alle hinreichend grossen $n$ im „$\varepsilon$-Schlauch“ rund um $f$ (also in dem Bereich zwischen den Graphen von $f - \varepsilon$ und $f + \varepsilon$ liegen.

**Applet 6.47** (Punktweise und Gleichmässige Konvergenz)**.** *Wir betrachten nochmals die Funktionenfolge aus Beispiel 6.42. Versuchen Sie anhand des Applets zu erkennen, dass durch Einschränkung auf geeignete Teilintervalle $[a, b] \subseteq [0, 1]$ man gleichmässige Konvergenz der Funktionenfolge erreichen kann. Können Sie dies auch formal beweisen?*

Gleichmässige Konvergenz hat für stetige Funktionen gute Eigenschaften.

**Satz 6.48** (Gleichmässige Konvergenz und Stetigkeit)**.** *Sei $D \subseteq \mathbb{C}$ und $f_n : D \to \mathbb{C}$ eine Funktionenfolge stetiger Funktionen. Falls $(f_n)_n$ gleichmässig gegen $f : D \to \mathbb{C}$ konvergiert, dann ist $f$ ebenso stetig.*

*Beweis.* Sei $x_0 \in D$ und $\varepsilon > 0$. Dann existiert ein $n \in \mathbb{N}$, so dass

$$|f_n(x) - f(x)| < \varepsilon$$

für alle $x \in D$. Da $f_n$ bei $x_0$ stetig ist, existiert ein $\delta > 0$, so dass

$$|x - x_0| < \delta \implies |f_n(x) - f_n(x_0)| < \varepsilon$$

für alle $x \in D$ gilt. Unter dem Strich gilt nun für alle $x \in D$ mit $|x - x_0| < \delta$, dass

$$|f(x) - f(x_0)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)| < 3\varepsilon.$$

Da $\varepsilon > 0$ beliebig war, ist $f$ bei $x_0$ stetig. Da $x_0 \in D$ beliebig war, folgt der Satz. □

VO: 23.11.17

Gleichmässige Konvergenz hat auch für die Integrierbarkeit gute Eigenschaften.

**Satz 6.49** (Gleichmässige Konvergenz und Riemann-Integrierbarkeit)**.** *Sei $[a,b]$ ein kompaktes Intervall und $f_n : [a,b] \to \mathbb{R}$ eine Funktionenfolge Riemann-integrierbarer Funktionen. Falls $(f_n)_n$ gleichmässig gegen $f : [a,b] \to \mathbb{R}$ konvergiert, dann ist $f$ Riemann-integrierbar und*

$$\lim_{n\to\infty} \int_a^b f_n \,\mathrm{d}x = \int_a^b \lim_{n\to\infty} f_n \,\mathrm{d}x = \int_a^b f \,\mathrm{d}x. \tag{6.4}$$

Bei gleichmässig konvergenten Folgen Riemann-integrierbarer Funktionen darf man also Integration und Grenzwert vertauschen. Dies gilt analog auch für Integrale von komplexwertigen oder vektorwertigen Funktionen (siehe Abschnitt 5.7.6), doch werden wir hier nur den Fall von reellwertigen Funktionen betrachten.

*Beweis.* Sei $\varepsilon > 0$. Dann gibt es ein $N$ mit $f - \varepsilon \leq f_n \leq f + \varepsilon$ für alle $n \geq N$. Da $f_n$ nach Annahme Riemann-integrierbar ist, gibt es Treppenfunktionen $u, o$ auf $[a,b]$ mit $u \leq f_n \leq o$ und $\int_a^b (u - o) \,\mathrm{d}x < \varepsilon$. Daraus folgt, dass

$$u' = u - \varepsilon \leq f_n - \varepsilon \leq f \leq f_n + \varepsilon \leq o + \varepsilon = o'$$

ist und

$$\int_a^b (o' - u') \,\mathrm{d}x = \int_a^b (o - u) \,\mathrm{d}x + 2\varepsilon(b-a) < \varepsilon(2b - 2a + 1)$$

ist. Da $\varepsilon > 0$ beliebig war, folgt die Riemann-Integrierbarkeit von $f$ aus Proposition 4.12.

Für die zweite Aussage sei wiederum $\varepsilon > 0$ und $N \in \mathbb{N}$, so dass $f - \varepsilon \leq f_n \leq f + \varepsilon$ für alle $n \geq N$ gilt. Aus der Monotonie des Riemann-Integrals in Satz 4.24 folgt nun

$$\int_a^b f \,\mathrm{d}x - \varepsilon(b-a) = \int_a^b (f - \varepsilon) \,\mathrm{d}x \leq \int_a^b f_n \,\mathrm{d}x \leq \int_a^b (f + \varepsilon) \,\mathrm{d}x = \int_a^b f \,\mathrm{d}x + \varepsilon(b-a),$$

was zu

$$\left| \int_a^b f \,\mathrm{d}x - \int_a^b f_n \,\mathrm{d}x \right| \leq \varepsilon(b-a)$$

äquivalent ist. Dies beweist die Konvergenz in Gleichung (6.4) und damit den Satz. □

**Übung 6.50** (Gleichmässige Konvergenz auf Teilmengen)**.** *Sei $(f_n)_n$ eine komplexwertige Funktionenfolge auf einer Menge $X$ und $f$ eine weitere komplexwertige Funktion auf $X$.*

*(i) Angenommen $X = X_1 \cup X_2$ für zwei Teilmengen, $f_n|_{X_1}$ strebt gleichmässig gegen $f|_{X_1}$ für $n \to \infty$ und $f_n|_{X_2}$ strebt gleichmässig gegen $f|_{X_2}$ für $n \to \infty$. Zeigen Sie, dass dann auch $f_n$ gleichmässig gegen $f$ strebt für $n \to \infty$.*

*(ii) Zeigen Sie, dass sich Teil (i) im Allgemeinen nicht für unendliche Vereinigungen $X = \bigcup_{k\in\mathbb{N}} X_k$ verallgemeinern lässt.*

**Übung 6.51** (Funktionenkonvergenz und Beschränktheit)**.** *Sei $(f_n)_n$ eine komplexwertige Funktionenfolge auf einer Menge $X$ und $f$ eine weitere komplexwertige Funktion auf $X$. Wir nehmen nun an, dass die Funktionen $f_n$ für jedes $n \in \mathbb{N}$ beschränkt ist.*

*(i) Zeigen Sie, dass falls $f_n$ gleichmässig gegen $f$ strebt für $n \to \infty$, dann ist auch $f$ eine beschränkte Funktion.*

*(ii) Finden Sie ein Beispiel für $X$ und beschränkte Funktionen, die punktweise gegen $f$ streben für $n \to \infty$, so dass $f$ nicht beschränkt ist.*

**Übung 6.52** (Funktionenkonvergenz und Auswertung entlang einer konvergenten Folge)**.** *Sei $D \subseteq \mathbb{C}$ und $f_n : D \to \mathbb{C}$ für $n \in \mathbb{N}$ eine Funktionenfolge stetiger Funktionen und $f : D \to \mathbb{C}$ eine weitere Funktion. Sei $z_0 \in D$ und $z_n \in D$ eine Folge mit $z_n \to z_0$ für $n \to \infty$.*

*(i) Zeigen Sie $f_n(z_n) \to f(z_0)$ für $n \to \infty$ unter der Annahme, dass $f_n$ gleichmässig gegen $f$ konvergiert für $n \to \infty$.*

*(ii) Finden Sie ein Beispiel, wo zwar $f_n$ punktweise gegen $f$ konvergiert für $n \to \infty$, aber $f_n(z_n)$ nicht gegen $f(z_0)$ konvergiert für $n \to \infty$.*

**Übung 6.53** (Gleichmässige Konvergenz und gleichmässige Stetigkeit)**.** *Sei $D \subseteq \mathbb{C}$ und $f_n : D \to \mathbb{C}$ für $n \in \mathbb{N}$ eine Funktionenfolge, die für $n \to \infty$ gleichmässig gegen $f : D \to \mathbb{C}$ strebt. Angenommen $f_n$ ist gleichmässig stetig für alle $n \in \mathbb{N}$. Zeigen Sie, dass $f$ ebenso gleichmässig stetig ist.*

## 6.4 Potenzreihen

**Definition 6.54** (Potenzreihe)**.** Für jedes $n \in \mathbb{N}_0$ sei $a_n \in \mathbb{C}$. Dann ist der formale Ausdruck

$$\sum_{n=0}^{\infty} a_n z^n \tag{6.5}$$

eine **Potenzreihe** in der Variable $z$.

Es drängt sich bei obiger Definition ein Vergleich zur Definition eines Polynoms in Definition 3.13 auf. Im Gegensatz zur Diskussion von Polynomen ist aber eine Potenzreihe vorerst nur ein formaler Ausdruck. Es ist nicht klar, bei welchen komplexen Zahlen man eine Potenzreihe auswerten darf. Insbesondere wissen wir (noch) nicht, ob wir diesem formalen Ausdruck überhaupt eine Funktion auf $\mathbb{C}$ oder einer bestimmten Teilmenge von $\mathbb{C}$ zuordnen können. Diese Frage hängt stark von den Koeffizienten $(a_n)_{n\in\mathbb{N}_0}$ ab und wird in Satz 6.56 beantwortet.

Wie schon bei Polynomen in Definition 3.11 ist auch hier die Definition (3.2) äusserst sinnvoll, damit die Potenzreihe (6.5) bei $z = 0$ auf jeden Fall konvergiert und den Wert $a_0$ hat.

### 6.4.1 Konvergenzradius

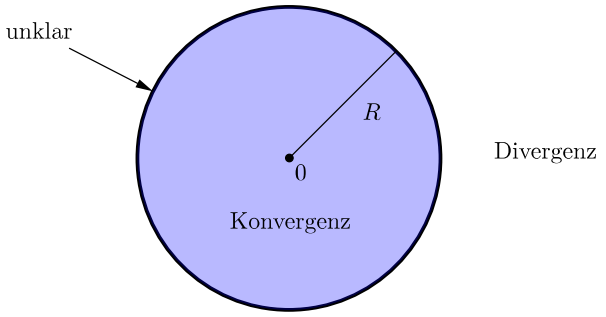
**Definition 6.55.** Sei $\sum_{n=0}^{\infty} a_n z^n$ eine Potenzreihe mit komplexen Koeffizienten $(a_n)_{n\in\mathbb{N}_0}$. Wir definieren den **Konvergenzradius** durch

$$R = \frac{1}{\limsup_{n\to\infty} \sqrt[n]{|a_n|}},$$

wobei wir $\frac{1}{+\infty} = 0$ setzen und hier (aber auch nur hier) die Vereinbarung $\frac{1}{0} = +\infty$ treffen.

**Satz 6.56** (Über den Konvergenzradius)**.** *Sei $\sum_{n=0}^{\infty} a_n z^n$ eine Potenzreihe und $R$ ihr Konvergenzradius. Dann konvergiert die Reihe $\sum_{n=0}^{\infty} a_n z^n$ für alle $z \in \mathbb{C}$ mit $|z| < R$ absolut und divergiert für alle $z \in \mathbb{C}$ mit $|z| > R$. Weiters konvergiert die Funktionenfolge $\sum_{n=0}^{N} a_n z^n$ gleichmässig gegen $\sum_{n=0}^{\infty} a_n z^n$ auf jeder Kreisscheibe der Form $B_S(0) = \{z \in \mathbb{C} \mid |z| < S\}$ für jedes $S \in (0, R)$. Insbesondere definiert die Potenzreihe die stetige Abbildung*

$$z \in B_R(0) \mapsto \sum_{n=0}^{\infty} a_n z^n \in \mathbb{C}.$$




*Beweis.* Wir verwenden das Wurzelkriterium aus Korollar 6.30 für ein beliebiges $z \in \mathbb{C}$ und die Reihe $\sum_{n=0}^{\infty} a_n z^n$ und berechnen deswegen

$$\limsup_{n\to\infty} \sqrt[n]{|a_n z^n|} = \limsup_{n\to\infty} \sqrt[n]{|a_n|}|z| = |z| \limsup_{n\to\infty} \sqrt[n]{|a_n|} = \frac{|z|}{R}.$$

Nach dem Wurzelkriterium konvergiert die Reihe also absolut für $\frac{|z|}{R} < 1$ und divergiert für $\frac{|z|}{R} > 1$. Die Fälle $R = 0$ und $R = +\infty$ ergeben sich aus dem gleichen Argument (wieso?).

Sei nun $S \in (0, R)$. Für den Beweis der gleichmässigen Konvergenz auf $B_S(0)$ bemerken wir, dass nach obigem bereits $\sum_{n=0}^{\infty} |a_n| S^n < \infty$ gilt. Daher existiert für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$ mit $\sum_{n=N}^{\infty} |a_n| S^n < \varepsilon$. Für alle $z \in B_S(0)$ und $n \geq N$ gilt damit

$$\left| \sum_{k=0}^{n} a_k z^k - \sum_{k=0}^{\infty} a_k z^k \right| = \left| \sum_{k=n+1}^{\infty} a_k z^k \right| \leq \sum_{k=N}^{\infty} |a_k| S^k < \varepsilon.$$

Dies beweist die gleichmässige Konvergenz der stetigen Funktionenfolge $\sum_{k=0}^{n} a_k z^k$ auf $B_S(0)$ gegen $\sum_{k=0}^{\infty} a_k z^k$ und damit die Stetigkeit von $z \in B_S(0) \mapsto \sum_{k=0}^{\infty} a_k z^k \in \mathbb{C}$ nach Satz 6.48.

Insbesondere ist die Funktion $z \in B_R(0) \mapsto \sum_{n=0}^{\infty} a_n z^n \in \mathbb{C}$ stetig an jedem Punkt, da es zu $z \in B_R(0)$ ein $S < R$ gibt, mit $z \in B_S(0)$ (wieso zeigt dies die Stetigkeit?). Dies beweist den Satz. □

**Beispiel 6.57** (Nicht gleichmässige Konvergenz)**.** *Man könnte denken, dass Satz 6.56 eigentlich sagt, dass die Partialsummen $\sum_{k=0}^{n} a_k z^k$ der Potenzreihe auf ganz $B_R(0)$ gleichmässig gegen die durch die Potenzreihe definierte Funktion $z \in B_R(0) \mapsto \sum_{k=0}^{\infty} a_k z^k$ streben, da ja $S < R$ beliebig ist. Dies ist aber nicht immer so (siehe auch Übung 6.50), wie wir hier kurz anhand der geometrischen Reihe zeigen wollen.*

*Für $\sum_{n=0}^{\infty} z^n$ ist der Konvergenzradius $R = 1$ und die mittels der Potenzreihe definierte Funktion ist $z \in B_1(0) \mapsto \frac{1}{1-z}$. Falls die Konvergenz auf ganz $B_1(0)$ gleichmässig wäre, dann gäbe es für $\varepsilon = 1$ ein $N$ so dass für alle $n \geq N$ und $z \in B_1(0)$ die Abschätzung*

$$\left| \sum_{k=0}^{n} z^k - \frac{1}{1-z} \right| < 1$$

*gelten würde. Wir setzen $n = N$ und erhalten mittels der Dreiecksungleichung daraus*

$$\left|\frac{1}{1-z}\right| < 1 + \left|\sum_{k=0}^{N} z^k\right| \leq 2 + N$$

*für alle $z \in B_1(0)$. Dies ist aber ein Widerspruch, da*

$$\lim_{x \nearrow 1} \frac{1}{1-x} = +\infty.$$

**Übung 6.58** (Konvergenzradien)**.** *Finden Sie für jedes $R \in [0, \infty) \cup \{\infty\}$ eine Potenzreihe mit Konvergenzradius $R$.*

**Übung 6.59.** *Berechnen Sie den Konvergenzradius $R$ der Potenzreihe*

$$\sum_{n=1}^{\infty} \frac{(\sqrt{n^2+n} - \sqrt{n^2+1})^n}{n^2} x^n$$

*und zeigen Sie Konvergenz der Potenzreihe bei den Punkten $-R, R \in \mathbb{R}$.*

**Lemma 6.60** (Konvergenzradius via Quotientenkriterium)**.** *Sei $\sum_{n=0}^{\infty} a_n z^n$ eine Potenzreihe mit $a_n \neq 0$ für alle $n \in \mathbb{N}$. Der Konvergenzradius $R$ ist gegeben durch*

$$R = \frac{1}{\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|}} = \lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|}$$

*falls dieser Grenzwert existiert.*

**Übung 6.61.** *Zeigen Sie Lemma 6.60.*

### 6.4.2 Addition und Multiplikation

**Proposition 6.62** (Summen und Produkte)**.** *Seien $\sum_{n=0}^{\infty} a_n z^n$ und $\sum_{n=0}^{\infty} b_n z^n$ zwei Potenzreihen mit Konvergenzradius $R_a$ respektive $R_b$. Dann gilt für alle $z \in \mathbb{C}$ mit $|z| < \min\{R_a, R_b\}$*

$$\sum_{n=0}^{\infty} a_n z^n + \sum_{n=0}^{\infty} b_n z^n = \sum_{n=0}^{\infty} (a_n + b_n) z^n$$
$$\left(\sum_{n=0}^{\infty} a_n z^n\right)\left(\sum_{n=0}^{\infty} b_n z^n\right) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_{n-k} b_k\right) z^n.$$

*Insbesondere ist der Konvergenzradius der Potenzreihen auf der rechten Seite mindestens $\min\{R_a, R_b\}$.*

*Beweis.* Die erste Eigenschaft folgt aus Linearität des Grenzwerts. Die zweite verwendet noch Korollar 6.37. □

**Beispiel 6.63.** *Falls* $\sum_{n=0}^{\infty} a_n z^n$ *mindestens Konvergenzradius* 1 *hat, so gilt*

$$\frac{1}{1-z}\sum_{n=0}^{\infty} a_n z^n = \sum_{n=0}^{\infty}(a_0 + \ldots + a_n)z^n. \tag{6.6}$$

*für alle* $z \in \mathbb{C}$ *mit* $|z| < 1$. *In der Tat hat die Potenzreihe* $\sum_{n=0}^{\infty} z^n$ *Konvergenzradius* 1 *und für* $z \in \mathbb{C}$ *mit* $|z| < 1$ *gilt* $\sum_{n=0}^{\infty} z^n = \frac{1}{1-z}$, *womit* (6.6) *aus Proposition 6.62 folgt.*

**Übung 6.64.** *Berechnen Sie* $\sum_{n=1}^{\infty} n2^{-n}$.

### 6.4.3 Stetigkeit bei Randpunkten

Wir wollen nun eine Potenzreihe $\sum_{n=0}^{\infty} a_n x^n$ betrachten, wobei wir $a_n \in \mathbb{C}$ für $n \in \mathbb{N}$ erlauben, aber nur reelle Zahlen $x$ einsetzen wollen. Nach Satz 6.56 existiert ein $R > 0$, so dass die Funktion

$$f : x \in (-R, R) \mapsto \sum_{n=0}^{\infty} a_n x^n \in \mathbb{C}$$

wohldefiniert und stetig ist, aber $\sum_{n=0}^{\infty} a_n x^n$ für alle $x \in \mathbb{R}$ mit $|x| > R$ divergiert. Wir nehmen nun weiter an, dass $R \in (0, \infty)$ und $\sum_{n=0}^{\infty} a_n R^n$ ebenfalls konvergiert. Satz 6.56 sagt in diesem Fall überhaupt nichts über die erweiterte Funktion

$$\bar{f} : x \in (-R, R] \mapsto \sum_{n=0}^{\infty} a_n x^n \in \mathbb{C}$$

aus.

**Satz 6.65** (Abelscher Grenzwertsatz)**.** *Unter obigen Annahmen ist auch* $\bar{f}$ *stetig. Das heisst,*

$$\sum_{n=0}^{\infty} a_n R^n = \bar{f}(R) = \lim_{x \nearrow R} \bar{f}(x) = \lim_{x \nearrow R} \sum_{n=0}^{\infty} a_n x^n$$

*Eine analoge Aussage gilt, falls* $\sum_{n=0}^{\infty} a_n(-R)^n$ *konvergiert.*

**Beispiel 6.66** (Zwei alternierende Potenzreihen)**.**

*(i) Für* $a_0 = 0$ *und* $a_n = \frac{(-1)^{n+1}}{n}$ *für alle* $n \in \mathbb{N}$ *ist der Konvergenzradius der Potenzreihe* $\sum_{n=0}^{\infty} a_n z^n$ *durch* $R = 1$ *gegeben und* $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$ *konvergiert, womit der Abelsche Grenzwertsatz (Satz 6.65) angewendet werden kann (sobald wir die Funktion* $f(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n$ *für* $|x| < 1$ *kennen).*

*(ii) Für* $a_n = (-1)^n$ *für alle* $n \in \mathbb{N}_0$ *ist* $R = 1$ *und der Abelsche Grenzwertsatz (Satz 6.65) kann nicht angewendet werden, da* $\sum_{n=0}^{\infty}(-1)^n$ *divergiert.*

*Bemerkung.* Hierzu eine historische Anmerkung: Euler (1707-1783) und seine Zeitgenossen hatten noch einen anderen Zugang zu Reihen und wiesen auf Grund der Gleichung $\sum_{n=0}^{\infty}(-1)^n x^n = \frac{1}{1+x}$ für $|x| < 1$ der Reihe $\sum_{n=0}^{\infty}(-1)^n = 1 - 1 + 1 - 1 + \cdots$ den Wert $\frac{1}{2}$ zu, was aber unserem modernerem Konvergenzbegriff (und insbesondere Proposition 6.2) widerspricht.

*Beweis des Abelschen Grenzwertsatzes.* Wir nehmen ohne Beschränkung der Allgemeinheit an, dass der Konvergenzradius 1 ist (sonst ersetzt man $a_n$ mit $a_n R^n$ für alle $n$). Nach Beispiel 6.63 gilt

$$\frac{1}{1-x}\sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty}(a_0 + \ldots + a_n)x^n.$$

für alle $x \in (-1,1)$. Wir definieren $A_n = a_0 + \ldots + a_n$, $A = \lim_{n\to\infty} A_n = \sum_{n=0}^{\infty} a_n$ (was nach Annahme existiert) und erhalten mit $b_n = A_n - A$ für $n \in \mathbb{N}$ die Gleichung

$$f(x) = \sum_{n=0}^{\infty} a_n x^n = (1-x)\sum_{n=0}^{\infty} A_n x^n = (1-x)\sum_{n=0}^{\infty}(b_n + A)x^n = (1-x)\sum_{n=0}^{\infty} b_n x^n + A$$

für alle $x \in (-1,1)$. Obige Formelmanipulationen mögen vielleicht vom Himmel gefallen sein; ab jetzt wird das Argument jedoch wenig Überraschungen bieten. Sei $\varepsilon > 0$. Dann existiert ein $N \in \mathbb{N}$ mit $|b_n| < \varepsilon$ für alle $n \geq N$. Daraus folgt für $x \in [0,1)$, dass

$$\begin{aligned} |f(x) - A| &= \left|(1-x)\sum_{n=0}^{\infty} b_n x^n\right| \leq \left|(1-x)\sum_{n=0}^{N} b_n x^n\right| + (1-x)\varepsilon \sum_{n=N+1}^{\infty} x^n \\ &\leq \left|(1-x)\sum_{n=0}^{N} b_n x^n\right| + \varepsilon. \end{aligned}$$

Da aber das Polynom $(1-x)\sum_{n=0}^{N} b_n x^n$ auf $\mathbb{R}$ stetig ist und bei 1 verschwindet, gibt es weiters ein $\delta > 0$, so dass

$$x \in (1-\delta, 1) \implies \left|(1-x)\sum_{n=0}^{N} b_n x^n\right| < \varepsilon.$$

Daher gilt $|f(x) - A| < 2\varepsilon$ für alle $x \in (1-\delta, 1)$ und der Satz folgt. □

**Übung 6.67.** *Wo wurde im obigen Beweis verwendet, dass $x$ reell ist?*

## 6.5 Die (komplexe) Exponentialabbildung

Wir haben in Abschnitt 5.3 die reelle Exponentiabbildung gesehen und ihre wichtigsten Eigenschaften gezeigt. Insbesondere haben wir verifiziert, dass

$$\exp(x) = \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = \lim_{n\to\infty}\sum_{k=0}^{n}\frac{1}{k!}x^k\prod_{\ell=0}^{k-1}\left(1-\frac{\ell}{n}\right).$$

für alle $x \in \mathbb{R}$. Wir zeigen nun, dass wir die Exponentialabbildung alternativ durch die Potenzreihe

$$\exp(x) = \sum_{k=0}^{\infty}\frac{1}{k!}x^k \tag{6.7}$$

mit unendlichem Konvergenzradius definieren können und dadurch auf die gesamte komplexe Ebene fortsetzen können. Die Darstellung der Exponentialabbildung als Potenzreihe (6.7) ist gewissermassen flexibler als die Darstellung als Grenzwert wie in Proposition 5.54. Sie wird uns später in diesem Kapitel beispielsweise dabei helfen, das Riemann-Integral der Exponentialfunktion zu berechnen (siehe Korollar 6.86).

### 6.5.1 Darstellung durch die Potenzreihe

Für ein festes $z \in \mathbb{C}$ gilt

$$\frac{\frac{1}{(k+1)!}|z|^{k+1}}{\frac{1}{k!}|z|^k} = \frac{|z|}{k+1} \to 0$$

für $k \to \infty$, was wegen dem Quotientenkriterium beweist, dass die Reihe $\sum_{k=0}^{\infty}\frac{1}{k!}z^k$ konvergiert. Da $z \in \mathbb{C}$ beliebig war, ist der Konvergenzradius der Potenzreihe $\sum_{k=0}^{\infty}\frac{1}{k!}z^k$ in der Tat unendlich (siehe Satz 6.56).

Sei nun $x \in \mathbb{R}$, $\varepsilon > 0$ und $N \in \mathbb{N}$, so dass

$$\sum_{k=N+1}^{\infty}\frac{1}{k!}|x|^k < \varepsilon.$$

Für dieses $N$ gilt dann

$$\left|\sum_{k=0}^{N}\frac{1}{k!}x^k - \sum_{k=0}^{\infty}\frac{1}{k!}x^k\right| \le \left|\sum_{k=N+1}^{\infty}\frac{1}{k!}x^k\right| \le \sum_{k=N+1}^{\infty}\frac{1}{k!}|x|^k < \varepsilon \tag{6.8}$$

und für $n \ge N$ ebenso

$$\left|\sum_{k=0}^{N}\frac{1}{k!}x^k - \sum_{k=0}^{n}\frac{1}{k!}x^k\prod_{\ell=0}^{k-1}\left(1-\frac{\ell}{n}\right)\right| \le \sum_{k=0}^{N}\frac{1}{k!}|x|^k\left(1-\prod_{\ell=0}^{k-1}\left(1-\frac{\ell}{n}\right)\right)+\varepsilon.$$

Wir fixieren nun $N$ und verwenden $\lim_{n\to\infty}\left(1-\prod_{\ell=0}^{k-1}\left(1-\frac{\ell}{n}\right)\right)=0$ für alle $k\in\{0,\dots,N\}$, woraus folgt, dass

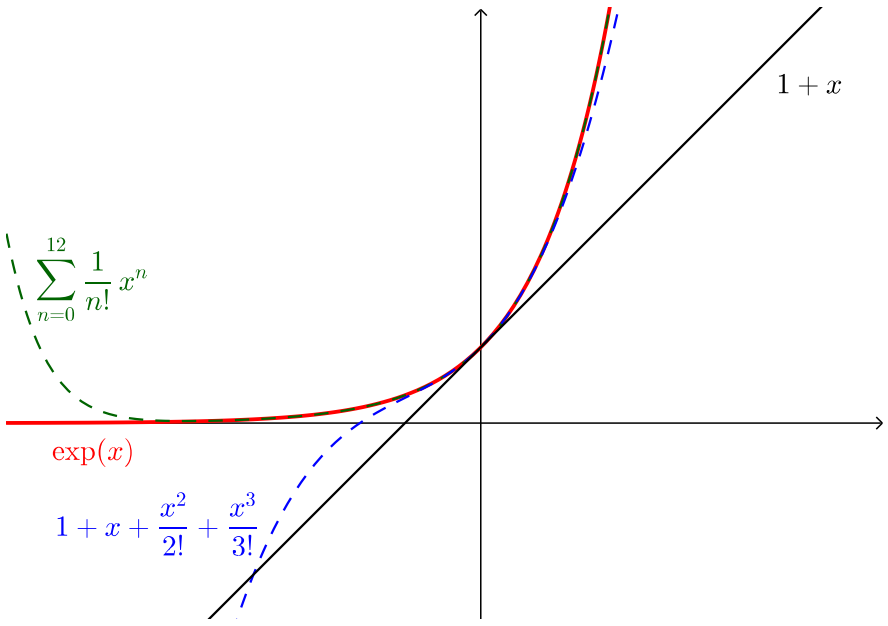
$$\left|\sum_{k=0}^{N}\frac{1}{k!}x^k-\exp(x)\right|\le\varepsilon.$$

Gemeinsam mit (6.8) erhalten wir

$$\left|\sum_{k=0}^{\infty}\frac{1}{k!}x^k-\exp(x)\right|\le 2\varepsilon.$$

Da $\varepsilon>0$ beliebig war, folgt (6.7).

Die Darstellung der Exponentialfunktion als Potenzreihe liefert durch Betrachten von nur endlich vielen Termen eine Approximation von exp durch Polynome.




### 6.5.2 Die komplexe Exponentialreihe

Wir verwenden nun die Potenzreihe aus (6.7), um die Exponentialfunktion $\exp:\mathbb{C}\to\mathbb{C}$ auf der komplexen Zahlenebene zu definieren.

**Satz 6.68** (Komplexe Exponentialabbildung)**.** *Für $z\in\mathbb{C}$ definieren wir*

$$\exp(z)=\sum_{k=0}^{\infty}\frac{1}{k!}z^k,$$

*womit eine stetige Erweiterung $\exp:\mathbb{C}\to\mathbb{C}$ der reellen Exponentialabbildung definiert wird. Des Weiteren gilt für alle $z,w\in\mathbb{C}$ die Additionsformel*

$$\exp(z+w)=\exp(z)\exp(w) \tag{6.9}$$

*und die Formel*

$$|\exp(z)| = \exp(\mathrm{Re}(z)) \tag{6.10}$$

*für den Absolutbetrag. Insbesondere gilt* $|\exp(\mathrm{i}y)| = 1$ *für alle* $y \in \mathbb{R}$.

Auf Grund der Diskussion in Abschnitt 6.5.1 ist der Konvergenzradius der Reihe $\sum_{k=0}^{\infty} \frac{1}{k!} z^k$ unendlich. Wiederum nach Satz 6.56 ist damit $\exp : \mathbb{C} \to \mathbb{C}$ eine stetige Funktion, welche wegen Abschnitt 6.5.1 die reelle Exponentialfunktion erweitert. (Wir verwenden zwar das gleiche Symbol exp für die reelle und komplexe Exponentialfunktion, doch müssen wir diese unterscheiden, wenn wir Eigenschaften von Funktionen wie zum Beispiel Injektivität besprechen wollen.)

Für eine positive Basis $a \in \mathbb{R}_{>0}$ und $z \in \mathbb{C}$ setzen wir des Weiteren

$$a^z = \exp(z \log(a)),$$

was wegen obigem mit der in Abschnitt 5.3.8 eingeführten Notation kompatibel ist. Insbesondere gilt $\mathrm{e}^z = \exp(z)$ für alle $z \in \mathbb{C}$.

**Übung 6.69** (Grenzwertformel)**.** *Zeigen Sie, dass*

$$\exp(z) = \lim_{n\to\infty} \left(1 + \frac{z}{n}\right)^n$$

*für alle* $z \in \mathbb{C}$.

### 6.5.3 Die Additionsformel

Wir wollen nun die Additionsformel (6.9) beweisen. In der Tat folgt für beliebige $z, w \in \mathbb{C}$ aus der Cauchy-Produktformel (Korollar 6.37), dass

$$\begin{aligned}
\exp(z)\exp(w) &= \left(\sum_{n=0}^{\infty} \frac{1}{n!} z^n\right)\left(\sum_{n=0}^{\infty} \frac{1}{n!} w^n\right) \\
&= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{1}{k!} \frac{1}{(n-k)!} z^k w^{n-k} \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \frac{n!}{k!\,(n-k)!} z^k w^{n-k} \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} z^k w^{n-k} \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} (z+w)^n = \exp(z+w).
\end{aligned}$$

### 6.5.4 Der Absolutbetrag der Exponentialabbildung

Es verbleibt für den Beweis von Satz 6.68 die Formel (6.10) für den Absolutbetrag zu beweisen. In der Tat gilt, da die Konjugation auf $\mathbb{C}$ stetig ist (wieso?), dass

$$\overline{\exp(z)} = \overline{\lim_{n\to\infty} \sum_{k=0}^{n} \frac{1}{k!} z^k} = \lim_{n\to\infty} \overline{\sum_{k=0}^{n} \frac{1}{k!} z^k} = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{1}{k!} \overline{z}^k = \exp(\overline{z}).$$

Insbesondere ist nach der Additionsformel

$$|\exp(z)|^2 = \exp(z)\overline{\exp(z)} = \exp(z)\exp(\overline{z}) = \exp(z+\overline{z}) = \exp(2\operatorname{Re}(z)) = \exp(\operatorname{Re}(z))^2$$

womit die Formel $|\exp(z)| = \exp(\operatorname{Re}(z))$ nach Wurzelziehen folgt.

**Übung 6.70.** *Zeigen Sie für alle $z \in \mathbb{C}$ mit $|z| < 1$ die Abschätzung $|\exp(z)| \leq \frac{1}{1-\operatorname{Re}(z)}$.*

**Applet 6.71** (Komplexe Exponentialabbildung)**.** *Wir stellen die komplexe Exponentialabbildung dar. Da der Graph dieser in $\mathbb{C}^2$ liegt, können wir den Graph wohl kaum auf dem Bildschirm darstellen. Stattdessen visualisieren wir die Abbildung anhand eines bewegbaren Punktes $z = s + \varphi\mathrm{i} \in \mathbb{C}$ und dessen Bildpunkt $\exp(z) = \mathrm{e}^s \exp(\varphi\mathrm{i}) \in \mathbb{C}$.*

VO: 27.11.17

## 6.6 Trigonometrische Funktionen

Wir definieren die **Sinusfunktion** bei $z \in \mathbb{C}$ durch

$$\sin(z) = z - \frac{z^3}{3!} + \frac{z^5}{5!} - + \ldots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1} \tag{6.11}$$

und die **Kosinusfunktion** bei $z \in \mathbb{C}$ durch

$$\cos(z) = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - + \ldots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n}. \tag{6.12}$$

Wir sagen, dass eine Funktion $f$ auf $\mathbb{C}$ **gerade** ist, wenn $f(-z) = f(z)$ für alle $z \in \mathbb{C}$ und **ungerade** ist, wenn $f(-z) = -f(z)$ für alle $z \in \mathbb{C}$. (Diese Begriffe werden analog für Funktionen mit Definitionsbereich $\mathbb{R}$ oder Intervallen der Form $[-a, a]$ für $a > 0$ verwendet.)

**Satz 6.72** (Sinus- und Kosinusfunktionen)**.** *Die Potenzreihe* (6.11) *definiert die ungerade stetige Sinusfunktion* $\sin : \mathbb{C} \to \mathbb{C}$ *und die Potenzreihe* (6.12) *definiert die gerade stetige Kosinusfunktion* $\cos : \mathbb{C} \to \mathbb{C}$*. Die Einschränkungen dieser Funktionen auf* $\mathbb{R}$ *sind reellwertig und werden ebenso also Sinusfunktion und Kosinusfunktion bezeichnet. Des Weiteren bestehen für alle* $z \in \mathbb{C}$ *die Beziehungen*

$$\begin{aligned} \exp(\mathrm{i}z) &= \cos(z) + \mathrm{i}\sin(z) \\ \sin(z) &= \frac{\mathrm{e}^{\mathrm{i}z} - \mathrm{e}^{-\mathrm{i}z}}{2\mathrm{i}} \\ \cos(z) &= \frac{\mathrm{e}^{\mathrm{i}z} + \mathrm{e}^{-\mathrm{i}z}}{2} \end{aligned}$$

*zu der Exponentialfunktion und es gelten die trigonometrischen Additionsformeln*

$$\sin(z+w) = \sin(z)\cos(w) + \cos(z)\sin(w) \tag{6.13}$$

$$\cos(z+w) = \cos(z)\cos(w) - \sin(z)\sin(w) \tag{6.14}$$

*für alle* $z, w \in \mathbb{C}$*.*

Eine kurze Rechnung (zum Beispiel wie schon in Abschnitt 6.5.1 unter Verwendung des Quotientenkriteriums) zeigt, dass die Reihen (6.11)–(6.12) für alle $z \in \mathbb{C}$ konvergieren. Nach Satz 6.56 haben diese Potenzreihen daher unendlichen Konvergenzradius, und Satz 6.56 besagt nun, dass sin und cos auf ganz $\mathbb{C}$ definiert und stetig sind. Für die Partialsummen $s_n(z)$ der Reihe in (6.11) gilt $s_n(-z) = -s_n(z)$ für alle $z \in \mathbb{C}$, und daher ist sin eine ungerade Funktion. Analog ergibt sich, dass cos eine gerade Funktion definiert. Da die Koeffizienten der Potenzreihen sin und cos in $\mathbb{R}$ liegen, gilt $\sin(\mathbb{R}) \subseteq \mathbb{R}$ und $\cos(\mathbb{R}) \subseteq \mathbb{R}$.

**Applet 6.73** (Potenzreihen)**.** *Wir betrachten die ersten Partialsummen der Potenzreihen, welche* exp*,* sin *und* cos *(beziehungsweise* sinh*,* cosh *vom nächsten Abschnitt) definieren. Durch Vergrössern des Ausschnittes können Sie die Qualität der Annäherungen der Partialsummen*

*überprüfen. Bei den trigonometrischen Funktionen kann man auch im Bild gut erkennen, dass die Potenzreihe alternierende Reihen bilden.*

Wir beweisen nun den Zusammenhang zur Exponentialabbildung. Für $z \in \mathbb{C}$ gilt

$$\mathrm{e}^{\mathrm{i}z} = 1 + \mathrm{i}z - \frac{z^2}{2!} - \mathrm{i}\frac{z^3}{3!} + \frac{z^4}{4!} + \mathrm{i}\frac{z^5}{5!} - \ldots = \cos(z) + \mathrm{i}\sin(z)$$

und analog

$$\mathrm{e}^{-\mathrm{i}z} = \cos(z) - \mathrm{i}\sin(z).$$

für alle $z \in \mathbb{C}$. Lösen wir diese beiden Gleichungen nach $\sin(z)$ und $\cos(z)$ auf, so ergeben sich die Formeln im Satz.

### 6.6.1 Additionsformeln

Wir wollen nun für alle $z, w \in \mathbb{C}$ die trigonometrischen Additionsformeln in Satz 6.72 beweisen. Hierzu multiplizieren wir $\mathrm{e}^{\mathrm{i}z}$ und $\mathrm{e}^{\mathrm{i}w}$ und erhalten auf Grund der Additionsformel (6.9) für die Exponentialabbildung

$$\begin{aligned}
\mathrm{e}^{\mathrm{i}(z+w)} &= \cos(z+w) + \mathrm{i}\sin(z+w) = \mathrm{e}^{\mathrm{i}z}\mathrm{e}^{\mathrm{i}w} \\
&= (\cos(z) + \mathrm{i}\sin(z))(\cos(w) + \mathrm{i}\sin(w)) \\
&= \cos(z)\cos(w) - \sin(z)\sin(w) \\
&\qquad + \mathrm{i}(\sin(z)\cos(w) + \sin(w)\cos(z))
\end{aligned}$$

und ebenso

$$\begin{aligned}
\mathrm{e}^{-\mathrm{i}(z+w)} &= \cos(z+w) - \mathrm{i}\sin(z+w) = \mathrm{e}^{-\mathrm{i}z}\mathrm{e}^{-\mathrm{i}w} \\
&= (\cos(z) - \mathrm{i}\sin(z))(\cos(w) - \mathrm{i}\sin(w)) \\
&= \cos(z)\cos(w) - \sin(z)\sin(w) \\
&\qquad - \mathrm{i}(\sin(z)\cos(w) + \sin(w)\cos(z)).
\end{aligned}$$

Auf Grund dieser beiden Gleichungen erhalten wir die Formeln (6.13)–(6.14). (Für $z, w \in \mathbb{R}$ würde die erste der beiden Gleichungen ausreichen indem wir Real- und Imaginärteile betrachten. Aber für $z, w \in \mathbb{C}$ erhalten wir zum Beispiel (6.14) durch Addition obiger Gleichungen.) Dies beendet den Beweis von Satz 6.72.

Für $z = w \in \mathbb{C}$ ergeben sich insbesondere die **Winkelverdoppelungsformeln**

$$\begin{aligned}
\sin(2z) &= 2\sin(z)\cos(z), \\
\cos(2z) &= \cos^2(z) - \sin^2(z),
\end{aligned}$$

wobei wir die Notation $\cos^2(z) = (\cos(z))^2$ und $\sin^2(z) = (\sin(z))^2$ für $z \in \mathbb{C}$ verwendet haben und auch für andere Funktionen und andere Potenzen des Öfteren verwenden werden. Des Weiteren folgt für $w = -z$ aus (6.14) und der Tatsache, dass der Sinus ungerade und der

Kosinus gerade ist, die **Kreisgleichung** für Sinus und Kosinus

$$1 = \cos^2(z) + \sin^2(z)$$

für alle $z \in \mathbb{C}$.

### 6.6.2 Die Kreiszahl

Wir können nun die Kreiszahl $\pi$ definieren.

**Satz 6.74** (Pi)**.** *Es gibt genau eine Zahl $\pi \in (0,4)$ mit $\sin(\pi) = 0$. Für diese Zahl gilt weiters*

$$\begin{aligned}
\mathrm{e}^{2\pi\mathrm{i}} &= \cos(2\pi) + \mathrm{i}\sin(2\pi) = 1,\\
\mathrm{e}^{\pi\mathrm{i}} &= \cos(\pi) = -1,\\
\mathrm{e}^{\frac{1}{2}\pi\mathrm{i}} &= \cos\left(\frac{\pi}{2}\right) + \mathrm{i}\sin\left(\frac{\pi}{2}\right) = \mathrm{i},\\
\mathrm{e}^{\frac{3}{2}\pi\mathrm{i}} &= \cos\left(\frac{3\pi}{2}\right) + \mathrm{i}\sin\left(\frac{3\pi}{2}\right) = -\mathrm{i},\\
\mathrm{e}^{\frac{1}{4}\pi\mathrm{i}} &= \cos\left(\frac{\pi}{4}\right) + \mathrm{i}\sin\left(\frac{\pi}{4}\right) = \frac{1}{\sqrt{2}} + \frac{\mathrm{i}}{\sqrt{2}}.
\end{aligned}$$

*Beweis.* Aus dem Leibniz-Kriterium (Proposition 6.25) und der Monotonie der Folge $(\frac{x^n}{n!})_n$ für alle $x \in (0,1]$ folgt, dass

$$\sin(x) \geq x - \tfrac{x^3}{3!} \geq 0$$

und

$$1 > 1 - \tfrac{x^2}{2} + \tfrac{x^4}{24} \geq \cos(x) \geq 1 - \tfrac{x^2}{2} > 0$$

für alle $x \in (0,1]$. Für $x = 0$ ist $\sin(0) = 0$ und für $x = 1$ ergibt sich

$$\sin^2(1) \geq (1 - \tfrac{1}{6})^2 = \tfrac{25}{36} > \frac{1}{2}$$

und wegen Positivität von $\sin(1)$ damit $\sin(1) > \frac{1}{\sqrt{2}}$. Daher existiert nach dem Zwischenwertsatz (Satz 3.59) eine Zahl $p \in (0,1)$ mit $\sin(p) = \frac{1}{\sqrt{2}}$. Wegen $\sin^2(p) + \cos^2(p) = 1$ und $\cos(p) > 0$ folgt ebenso $\cos(p) = \frac{1}{\sqrt{2}}$. Wir definieren $\pi = 4p$. Die weiteren Formeln im Satz folgen aus den Additionsformeln (und den Winkelverdopplungsformeln).

Es verbleibt die Eindeutigkeit von $\pi$ wie im Satz zu zeigen. Für dies zeigen wir zuerst, dass es keine Nullstelle $r \in (0,2]$ von der Sinusfunktion geben kann. Wir nehmen indirekt an, dass $r \in (0,2]$ die Gleichung $\sin(r) = 0$ erfüllt. Dann gilt $\cos(r) = \mathrm{e}^{r\mathrm{i}} \in \{1,-1\}$ und somit $\mathrm{e}^{\frac{r}{2}\mathrm{i}} \in \{1,-1,\mathrm{i},-\mathrm{i}\}$. Da aber $\frac{r}{2}$ in $(0,1]$ liegt, verfügen wir bereits über die Ungleichungen $0 < \cos(\frac{r}{2}) < 1$, was einen Widerspruch darstellt. Insbesondere gilt daher $\pi \in (2,4)$.

Angenommen $s \in (0,4) \setminus \{\pi\}$ erfüllt ebenso $\sin(s) = 0$. Dann gilt nach obigem, dass $s \in (2,4)$. Wir definieren

$$r = \begin{cases} \pi - s & \text{falls } 2 < s < \pi \\ s - \pi & \text{falls } 2 < \pi < s. \end{cases}$$

Dann gilt $r \in (0,2]$ und $\sin(r) = 0$ (wieso?), was aber obigem Argument widerspricht. Daher ist $\pi \in (0,4)$ durch die Gleichung $\sin(\pi) = 0$ eindeutig bestimmt. □

Aus den Additionsformeln erhält man folgendes Korollar.

**Korollar 6.75** (Periodizität)**.** *Es gelten*

$$\begin{aligned} \sin(z+\pi) &= -\sin(z), \\ \cos(z+\pi) &= -\cos(z), \\ \sin(z+2\pi) &= \sin(z), \\ \cos(z+2\pi) &= \cos(z), \end{aligned}$$

*und* $\cos(z) = \sin\left(z + \frac{\pi}{2}\right)$ *für alle* $z \in \mathbb{C}$.

Insbesondere sind die Werte des Sinus und des Kosinus auf $\mathbb{R}$ durch die Werte auf dem Intervall $[0, \frac{\pi}{2}]$ eindeutig bestimmt (wieso?). Dieses Prinzip überträgt sich unter anderem auf die Nullstellen von Sinus und Kosinus.

**Wichtige Übung 6.76** (Nullstellen von sin und cos)**.** *Zeigen Sie, dass die Nullstellen von* sin *genau die Punkte in* $\pi\mathbb{Z} \subseteq \mathbb{C}$ *sind und dass die Nullstellen von* cos *genau die Punkte in* $\pi\mathbb{Z} + \frac{\pi}{2}$ *sind.*

**Übung 6.77** (Monotonie von Sinus)**.** *Zeigen Sie die Formel*

$$\sin(z) - \sin(w) = 2\cos\left(\frac{z+w}{2}\right)\sin\left(\frac{z-w}{2}\right)$$

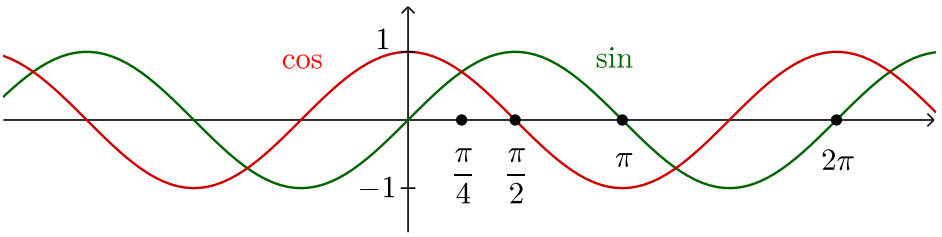
*für alle* $z, w \in \mathbb{C}$. *Verwenden Sie dies, um zu zeigen, dass der eingeschränkte Sinus*

$$\sin : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \to [-1, 1]$$

*strikt monoton wachsend und bijektiv ist.*

**Übung 6.78.** *Zeigen Sie, dass die Zahl* $\cos\left(\frac{\pi}{7}\right)$ *algebraisch ist.*

**Übung 6.79** (Abschätzungen für $\pi$)**.** *Zeigen Sie* $\pi \in (3,4)$ *oder sogar* $3.1 < \pi < 3.2$ *(wobei die Verwendung eines elektronischen Hilfsmittels zur Berechnung von gewissen rationalen Zahlen hilfreich sein könnte).*




### 6.6.3 Tangens und Cotangens

Wir definieren die **Tangensfunktion** tan durch

$$\tan(z) = \frac{\sin(z)}{\cos(z)}$$

für alle $z \in \mathbb{C}$ mit $\cos(z) \neq 0$, was nach Übung 6.76 gerade alle $z \in \mathbb{C} \setminus (\pi\mathbb{Z} + \frac{\pi}{2})$ sind. Analog ist die **Cotangensfunktion** cot durch

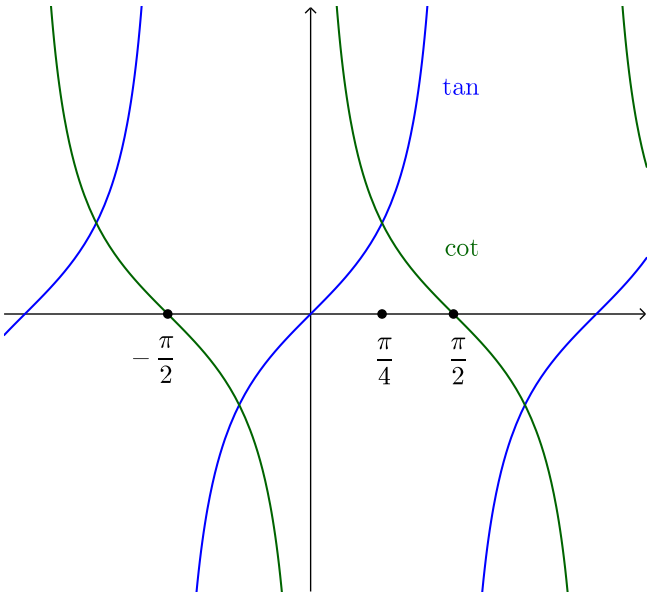
$$\cot(z) = \frac{\cos(z)}{\sin(z)}$$

für alle $z \in \mathbb{C}$ mit $\sin(z) \neq 0$ (oder äquivalent alle $z \in \mathbb{C} \setminus \pi\mathbb{Z}$ nach Übung 6.76) definiert.

**Übung 6.80** (Additionsformel für den (Co-)Tangens)**.** *Zeigen Sie, dass für $z, w \in \mathbb{C}$ die Additionsformel*

$$\tan(z + w) = \frac{\tan(z) + \tan(w)}{1 - \tan(z)\tan(w)}$$

*gilt, wo definiert. Finden und beweisen Sie eine analoge Additionsformel für den Cotangens.*



### 6.6.4 Polarkoordinaten und Multiplikation auf den komplexen Zahlen

Die Beschreibung eines Punktes $z \in \mathbb{C}$ durch Polarkoordinaten besteht aus einem Radius $r \geq 0$ und einem „Winkel“ $\theta \in \mathbb{R}$ mit

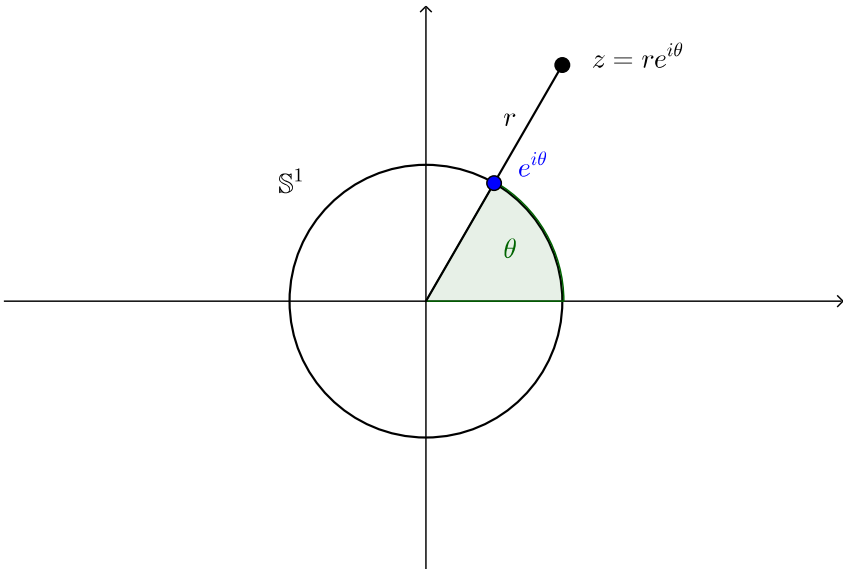
$$z = r\mathrm{e}^{\mathrm{i}\theta} = r(\cos(\theta) + \mathrm{i}\sin(\theta)).$$

Wir bezeichnen hier den reellen Parameter $\theta$ als Winkel, obwohl wir diesem noch keine geometrische Bedeutung formal zuweisen. Wir werden dies korrigieren, sobald wir die Definition der Bogenlänge einer Kurve kennen.

Ist eine derartige Darstellung von $z \in \mathbb{C}$ gegeben, so gilt $r = |r| = |z|$ womit $r$ den Abstand von $z$ zum Ursprung darstellt. Die Menge der Elemente mit Absolutbetrag Eins

$$\mathbb{S}^1 = \{z \in \mathbb{C} \mid |z| = 1\} = \left\{\mathrm{e}^{\mathrm{i}\theta} \mid \theta \in \mathbb{R}\right\}$$

wird als der **Einheitskreis** in $\mathbb{C}$ bezeichnet (wobei $\mathbb{S}^1$ auch für die 1-dimensionale Sphäre steht).




**Lemma 6.81** (Existenz von Polarkoordinaten)**.** *Für alle $z \in \mathbb{C}$ existiert ein eindeutig bestimmtes $r \geq 0$ und ein Winkel $\theta \in [0, 2\pi)$ mit $z = r\mathrm{e}^{\mathrm{i}\theta}$. Des Weiteren ist der Winkel $\theta$ eindeutig bestimmt, falls $z \neq 0$.*

*Beweis.* Für $z = 0$ gibt es nichts zu zeigen. Also nehmen wir $z \neq 0$ an. Nach Division mit $|z|$ können wir des Weiteren $r = |z| = 1$ annehmen. Wir betrachten zuerst den Fall $\mathrm{Im}(z) \geq 0$ und behaupten, dass ein $\theta \in [0, \pi]$ existiert, so dass $z = \mathrm{e}^{\mathrm{i}\theta}$. In der Tat ist $z = x + y\mathrm{i}$ für ein $x \in [-1, 1]$ und es gilt $\cos(0) = 1$ und nach Satz 6.74 auch $\cos(\pi) = -1$. Insbesondere existiert nach dem Zwischenwertsatz ein $\theta \in [0, \pi]$ mit $\mathrm{Re}(z) = x = \cos(\theta)$. Nun gilt aber auch

$$\sin(\theta) = \sqrt{1 - \cos^2(\theta)} = \sqrt{1 - \mathrm{Re}(z)^2} = y = \mathrm{Im}(z)$$

da wir $y = \mathrm{Im}(z) \geq 0$ angenommen haben, und somit

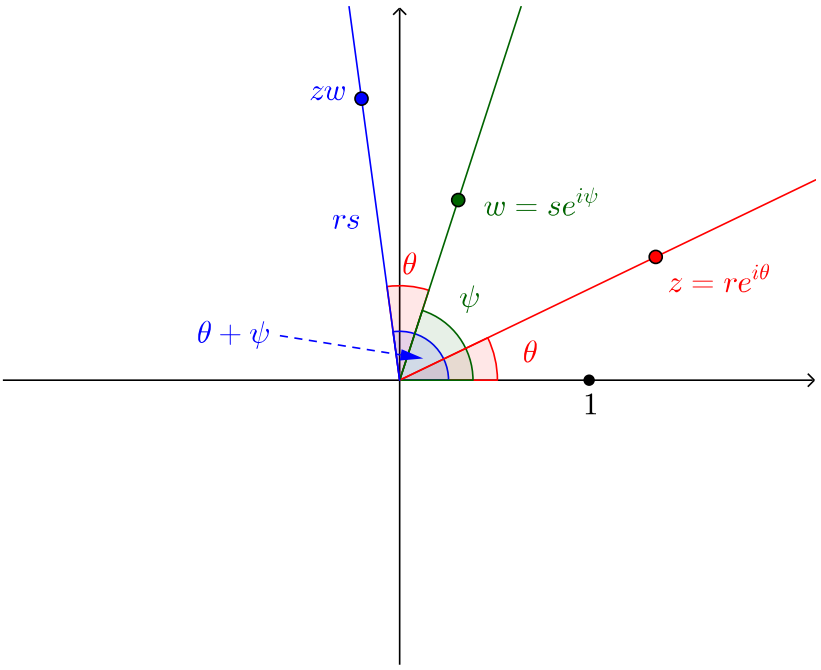
$$z = \mathrm{Re}(z) + \mathrm{Im}(z)\mathrm{i} = \cos(\theta) + \sin(\theta)\mathrm{i} = \mathrm{e}^{\mathrm{i}\theta}.$$

Falls $\mathrm{Im}(z) < 0$ ist, dann wenden wir obiges Argument für $-z$ an und finden ein $\theta \in (0, \pi)$ mit $-z = \mathrm{e}^{\mathrm{i}\theta}$, womit $z = -\mathrm{e}^{\mathrm{i}\theta} = \mathrm{e}^{\pi\mathrm{i}}\mathrm{e}^{\mathrm{i}\theta} = \mathrm{e}^{\mathrm{i}(\pi+\theta)}$ und $\pi + \theta \in (\pi, 2\pi)$.

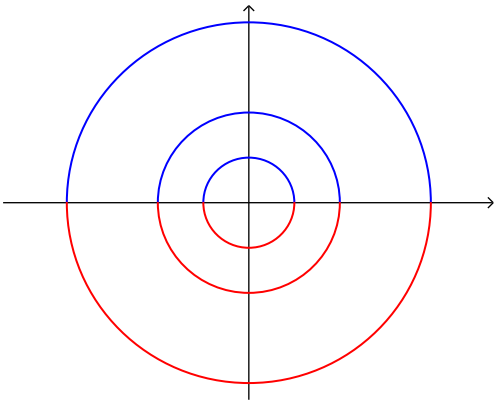
Sind $\theta, \theta' \in [0, 2\pi)$ mit $z = \mathrm{e}^{\mathrm{i}\theta} = \mathrm{e}^{\mathrm{i}\theta'}$, dann gilt $\mathrm{e}^{\mathrm{i}(\theta-\theta')} = 1$ und daher auch $\sin(\theta - \theta') = 0$. Aus der Eindeutigkeit von $\pi$ in Satz 6.74 und der Formel $\sin(x + \pi) = -\sin(x)$ für alle $x \in \mathbb{R}$ in Korollar 6.75 ergibt sich daraus, dass $\theta - \theta' \in \{-\pi, 0, \pi\}$. Falls $\theta - \theta' \in \{-\pi, \pi\}$, dann ist aber $\mathrm{e}^{\mathrm{i}(\theta-\theta')} = -1$, und daher muss $\theta = \theta'$ gelten. □

In Polarkoordinaten lässt sich die Multiplikation auf $\mathbb{C}$ neu interpretieren. Sind $z = r\mathrm{e}^{\mathrm{i}\theta}$ und $w = s\mathrm{e}^{\mathrm{i}\psi}$ in $\mathbb{C}$, dann ist $zw = rs\mathrm{e}^{\mathrm{i}(\theta+\psi)}$. Multiplikation mit $z$ streckt also um den Faktor $r$ und rotiert um den Winkel $\theta$, oder anders formuliert bei Multiplikation der Vektoren $z, w \in \mathbb{C}$ multiplizieren sich die Längen der Vektoren und addieren sich die Winkel. Diese geometrische Erklärung der Multiplikation hat in der Geschichte der komplexen Zahlen die komplexen Zahlen ausgehend von einem etwas mysteriösen rein algebraischem Objekt zu einem eigenständigen Zahlenbegriff verwandelt. Daraufhin hat dieser Zahlenbegriff schnell Einzug in die weitere Entwicklung der Mathematik gefunden. Wir verweisen hierfür auf einen weiteren BBC Podcast.

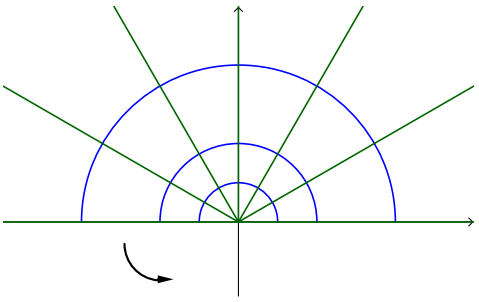
**Applet 6.82** (Geometrische Bedeutung der Komplexen Zahlen)**.** *Wir können anhand der eingezeichneten Polarkoordinatenlinien die geometrische Bedeutung der Multiplikation von komplexen Zahlen und der Inversen und der Wurzeln einer vorgegebenen Zahl erkennen. Bei Bewegung von $z$ um den Ursprung im Wurzelmodus ist ersichtlich, warum eine stetige Definition von Wurzelfunktionen nicht möglich ist.*

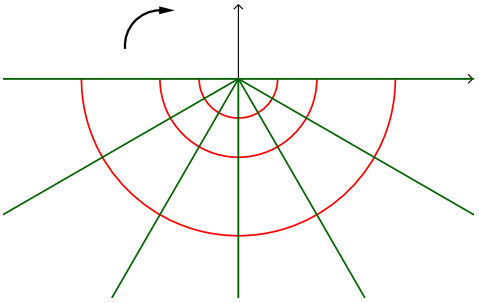


**Darstellung der Abbildung $z \in \mathbb{C} \mapsto z^2 \in \mathbb{C}$.** Wir möchten illustrieren, wie sich die Abbildung $z \in \mathbb{C} \mapsto z^2 \in \mathbb{C}$ auf Kreisen in der komplexen Ebene verhält. Dazu teilen wir die komplexe Ebene in die obere und die untere Halbebene auf und färben die Kreise (hier für Radien in $\left\{\frac{1}{2}, 1, 2\right\}$) entsprechend ein.

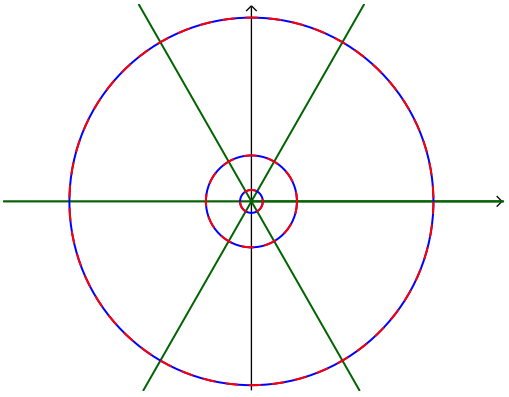


Wir betrachten die beiden Hälften nun gesondert. Zusätzlich zu den bestehenden Halbkreisen zeichnen wir von Null ausgehende Strahlen zu verschiedenen Winkeln ein.





Die Abbildung $z \mapsto z^2$ öffnet nun obige Bilder, die an Fächern erinnern, in die eingezeichnete Richtung auf. Genauer werden die Radien quadriert (und liegen somit in $\left\{\frac{1}{4}, 1, 4\right\}$) und die Winkel in die eingezeichnete Richtung verdoppelt. Das erhaltene Bild ist in beiden Fällen dasselbe; das Bild der gesamten Abbildung stellt sich also wie folgt dar. Insbesondere hat also jeder Punkt in $\mathbb{C}^\times$ genau zwei Quadratwurzeln.



Abgesehen von dieser geometrischen Interpretation der Multiplikation können Polarkoordinaten verwendet werden, um Nullstellen von Polynomen zu bestimmen. Dies trifft in einem

gewissen Sinn auch für allgemeine Polynome im Beweis des Fundamentalsatzes der Algebra zu, doch wollen wir hier zuerst nur folgende Spezialfälle betrachten.

**Übung 6.83** (Wurzeln)**.** *Sei $w = re^{i\theta}$. Zeigen Sie, dass die Nullstellen des Polynoms $p(z) = z^n - w$ (die $n$-ten Wurzeln von $w$) gerade durch*

$$\sqrt[n]{r}e^{i\frac{\theta}{n}},\ \ \sqrt[n]{r}e^{i\frac{\theta}{n}+i\frac{2\pi}{n}},\ \ldots,\ \sqrt[n]{r}e^{i\frac{\theta}{n}+i\frac{2\pi(n-1)}{n}}$$

*gegeben sind und dass diese paarweise verschieden sind. Die $n$-ten Wurzeln von 1 nennen sich die **$n$-ten Einheitswurzeln**.*

*Zeigen Sie des Weiteren für alle natürlichen Zahlen $n \geq 2$ die Identität*

$$\sum_{k=0}^{n-1} e^{i\frac{2\pi k}{n}} = 0.$$

**Übung 6.84** (Quadratische Formel)**.** *Verallgemeinern Sie Übung 3.18 und zeigen Sie, dass jedes komplexe Polynom vom Grad 2 als Produkt zweier komplexen Polynomen mit Grad 1 geschrieben werden kann.*

### 6.6.5 Zwei Logarithmen auf der komplexen Zahlenebene

Wir erinnern daran, dass wir den reellen Logarithmus als die inverse Abbildung von der bijektiven Abbildung $\exp : \mathbb{R} \to \mathbb{R}_{>0}$ definierten. Wir wollen dies nun wiederholen und den Logarithmus für komplexe Zahlen definieren. Leider gibt es hier aber ein fundamentales Problem: die Exponentialabbildung ist auf der komplexen Zahlenebene ganz und gar nicht bijektiv, da beispielsweise $\exp(2n\pi i) = 1$ für alle $n \in \mathbb{Z}$.

Aus diesem Grund müssen wir die Exponentialabbildung auf eine geeignete Teilmenge $D$ von $\mathbb{C}$ einschränken, so dass die eingeschränkte Abbildung $\exp|_D : D \to \mathbb{C}^\times$ bijektiv ist. Dies kann durch viele verschiedene Teilmengen erreicht werden. Wir wollen hier zwei Möglichkeiten ansprechen.

Zum Beispiel könnten wir die Teilmenge

$$D_{\text{Polar}} = \mathbb{R} \times [0, 2\pi) = \{x + yi \mid x \in \mathbb{R},\ y \in [0, 2\pi)\}$$

betrachten. In diesem Fall entsprechen $x \in \mathbb{R}$ und $y \in [0, 2\pi)$ mit $\exp(x + yi) = z \in \mathbb{C}^\times$ der Polarkoordinatendarstellung von $z = re^{\theta i} \in \mathbb{C}^\times$ in Lemma 6.81 mit $r = \exp(x)$ und $\theta = y$. Da die Polarkoordinatendarstellung von jedem $z \in \mathbb{C}^\times$ eindeutig bestimmt ist, ist die Einschränkung $\exp|_{D_{\text{Polar}}} : D_{\text{Polar}} \to \mathbb{C}^\times$ bijektiv. Der Nachteil dieser Abbildung ist, dass die entsprechende Umkehrabbildung bei jeder positiven Zahl in $(0, \infty)$ unstetig ist (wieso?).

Die Unstetigkeit des komplexen Logarithmus lässt sich zwar nicht vermeiden (zumindest dann nicht, wenn man Bijektivität der Logarithmusabbildung bewahren will), doch können wir diese mit der anderen Wahl $D$ "etwas besser verstecken". Wir definieren

$$D_{\text{Haupt}} = \mathbb{R} \times (-\pi, \pi] = \{x + yi \mid x \in \mathbb{R},\ y \in (-\pi, \pi]\}.$$

Man sieht auf die gleiche Weise wie oben, dass

$$\exp|_{D_{\text{Haupt}}} : D_{\text{Haupt}} \to \mathbb{C}^\times$$

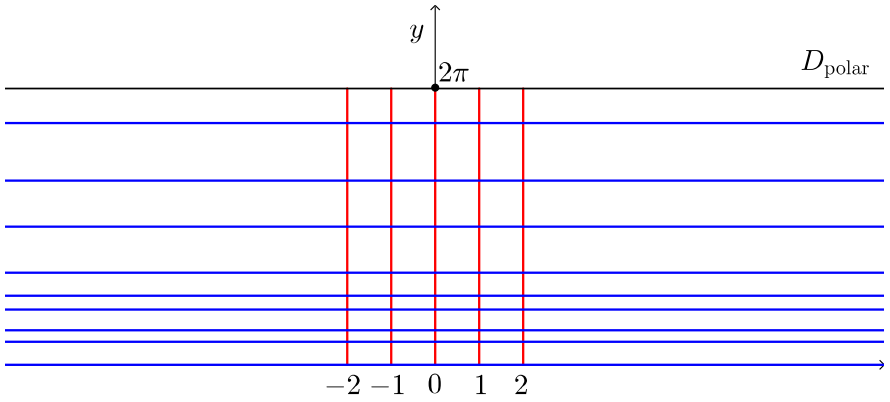
bijektiv ist. In der Tat entspricht dies leicht veränderten Polarkoordinaten, wo wir einen Winkel $\theta \in [0, \pi]$ unverändert lassen aber einen Winkel $\theta \in (\pi, 2\pi)$ durch $\theta - 2\pi \in (-\pi, 0)$ ersetzen (welcher wegen $\mathrm{e}^{(\theta-2\pi)\mathrm{i}} = \mathrm{e}^{\theta\mathrm{i}}$ den gleichen Zweck wie $\theta$ erfüllt). Die entsprechende Umkehrabbildung wird der Hauptzweig des Logarithmus

$$\log : \mathbb{C}^\times \to D_{\text{Haupt}} = \{x + y\mathrm{i} \mid x \in \mathbb{R},\ y \in (-\pi, \pi]\}$$

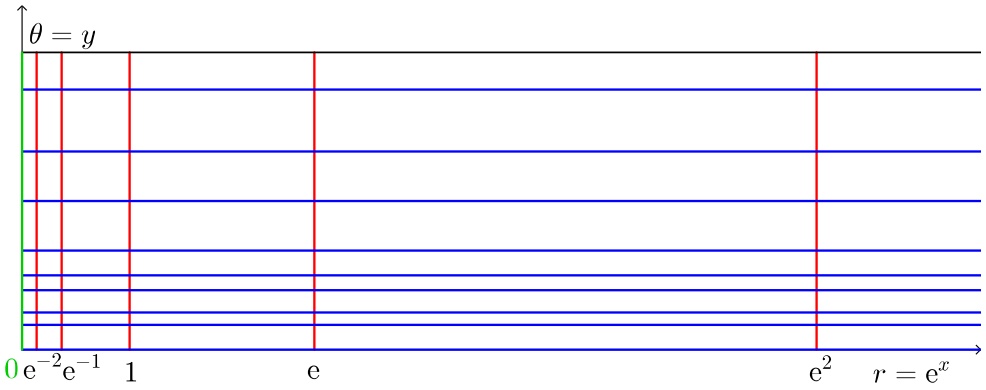
genannt. Er ist auf der negativen Halbachse $\mathbb{R}_{<0}$ unstetig, aber auf $\mathbb{C}^\times \setminus \mathbb{R}_{<0}$ stetig. (Diese Stetigkeitsbehauptungen kann man auch jetzt schon zeigen, werden aber klarer sein, wenn wir die trigonometrischen Umkehrfunktionen besprochen haben.)

Wir bemerken noch, dass wir keine Potenzen $z^w$ mit komplexer (oder auch nur negativer) Basis $z$ und komplexen (oder auch nur reellen) Exponenten $w$ definieren. Der Grund dafür ist, dass diese Definition von einer Wahl eines geeigneten Logarithmus abhängen würde und aus diesem Grunde nicht natürlich wäre.

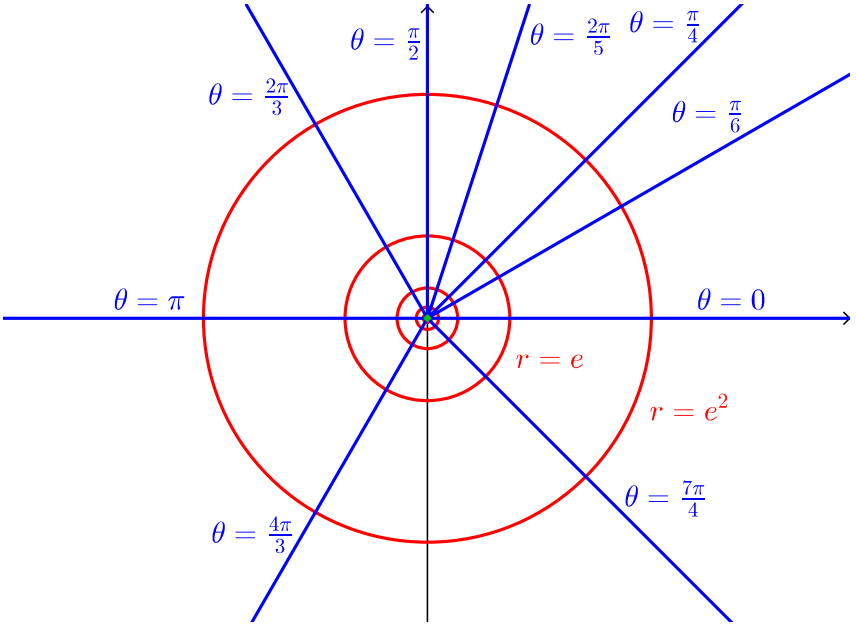
**Die Abbildung $\exp|_{D_{\text{Polar}}}$ in Bildern** Wir stellen zuerst die Teilmenge $D_{\text{polar}}$ dar.




In obigem Bild sind die blauen Linien bei $y \in \left\{0, \frac{\pi}{6}, \frac{\pi}{4}, \frac{2\pi}{5}, \frac{\pi}{2}, \frac{2\pi}{3}, \pi, \frac{4\pi}{3}, \frac{7\pi}{4}\right\}$ eingezeichnet. Wir wenden nun auf der $x$-Koordinate die Exponentialabbildung an und erhalten folgendes Bild in den Koordinaten $r = \mathrm{e}^x,\ \theta = y$.



Man beachte, dass die grüne Linie von der Abbildung nicht getroffen wird. Jetzt weisen wir jedem Punkt mit Koordinaten $(r, \theta)$ die komplexe Zahl zu, die diese Polarkoordinaten hat. Die blaue Linie beim Wert $\theta$ wird damit auf den von 0 ausgehenden Strahl mit Winkel $\theta$ abgebildet. Die rote Linie beim Wert $r$ wird auf den Kreis mit Radius $r$ abgebildet. Die grüne Linie wird auf den Punkt 0 kollabiert.



## 6.7 Integration von Potenzreihen

Wir haben in den Abschnitten 6.5 und 6.6 einige sehr wichtige, durch Potenzreihen definierte Funktionen gesehen. Daher drängt sich die Frage auf, ob wir das Riemann-Integral über diese Funktionen berechnen können. Die positive Antwort ist eine direkte Anwendung der Resultate in Abschnitt 6.3.

**Satz 6.85** (Integration von Potenzreihen)**.** *Sei $f(x) = \sum_{n=0}^{\infty} c_n x^n$ eine Potenzreihe mit komplexen Koeffizienten $(c_n)_{n\in\mathbb{N}_0}$ und Konvergenzradius $R$. Dann definiert $F(x) = \sum_{n=0}^{\infty} \frac{c_n}{n+1} x^{n+1}$ eine Potenzreihe mit demselben Konvergenzradius $R$ und es gilt*

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a)$$

*für alle $a, b \in (-R, R)$.*

Man spricht in diesem Zusammenhang auch von gliedweiser Integration der Potenzreihe, da wir einfach jeden Summanden getrennt integrieren und damit das richtige Ergebnis erhalten.

*Beweis.* Wir beweisen zuerst die Behauptung über den Konvergenzradius $R$. Sei $S$ der Konvergenzradius der Reihe $\sum_{\ell=1}^{\infty} \frac{c_{\ell-1}}{\ell} x^\ell = \sum_{n=0}^{\infty} \frac{c_n}{n+1} x^{n+1}$. Falls $\sum_{n=0}^{\infty} c_n x^n$ absolut konvergiert, dann konvergiert $\sum_{n=0}^{\infty} \frac{c_n}{n+1} x^{n+1}$ ebenfalls absolut, da

$$\left|\frac{c_n}{n+1} x^{n+1}\right| \leq |x||c_n x^n|$$

für alle $n \in \mathbb{N}_0$. Dies beweist $S \geq R$ auf Grund der Eigenschaften des Konvergenzradiuses in Satz 6.56.

Falls umgekehrt $\sum_{n=0}^{\infty} \frac{c_n}{n+1} x^{n+1}$ konvergiert, dann gibt es ein $M > 0$ mit $\left|\frac{c_n}{n+1} x^n\right| \leq M$ für alle $n \in \mathbb{N}_0$. Für $y \in (-x, x)$ ist dann

$$|c_n y^n| = \left|\frac{c_n}{n+1} x^n\right| (n+1) \left|\frac{y}{x}\right|^n \leq M(n+1) \left|\frac{y}{x}\right|^n$$

und daher

$$\limsup_{n\to\infty} \sqrt[n]{|c_n| y^n} \leq \lim_{n\to\infty} \sqrt[n]{M(n+1) \left|\frac{y}{x}\right|^n} = \left|\frac{y}{x}\right| < 1.$$

Aus dem Wurzelkriterium (Korollar 6.30) folgt daher die Konvergenz von $\sum_{n=0}^{\infty} c_n y^n$. Dies beweist die umgekehrte Ungleichung $S \leq R$ (siehe wiederum Satz 6.56). Daher gilt $S = R$ und der erste Teil des Satzes ist bewiesen.

Seien nun $a, b \in (-R, R)$ mit $a < b$. Wir betrachten die Polynome $f_N(x) = \sum_{n=0}^{N} c_n x^n$ und verwenden Satz 4.37, um

$$\int_a^b f_N(x)\,\mathrm{d}x = \sum_{n=0}^{N} \frac{c_n}{n+1} b^{n+1} - \sum_{n=0}^{N} \frac{c_n}{n+1} a^{n+1}$$

zu erhalten. Auf Grund von Satz 6.56 konvergiert die Funktionenfolge $f_N(x)$ der Partialsummen der Potenzreihe gleichmässig gegen $f(x)$ auf $[a,b]$. Lässt man nun $N$ gegen Unendlich gehen, ergibt sich unter Verwendung von Satz 6.49, dass

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{N\to\infty}\int_a^b f_N(x)\,\mathrm{d}x = F(b) - F(a). \qquad \square$$

Wenden wir obigen Satz auf die uns mittlerweile wohlbekannten Potenzreihen an, erhalten wir folgende Formeln.

**Korollar 6.86.** *Für alle $a, b \in \mathbb{R}$ gelten die Integrationsformeln*

$$\begin{aligned}
\int_a^b \exp(x)\,\mathrm{d}x &= [\exp(x)]_a^b = \exp(b) - \exp(a)\\
\int_a^b \sin(x)\,\mathrm{d}x &= [-\cos(x)]_a^b = -\cos(b) + \cos(a)\\
\int_a^b \cos(x)\,\mathrm{d}x &= [\sin(x)]_a^b = \sin(b) - \sin(a).
\end{aligned}$$

*Beweis.* Wir verwenden Satz 6.85 und erhalten zum Beispiel

$$\begin{aligned}
\int_a^b \exp(x)\,\mathrm{d}x &= \sum_{n=0}^{\infty} \frac{\frac{1}{n!}}{n+1} b^{n+1} - \sum_{n=0}^{\infty} \frac{\frac{1}{n!}}{n+1} a^{n+1}\\
&= \sum_{n=0}^{\infty} \frac{1}{(n+1)!} b^{n+1} - \sum_{n=0}^{\infty} \frac{1}{(n+1)!} a^{n+1} = \exp(b) - \exp(a).
\end{aligned}$$

Die Beweise der anderen beiden Formeln sind ähnlich (siehe Übung 6.87). $\square$

**Übung 6.87.** *Beweisen Sie die letzten beiden Formeln in Korollar 6.86.*

**Applet 6.88** (Integration von transzendenten Funktionen)**.** *Wir stellen obige Integrationsgesetze (und jene vom nächsten Unterabschnitt) grafisch dar.*

### 6.7.1 Die hyperbolischen Funktionen

In Analogie zu den Winkelfunktionen sind manchmal auch folgende Funktionen, die in engem Zusammenhang mit der Hyperbel $\{(x,y) \in \mathbb{R}^2 \mid x^2 - y^2 = 1\}$ stehen, nützlich. Wir definieren den **Sinus Hyperbolicus** durch

$$\sinh(z) = \frac{e^z - e^{-z}}{2} = z + \frac{z^3}{3!} + \frac{z^5}{5!} + \ldots = \sum_{k=0}^{\infty} \frac{1}{(2k+1)!} z^{2k+1}$$

für alle $z \in \mathbb{C}$, den **Kosinus Hyperbolicus** durch

$$\cosh(z) = \frac{e^z + e^{-z}}{2} = 1 + \frac{z^2}{2!} + \frac{z^4}{4!} + \ldots = \sum_{k=0}^{\infty} \frac{1}{(2k)!} z^{2k}$$

für alle $z \in \mathbb{C}$ und den **Tangens Hyperbolicus** durch

$$\tanh(z) = \frac{\sinh(z)}{\cosh(z)} = \frac{e^z - e^{-z}}{e^z + e^{-z}}$$

für alle $z \in \mathbb{C}$.

Die Funktionen sinh und tanh sind ungerade und cosh ist gerade. Es gelten die Additionsformeln

$$\begin{aligned}\sinh(z+w) &= \sinh(z)\cosh(w) + \cosh(z)\sinh(w),\\ \cosh(z+w) &= \cosh(z)\cosh(w) + \sinh(z)\sinh(w)\end{aligned}$$

für alle $z, w \in \mathbb{C}$ und weiters

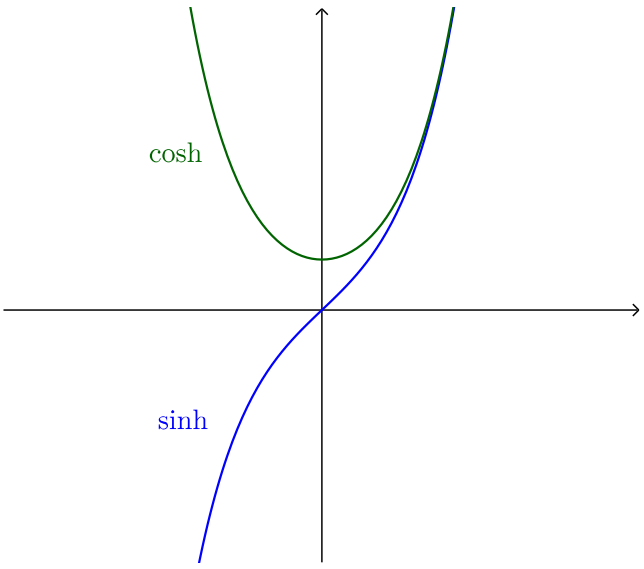
$$\begin{aligned}\cosh^2(z) - \sinh^2(z) &= 1,\\ \cosh(z) + \sinh(z) &= e^z\end{aligned}$$

für alle $z \in \mathbb{C}$.

Auf Grund der Definitionen von Sinus und Kosinus Hyperbolicus und Korollar 6.86 ergibt sich nun

$$\begin{aligned}\int_a^b \sinh(x)\,\mathrm{d}x &= [\cosh(x)]_a^b = \cosh(b) - \cosh(a)\\ \int_a^b \cosh(x)\,\mathrm{d}x &= [\sinh(x)]_a^b = \sinh(b) - \sinh(a).\end{aligned}$$

**Übung 6.89.** *Beweisen Sie ausgehend von den Definitionen des hyperbolischen Sinus und des hyperbolischen Kosinus die obigen Formeln.*




VO: 29.11.17, Abschnitt 6.8 wird nicht behandelt

## 6.8 Ziffernentwicklungen und fraktale Konstruktionen*

Für jede natürliche Zahl $p \geq 2$ und jede reelle positive Zahl $x > 0$ gibt es eine **Ziffernentwicklung** zur Basis $p$. Genauer formuliert existiert ein $\ell \in \mathbb{Z}$ mit $x < p^{-\ell+1}$ und eine Folge $(d_n)_{n\geq\ell}$ mit $d_n \in \{0, 1, \ldots, p-1\}$, so dass

$$x = \sum_{n=\ell}^{\infty} d_n p^{-n}. \tag{6.15}$$

Um dies zu sehen, nehmen wir, da die Ziffernentwicklung auf $\mathbb{N}_0$ ja bereits bekannt ist (siehe Übung 3.7), ohne Beschränkung der Allgemeinheit an, dass $\ell = 1$ und also $x = x_0 \in (0, 1)$. Wir können dann die rekursiv definierten Zahlen

$$\begin{aligned} d_1 &= \lfloor px_0 \rfloor \in \{0, 1, 2, \ldots, p-1\}, \qquad & x_1 &= \{px_0\} \in (0, 1) \\ d_2 &= \lfloor px_1 \rfloor \in \{0, 1, 2, \ldots, p-1\}, \qquad & x_2 &= \{px_1\} \in (0, 1) \end{aligned}$$

und weiter

$$d_{n+1} = \lfloor px_n \rfloor \in \{0, 1, 2, \ldots, p-1\}, \qquad x_{n+1} = \{px_n\} \in (0, 1)$$

für jedes $n \in \mathbb{N}$ betrachten. Insbesondere ist $px_n = d_{n+1} + x_{n+1}$ für alle $n \in \mathbb{N}$ und damit ergibt sich mittels Induktion

$$x - \sum_{n=1}^{N} d_n p^{-n} = p^{-N} x_N$$

für alle $N \in \mathbb{N}$. Daraus folgt die Fehlerabschätzung

$$\left| x - \sum_{n=1}^{N} d_n p^{-n} \right| \leq p^{-N} |x_N| < p^{-N}.$$

und mit dem Grenzübergang $N \to \infty$ auch (6.15). Wir möchten an dieser Stelle erwähnen, dass obiges Vorgehen eine explizite Variante (einen Algorithmus) liefert, um die Ziffernentwicklung einer Zahl zur Basis $p$ zu finden.

**Übung 6.90.** *Finden Sie eine Ziffernentwicklung von $\frac{1}{7}$ zur Basis* 3.

In den meisten, aber nicht allen Fällen ist die Ziffernentwicklung eindeutig. Es ist zum Beispiel

$$\sum_{n=1}^{\infty} (p-1)p^{-n} = \frac{p-1}{p} \sum_{n=0}^{\infty} p^{-n} = \frac{p-1}{p} \frac{1}{1-p^{-1}} = 1.$$

Im Spezialfall $p = 10$ kennt man dies in der Form $0.\dot{9} = 1$.

Wir behaupten, dass diese Rechnung gewissermassen der einzige Grund für eine Zweideutigkeit in der Ziffernentwicklung sein kann. Angenommen

$$x = \sum_{n=\ell}^{\infty} d_n p^{-n} = \sum_{n=\ell'}^{\infty} d'_n p^{-n}$$

sind zwei verschiedene Ziffernentwicklungen einer Zahl $x > 0$ zur Basis $p$. Dann können wir, notfalls unter Hinzufügen von Nullen, annehmen, dass $\ell = \ell'$ ist. Falls $k > \ell$ die kleinste Zahl ist mit $d_n = d'_n$ für $\ell \leq n < k$ und $d_k \neq d'_k$, dann können wir die endliche Summe $\sum_{n=\ell}^{k} d_n p^{-n}$ von $x$ abziehen. Daher genügt es den Fall $d_\ell < d'_\ell$ zu betrachten. Dann gilt

$$x = \sum_{n=\ell}^{\infty} d_n p^{-n} \leq d_\ell p^{-\ell} + \sum_{n=\ell+1}^{\infty} (p-1)p^{-n} = (d_\ell + 1)p^{-\ell} \leq \sum_{n=\ell}^{\infty} d'_n p^{-n} = x,$$

was aber $d_\ell + 1 = d'_\ell$, $d_n = p-1$ und $d'_n = 0$ für alle $n \geq \ell+1$ impliziert (wieso?). Insbesondere sind somit die Zahlen mit nicht eindeutiger Ziffernentwicklung genau jene (rationale Zahlen), die eine abbrechende Ziffernentwicklung besitzen.

**Übung 6.91** (Zahlen mit schliesslich periodischer Ziffernentwicklung)**.** *Eine Folge $(d_n)_{n\geq\ell}$ für $\ell \in \mathbb{N}$ heisst **periodisch**, falls ein $N$ existiert mit $d_{n+N} = d_n$ für alle $n \geq \ell$. Sie heisst **schliesslich periodisch**, falls es ein $k \geq \ell$ und ein $N \in \mathbb{N}$ gibt mit $d_{n+N} = d_n$ für alle $n \geq k$. Wir wollen hier zeigen, dass die Zahlen mit einer schliesslich periodischen Ziffernentwicklung gerade die rationalen Zahlen sind. Sei $x > 0$ eine reelle und $p \geq 2$ eine natürliche Zahl. Unterscheiden Sie folgende Fälle:*

*(a) Angenommen $x$ hat eine schliesslich periodische Ziffernentwicklung. Zeigen Sie, dass $x$ rational ist.*

*(b) Zeigen Sie, dass $x$ genau dann eine abbrechende (und damit schliesslich periodische) Ziffernentwicklung hat, wenn $x$ von der Form $\frac{n}{p^k}$ für $k, n \in \mathbb{N}_0$ ist (und damit rational ist).*

*(c) Angenommen $x \in \mathbb{Q}$ hat keine abbrechende Ziffernentwicklung. Dann ist insbesondere die Ziffernentwicklung von $x$ eindeutig bestimmt. Sei o.B.d.A. $x \in (0,1)$ und seien $x_k, d_k$ für $k \in \mathbb{N}$ wie in der Konstruktion der Ziffernentwicklung nach (6.15). Nach Annahme ist $x = \frac{m}{n}$ für $(m,n) \in \mathbb{N}^2$ teilerfremd und $m \leq n$. Zeigen Sie nun, dass $x_k \in \left\{\frac{1}{n}, \ldots, \frac{n-1}{n}\right\}$ ist für alle $k \in \mathbb{N}$ und wenden Sie das Schubfachprinzip an.*

**Übung 6.92** (Mächtigkeit der reellen Zahlen)**.** *Zeigen Sie $[0,1] \sim \mathcal{P}(\mathbb{N})$ sowie $\mathbb{R} \sim \mathcal{P}(\mathbb{N})$.*

### 6.8.1 Die Cantor-Menge

Wir haben die Cantor-Menge bereits in Abschnitt 2.6.5 kurz besprochen. Wir möchten nun die Cantor-Menge neu, jetzt deutlich einfacher, mit Hilfe der Ziffernentwicklung zur Basis 3 charakterisieren. Die Cantor-Menge ist

$$C = \left\{ x \in [0,1] \mid \text{es gibt } (d_n)_n \in \{0,2\}^{\mathbb{N}} \text{ mit } x = \sum_{n=1}^{\infty} d_n 3^{-n} \right\}.$$

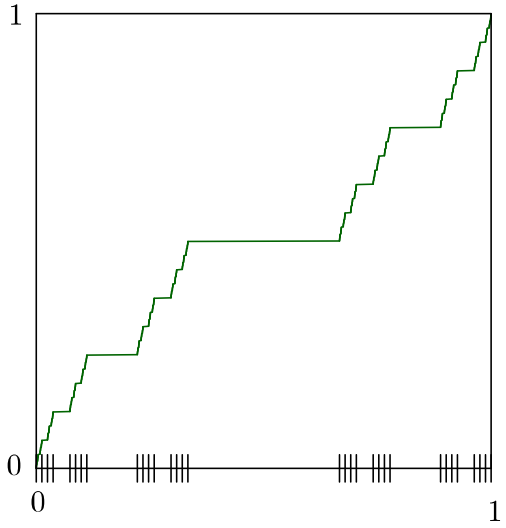
**Übung 6.93** (Cantor-Menge in Basis 3)**.** *Zeigen Sie, dass diese Beschreibung der Cantor-Menge zutrifft.*

Insbesondere hat jeder Punkt $x \in C$ eine eindeutig bestimmte Ziffernentwicklung zur Basis 3 mit Ziffern in $\{0, 2\}$. Jede Ziffernentwicklung $(d_n)_n \in \{0, 2\}^{\mathbb{N}}$ bestimmt einen eindeutig bestimmten Punkt $x \in C$.

**Übung 6.94** („Gesamtlänge" der Cantor-Menge)**.** *Zeigen Sie, dass die charakteristische Funktion der Cantor-Menge $\mathbb{1}_C$ auf $[0, 1]$ Riemann-integrierbar ist und $\int_0^1 \mathbb{1}_C \, \mathrm{d}x = 0$ ist. In diesem Sinne hat $C$ sozusagen „Gesamtlänge" 0 und $[0, 1] \setminus C$ hat „Gesamtlänge" 1.*

### 6.8.2 Cantors Teufelstreppe

Wir wollen hier kurz die **Cantor Funktion** $f : [0, 1] \to [0, 1]$ besprechen, die monoton wachsend und surjektiv ist. Der Graph von $f$ wird auch die **Teufelstreppe** genannt und ist unten im Bild dargestellt.




Die teuflische Eigenheit dieser Funktion ist die Tatsache, dass jeder Punkt in $[0, 1] \setminus C$ eine Umgebung besitzt, auf der $f$ konstant ist und $[0, 1] \setminus C$ in einem gewissen Sinne fast das ganze Intervall ausmacht (siehe Übung 6.94). Trotzdem schafft es die Cantor Funktion auf stetige Weise von 0 zu 1 anzuwachsen.

Wir definieren $f$ zuerst auf der Cantormenge $C$ und beschreiben dann, wie $f$ von $C$ auf ganz $[0, 1]$ fortgesetzt werden kann. Wie in Abschnitt 6.8.1 beschrieben wurde, hat jede Zahl $x \in C$ eine eindeutig bestimmte Ziffernentwicklung $x = \sum_{n=1}^{\infty} d_n 3^{-n}$ zur Basis 3 mit $d_n \in \{0, 2\}$ für alle $n \in \mathbb{N}$. Auf dem Punkt $x$ definieren wir nun den Wert von $f$ mittels der Ziffernentwicklung als

$$f(x) = \sum_{n=1}^{\infty} \frac{d_n}{2} 2^{-n}.$$

Das heisst, wir ersetzen jede Ziffer 2 durch die Ziffer 1 und ersetzen die Basis 3 durch die Basis 2. Daraus folgt, dass $f : C \to [0,1]$ surjektiv ist, da jede Zahl $y \in [0,1]$ eine Ziffernentwicklung $y = \sum_{n=1}^{\infty} e_n 2^{-n}$ zur Basis 2 hat und wir durch $d_n = 2e_n$ für alle $n \in \mathbb{N}$ und $x = \sum_{n=1}^{\infty} d_n 3^{-n}$ einen Punkt mit $f(x) = y$ finden können.

Per Definition ist $f : C \to [0,1]$ monoton wachsend. Des Weiteren ist $f$ nicht streng monoton wachsend (wieso?), aber da eine Zahl in $[0,1]$ höchstens zwei verschiedene Ziffernentwicklungen zur Basis 2 besitzt, gibt es zu $x \in C$ höchstens einen weiteren Punkt $x' \in C$ mit $f(x') = f(x)$.

Um $f$ auf ganz $[0,1]$ zu definieren, finden wir zu einem Punkt $x \notin C$ den grössten Punkt $x_-$ links von $x$ in der Cantormenge und den kleinsten Punkt $x_+$ rechts von $x$ in der Cantormenge. Wir werden unten zeigen, dass die Punkte $x_-$ und $x_+$ eindeutig durch $x$ bestimmt sind und $f(x_-) = f(x_+)$ gilt. Nun definieren wir $f(x) = f(x_-) = f(x_+)$. Dies erweitert also $f$ zu einer Abbildung $[0,1] \to [0,1]$. Des Weiteren ist $f$ für jedes $x \in [0,1] \setminus C$ konstant in der Umgebung $[x_-, x_+]$ von $x$ (und damit auch stetig auf $(x_-, x_+)$).

Formal können wir dies wie folgt beschreiben. Sei

$$x = \sum_{n=1}^{k-1} d_n 3^{-n} + 1 \cdot 3^{-k} + \sum_{n=k+1}^{\infty} d_n 3^{-n} \notin C,$$

wobei $k \in \mathbb{N}$ die kleinste Zahl mit $d_k = 1$ ist. Dann setzen wir

$$\begin{aligned} x_- &= \sum_{n=1}^{k-1} d_n 3^{-n} + 1 \cdot 3^{-k} = \sum_{n=1}^{k-1} d_n 3^{-n} + \sum_{n=k+1}^{\infty} 2 \cdot 3^{-n} \in C \\ x_+ &= \sum_{n=1}^{k-1} d_n 3^{-n} + 2 \cdot 3^{-k} \in C, \end{aligned}$$

womit $x_- \leq x \leq x_+$ erfüllt ist. Damit gilt

$$f(x_-) = \sum_{n=1}^{k-1} \frac{d_n}{2} 2^{-n} + \sum_{n=k+1}^{\infty} 2^{-n} = \sum_{n=1}^{k-1} \frac{d_n}{2} 2^{-n} + 2^{-k} = f(x_+)$$

wie gewünscht.

**Übung 6.95** (Cantor-Funktion)**.** *Zeigen Sie, dass die Cantor-Funktion $f : [0,1] \to [0,1]$ monoton wachsend und stetig ist.*

### 6.8.3 Peanos raumfüllende Kurve

Wir wollen nun dem ursprünglichen Argument von Peano ([Pea90]) folgend eine stetige, surjektive Funktion (die sogenannte **Peano-Kurve**)

$$f : [0,1] \to [0,1]^2$$

konstruieren. Dabei werden wir von der Cantormenge ausgehen.

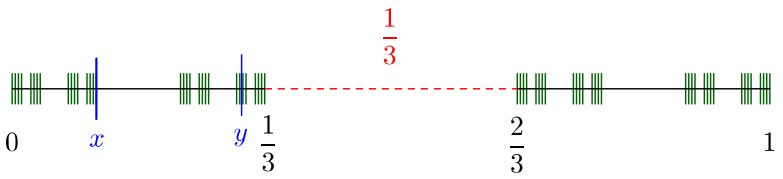
Für $x = \sum_{n=1}^{\infty} d_n(x)3^{-n} \in C$ mit $d_n(x) \in \{0,2\}$ für alle $n \in \mathbb{N}$ definieren wir

$$\mathbf{f}_C(x) = (f_1(x), f_2(x))^t = \left( \sum_{n=1}^{\infty} \frac{d_{2n-1}(x)}{2} 2^{-n}, \sum_{n=1}^{\infty} \frac{d_{2n}(x)}{2} 2^{-n} \right)^t. \tag{6.16}$$

Wir behaupten nun, dass die Abbildung $\mathbf{f}_C : C \to [0,1]^2$ surjektiv ist. Für $(y_1, y_2)^t \in [0,1]^2$ existiert eine Ziffernentwicklung der Komponenten $y_j = \sum_{n=1}^{\infty} e_{j,n} 2^{-n}$ für $j \in \{1,2\}$ zur Basis 2, wobei $e_{j,n} \in \{0,1\}$ für alle $n \in \mathbb{N}$ und $j \in \{1,2\}$. Wir definieren nun $d_{2n-1} = 2e_{1,n}$ und $d_{2n} = 2e_{2,n}$ für alle $n \in \mathbb{N}$, womit $d_n \in \{0,2\}$ für alle $n \in \mathbb{N}$ gilt. Sei $x = \sum_{n=1}^{\infty} d_n 3^{-n} \in C$ das Element der Cantormenge mit diesen Ziffern. Dann gilt $(y_1, y_2)^t = \mathbf{f}_C(x) \in \mathbf{f}_C(C)$ wie gewünscht. Da $(y_1, y_2)^t \in [0,1]^2$ beliebig war, ist $\mathbf{f}_C$ surjektiv.

**Lemma 6.96** (Stetigkeit)**.** *Für jedes $n \in \mathbb{N}$ ist die Abbildung $x \in C \mapsto d_n(x) \in \{0,2\}$ stetig. Insbesondere ist die Abbildung $\mathbf{f}_C : C \to [0,1]^2$ stetig.*

*Beweis.* Für $n = 1$, $x, y \in C$ und $|x - y| < \frac{1}{3}$ folgt $d_1(x) = d_1(y)$, siehe folgendes Bild.




Figur 6.4: Zwei Punkte $x, y \in C$ mit $|x - y| < \frac{1}{3}$ sind entweder beide links (und erfüllen also $d_1(x) = d_2(x) = 0$) oder beide rechts (und erfüllen $d_1(x) = d_2(x) = 1$) in der Cantormenge.

Dies zeigt die Stetigkeit von $d_1 : C \to \{0,2\}$ ($\delta = \frac{1}{3}$ tut's für jedes $\varepsilon > 0$). Mittels Induktion folgt damit, dass für $n > 1$ die Abbildung

$$x \mapsto d_n(x) = d_1\left(3^{n-1}(x - (d_1(x)3^{-1} + \ldots + d_{n-1}(x)3^{-(n-1)})\right)$$

als Verknüpfung von stetigen Funktionen wiederum stetig ist (siehe Proposition 3.51 und Proposition 3.53). Für jedes $N \in \mathbb{N}$ sind die Abbildungen

$$x \in C \mapsto \sum_{n=1}^{N} \frac{d_{2n-1}(x)}{2} 2^{-n}, \quad x \in C \mapsto \sum_{n=1}^{N} \frac{d_{2n}(x)}{2} 2^{-n}$$

als Verknüpfung von stetigen Funktionen stetig. Wegen

$$\left| \sum_{n=1}^{N} \frac{d_{2n-1}(x)}{2} 2^{-n} - \sum_{n=1}^{\infty} \frac{d_{2n-1}(x)}{2} 2^{-n} \right| = \left| \sum_{n=N+1}^{\infty} \frac{d_{2n-1}(x)}{2} 2^{-n} \right| \leq \sum_{n=N+1}^{\infty} 2^{-n} = 2^{-N}$$

für alle $x \in C$, konvergieren die Abbildungen $x \in C \mapsto \sum_{n=1}^{N} \frac{d_{2n-1}(x)}{2} 2^{-n}$ gleichmässig gegen $f_1 : C \to [0,1]$. Damit folgt aus Satz 6.48, dass $f_1$ stetig ist. Analog zeigt man, dass $f_2$ stetig ist und somit ist auch $\mathbf{f}_C$ stetig. $\square$

Wir wollen nun $\mathbf{f}_C$ zu einer stetigen Abbildung $\mathbf{f}_P : [0,1] \to [0,1]^2$ fortsetzen, wobei wir stückweise affine Abbildungen auf $[0,1] \setminus C$ verwenden wollen. Das heisst, für $x \in C$ definieren wir $\mathbf{f}_P(x) = \mathbf{f}_C(x)$ und für $x \notin C$ setzen wir $\mathbf{f}_P(x)$ wie folgt fest. Ist $x_- \in C$ der eindeutige Punkt in der Cantormenge links neben $x$ und $x_+ \in C$ der eindeutige Punkt in der Cantormenge rechts neben $x$ wie in Abschnitt 6.8.2, so setzen wir $s(x) = \frac{x-x_-}{x_+-x_-} \in (0,1)$ und

$$\begin{aligned} \mathbf{f}_P(x) &= \mathbf{f}_C(x_-) + s(x)(\mathbf{f}_C(x_+) - \mathbf{f}_C(x_-)) \\ &= (x_+ - x_-)^{-1}((x_+ - x)\mathbf{f}_C(x_-) + (x - x_-)\mathbf{f}_C(x_+)) \end{aligned} \tag{6.17}$$

unter Verwendung von $x = x_- + s(x)(x_+ - x_-)$.

**Proposition 6.97** (Existenz einer raumfüllenden Kurve)**.** *Die Abbildung* $\mathbf{f}_P : [0,1] \to [0,1]^2$ *ist stetig und surjektiv.*

*Beweis.* Da $\mathbf{f}_C : C \to [0,1]^2$ surjektiv ist und $\mathbf{f}_P|_C = \mathbf{f}_C$ gilt, ist auch $\mathbf{f}_P$ surjektiv. Daher müssen wir nur noch die Stetigkeit von $\mathbf{f}_P$ überprüfen.

Für Punkte, die nicht in $C$ liegen, ist dies einfacher zu sehen. Intuitiv ausgedrückt liegt dies daran, dass die Funktionen $x \mapsto x_+$ und $x \mapsto x_-$ auf dem Komplement von $C$ „lokal konstant“ sind. Nun genauer. Sei also $x_0 \notin [0,1] \setminus C$ und seien $x_{0,-}, x_{0,+} \in C$ wie oben. Dann gilt $(x_{0,-}, x_{0,+}) \subseteq [0,1] \setminus C$, wie in Abschnitt 6.8.2 erklärt wurde. Für alle $x \in (x_{0,-}, x_{0,+})$ erfüllen $x_{0,-}$ und $x_{0,+}$ dieselbe Rolle für $x$ wie für $x_0$, das heisst $x_- = x_{0,-}$ und $x_+ = x_{0,+}$. Damit ist $x \mapsto s(x) = \frac{x-x_-}{x_+-x_-} = \frac{x-x_{0,-}}{x_{0,+}-x_{0,-}}$ stetig auf dem Intervall $(x_{0,-}, x_{0,+})$. Es folgt die Stetigkeit von $\mathbf{f}_P$ bei $x_0$.

Sei nun $x_0 \in C$. Falls $x_0$ ein „rechter Endpunkt“ von $C$ ist, also von der Form $x_0 = y_-$ für ein $y \in [0,1] \setminus C$ ist, dann gilt

$$\mathbf{f}_P(x) = (y_+ - y_-)^{-1}((y_+ - x)\mathbf{f}_C(x_0) + (x - x_0)\mathbf{f}_C(y_+))$$

für alle $x \in (x_0, y_+)$. Daraus folgt gemeinsam mit $\mathbf{f}_P(x_0) = \mathbf{f}_C(x_0)$ die rechtsseitige Stetigkeit von $\mathbf{f}_P$ bei $x_0$.

Sei also nun $x_0$ kein rechter Endpunkt von $C$, womit eine streng monoton fallende Folge $(x_n)_n$ in $C$ existiert, die gegen $x_0$ konvergiert. Sei $\varepsilon > 0$, dann gibt es wegen der Stetigkeit von $\mathbf{f}_C$ ein $\delta > 0$, so dass für alle $y \in C$

$$|y - x| < \delta \implies |\mathbf{f}_C(y) - \mathbf{f}_C(x)| < \varepsilon$$

gilt. Weiter gibt es ein $x_n \in C \cap (x_0, x_0 + \delta)$. Für $x \in (x_0, x_n)$ unterscheiden wir zwei Fälle. Wenn $x$ in $C$ liegt, gilt $|\mathbf{f}_P(x) - \mathbf{f}_P(x_0)| = |\mathbf{f}_C(x) - \mathbf{f}_C(x_0)| < \varepsilon$. Wenn $x$ nicht in $C$ liegt, dann

gilt $x,x_+ \in (x_0, x_n]$ und

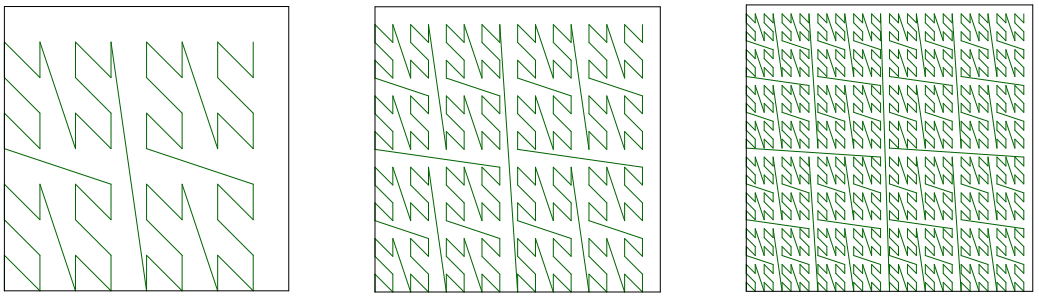
$$\begin{aligned}
|\mathbf{f}_P(x) - \mathbf{f}_P(x_0)| &= \big|(x_+ - x_-)^{-1}((x_+ - x)\mathbf{f}_C(x_-) + (x - x_-)\mathbf{f}_C(x_+)) - \mathbf{f}_C(x_0)\big| \\
&= \frac{1}{x_+ - x_-}\big|(x_+ - x)\mathbf{f}_C(x_-) + (x - x_-)\mathbf{f}_C(x_+) - (x_+ - x_-)\mathbf{f}_C(x_0)\big| \\
&= \frac{1}{x_+ - x_-}\big|(x_+ - x)(\mathbf{f}_C(x_-) - \mathbf{f}_C(x_0)) + (x - x_-)(\mathbf{f}_C(x_+) - \mathbf{f}_C(x_0))\big| \\
&< \frac{1}{x_+ - x_-}((x_+ - x)\varepsilon + (x - x_-)\varepsilon) = \varepsilon.
\end{aligned}$$

Somit gilt nun $|\mathbf{f}_P(x) - \mathbf{f}_P(x_0)| < \varepsilon$ für alle $x \in (x_0, x_n)$ und da $\varepsilon > 0$ beliebig war, ist $\mathbf{f}_P$ rechsseitig stetig bei $x_0$. Das Argument für die linksseitige Stetigkeit ist analog. □

*Bemerkung* (Approximative Darstellung der Peano-Kurve). Auf Grund der Surjektivität der Kurve $\mathbf{f}_P$ macht es wenig Sinn, diese darstellen zu wollen. Stattdessen stellt man eine Approximation der Peano-Kurve dar. Für $N \in \mathbb{N}$ setzt man

$$\mathbf{f}_{\mathbf{C}}^{(\mathbf{N})}(x) = (f_1^{(N)}(x), f_2^{(N)}(x))^t = \left(\sum_{n:2n-1\le N} \frac{d_{2n-1}(x)}{2} 2^{-n}, \sum_{n:2n\le N} \frac{d_{2n}(x)}{2} 2^{-n}\right)^t$$

für alle $x \in C$ und erweitert anschliessend die Funktion $\mathbf{f}_{\mathbf{C}}^{(\mathbf{N})} : C \to [0,1]^2$ zu einer stetigen Funktion $\mathbf{f}_P : [0,1] \to [0,1]^2$ in Analogie zu (6.17). Die Funktionenfolge $(\mathbf{f}_{\mathbf{C}}^{(\mathbf{N})})_N$ konvergiert, wie man zeigen kann, gleichmässig gegen $\mathbf{f}_C$.



Figur 6.5: Das Bild der approximativen Peano-Kurven $\mathbf{f}_{\mathbf{C}}^{(\mathbf{N})}$ für $N \in \{6, 8, 10\}$.

Wir bemerken noch, dass eine stetige, surjektive Funktion $\mathbf{f} : [0,1] \to [0,1]^2$ nie injektiv sein kann. Dies ist eine Manifestation der Tatsache, dass $[0,1]$ und $[0,1]^2$ geometrisch verschiedene Objekte sind. Aussagen dieser Form sind Teil der (algebraischen) Topologie, doch der vorliegende konkrete Fall benötigt keine allzu grosse Theorie (siehe Übung 6.98).

**Übung 6.98** (Kein Homöomorphismus)**.** *Wir wollen hier zeigen, dass es keine stetige und bijektive Funktion von* $[0,1]$ *nach* $[0,1]^2$ *gibt. Wir nehmen indirekt an, dass* $\mathbf{f} : [0,1] \to [0,1]^2$ *stetig und bijektiv sei.*

*(i) Wir wissen bereits, dass* $\mathbf{f}$ *konvergente Teilfolgen auf konvergente Teilfolgen abbildet. Zeigen Sie, dass* $\mathbf{f}^{-1} : [0,1]^2 \to [0,1]$ *dies ebenfalls tut.*

*(ii) Finden Sie eine stetige Abbildung (eine Kurve)* $\mathbf{g} : [0,1] \to [0,1]^2 \setminus \{\mathbf{f}(\frac{1}{2})\}$ *mit* $\mathbf{g}(0) = \mathbf{f}(\frac{1}{4})$ *und* $\mathbf{g}(1) = \mathbf{f}(\frac{3}{4})$.

*(iii) Zeigen Sie, dass die Abbildung* $\mathbf{f}^{-1} \circ \mathbf{g} : [0,1] \to [0,1] \setminus \{\frac{1}{2}\}$ *stetig ist und schliessen Sie mit dem Zwischenwertsatz auf einen Widerspruch.*

**Übung 6.99** (Stetiges Füllen der Ebene)**.** *Finden Sie eine stetige, surjektive Funktion* $f : \mathbb{R} \to \mathbb{R}^2$.

Die hier besprochene Konstruktion einer fraktalen raumfüllenden Kurve mag auf dem ersten Blick unnatürlich und auf jeden Fall weltfremd anmuten. Doch enthält die mathematischen Modellierung der Brownschen Bewegung ähnliche fraktale Kurven, die typischerweise nicht raumfüllend aber genauso wie die Peano-Kurve stetig und "sehr zittrig" sind. Wie man das Gegenteil "schön glatt" von "sehr zittrig" mathematisch formulieren kann, besprechen wir im nächsten Kapitel.

## 6.9 Weitere Lernmaterialien

### 6.9.1 Verwendung des Kapitels

Wie wir gesehen haben, sind Potenzreihen, deren Konvergenzradius und Konvergenzverhalten fundamentale Werkzeuge für die Definition von vielen Ihnen bereits bekannten Funktionen (und auch weiteren). Wir werden also ab nun sowohl die komplexe Exponentialfunktion, die trigonometrischen Funktionen auf $\mathbb{R}$ und auf $\mathbb{C}$, als auch die hyperbolischen Funktionen gemeinsam mit den wichtigsten Eigenschaften dieser Funktionen (meist ohne Verweise) verwenden. (Die Umkehrfunktionen der trigonometrischen Funktionen werden wir erst im nächsten Kapitel einführen.)

Für den Begriff der Potenzreihe benötigten wir die grundlegenden Definitionen der Reihe und der Funktionenfolgen. Für Reihen ist die Unterscheidung der bedingten und absoluten Konvergenz fundamental, da gewisse Operationen (Umordnen, Cauchy-Produkt) nur für den letzteren Konvergenzbegriff erlaubt sind. Dabei ist es sehr hilfreich, dass für Potenzreihen im Inneren des Konvergenzbereichs absolute Konvergenz vorliegt und damit alle Operationen erlaubt sind. Die folgenden Konvergenzkriterien sind für Beispiele aber auch für die Theorie unabdingbar:

- die geometrische Reihe in Beispiel 6.3,
- Majoranten- und Minorantenkriterium für Reihen mit positiven Gliedern in Korollar 6.12 und Korollar 6.29,
- Verdichtung in Proposition 6.16,
- $p$-Test in Beispiel 6.17,
- Leibniz-Kriterium in Proposition 6.25 (welches vor allem für bedingt konvergente aber wegen der Fehlerabschätzung auch für absolut konvergente Reihen nützlich sein kann),
- Cauchy-Kriterium in Satz 6.26 (meist als theoretisches Hilfsmittel),
- Wurzelkriterium in Korollar 6.30 (als theoretisches und praktisches Hilfsmittel),
- Quotientenkriterium in Korollar 6.32 (meist als praktisches Hilfsmittel, da es oft einfacher anwendbar ist, aber im Gegensatz zu dem Wurzelkriterium zum Beispiel für Potenzreihen weniger allgemein einsetzbar ist),
- aber wenn sonst nichts zum Erfolg führt, sollte man nicht vergessen, dass auf Grund von Proposition 6.2 die Folgenglieder einer konvergenten Reihe eine Nullfolge bilden.

Wir bemerken noch, dass diese Kriterien sehr hilfreich sind für die Entscheidung ob Konvergenz oder Divergenz bei einer Reihe vorliegt, doch haben wir sehr wenige allgemeine Gesetze um den Grenzwert von Reihen zu bestimmen.

Wie bereits erwähnt war der Begriff der Funktionenfolge auch für die Besprechung der Potenzreihen notwendig. Für Funktionenfolgen haben wir zwei unterschiedliche Konvergenzbegriffe besprochen. Der Begriff der punktweisen Konvergenz mag zwar als der natürliche Konvergenzbegriff für Funktionen betrachtet werden, doch hat dieser keine guten Eigenschaften (weder für Stetigkeit noch für das Riemann-Integral). Sie sollten die entsprechenden Gegenbeispiele im Gedächtnis behalten. Dies motivierte die Definition der gleichmässigen Konvergenz, welche wegen den guten Eigenschaften für Stetigkeit und das Riemann-Integral für uns immer wieder wichtig sein wird. Die Unterscheidung dieser Konvergenzbegriffe ist wohlgemerkt keine Spitzfindigkeit.

### 6.9.2 Übungen

**Übung.** *Sei $\sum_{k=1}^{\infty} a_k$ eine konvergente Reihe. Falls $(a_k)_k$ eine monoton fallende Folge ist, so ist nicht nur $(a_k)_k$, sondern auch $(ka_k)_k$ eine Nullfolge. Beweisen Sie dies.*

**Übung.** *Sei $f : \mathbb{N} \to \mathbb{N}$ eine bijektive Abbildung. Konvergiert die Reihe $\sum_{n=1}^{\infty} \frac{f(n)}{n^2}$?*

**Übung** (Raabes Quotientenkriterium)**.** *Sei $(a_n)_n$ eine Folge komplexer Zahlen mit $a_n \neq 0$ für alle $n \in \mathbb{N}$, so dass $\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = 1$ und*

$$Q = \lim_{n\to\infty} n \left(1 - \frac{|a_{n+1}|}{|a_n|}\right)$$

*existiert. Zeigen Sie, dass die Reihe $\sum_{n=1}^{\infty} a_n$ konvergiert, falls $Q > 1$. (J. Raabe war einer der ersten Mathematikprofessoren an der ETH Zürich.)*

*Hinweis: Finden Sie ein $p > 1$, so dass für alle bis auf endlich viele $n$*

$$\frac{|a_{n+1}|}{|a_n|} < 1 - \frac{p}{n}$$

*gilt. Verwenden Sie die kontinuierliche Bernoulli-Ungleichung aus Übung 5.61, um zu zeigen, dass*

$$\frac{|a_{n+1}|}{|a_n|} \leq \left(1 - \frac{1}{n}\right)^p = \frac{(n-1)^p}{n^p}.$$

*Schliessen Sie nun auf die Aussage unter Verwendung von Korollar 6.29 und Beispiel 6.17.*

**Übung** (Vertauschung der Summationsreihenfolge)**.** *Wie schon vor dem Beweis des Produktsatzes (Satz 6.36) angedeutet, ist der Produktsatz stark mit Vertauschbarkeit von Summationsreihenfolge verwandt. Wir wollen dies hier genauer formulieren. Dazu betrachten wir eine doppelt indizierte Folge $(a_{(m,n)})_{(m,n)\in\mathbb{N}^2}$.*

a) *Um zu sehen, dass die Vertauschbarkeit der Summationsreihenfolge für Reihen nicht immer gilt, definieren wir*

$$a_{(m,n)} = \begin{cases} 1 & \text{falls } m = n \\ -1 & \text{falls } m+1 = n \\ 0 & \text{sonst .} \end{cases}$$

*für alle* $(m,n) \in \mathbb{N}^2$. *Zeigen Sie, dass die Doppelreihen*

$$\sum_{m=1}^{\infty}\left(\sum_{n=1}^{\infty} a_{(m,n)}\right), \quad \sum_{n=1}^{\infty}\left(\sum_{m=1}^{\infty} a_{(m,n)}\right)$$

*konvergieren, aber verschieden sind.*

b) *Angenommen* $(a_{(m,n)})_{(m,n)\in\mathbb{N}^2}$ *erfüllt* $\sum_{m=1}^{\infty}\left(\sum_{n=1}^{\infty}|a_{(m,n)}|\right) < \infty$. *Zeigen Sie, dass die beiden Reihen* $\sum_{m=1}^{\infty}\left(\sum_{n=1}^{\infty} a_{(m,n)}\right)$ *und* $\sum_{n=1}^{\infty}\left(\sum_{m=1}^{\infty} a_{(m,n)}\right)$ *konvergent sind und den gleichen Wert haben.*

**Übung** (Zerlegung in gerade und ungerade Funktionen)**.** *Sei* $f : \mathbb{C} \to \mathbb{C}$ *eine Funktion. Zeigen Sie, dass sich* $f$ *als Summe einer geraden und einer ungeraden Funktion schreiben lässt.*

**Übung.** *Seien* $a, b \in \mathbb{R}$ . *Zeigen Sie, dass es ein* $\theta \in \mathbb{R}$ *gibt, so dass*

$$a\sin(\varphi) + b\cos(\varphi) = \sqrt{a^2+b^2}\,\sin(\varphi+\theta)$$

*für alle* $\varphi \in \mathbb{R}$. *Sie können dies auch in diesem Applet sehen.*

**Übung** (Irrationalität der Eulerschen Zahl)**.** *Zeigen Sie, dass* $\mathrm{e} = \sum_{k=0}^{\infty}\frac{1}{k!}$ *irrational. Nehmen Sie indirekt an, dass* e *rational ist und verwenden Sie die Exponentialreihe.*

**Übung.** *Wir wollen einen (spielsüchtigen) Kasinobesucher und ein unübliches, den Spieler bevorzugendes, Spiel betrachten. Das Spiel besteht aus einem einfachen Münzwurf mit zwei möglichen gleich wahrscheinlichen Ergebnissen, nämlich Kopf und Zahl. Bei Kopf gewinnt das Kasino den Einsatz des Spielers und bei Zahl gewinnt der Spieler das Vierfache seines Einsatzes. Wir wollen die verschiedenen Ergebnisse des iterierten Spieles anhand einer Funktion auf* $[0,1]$ *beschreiben. Hierbei verwenden wir die binäre Zifferndarstellung reeller Zahlen, wobei wir die Nichteindeutigkeit einfach ignorieren, da diese für sich gesehen extrem unwahrscheinlichen Ergebnissen des iterierten Spieles entsprechen.*[4] *Wir interpretieren die Ziffer* 0 *als Kopf und die Ziffer* 1 *als Zahl, womit die reelle Zahl* $0.100101\ldots$ *für die wiederholten Würfe der Münze mit den Ergebnissen Zahl-Kopf-Kopf-Zahl-Kopf-Zahl und so weiter steht (und die Null vor dem Komma ignoriert wird).*

(i) *Bestimmen Sie eine Funktion* $f_n : [0,1] \to \mathbb{R}$, *die den Gewinn nach* $n$ *Wiederholungen des Spiels beschreibt, wobei der Spieler mit einem Franken das Spiel beginnt und bei Gewinn jeweils sein Gesamtvermögen im nächsten Spiel wieder einsetzt.*

[4] Dies genauer und mathematisch exakt zu formulieren würde hier zu weit führen.

*(ii) Bestimmen Sie den Erwartungswert für den Spieler, wenn dieser das Spiel $n$ mal wiederholt (also das Integral von $f_n$).*

*(iii) Wir nehmen nun an, dass der Spieler spielsüchtig ist und auch bei Gewinn von wirklich grossen Summen nicht aufhören kann zu spielen. Bestimmen Sie die Funktion, die den Gewinn des Spielers beschreibt. Sie sollten bemerken, dass dies der punktweise Grenzwert der Funktion $f_n$ ist.*

*(iv) Bestimmen Sie den Erwartungswert des unbeschränkt langen Spiels für den spielsüchtigen Spieler.*

**Übung** (Der Fall der absoluten Konvergenz am Rand)**.** *Sei $\sum_{n=0}^{\infty} a_n z^n$ eine Potenzreihe mit Konvergenzradius $R \in (0, \infty)$. Sei des Weiteren $\sum_{n=0}^{\infty} |a_n| R^n < \infty$. Zeigen Sie, dass in diesem Fall die Funktion*

$$z \in \overline{B_R(0)} = \{z \in \mathbb{C} \mid |z| \leq R\} \mapsto \sum_{n=0}^{\infty} a_n z^n$$

*wohldefiniert und stetig ist.*

**Übung** (Bestimmte Divergenz am Rand)**.** *Sei $\sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihe mit Konvergenzradius $R \in (0, \infty)$ und nicht-negativen Koeffizienten $a_n \geq 0$ für alle $n \in \mathbb{N}$. Sei des Weiteren $\sum_{n=0}^{\infty} a_n R^n = \infty$. Zeigen Sie, dass in diesem Fall*

$$\lim_{x \nearrow R} \sum_{n=0}^{\infty} a_n x^n = \infty.$$

**Übung** (Challenge)**.** *Wir betrachten das Gitter $\mathbb{N}^2$ in der Ebene $\mathbb{R}^2$ und fixieren uns ein Quadrat darin, $Q = [1, 15]^2 \cap \mathbb{N}^2$. Auf den Punkten $(1, 1)$, $(2, 1)$, $(3, 1)$, $(1, 2)$, $(2, 2)$, $(1, 3)$ platzieren wir nun jeweils eine Münze und beginnen dann folgendes Spiel. Sie als Spieler dürfen jeweils eine der Münzen entfernen, worauf Sie eine Münze oberhalb und eine Münze rechts von der entfernten Münze platzieren. Dieser Zug ist aber nur dann erlaubt, wenn die Plätze rechts und oberhalb der zu entfernenden Münze noch frei sind. Ihre Aufgabe besteht nun darin, nach endlich vielen Zügen keine Münzen mehr innerhalb des Quadrat $Q$ liegen zu haben. Ist das möglich?*

### 6.9.3 Lernkarten

Sie können wiederum die Lernkarten oder den Graphen für Ihre Wiederholung der Themen des Kapitels verwenden.

# Kapitel 7

# Differentialrechnung

Wir beginnen nun mit der Differentialrechnung in einer Variablen. Wie wir sehen werden, ist diese für das Verständnis von Funktionen auf Intervallen in $\mathbb{R}$, aber auch für die Berechnung des Riemann-Integrals von fundamentaler Bedeutung.

## 7.1 Die Ableitung

### 7.1.1 Definition und geometrische Interpretation

Eine (nicht-vertikale) **Gerade** im $\mathbb{R}^2$ ist eine Teilmenge der Form $\{(x, y) \mid y = mx + q\}$ für Parameter $m, q \in \mathbb{R}$ oder alternativ ausgedrückt der Graph der (affinen) Abbildung $x \in \mathbb{R} \mapsto mx + q \in \mathbb{R}$. Meist nennt man Funktionen dieser Form ebenfalls Geraden. Der Parameter $m$ der Geraden $y = mx + q$ wird auch die **Steigung** der Geraden genannt. Wir möchten uns nun mit Funktionen beschäftigen, die sich um einen Punkt im Definitionsbereich durch Geraden approximieren lassen.

**Definition 7.1** (Differenzierbarkeit)**.** Sei $D \subseteq \mathbb{R}$ eine Teilmenge, $f : D \to \mathbb{R}$ eine Funktion und $a \in D$ ein Häufungspunkt von $D$. Wir sagen, dass $f$ bei $a$ **differenzierbar** ist, falls der Grenzwert

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{h \to 0} \frac{f(a + h) - f(a)}{h} \tag{7.1}$$

existiert. In diesem Fall nennen wir $f'(a)$ die **Ableitung** von $f$ bei $a$. Falls $f$ bei jedem Häufungspunkt von $D$ in $D$ differenzierbar ist, dann sagen wir auch, dass $f$ (auf $D$) **differenzierbar** ist und nennen die Funktion $a \mapsto f'(a)$ definiert auf den Häufungspunkten von $D$ in $D$ die **Ableitung** von $f$.

Falls $a \in D$ ein rechtseitiger Häufungspunkt von $D$ ist, dann ist $f$ bei $a$ **rechtsseitig differenzierbar**, wenn die **rechtsseitige Ableitung**

$$f'_+(a) = \lim_{x \searrow a} \frac{f(x) - f(a)}{x - a} = \lim_{h \searrow 0} \frac{f(a + h) - f(a)}{h}$$

existiert. **Linksseitige Differenzierbarkeit** und die **linksseitige Ableitung** $f'_-(a)$ werden analog über die Bewegung $x \nearrow a$ definiert.

Wir nennen $\triangle x = x - a = h$ im Zusammenhang mit der Definition in (7.1) auch das **Inkrement des Arguments** oder der **unabhängigen Variablen** $x$, $\triangle f = f(x) - f(a) = f(a+h) - f(a)$ das **Inkrement der Funktion** und $\frac{\triangle f}{\triangle x}$ den **Differenzenquotienten**. Die Ableitung von $f$ bei $a$, welche in dieser Formulierung der Grenzwert des Differenzenquotienten $\frac{\triangle f}{\triangle x}$ für $\triangle x \to 0$ ist, schreibt man auch als $\frac{\mathrm{d}f}{\mathrm{d}x}(a) = f'(a)$ und nennt dies den **Differentialquotienten** (in der Leibniz-Notation). Weiters nennt man $f' = \frac{\mathrm{d}f}{\mathrm{d}x}$ auch die **Ableitung nach** $x$, was vor allem dann nützlich ist, wenn $f$ auch von weiteren Parametern abhängen darf.

Wir möchten aber betonen, dass $\frac{\mathrm{d}f}{\mathrm{d}x}(a)$ nicht als Quotient, sondern nur als Grenzwert von Quotienten definiert wurde. Falls die unabhängige Variable $t$ (für Zeit) und nicht $x$ ist, dann verwendet man manchmal auch die Notation $\dot{x}, \dot{y}$ für die Ableitung von Funktionen $x : D \to \mathbb{R}$, $y : D \to \mathbb{R}$.

Eine weitere Schreibweise der Definition in (7.1) ist in der Landau-Notation (siehe Abschnitt 5.6)

$$\frac{f(x) - f(a)}{x - a} = f'(a) + o(1)$$

für $x \to a$ oder äquivalenterweise

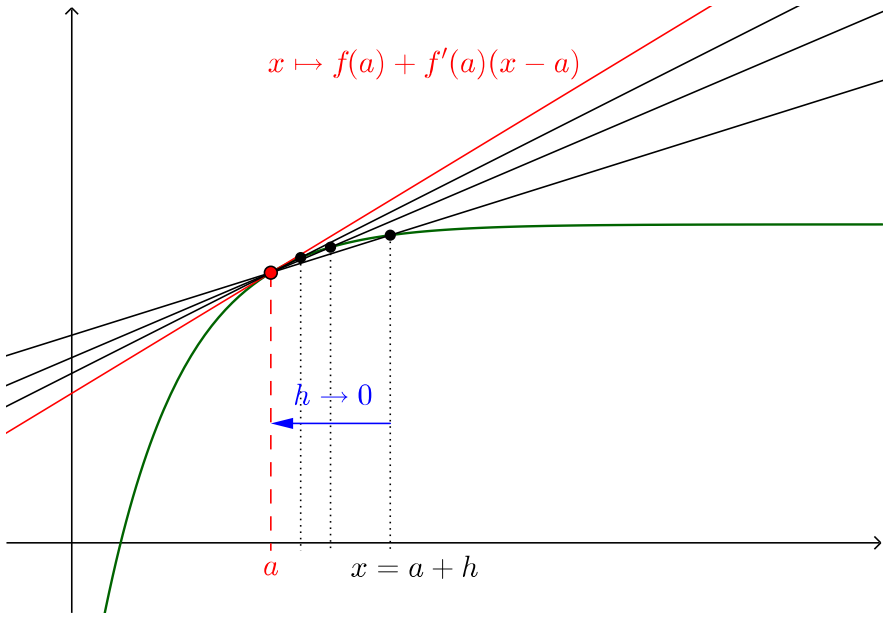
$$f(x) = f(a) + f'(a)(x - a) + o(x - a) \tag{7.2}$$

für $x \to a$. Hierbei wird die Funktion $x \mapsto f'(a)(x - a)$ das **Differential** von $f$ bei $a$ genannt und die Gerade $x \mapsto f(a) + f'(a)(x - a)$ die **affine** oder **lineare Approximation** von $f$ bei $a$ oder die **Tangente** von $f$ bei $a$, siehe auch Figur 7.1. Wir erinnern daran, dass wir in (7.2) $o(x - a)$ als Platzhalter einer Funktion (welcher?) interpretieren, die für $x \to a$ schneller abfällt als $x - a$. Insbesondere ist wegen (7.2)

$$\lim_{x \to a} f(x) = f(a) + \lim_{x \to a} \big(f'(a)(x - a) + o(x - a)\big) = f(a)$$

und $f$ ist bei $a$ stetig, wenn $f$ bei $a$ differenzierbar ist.

**Applet 7.2** (Bewegung der Sekante)**.** *Wir sehen den Graphen einer Funktion und wie die Sekante zwischen $x_0$ und $x_0 + h$ sich bei den meisten Fusspunkten $x_0$ der Tangente bei $x_0$ nähert falls $h \to 0$.*




Figur 7.1: Die geometrische Interpretation der Ableitung einer reellwertigen Funktion $f$ bei $a$ ist die Steigung der Tangenten des Graphen bei $a$. Denn wenn $x$ gegen $a$ strebt, wird die Sekante, die durch $(a, f(a))$ und $(x, f(x))$ geht und den Differenzenquotienten als Steigung besitzt, immer mehr zur Tangente des Graphen bei $a$.

Häufig wird in diesem Kapitel (und dem nächsten) der Definitionsbereich $D$ der betrachteten Funktion $f : D \to \mathbb{R}$ ein Intervall $D = I$ mit Endpunkten $a < b$ sein. Dies hat den Vorteil, dass jeder Punkt in $I$ ein Häufungspunkt ist (wieso?) und es somit für jeden Punkt in $I$ Sinn macht, nach der Differenzierbarkeit von $f$ bei diesem Punkt zu fragen. Wir wollen dies aber weder in der Definition noch in den zu besprechenden Ableitungsregeln voraussetzen, damit wir beispielsweise auch von der Ableitung der Funktion $x \in \mathbb{R} \setminus \{0\} \mapsto \frac{1}{x} \in \mathbb{R}$ sprechen können.

Meist werden wir reellwertige Funktionen betrachten. Doch wird es teilweise nützlich sein, den Begriff der Ableitung und manche der Gesetze auch für komplexwertige Funktionen verwenden zu können. Wir bemerken also, dass Definition 7.1 analog auch für komplexwertige Funktionen verwendet werden kann. Wie in Abschnitt 5.7.4 läuft dies darauf hinaus, dass sowohl Real- als auch Imaginärteil differenzierbar sein sollten.

Schlussendlich wollen wir noch anmerken, dass die Ableitung eine rein lokale Operation darstellt. Genauer gesagt, angenommen $a \in D$ ist ein Häufungspunkt von $D$ und $f, g : D \to \mathbb{R}$ sind bei $a$ differenzierbare Funktionen, so dass es ein $\delta > 0$ gibt mit $f(x) = g(x)$ für alle $x \in D \cap (a - \delta, a + \delta)$. Dann gilt $f'(a) = g'(a)$. Dies ergibt sich unmittelbar aus der Definition der Grenzwerte, die $f'(a)$ und $g'(a)$ definieren (wieso?). Wir werden dies im Folgenden teils implizit verwenden.

### 7.1.2 Beispiele und Ableitungsregeln

Wir wollen nun zeigen, dass viele der uns geläufigen Funktionen differenzierbar sind und dass wir die Ableitung (meistens) mittels einigen konkreten Gesetzen bestimmen können. Wir beginnen aber zuerst mit elementaren Beispielen.

**Beispiel 7.3** (Erste Beispiele differenzierbarer Funktionen)**.**

*(i) Konstante Funktionen sind überall differenzierbar und haben die Nullfunktion als Ableitung (wieso?).*

*(ii) Die Identitätsfunktion $f : x \in \mathbb{R} \mapsto x \in \mathbb{R}$ ist differenzierbar und ihre Ableitung ist die konstante 1-Funktion, denn*

$$f'(a) = \lim_{x \to a} \frac{x-a}{x-a} = 1$$

*für alle $a \in \mathbb{R}$.*

*(iii) Die Exponentialfunktion $\exp : \mathbb{R} \to \mathbb{R}_{>0}$ ist differenzierbar und ihre Ableitung ist die Exponentialfunktion. Allgemeiner behaupten wir, dass für ein festes $\alpha \in \mathbb{R}$ (oder $\alpha \in \mathbb{C}$) die Ableitung von $f : x \in \mathbb{R} \mapsto \exp(\alpha x) \in \mathbb{R}$ durch $f'(a) = \alpha \exp(\alpha a)$ für alle $a \in \mathbb{R}$ gegeben ist. In der Tat gilt für $a \in \mathbb{R}$, dass*

$$\begin{aligned} f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h} &= \lim_{h \to 0} \frac{\exp(\alpha a)\exp(\alpha h) - \exp(\alpha a)}{h} \\ &= \exp(\alpha a) \lim_{h \to 0} \frac{\exp(\alpha h) - 1}{h} \\ &= \exp(\alpha a) \lim_{h \to 0} \frac{\sum_{k=0}^{\infty} \frac{1}{k!}(\alpha h)^k - 1}{h} \\ &= \exp(\alpha a) \lim_{h \to 0} \sum_{k=1}^{\infty} \frac{1}{k!} \alpha^k h^{k-1} \\ &= \exp(\alpha a) \lim_{h \to 0} \sum_{\ell=0}^{\infty} \frac{\alpha^{\ell+1}}{(\ell+1)!} h^{\ell} \\ &= \exp(\alpha a)\alpha, \end{aligned}$$

*da die Abbildung $h \in \mathbb{R} \mapsto \sum_{\ell=0}^{\infty} \frac{\alpha^{\ell+1}}{(\ell+1)!} h^{\ell}$ nach Satz 6.56 stetig ist.*

Wir besprechen weitere Beispiele von differenzierbaren Funktionen und ein Beispiel einer nicht-differenzierbaren Funktion in der folgenden Übung.

**Wichtige Übung 7.4** (Weitere differenzierbare Funktionen)**.**

*(i) Zeigen Sie*

$$\lim_{h \to 0} \frac{\sin(h)}{h} = 1, \quad \lim_{h \to 0} \frac{\cos(h) - 1}{h} = 0.$$

*(ii) Verwenden Sie die Additionstheoreme aus Abschnitt 6.6.1 (oder Beispiel 7.3 (iii)), um zu zeigen, dass der Sinus und der Kosinus differenzierbare Funktionen sind und die Ableitungsregeln*

$$\sin'(x) = (\sin(x))' = \cos(x), \quad \cos'(x) = (\cos(x))' = -\sin(x)$$

*für alle $x \in \mathbb{R}$ gelten.*

*(iii) Zeigen Sie, dass die Funktionen* sinh *und* cosh *differenzierbar sind und verifizieren Sie die Ableitungsregeln*

$$\sinh'(x) = (\sinh(x))' = \cosh(x), \quad \cosh'(x) = (\cosh(x))' = \sinh(x).$$

*für alle $x \in \mathbb{R}$.*

*(iv) Zeigen Sie, dass die Betragsfunktion $x \in \mathbb{R} \mapsto |x| \in \mathbb{R}_{>0}$ nicht differenzierbar ist und bestimmen Sie bei jedem Punkt in $\mathbb{R}$ die linksseitige und die rechtsseitige Ableitung.*

Wie in (ii) und (iii) von Übung 7.4 schon verwendet, wollen wir für Funktionen wie zum Beispiel die Funktion $x \in \mathbb{R} \setminus \{1\} \mapsto \frac{x}{x-1}$, die durch Formeln gegeben sind, nicht immer einen Namen einführen, um die Ableitung hinschreiben zu können. Stattdessen schreiben wir

$$\left(\frac{x}{x-1}\right)' = -\frac{1}{(x-1)^2}$$

und meinen damit, dass die Funktion $x \mapsto \frac{x}{x-1}$ auf ihrem maximalen Definitionsbereich differenzierbar ist und dass ihre Ableitung bei $x$ durch $-\frac{1}{(x-1)^2}$ gegeben ist. Insbesondere ist $x$ in obiger Gleichung nicht als Zahl, sondern vielmehr als Argument der Funktion und der Ableitung zu erachten.

Wie schon bei stetigen und Riemann-integrierbaren Funktionen möchten wir nicht immer von Hand zeigen müssen, dass eine gegebene Funktion differenzierbar ist. Stattdessen wollen wir allgemeine Regeln beweisen, auf die sich die Differenzierbarkeit verschiedener Funktionen zurückführen lässt.

**Proposition 7.5** (Summen und Produkte differenzierbarer Funktionen)**.** *Sei $D \subseteq \mathbb{R}$ eine Teilmenge und $a \in D$ ein Häufungspunkt von $D$. Seien $f, g : D \to \mathbb{R}$ bei $a$ differenzierbar. Dann sind $f + g$ und $f \cdot g$ bei $a$ differenzierbar und es gilt*

$$\begin{aligned}(f+g)'(a) &= f'(a) + g'(a),\\ (fg)'(a) &= f'(a)g(a) + f(a)g'(a).\end{aligned}$$

*Insbesondere ist jedes skalare Vielfache von $f$ bei $a$ differenzierbar und $(\alpha f)'(a) = \alpha f'(a)$ für alle $\alpha \in \mathbb{R}$. Dies gilt ebenso für komplexwertige Funktionen.*

Somit bilden die bei $a \in D$ differenzierbaren reellwertigen Funktionen einen Unterraum des Vektorraums $\mathcal{F}(D)$ der reellwertigen Funktionen von $D$ nach $\mathbb{R}$ und die Ableitung bei $a$ ist eine lineare Abbildung von diesem Unterraum nach $\mathbb{R}$. Die Ableitungsregel für das Produkt zweier Funktionen wird auch die **Produktregel** genannt.

*Beweis.* Wir berechnen unter Verwendung der Eigenschaften des Grenzwerts in Abschnitt 5.4.1

$$\lim_{x \to a} \frac{(f+g)(x) - (f+g)(a)}{x-a} = \lim_{x \to a} \frac{f(x) - f(a)}{x-a} + \frac{g(x) - g(a)}{x-a} = f'(a) + g'(a)$$

und

$$\begin{aligned}\lim_{x\to a}\frac{(f\cdot g)(x)-(f\cdot g)(a)}{x-a} &= \lim_{x\to a}\frac{(f(x)-f(a))g(x)+f(a)(g(x)-g(a))}{x-a}\\ &= \lim_{x\to a}\frac{f(x)-f(a)}{x-a}g(x)+f(a)\frac{g(x)-g(a)}{x-a}\\ &= f'(a)g(a)+f(a)g'(a),\end{aligned}$$

da $g$ bei $a$ stetig ist. □

**Korollar 7.6** (Differenzierbarkeit von Polynomen)**.** *Reelle Polynome sind auf ganz $\mathbb{R}$ differenzierbar und es gilt*

$$(1)' = 0, \quad (x^n)' = nx^{n-1} \tag{7.3}$$

*für alle $n \in \mathbb{N}$.*

Nach Proposition 7.5 und Korollar 7.6 ist insbesondere die Ableitung eines Polynoms wieder ein Polynom. Weiters ist $f \in \mathbb{R}[x] \mapsto f' \in \mathbb{R}[x]$ eine lineare Abbildung.

*Beweis von Korollar 7.6.* Die Fälle $n = 0$ und $n = 1$ wurden bereits in Beispiel 7.3 besprochen. Wir beweisen (7.3) per Induktion nach $n$. Angenommen für $n \in \mathbb{N}$ gilt $(x^n)' = nx^{n-1}$. Dann folgt aus Proposition 7.5, dass $x^{n+1} = xx^n$ differenzierbar ist und

$$(x^{n+1})' = (xx^n)' = 1x^n + x(nx^{n-1}) = (n+1)x^n$$

erfüllt, was den Induktionsbeweis abschliesst. Differenzierbarkeit eines beliebigen Polynoms folgt nun aus der Linearität der Ableitung in Proposition 7.5. □

**Übung 7.7** (Potenzregel mittels Binomialsatz)**.** *Zeigen Sie Korollar 7.6 direkt unter Verwendung des Binomialsatzes. Beweisen Sie des Weiteren, dass der Kern der Abbildung $f \in \mathbb{R}[x] \mapsto f' \in \mathbb{R}[x]$ aus den konstanten Polynomen besteht. Später werden wir sehen, dass nicht nur Polynome mit Ableitung Null, sondern auch differenzierbare Funktionen auf $\mathbb{R}$ mit Ableitung Null konstant sein müssen.*

Wieder in Analogie zur Diskussion von stetigen Funktionen (genauer Proposition 3.53) wollen wir zeigen, dass die Verknüpfung zweier differenzierbaren Funktionen auch differenzierbar ist.

**Satz 7.8** (Kettenregel)**.** *Seien $D, E \subseteq \mathbb{R}$ Teilmengen und sei $x_0 \in D$ ein Häufungspunkt. Sei $f : D \to E$ eine bei $x_0$ differenzierbare Funktion, so dass $y_0 = f(x_0)$ ein Häufungspunkt von $E$ ist, und sei $g : E \to \mathbb{R}$ eine bei $y_0$ differenzierbare Funktion. Dann ist $g \circ f : D \to \mathbb{R}$ in $x_0$ differenzierbar und*

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

Wir bemerken, dass man zwar versucht sein mag, für den Beweis der Kettenregel den Differenzenquotienten

$$\frac{(g \circ f)(x) - (g \circ f)(x_0)}{x - x_0}$$

mit $f(x) - f(x_0)$ zu erweitern. Dies ist im Allgemeinen aber nicht erlaubt, da wir nicht ausschliessen können, dass $f(x) = f(x_0)$ für gewisse Punkte $x$ nahe bei $x_0$ ist.

*Beweis.* Wir verwenden stattdessen die Umformulierung

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + o(x - x_0)$$

für $x \to x_0$, oder genauer formuliert

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \varepsilon_f(x)(x - x_0),$$

wobei die Funktion $\varepsilon_f$ auf $D$ durch

$$\varepsilon_f(x) = \begin{cases} \frac{f(x)-f(x_0)}{x-x_0} - f'(x_0) & \text{falls } x \in D \setminus \{x_0\} \\ 0 & \text{falls } x = x_0 \end{cases}$$

für alle $x \in D$ gegeben ist und bei $x_0$ stetig ist. Ebenso gilt

$$g(y) = g(y_0) + g'(y_0)(y - y_0) + \varepsilon_g(y)(y - y0),$$

wobei die bei $y_0$ stetige Funktion $\varepsilon_g$ auf $E$ durch

$$\varepsilon_g(y) = \begin{cases} \frac{g(y)-g(y_0)}{y-y_0} - g'(y_0) & \text{falls } y \in E \setminus \{y_0\} \\ 0 & \text{falls } y = y_0 \end{cases}$$

für alle $y \in E$ gegeben ist. Zusammen ergibt sich durch Einsetzen von $y = f(x)$

$$\begin{aligned} g(f(x)) &= g(f(x_0)) + g'(f(x_0))(f(x) - f(x_0)) + \varepsilon_g(f(x))(f(x) - f(x_0)) \\ &= g(f(x_0)) + g'(f(x_0))f'(x_0)(x - x_0) \\ &\qquad + \big(g'(f(x_0))\varepsilon_f(x) + \varepsilon_g(f(x))(f'(x_0) + \varepsilon_f(x))\big)(x - x_0), \end{aligned}$$

für alle $x \in D$, womit

$$\begin{aligned} &\lim_{x \to x_0} \frac{(g \circ f)(x) - (g \circ f)(x_0)}{x - x_0} \\ &\quad = \lim_{x \to x_0} \big(g'(f(x_0))f'(x_0) + g'(f(x_0))\varepsilon_f(x) + \varepsilon_g(f(x))(f'(x_0) + \varepsilon_f(x))\big) \\ &\quad = g'(f(x_0))f'(x_0) \end{aligned}$$

wie gewünscht. □

Abgesehen von Summen, Produkten und Verknüpfungen von differenzierbaren Funktionen, möchten wir zeigen, dass Quotienten von differenzierbaren Funktionen differenzierbar sind. Wir beginnen dazu mit einem wichtigen Beispiel.

VO: 30.11.17

**Beispiel 7.9** (Kehrwert)**.** *Sei $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}, x \mapsto \frac{1}{x}$. Dann ist $f$ differenzierbar und es gilt $f'(x) = -\frac{1}{x^2}$ für alle $x \in \mathbb{R} \setminus \{0\}$. In der Tat ist*

$$f'(x) = \lim_{h\to 0} \frac{\frac{1}{x+h} - \frac{1}{x}}{h} = \lim_{h\to 0} \frac{x-(x+h)}{(x+h)xh} = -\lim_{h\to 0} \frac{1}{(x+h)x} = -\frac{1}{\lim_{h\to 0}(x+h)x} = -\frac{1}{x^2}$$

*wegen der Stetigkeit von $h \mapsto (x+h)x$ bei $0$.*

**Übung 7.10** (Negative Potenzen)**.** *Berechnen Sie $(x^{-n})'$ für alle $n \in \mathbb{N}$.*

Unter Kombination der Kettenregel und Beispiel 7.9 erhält man nun folgendes Korollar.

**Korollar 7.11** (Quotientenregel)**.** *Sei $D \subseteq \mathbb{R}$ eine Teilmenge, $a \in D$ ein Häufungspunkt und seien $f, g : D \to \mathbb{R}$ bei $a$ differenzierbar. Falls $g(a) \neq 0$ ist, dann ist auch $\frac{f}{g}$ bei $a$ differenzierbar und es gilt*

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}.$$

Man beachte, dass der (natürliche) Definitionsbereich der Funktion $\frac{f}{g}$, der in obigem Korollar nicht erwähnt wurde, die Teilmenge $E = \{x \in D \mid g(x) \neq 0\}$ ist. Da $g$ beim Punkt $a$ differenzierbar ist, ist $g$ bei $a$ stetig. Insbesondere ist, da $g(a) \neq 0$ ist, $g(x) \neq 0$ für alle $x$ nahe genug bei $a$ und $a$ ist ein Häufungspunkt von $E$. Damit macht es auch Sinn, von Differenzierbarkeit von $\frac{f}{g}$ bei $a$ zu sprechen.

Eine direkte Konsequenz von Korollar 7.11 ist, dass rationale Funktionen differenzierbar sind, wo definiert. Wir erinnern daran, dass eine rationale Funktion eine Funktion der Form $\frac{f(x)}{g(x)}$ ist, wobei $f(x)$ und $g(x)$ reelle Polynome sind und $g(x)$ nicht das Nullpolynom ist.

*Beweis von Korollar 7.11.* Es bezeichne $\psi$ die Funktion $y \in \mathbb{R}\setminus\{0\} \mapsto \frac{1}{y} \in \mathbb{R}$, welche nach Beispiel 7.9 differenzierbar ist. Wir kombinieren dies mit der Kettenregel (Satz 7.8) und erhalten, dass die Funktion $\frac{1}{g} = \psi \circ g$ bei $a$ differenzierbar ist mit Ableitung

$$\left(\frac{1}{g}\right)'(a) = -\frac{1}{g(a)^2} g'(a).$$

Verwenden wir nun die Produktregel in Proposition 7.5, so ergibt sich, dass $\frac{f}{g} = f \cdot \frac{1}{g}$ bei $a$ differenzierbar ist und

$$\left(\frac{f}{g}\right)'(a) = \left(f \cdot \frac{1}{g}\right)'(a) = f'(a)\frac{1}{g(a)} - f(a)\frac{g'(a)}{g(a)^2} = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}$$

erfüllt, was zu zeigen war. □

Die Kettenregel erlaubt uns die Berechnung der Ableitung von beliebig kompliziert anmutenden konkreten Beispielen, wobei man stur von aussen nach innen vorgeht wie in folgendem Beispiel.

**Beispiel 7.12** (Vierfach verschachtelte Funktionen)**.** *Wir bestimmen die Ableitung der Funktion*

$$f : x \in \mathbb{R} \mapsto \exp(\sin(\sin(x^2)))$$

*mittels mehrmaligem Anwenden der Kettenregel (Satz 7.8). Da* $\exp' = \exp$ *erhalten wir*

$$f'(x) = \exp(g(x))g'(x),$$

*wobei* $g(x) = \sin(\sin(x^2))$*. Ebenso ist wegen* $\sin' = \cos$

$$g'(x) = \cos(h(x))h'(x),$$

*wobei* $h(x) = \sin(x^2)$ *und* $h'(x) = \cos(x^2)2x$*. Dadurch erhalten wir*

$$f'(x) = \exp(\sin(\sin(x^2)))\cos(\sin(x^2))\cos(x^2)2x$$

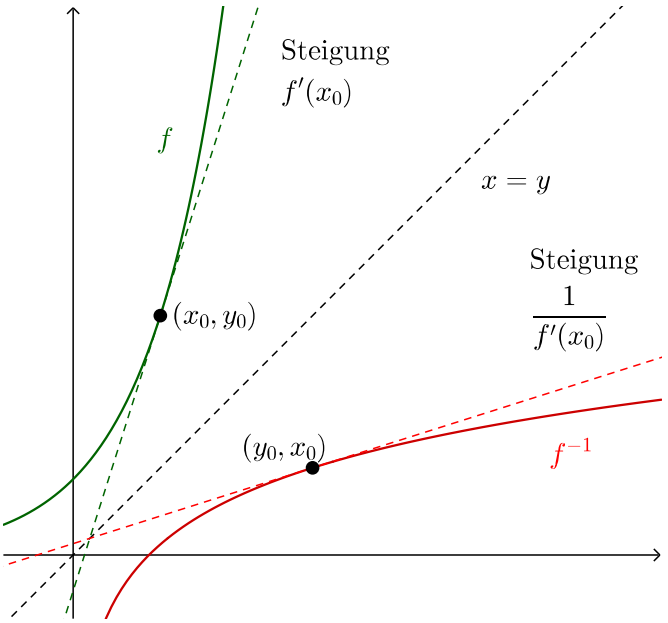
*für alle* $x \in \mathbb{R}$*.*

**Übung 7.13** (Nochmals vierfach verschachtelt)**.** *Bestimmen Sie die Ableitung von der Funktion* $x \in \mathbb{R} \mapsto \cos((\sin(\exp(x)))^3)$*.*

Unsere vorläufig letzte allgemeine Ableitungsregel betrifft die Ableitung der Umkehrabbildung (siehe dazu auch Satz 3.65 über die Existenz einer stetigen Umkehrabbildung).

**Satz 7.14** (Differenzierbarkeit der inversen Funktion)**.** *Seien* $D, E \subseteq \mathbb{R}$ *Teilmengen und sei* $f : D \to E$ *eine stetige, bijektive Abbildung, deren inverse Abbildung* $f^{-1} : E \to D$ *ebenfalls stetig ist. Falls* $f$ *in dem Häufungspunkt* $x_0 \in D$ *differenzierbar ist und* $f'(x_0) \neq 0$ *gilt, dann ist* $f^{-1}$ *in* $y_0 = f(x_0)$ *differenzierbar und es gilt*

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)}$$




Figur 7.2: Eine intuitive Darstellung von Satz 7.14. Spiegelt man den Graphen von $f$ und die Tangente beim Punkt $(x_0, y_0)$ um die Gerade $x = y$ in $\mathbb{R}^2$, so erhält man den Graphen von $f^{-1}$ und, das ist die Behauptung, die Tangente bei $(y_0, x_0)$. Eine kurze Rechnung zeigt, dass die Spiegelung einer Gerade mit Steigung $m$ um $x = y$ Steigung $\frac{1}{m}$ hat.

*Beweis.* Wir bemerken zuerst, dass $y_0$ ein Häufungspunkt von $E$ ist, womit man von Differenzierbarkeit bei $y_0$ sprechen darf. Tatsächlich ist nach Annahme $x_0$ ein Häufungspunkt und es existiert eine Folge $(x_n)_n$ in $D \setminus \{x_0\}$ mit $x_n \to x_0$ für $n \to \infty$. Da $f$ stetig ist, gilt $f(x_n) \to f(x_0) = y_0$ für $n \to \infty$ und da $f$ bijektiv ist, gilt $f(x_n) \neq y_0$ für alle $n \in \mathbb{N}$.

Sei nun $(y_n)_n$ eine Folge in $D \setminus \{y_0\}$, die gegen $y_0$ konvergiert. Dann strebt $x_n = f^{-1}(y_n)$ in $D \setminus \{x_0\}$ gegen $x_0$, da $f^{-1}$ per Annahme stetig ist, und es gilt

$$\lim_{n\to\infty} \frac{f^{-1}(y_n) - f^{-1}(y_0)}{y_n - y_0} = \lim_{n\to\infty} \frac{x_n - x_0}{y_n - y_0} = \lim_{n\to\infty} \left( \frac{f(x_n) - f(x_0)}{x_n - x_0} \right)^{-1} = (f'(x_0))^{-1}$$

nach der Charakterisierung der Konvergenz einer Funktion mittels Folgen in Lemma 5.65. Da dies aber für jede Folge $(y_n)_n$ wie oben gilt, folgt der Satz wiederum aus Lemma 5.65. □

**Beispiel 7.15** (Differenzierbarkeit des Logarithmus und der Potenzfunktionen)**.**

*(i) Die Funktion $g : y \in \mathbb{R} \setminus \{0\} \mapsto \log(|y|) \in \mathbb{R}$ ist differenzierbar mit Ableitung $g'$ gegeben durch $g'(y) = \frac{1}{y}$ für alle $y \in \mathbb{R} \setminus \{0\}$. Denn die Abbildung $\log : y \in \mathbb{R}_{>0} \mapsto \log(y) = g(y)$ ist die Umkehrabbildung von $\exp : \mathbb{R} \to \mathbb{R}_{>0}$ und damit folgt aus Satz 7.14, dass $g$ bei allen Punkten $y > 0$ differenzierbar ist mit $g'(y) = \frac{1}{f'(x)}$, wobei $x = g(y) = \log(y)$. Da $\exp' = \exp$ folgt nun*

$$g'(y) = \log'(y) = \frac{1}{\exp(x)} = \frac{1}{\exp(\log(y))} = \frac{1}{y}.$$

*Für $y < 0$ ist $g(y) = \log(-y)$. Also folgt Differenzierbarkeit von $g$ bei $y$ sowie die Formel $g'(y) = -\log'(-y) = -\frac{1}{-y} = \frac{1}{y}$ aus der Kettenregel (Satz 7.8).*

*(ii) Für ein beliebiges $s \in \mathbb{C}$ ist die Abbildung $x \in \mathbb{R}_{>0} \mapsto x^s$ differenzierbar und es gilt*

$$(x^s)' = sx^{s-1}.$$

*In der Tat gilt $x^s = \exp(s\log(x))$ für alle $x > 0$ per Definition beliebiger Potenzen in Abschnitt 6.5.2. Aus Beispiel 7.3 und der Ableitung der Logarithmusabbildung folgt somit*

$$(x^s)' = \exp(s\log(x))' = \exp(s\log(x))s\frac{1}{x} = x^s sx^{-1} = sx^{s-1}$$

*für alle $x > 0$.*

### 7.1.3 Extremwerte

Wie wir in diesem Abschnitt sehen werden, ist die Ableitung auch nützlich, um Punkte zu finden, bei denen eine Funktion $f$ ihre Maxima und ihre Minima annimmt. Genauer kann man damit die in folgender Definition eingeführten Punkte finden.

**Definition 7.16** (Lokale Extremwerte)**.** Sei $D \subseteq \mathbb{R}$ eine Teilmenge und $x_0 \in D$. Wir sagen, dass eine Funktion $f : D \to \mathbb{R}$ in $x_0$ ein **lokales Maximum** hat, falls es eine Umgebung $U$ von $x_0$ in $D$ gibt, auf der $f$ durch $f(x_0)$ beschränkt ist. Genauer formuliert heisst dies, dass es ein $\delta > 0$ gibt, so dass für alle $x \in D \cap (x_0 - \delta, x_0 + \delta)$ gilt $f(x) \leq f(x_0)$. Falls es sogar ein $\delta > 0$ gibt, so dass $f(x) < f(x_0)$ für alle $x \in D \cap (x_0 - \delta, x_0 + \delta) \setminus \{x_0\}$ gilt, dann hat $f$ in $x_0$ ein **isoliertes lokales Maximum**. Der Wert $f(x_0)$ wird auch ein **lokaler Maximalwert** von $f$ genannt. Ein **lokales Minimum**, ein **isoliertes lokales Minimum** und ein **lokaler Minimalwert** von $f$ sind analog definiert.

Des Weiteren nennen wir $x_0$ ein **lokales Extremum** von $f$ und $f(x_0)$ einen **lokalen Extremwert** von $f$, falls $f$ ein lokales Minimum oder ein lokales Maximum in $x_0$ hat.

**Proposition 7.17** (Notwendige Bedingung für Extremum)**.** *Sei $D \subseteq \mathbb{R}$ eine Teilmenge und $f$ eine reellwertige Funktion auf $D$. Angenommen $f$ ist in einem lokalen Extremum $x_0 \in D$ differenzierbar und $x_0$ ist sowohl ein rechtsseitiger als auch ein linksseitiger Häufungspunkt von $D$. Dann gilt $f'(x_0) = 0$.*

Die Annahme in Proposition 7.17, dass sich die Menge $D$ dem Extremum $x_0 \in D$ sowohl von links als auch von rechts nähert, ist notwendig, da wir $f'(x_0)$ von links und von rechts mit Differenzenquotienten approximieren wollen. In konkreten Rechenbeispielen ist sie jedoch meist erfüllt. Beispielsweise ist dies in allen Punkten eines Intervalls abgesehen von den Endpunkten erfüllt.

*Beweis.* Ohne Beschränkung der Allgemeinheit nehmen wir an, dass $f$ ein lokales Maximum in $x_0 \in D$ hat (sonst ersetzt man $f$ durch $-f$). Da $f$ bei $x_0$ differenzierbar ist und $x_0$ von links und rechts angenähert werden kann, existieren sowohl der linksseitige als auch der rechtsseitige Grenzwert der Differenzenquotienten bei $x_0$ und beide sind gleich $f'(x_0)$. Dann ist

$$f'(x_0) = f'_+(x_0) = \lim_{x \searrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \leq 0,$$

da $f(x) \leq f(x_0)$ für alle $x$ hinreichend nahe bei $x_0$ gilt und $x > x_0$ für die Bewegung $x \searrow x_0$ erfüllt ist. Weiters ist aber auch

$$f'(x_0) = f'_-(x_0) = \lim_{x \nearrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \geq 0,$$

da wiederum $f(x) \leq f(x_0)$ für alle $x$ hinreichend nahe bei $x_0$ gilt und $x < x_0$ für die Bewegung $x \nearrow x_0$ erfüllt ist. Unter dem Strich erhalten wir $f'(x_0) = 0$. □

Falls der Definitionsbereich $D$ ein Intervall ist, so besagt Proposition 7.17 das Folgende.

**Korollar 7.18** (Lokale Extremwerte)**.** *Sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$. Sei $x_0 \in I$ ein lokales Extremum von $f$. Dann bestehen genau folgende Möglichkeiten:*

*(i) $x_0$ ist ein in $I$ enthaltener Endpunkt von $I$,*

*(ii) $f$ ist bei $x_0$ nicht differenzierbar oder*

*(iii) $f$ ist bei $x_0$ differenzierbar und $f'(x_0) = 0$.*

*Insbesondere sind alle lokalen Extrema einer differenzierbaren Funktion auf einem offenen Intervall Nullstellen der Ableitung.*

Man beachte, dass alle Fälle in obigem Korollar eintreten können (wieso?). Des Weiteren ist die Umkehrung von Proposition 7.17 nicht richtig, wie wir in folgender Übung zeigen wollen.

**Übung 7.19.** *(a) Finden Sie alle lokalen Extrema des Polynoms $f(x) = x^3 - x$ auf $\mathbb{R}$.*

*(b) Finden Sie alle lokalen Extrema der Funktion $|f|$ auf $[-3, 3]$.*

### 7.1.4 Stetige Differenzierbarkeit

Sei $D \subseteq \mathbb{R}$ eine Teilmenge, so dass jeder Punkt in $D$ ein Häufungspunkt von $D$ ist (wie zum Beispiel bei einem Intervall mit Endpunkten $a < b$ in $\overline{\mathbb{R}}$). Falls $f : D \to \mathbb{R}$ eine differenzierbare Funktion ist (also bei jedem Punkt in $D$ differenzierbar ist), können wir die Ableitung

$$f' : x \in D \to f'(x)$$

als eine neue Funktion betrachten. Ist $f'$ stetig, so nennen wir $f$ **stetig differenzierbar**.

Man beachte, dass eine differenzierbare Funktion nicht zwingend stetig differenzierbar sein muss. Wir illustrieren dies in einem Beispiel.

**Beispiel 7.20** (Unstetige Ableitung)**.** *Wir betrachten zu $p > 0$ die Abbildung $f : \mathbb{R} \to \mathbb{R}$ definiert durch*

$$f_p(x) = \begin{cases} |x|^p \sin\left(\frac{1}{x}\right) & \text{falls } x \neq 0 \\ 0 & \text{falls } x = 0 \end{cases}$$

*für alle $x \in \mathbb{R}$. Da $p > 0$ ist, ist $f_p$ auch bei $0$ stetig (siehe das Sandwich Lemma 5.74). Die Ableitung von $f$ bei $x \neq 0$ existiert und ist durch*

$$f_p'(x) = p|x|^{p-1}\operatorname{sgn}(x)\sin\left(\tfrac{1}{x}\right) - |x|^{p-2}\cos\left(\tfrac{1}{x}\right)$$

*gegeben. Hierbei verwendeten wir auch, dass die Ableitung von $x \in \mathbb{R}^\times \mapsto |x|$ durch $x \in \mathbb{R}^\times \mapsto \operatorname{sgn}(x)$ gegeben ist. Für die Ableitung von $f$ bei $0$ können wir keine allgemeine Ableitungsregel verwenden und manipulieren deswegen den Grenzwert*

$$\lim_{x\searrow 0}\frac{f_p(x)-f_p(0)}{x-0} = \lim_{x\searrow 0}\frac{x^p\sin\left(\frac{1}{x}\right)}{x} = \lim_{x\searrow 0} x^{p-1}\sin\left(\tfrac{1}{x}\right) = \lim_{t\to\infty}\sin(t)t^{1-p}$$

*wobei wir $t = \frac{1}{x}$ gesetzt haben. Falls $p \leq 1$ ist, dann existiert wegen $1-p \geq 0$ der Grenzwert $\lim_{t\to\infty}\sin(t)t^{1-p}$ nicht und somit ist $f_p$ nicht differenzierbar. Wir nehmen nun $p > 1$ an, womit $f$ (ebenso auf Grund des Sandwich Lemmas) eine rechtsseitige Ableitung $(f_p)'_+(0) = 0$ besitzt. Analog können wir den Grenzwert mit $(f_p)'_-(0) = 0$ berechnen und erhalten drei Fälle.*

- *Falls $p < 2$ ist, ist $f_p'$ in jeder Umgebung von $0$ unbeschränkt und insbesondere nicht stetig bei $0$. Somit ist $f_p$ nicht stetig differenzierbar.*
- *Falls $p = 2$ ist, ist $f_p'$ beschränkt, aber nicht stetig bei $0$, da der Grenzwert $\lim_{x\to 0}\cos\left(\frac{1}{x}\right)$ nicht existiert.*
- *Falls $p > 2$ ist, ist $f_p'$ stetig und $f_p$ ist stetig differenzierbar. Man beachte aber, dass $f_p'$ nicht differenzierbar sein muss, da für $p \leq 3$ der Grenzwert*

  $$\lim_{x\searrow 0}\frac{f_p'(x)-f_p'(0)}{x-0} = \lim_{x\searrow 0}\frac{px^{p-1}\sin\left(\frac{1}{x}\right) - x^{p-2}\cos\left(\frac{1}{x}\right)}{x} = \lim_{t\to\infty}\left(pt^{2-p}\sin(t) - t^{3-p}\cos(t)\right)$$

  *nicht existiert.*

Das Beispiel 7.20 lässt sich mit fraktalen Konstruktionen stark verschärfen. In der Tat kann man eine differenzierbare Funktion auf dem Intervall $[0,1]$ finden, deren Ableitung überabzählbar viele Unstetigkeitsstellen besitzt (beispielsweise auf der Cantor-Menge). Eine Konstruktion dieser Art finden Sie in Abschnitt 7.6.2.

### 7.1.5 Ableitungen höherer Ordnung

Sei $D \subseteq \mathbb{R}$ eine Teilmenge, so dass jeder Punkt in $D$ ein Häufungspunkt ist, und sei $f : D \to \mathbb{R}$ eine Funktion. Falls $f'$ existiert und differenzierbar ist, nennen wir $f$ **zweimal differenzierbar**. Die Funktion $(f')'$ ist die **zweite Ableitung** von $f$ und wird auch mit $f''$, $f^{(2)}$ oder $\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}$ bezeichnet. Falls die unabhängige Variable $t$ ist, schreiben wir $\ddot{f} = (\dot{f})^{\cdot} = \frac{\mathrm{d}^2 f}{\mathrm{d}t^2}$ für die zweite Ableitung nach $t$. Insbesondere erhalten wir, dass eine zweimal differenzierbare Funktion $f$ stetig differenzierbar ist.

Induktiv kann man nun höhere Differenzierbarkeit und höhere Ableitungen definieren. Formal definieren wir also die Ableitungen

$$f^{(0)} = f, \quad f^{(1)} = \frac{\mathrm{d}f}{\mathrm{d}x} = f', \quad f^{(2)} = \frac{\mathrm{d}^2 f}{\mathrm{d}x^2} = f'', \quad \dots \quad , \quad f^{(n+1)} = \frac{\mathrm{d}^{(n+1)} f}{\mathrm{d}x^{n+1}} = (f^{(n)})'$$

für alle $n \in \mathbb{N}$. Falls $f^{(n)}$ für ein $n \in \mathbb{N}$ (auf ganz $D$) existiert, heisst $f$ **$n$-mal differenzierbar**. Falls die **$n$-te Ableitung** $f^{(n)}$ zusätzlich stetig ist, heisst $f$ **$n$-mal stetig differenzierbar**. Die Menge der $n$-mal stetig differenzierbaren Funktionen auf $D$ bezeichnen wir mit $C^n(D)$.

Für jedes $n \in \mathbb{N}$ kann man eine Funktion finden, die zwar $n$-mal differenzierbar, aber nicht $(n+1)$-mal differenzierbar ist.

**Übung 7.21.** *Sei $n \in \mathbb{N}$. Zeigen Sie, dass die Funktion $x \in \mathbb{R} \mapsto x^n|x| \in \mathbb{R}$ $n$-mal stetig differenzierbar, aber nicht $(n+1)$-mal differenzierbar ist.*

Wir sagen, dass $f$ **glatt** oder **beliebig oft differenzierbar** ist, falls $f$ für jedes $n \in \mathbb{N}$ $n$-mal differenzierbar ist. Ist $f$ glatt, so sind insbesondere alle Ableitungen von $f$ stetig ($f$ ist also beliebig oft stetig differenzierbar). Die Menge der glatten Funktionen auf $D$ bezeichnen wir mit $C^\infty(D)$.

Wir kennen bereits einige Beispiele glatter Funktionen. Dazu gehören die Polynome, da diese nach Korollar 7.6 differenzierbar sind und da deren Ableitung ein Polynom ist, womit die Aussage aus Induktion folgt. Ebenfalls glatt sind die Funktion $\exp, \sin, \cos, \sinh, \cosh$ nach Beispiel 7.3 und Übung 7.4. Etwas interessanter, aber nicht ganz unerwartet ist vermutlich folgendes Beispiel.

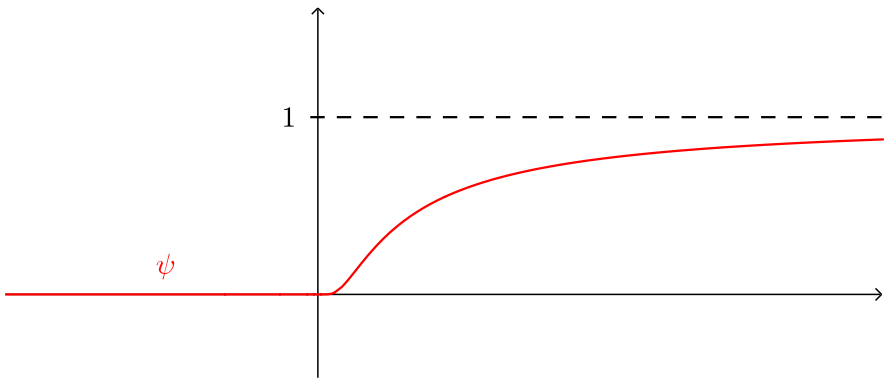
**Beispiel 7.22** (Logarithmusfunktion)**.** *Der Logarithmus $f = \log : (0, \infty) \to \mathbb{R},\ x \to \log(x)$ ist glatt. In der Tat gilt $f'(x) = \frac{1}{x}$, $f''(x) = -\frac{1}{x^2}$, $f^{(3)}(x) = \frac{2}{x^3}$ oder allgemein $f^{(n)}(x) = (-1)^{n-1}(n-1)!x^{-n}$, was sich mit vollständiger Induktion beweisen lässt.*

Ein überraschenderes Beispiel einer glatten Funktion ist vielleicht das folgende.

**Beispiel 7.23** (Glattes Abklingen)**.** *Die Funktion $\psi : \mathbb{R} \to \mathbb{R}$ definiert durch*

$$\psi(x) = \begin{cases} 0 & \text{falls } x \leq 0 \\ \exp\left(-\frac{1}{x}\right) & \text{falls } x > 0 \end{cases}$$

*für alle $x \in \mathbb{R}$ ist glatt und demnach auch beliebig oft stetig differenzierbar, siehe das folgende Bild.*



*Für $x < 0$ gibt es nichts zu zeigen, da die Ableitung der Nullfunktion die Nullfunktion ist. Für $x > 0$ ergibt sich dies mittels Induktion, der Kettenregel (Satz 7.8), Beispiel 7.9, der Produktregel in Proposition 7.5 und Korollar 7.6. In der Tat gilt für $x > 0$, dass*

$$\psi'(x) = \exp\left(-\frac{1}{x}\right)\frac{1}{x^2}, \quad \psi''(x) = \exp\left(-\frac{1}{x}\right)\frac{1}{x^2}\frac{1}{x^2} + \exp\left(-\frac{1}{x}\right)\frac{-2}{x^3}$$

*und (da die konkrete Formel für $\psi^{(n)}$ schnell kompliziert wird) allgemeiner*

$$\psi^{(n)}(x) = \exp\left(-\frac{1}{x}\right) f_n\left(\frac{1}{x}\right) \tag{7.4}$$

*für gewisse Polynome $f_n$ und jedes $n \in \mathbb{N}$. Für $n = 1$ und $n = 2$ haben wir diese Darstellung der Ableitung bereits bewiesen, wobei $f_1(t) = t^2$ und $f_2(t) = t^4 - 2t^3$. Für den Induktionsschritt nehmen wir (7.4) für $n \in \mathbb{N}$ an und erhalten*

$$\begin{aligned}\psi^{(n+1)}(x) &= \left(\exp\left(-\frac{1}{x}\right) f_n\left(\frac{1}{x}\right)\right)' = \exp\left(-\frac{1}{x}\right)\frac{1}{x^2} f_n\left(\frac{1}{x}\right) + \exp\left(-\frac{1}{x}\right) f_n'\left(\frac{1}{x}\right)\frac{-1}{x^2} \\ &= \exp\left(-\frac{1}{x}\right) f_{n+1}\left(\frac{1}{x}\right),\end{aligned}$$

*wobei das Polynom $f_{n+1}$ als $f_{n+1}(t) = t^2(f_n(t) - f_n'(t))$ gewählt wurde.*

*Es bleibt noch zu zeigen, dass $\psi$ auch in $x = 0$ beliebig oft differenzierbar ist. Dabei können wir nicht auf unsere Ableitungsregeln zurückgreifen, sondern müssen dies direkt mit der Definition der Ableitung überprüfen. Wir behaupten, dass $\psi^{(n)}(0) = 0$ für alle $n \in \mathbb{N}$.*

*Für den Beweis der Behauptung zeigen wir zuerst, dass für jedes Polynom $f$*

$$\lim_{x\to 0} \psi(x) f(\tfrac{1}{x}) = 0 \tag{7.5}$$

*ist. Auf Grund der Linearität des Grenzwerts und da $\psi(x) = 0$ für $x < 0$ gilt, genügt es zu zeigen, dass $\lim_{x\searrow 0} \psi(x)x^{-n} = 0$ für alle $n \in \mathbb{N}$ gilt. Setzen wir $y = \frac{1}{x}$, so erhalten wir, dass diese Behauptung wiederum zu*

$$\lim_{y\to\infty} \frac{y^n}{\exp(y)} = 0$$

*äquivalent ist. Dies folgt aber mit dem Sandwich-Lemma aus der Ungleichung $(1 + \frac{y}{n+1})^{n+1} \leq \exp(y)$ für alle $y \geq 0$ und $n \in \mathbb{N}$ (siehe Abschnitt 5.3).*

*Wir zeigen nun $\psi^{(n)}(0) = 0$ für alle $n \in \mathbb{N}$ per Induktion. Verwenden wir (7.5), so erhalten wir*

$$\psi'(0) = \lim_{x\to 0} \frac{\psi(x) - 0}{x} = \lim_{x\to 0} \psi(x)\frac{1}{x} = 0.$$

*Falls wir bereits $\psi^{(n)}(0) = 0$ für ein $n \in \mathbb{N}$ wissen, dann folgt ebenso*

$$\psi^{(n+1)}(0) = \lim_{x\to 0} \frac{\psi^{(n)}(x) - \psi^{(n)}(0)}{x - 0} = \lim_{x\to 0} \frac{\psi(x) f_n(\frac{1}{x}) - 0}{x} = \lim_{x\to 0} \psi(x) f_n\left(\frac{1}{x}\right)\frac{1}{x} = 0.$$

*Wir haben nun also gezeigt, dass alle Ableitungen von $\psi$ auf ganz $\mathbb{R}$ existieren und somit ist $\psi$ glatt.*

**Übung 7.24** (Hutfunktion)**.** *Finden Sie für beliebige reelle Zahlen $a < b < c < d$ eine glatte Funktion $\varphi$ auf $\mathbb{R}$, so dass $\varphi$ gleich Null ist ausserhalb des Intervalls $(a, d)$ und gleich $1$ ist auf dem Intervall $[b, c]$.*

Wir wenden uns nun wieder allgemeinen Aussagen im Stile von Abschnitt 7.1.2 zu. Aus Proposition 7.5 lässt sich folgendes Korollar deduzieren.

**Korollar 7.25** (Summen und Produkte bei höherer Differenzierbarkeit)**.** *Sei $D \subseteq \mathbb{R}$ eine Teilmenge, so dass jeder Punkt in $D$ ein Häufungspunkt von $D$ ist. Seien $f, g : D \to \mathbb{R}$ $n$-mal differenzierbar. Dann sind $f + g$ und $f \cdot g$ ebenso $n$-mal differenzierbar und es gilt $f^{(n)} + g^{(n)} = (f + g)^{(n)}$ sowie*

$$(fg)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}.$$

*Insbesondere ist jedes skalare Vielfache $n$-mal differenzierbar und $(\alpha f)^{(n)} = \alpha f^{(n)}$ für alle $\alpha \in \mathbb{R}$.*

Die obige Produktregel für höhere Ableitungen nennt sich auch **Leibniz-Regel**.

Natürlich sind auch Verknüpfungen von $n$-mal differenzierbaren Funktionen $n$-mal differenzierbar. Allerdings ist es im Gegensatz zum Produkt deutlich schwerer, hier eine explizite Formel anzugeben. Wir beschränken uns deswegen darauf, nur die Differenzierbarkeit zu formulieren.

**Korollar 7.26** (Verknüpfungen und höhere Differenzierbarkeit)**.** *Seien $D, E \subseteq \mathbb{R}$ Teilmengen, so dass jeder Punkt in $D$ respektive $E$ ein Häufungspunkt von $D$ respektive $E$ ist. Sei des Weiteren $f : D \to E$ eine $n$-mal differenzierbare Funktion und sei $g : E \to \mathbb{R}$ eine $n$-mal differenzierbare Funktion. Dann ist $g \circ f : D \to \mathbb{R}$ $n$-mal differenzierbar.*

**Übung 7.27.** *Beweisen Sie die Korollare 7.25 und 7.26.*

VO: 04.12.17

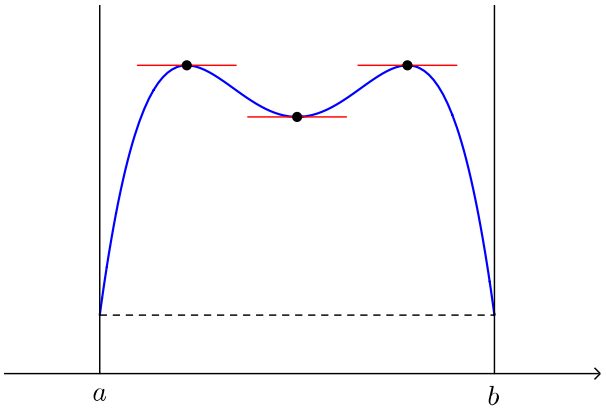
## 7.2 Zentrale Sätze der Differentialrechnung

### 7.2.1 Der Mittelwertsatz

Wir wenden uns nun allgemeinen Sätzen der Differentialrechnung und deren Konsequenzen zu. Unsere erste Frage wird sein, ob die Ableitung einer differenzierbaren Funktion die Steigung gewisser Sekanten annimmt, wobei folgender Satz unser Ausgangspunkt sein wird.

**Satz 7.28** (Rolle)**.** *Sei $[a, b]$ ein kompaktes Intervall mit $a < b$ und $f : [a, b] \to \mathbb{R}$ eine stetige Funktion, die auf dem offenen Intervall $(a, b)$ differenzierbar ist. Falls $f(a) = f(b)$ gilt, so existiert ein $\xi \in (a, b)$ mit $f'(\xi) = 0$.*

In Worten besagt der Satz von Rolle also, dass wenn eine „schöne“ Funktion auf einem Intervall an den Endpunkten den selben Wert annimmt, die Steigung irgendwo (strikt) zwischen den Endpunkten Null sein muss. Wir veranschaulichen dies in folgendem Bild, das bereits einen Hinweis enthält, wie wir im Beweis vorgehen wollen.




*Beweis.* Nach dem Extremwertsatz (Korollar 3.72) werden Minimum und Maximum von $f$ auf $[a, b]$ angenommen. Das heisst, es existieren $x_{\min}, x_{\max} \in [a, b]$ mit

$$f(x_{\min}) = \min f([a, b]), \quad f(x_{\max}) = \max f([a, b]).$$

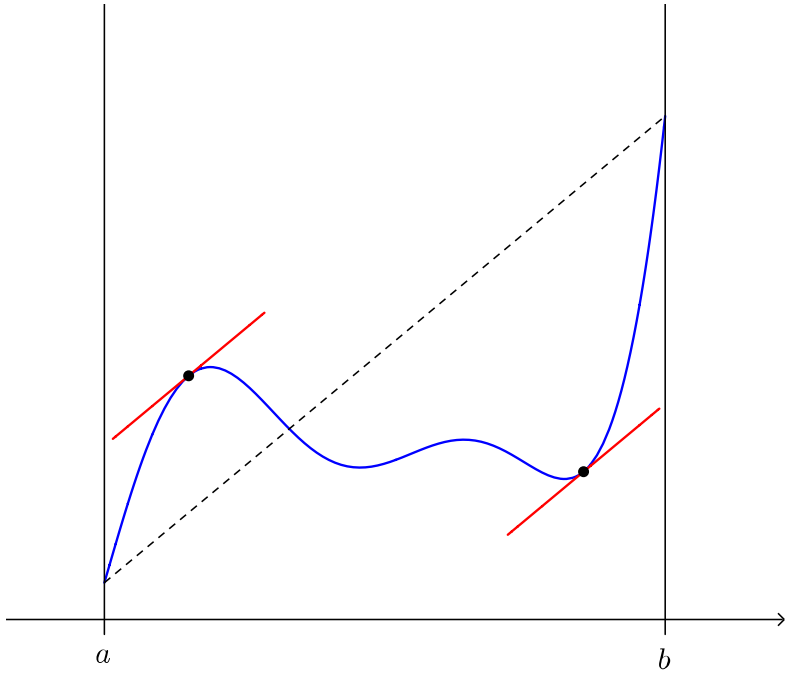
Nach Proposition 7.17 muss die Ableitung von $f$ bei allen Extrema in $(a, b)$ Null sein. Falls also $x_{\min} \in (a, b)$ oder $x_{\max} \in (a, b)$ gilt, dann haben wir bereits ein $\xi \in (a, b)$ gefunden mit $f'(\xi) = 0$ (wobei $\xi = x_{\min}$ oder $\xi = x_{\max}$ ist). Falls aber $x_{\min}$ und $x_{\max}$ Endpunkte des Intervalles sind, dann muss wegen $f(a) = f(b)$ auch $f(x_{\min}) = f(x_{\max})$ gelten, womit die Funktion $f$ konstant und $f'(x) = 0$ für alle $x \in (a, b)$ ist. □

Der Satz von Rolle führt unmittelbar zu folgendem fundamentalen Satz.

**Theorem 7.29** (Mittelwertsatz)**.** *Sei $[a, b]$ ein kompaktes Intervall mit $a < b$ und $f : [a, b] \to \mathbb{R}$ eine stetige Funktion, die auf dem offenen Intervall $(a, b)$ differenzierbar ist. Dann gibt es ein $\xi \in (a, b)$ mit*

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

Somit gibt es also mindestens einen Punkt $\xi$, an dem die Steigung $f'(\xi)$ der durchschnittlichen Steigung $\frac{f(b)-f(a)}{b-a}$, also der Steigung der Sekante durch $(a, f(a))$ und $(b, f(b))$ entspricht. Wir stellen dies in einem Bild dar.




Der Beweis des Mittelwertsatzes adaptiert die Werte von $f$ auf eine (affin) lineare Weise, um danach den Satz von Rolle (Satz 7.28) anwenden zu können. In der Tat behandelt der Satz von Rolle genau den Spezialfall $f(a) = f(b)$. Wenn wir in obigem Bild die Funktion rechts „nach unten ziehen“, können wir genau diesen Spezialfall verwenden.

*Beweis.* Wir definieren eine Funktion $F : [a, b] \to \mathbb{R}$ durch

$$F(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a)$$

für alle $x \in [a, b]$. Dann gilt $F(a) = f(a)$ und $F(b) = f(b) - (f(b) - f(a)) = f(a)$. Des Weiteren ist $F$ stetig an den Endpunkten und differenzierbar auf $(a, b)$ nach Proposition 7.5. Nach dem Satz von Rolle (Satz 7.28) existiert also ein $\xi \in (a, b)$ mit

$$0 = F'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{b - a}$$

wie gewünscht. □

**Übung 7.30** (Intervall als Voraussetzung)**.** *Zeigen Sie anhand eines Beispiels die Notwendigkeit der Voraussetzung im Mittelwertsatz (Theorem 7.29), dass der Definitionsbereich $D$ der Funktion $f : D \to \mathbb{R}$ ein Intervall ist.*

**Übung 7.31** (Lipschitz-Stetigkeit differenzierbarer Funktionen)**.** *Sei $[a, b]$ ein kompaktes Intervall mit Endpunkten $a < b$ und sei $f : [a, b] \to \mathbb{R}$ stetig differenzierbar. Zeigen Sie, dass $f$ Lipschitz-stetig ist. Was geschieht, wenn man Kompaktheit fallen lässt, das heisst, wenn man $a = -\infty$ oder $b = \infty$ zulässt?*

Der Mittelwertsatz (Theorem 7.29) wird zu einem zentralen Tool für die folgenden Diskussionen werden. Da wir die Ableitung auch für komplexwertige Funktionen definiert haben und in Beispiel 7.3(iii) auch schon für eine spezielle Funktion berechnet haben, wollen wir in folgendem Beispiel zeigen, dass der Mittelwertsatz *nur für reellwertige Funktionen* und nicht für komplexwertige Funktionen zutrifft.

**Beispiel 7.32** (Kreisparametrisierung)**.** *Sei $\gamma : [0, 2\pi] \to \mathbb{C}$ die Abbildung (besser: die Kurve) gegeben durch*

$$\gamma : t \in [0, 2\pi] \mapsto e^{t\mathrm{i}} = \cos(t) + \mathrm{i}\sin(t)$$

*Nach Beispiel 7.3(iii) ist die Ableitung gegeben durch $\gamma'(t) = \mathrm{i}e^{t\mathrm{i}}$. An den Endpunkten des Intervalles $[0, 2\pi]$ gilt $\gamma(0) = \gamma(2\pi) = 1$. Die Ableitung von $\gamma$ nimmt jedoch nie den Wert Null an, denn es gilt $|\gamma'(\xi)| = 1$ für alle $\xi \in [0, 2\pi]$. Somit können die Aussagen des Satzes von Rolle (Satz 7.28) und des Mittelwertsatzes (Theorem 7.29) für komplexwertige Funktionen in dieser Allgemeinheit nicht zutreffen.*

### 7.2.2 Korollare des Mittelwertsatzes und Kurvendiskussion

Der Mittelwertsatzes erlaubt es uns nun, uns bekannte Eigenschaften von Funktionen mittels der Ableitung zu charakterisieren.

**Korollar 7.33** (Kriterium für Konstanz)**.** *Sei $I \subseteq \mathbb{R}$ ein Intervall mit Endpunkten $a < b$ und $f : I \to \mathbb{R}$ eine Funktion. Dann ist $f$ genau dann konstant, wenn $f$ differenzierbar ist und $f'(x) = 0$ für alle $x \in I$ gilt.*

*Beweis.* Wir wissen bereits, dass die Ableitung einer konstanten Funktion die Nullfunktion ist. Also angenommen es gilt $f'(x) = 0$ für alle $x \in I$. Seien $x_1 < x_2$ in $I$. Dann folgt aus dem Mittelwertsatz (Theorem 7.29) angewendet auf $f|_{[x_1,x_2]} : [x_1, x_2] \to \mathbb{R}$, dass es ein $\xi \in (x_1, x_2) \subseteq I$ gibt mit $f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1)$. Es gilt aber $f'(\xi) = 0$ und damit $f(x_1) = f(x_2)$. Die Punkte $x_1, x_2 \in I$ waren jedoch beliebig, also folgt die Aussage. □

**Übung 7.34** (Charakterisierung von Polynomen)**.** *Sei $I \subseteq \mathbb{R}$ ein Intervall mit Endpunkten $a < b$ und $f : I \to \mathbb{R}$ eine Funktion. Zeigen Sie, dass $f$ genau dann ein Polynom ist, wenn $f$ glatt ist und es ein $n \in \mathbb{N}$ gibt mit $f^{(n)} = 0$.*

Im Folgenden werden wir den Mittelwertsatz, wie schon im obigen Beweis von Korollar 7.33, oft auf zwei verschiedene Punkte $x_1, x_2$ in einem Intervall anwenden, was jeweils zu einem neuen Punkt $\xi$ strikt zwischen $x_1$ und $x_2$ mit $f'(\xi) = \frac{f(x_2)-f(x_1)}{x_2-x_1} = \frac{f(x_1)-f(x_2)}{x_1-x_2}$ führt.

**Korollar 7.35** (Kriterium für Monotonie). *Sei $I \subseteq \mathbb{R}$ ein nicht-leeres Intervall und $f : I \to \mathbb{R}$ eine differenzierbare Funktion. Dann gilt*

$$\begin{aligned}
\big(\forall x \in I : f'(x) \geq 0\big) &\implies f \text{ ist monoton wachsend} \\
\big(\forall x \in I : f'(x) > 0\big) &\implies f \text{ ist streng monoton wachsend} \\
\big(\forall x \in I : f'(x) \leq 0\big) &\implies f \text{ ist monoton fallend} \\
\big(\forall x \in I : f'(x) < 0\big) &\implies f \text{ ist streng monoton fallend.}
\end{aligned}$$

*Beweis.* Dies folgt unmittelbar aus dem Mittelwertsatz (Theorem 7.29), da für zwei Punkte $x_1 < x_2$ in $I$ gilt $f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1)$ für ein $\xi \in I$. □

Zwei (aber nur zwei) dieser Implikationen sind sogar Äquivalenzen, wie folgende Übung zeigt.

**Wichtige Übung 7.36** (Exakte Charakterisierung von Monotonie). *Sei $I \subseteq \mathbb{R}$ ein nicht-leeres Intervall und $f : I \to \mathbb{R}$ eine differenzierbare Funktion. Zeigen Sie, dass*

$$\begin{aligned}
\big(\forall x \in I : f'(x) \geq 0\big) &\iff f \text{ ist monoton wachsend} \\
\big(\forall x \in I : f'(x) \leq 0\big) &\iff f \text{ ist monoton fallend}
\end{aligned}$$

*Untersuchen Sie weiter das Beispiel $f : x \in \mathbb{R} \mapsto x^3 \in \mathbb{R}$.*

**Korollar 7.37** (Hinreichende Kriterien für lokale Extrema). *Sei $I \subseteq \mathbb{R}$ ein Intervall mit Endpunkten $a < b$ und $x_0 \in I$ kein Endpunkt von $I$. Sei $f : I \to \mathbb{R}$ stetig und zumindest auf $I \setminus \{x_0\}$ stetig differenzierbar.*

- *Angenommen der linke Endpunkt $a$ liegt in $I$.*
  - *Falls $f'(a) > 0$ erfüllt ist, dann hat $f$ in $a$ ein isoliertes lokales Minimum.*
  - *Falls $f'(a) < 0$ erfüllt ist, dann hat $f$ in $a$ ein isoliertes lokales Maximum.*
- *Beim Punkt $x_0 \in I$ gelten folgende Kriterien.*
  - *Falls ein $\delta > 0$ existiert, so dass $f'(x) > 0$ für alle $x \in (x_0 - \delta, x_0 + \delta) \setminus \{x_0\}$ (oder $f'(x) < 0$ für alle $x \in (x_0 - \delta, x_0 + \delta) \setminus \{x_0\}$), dann ist $x_0$ kein lokales Extremum von $f$.*
  - *Falls ein $\delta > 0$ existiert, so dass $f'(x) > 0$ für alle $x \in (x_0 - \delta, x_0)$ und $f'(x) < 0$ für alle $x \in (x_0, x_0 + \delta)$, dann nimmt $f$ in $x_0$ ein isoliertes lokales Maximum an.*
  - *Falls ein $\delta > 0$ existiert, so dass $f'(x) < 0$ für alle $x \in (x_0 - \delta, x_0)$ und $f'(x) > 0$ für alle $x \in (x_0, x_0 + \delta)$, dann nimmt $f$ in $x_0$ ein isoliertes lokales Minimum an.*
  - *Falls $f$ auf ganz $I$ zweimal stetig differenzierbar ist und $f'(x_0) = 0$ sowie $f''(x_0) < 0$ gilt, dann nimmt $f$ in $x_0$ ein isoliertes lokales Maximum an.*
  - *Falls $f$ auf ganz $I$ zweimal stetig differenzierbar ist und $f'(x_0) = 0$ sowie $f''(x_0) > 0$ gilt, dann nimmt $f$ in $x_0$ ein isoliertes lokales Minimum an.*

- *Angenommen der rechte Endpunkt $b$ liegt in $I$.*
  - *Falls $f'(b) > 0$ erfüllt ist, dann hat $f$ in $b$ ein isoliertes lokales Maximum.*
  - *Falls $f'(b) < 0$ erfüllt ist, dann hat $f$ in $b$ ein isoliertes lokales Minimum.*

Auf Grund der Vielzahl obiger Kriterien möchten wir hier erwähnen, dass man sich dieses Korollar viel eher merken kann, wenn man es (anhand eines Bildes und des Beweises) verstanden hat, als wenn man es auswendig lernen möchte.

*Beweis.* Die Beweise dieser Aussagen sind alle sehr ähnlich und beruhen in sämtlichen Fällen auf dem Mittelwertsatz. Sei $a \in I$ der linke Endpunkt und angenommen $f'(a) > 0$. Dann existiert wegen der Stetigkeit von $f'$ ein $\delta > 0$, so dass für alle $x \in [a, a+\delta)$ ebenso $f'(x) > 0$ gilt. Für jedes $x \in [a, a+\delta)$ existiert nun nach dem Mittelwertsatz (Theorem 7.29) ein $\xi \in (a, x)$ mit

$$f(x) - f(a) = f'(\xi)(x - a) > 0.$$

Dies zeigt, dass $f$ in $a$ ein lokales Minimum annimmt. Der Fall $f'(a) < 0$ ergibt sich analog. Ebenfalls analog ist das Argument für die Aussagen beim rechten Endpunkt.

Für $x_0$ gilt auf Grund des gleichen Arguments für jedes $\delta > 0$

$$\begin{aligned}
\forall x \in (x_0 - \delta, x_0) : f'(x) > 0 &\implies \forall x \in (x_0 - \delta, x_0) : f(x) < f(x_0) \\
\forall x \in (x_0, x_0 + \delta) : f'(x) > 0 &\implies \forall x \in (x_0, x_0 + \delta) : f(x) > f(x_0) \\
\forall x \in (x_0 - \delta, x_0) : f'(x) < 0 &\implies \forall x \in (x_0 - \delta, x_0) : f(x) > f(x_0) \\
\forall x \in (x_0, x_0 + \delta) : f'(x) < 0 &\implies \forall x \in (x_0, x_0 + \delta) : f(x) < f(x_0)
\end{aligned}$$

Durch Kombination dieser Aussagen ergeben sich das erste, zweite und dritte Kriterium für $x_0$.

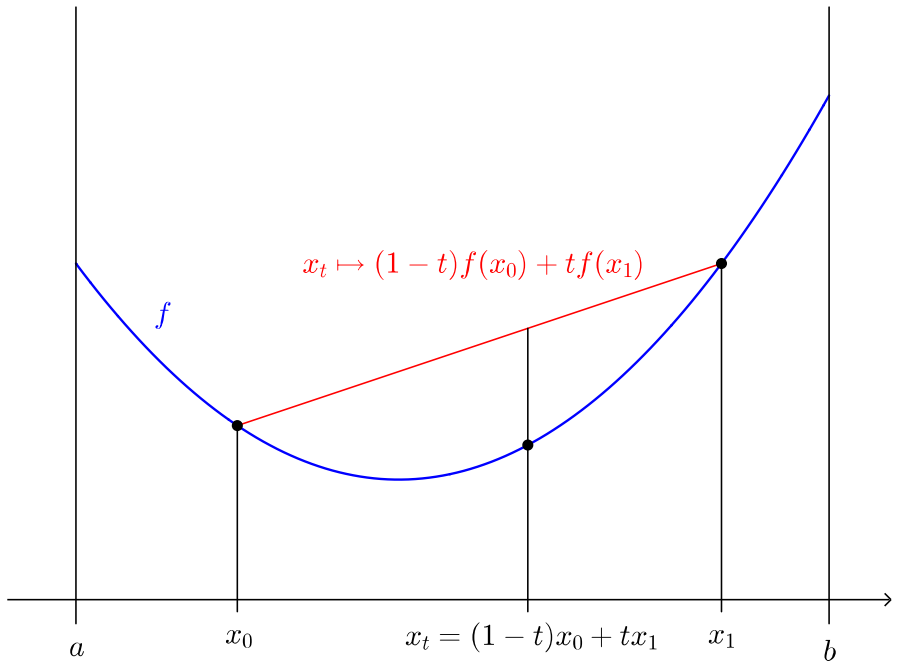
Angenommen $f$ ist auf ganz $I$ zweimal stetig differenzierbar und $f'(x_0) = 0$ sowie $f''(x_0) < 0$ sind erfüllt. Dann existiert ein $\delta > 0$ so dass $f''(x) < 0$ für alle $x \in (x_0 - \delta, x_0 + \delta)$. Nach Korollar 7.35 ist $f'|_{(x_0-\delta, x_0+\delta)}$ streng monoton fallend und somit nimmt $f$ nach dem zweiten Kriterium für $x_0$ in $x_0$ ein isoliertes lokales Maximum an. Das fünfte Kriterium für $x_0$ folgt wiederum analog. □

### 7.2.3 Konvexität

**Definition 7.38** (Konvexität und Konkavität)**.** Sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ eine Funktion. Dann heisst $f$ **konvex**, falls

$$f\big((1-t)x_0 + tx_1\big) \leq (1-t)f(x_0) + tf(x_1) \tag{7.6}$$

für alle $x_0, x_1 \in I$ und für alle $t \in [0, 1]$. Wir sagen, dass $f$ **streng konvex** ist, falls in (7.6) eine strikte Ungleichung gilt, wenn immer $x_0 \neq x_1$ und $t \in (0, 1)$ (in diesem Falls ist $(1-t)x_0 + tx_1$ echt zwischen $x_0$ und $x_1$). Eine Funktion $g : I \to \mathbb{R}$ heisst (**streng**) **konkav**, wenn $f = -g$ (streng) konvex ist.




Figur 7.3: Anschaulich formuliert ist eine Funktion konvex $x_0$ , wenn ihr Graph jeweils unterhalb der Strecke zwischen zwei Punkten auf dem Graphen bleibt. Wie wir in Kürze sehen werden, kann man die konvexen Funktionen als solche sehen, deren Graphen aufwärts gekrümmt sind, wie ebenfalls in diesem Bild ersichtlich ist.

Konvexe Funktionen sind unter anderem nützlich, weil sie bei verschiedenen Ungleichungen zum Vorschein treten. Wir bemerken, dass auf Grund der Definition von Konkavität die Resultate dieses Unterabschnitts für konvexe Funktionen auf ähnlichen Weise auf konkave Funktionen zutreffen. Nun beginnen wir damit Konvexität auf eine andere Art zu charakterisieren.

**Lemma 7.39** (Konvexität via Steigung von Sekanten)**.** *Sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ eine Funktion. Die Funktion $f$ ist genau dann konvex, wenn für alle $x, x_0, x_1 \in I$ gilt*

$$x_0 < x < x_1 \implies \frac{f(x) - f(x_0)}{x - x_0} \leq \frac{f(x_1) - f(x)}{x_1 - x}.$$

*Des Weiteren ist $f$ genau dann streng konvex, wenn*

$$x_0 < x < x_1 \implies \frac{f(x) - f(x_0)}{x - x_0} < \frac{f(x_1) - f(x)}{x_1 - x}.$$

*gilt.*

In Worten ausgedrückt besagt das Lemma insbesondere, dass für eine konvexe Funktion die Steigung der Sekanten zwischen Punkten $x_0 < x$ kleiner (gleich) ist als die Steigung der Sekanten zwischen $x < x_1$. Veranschaulichen Sie sich dies an Figur 7.3.

*Beweis.* Für Punkte $x_0 < x_1$ in $I$ hat $t \in (0,1) \mapsto x_t = (1-t)x_0 + tx_1 \in (x_0, x_1)$ die Umkehrabbildung $x \in (x_0, x_1) \mapsto t_x = \frac{x - x_0}{x_1 - x_0} \in (0,1)$. Mit dieser Notation gilt (7.6) für alle

$t \in (0,1)$ genau dann, wenn

$$f(x) \leq \frac{x_1 - x}{x_1 - x_0} f(x_0) + \frac{x - x_0}{x_1 - x_0} f(x_1).$$

für alle $x \in (x_0, x_1)$ gilt. Letzteres ist wiederum äquivalent zu

$$(x_1 - x_0) f(x) \leq (x_1 - x) f(x_0) + (x - x_0) f(x_1)$$

für alle $x \in (x_0, x_1)$, was genau dann gilt, wenn

$$(x_1 - x)(f(x) - f(x_0)) \leq (x - x_0)(f(x_1) - f(x)).$$

für alle $x \in (x_0, x_1)$. Dies ist aber zu $\frac{f(x)-f(x_0)}{x-x_0} \leq \frac{f(x_1)-f(x)}{x_1-x}$ äquivalent, wie gewünscht.

Im Falle der strengen Konvexität können wir alle „$\leq$“ in obigem Beweis durch „$<$“ ersetzen. □

Für differenzierbare Funktion existiert folgende, sehr direkte Charakterisierung der Konvexität, welche erklärt, warum (differenzierbare) konvexe Funktionen aufwärts gekrümmte Graphen besitzen.

**Proposition 7.40** (Kriterium für Konvexität)**.** *Sei $I \subseteq \mathbb{R}$ ein Intervall mit Endpunkten $a < b$ und $f : I \to \mathbb{R}$ eine differenzierbare Funktion. Dann ist $f$ genau dann (streng) konvex, wenn $f'$ (streng) monoton wachsend ist.*

*Beweis.* Der Beweis beruht auf dem Mittelwertsatz (Theorem 7.29) und der Charakterisierung von konvexen Funktionen in Lemma 7.39. Denn für drei Punkte $x_0 < x < x_1$ in $I$ existieren $\xi_1 \in (x_0, x)$ und $\xi_2 \in (x, x_1)$ mit

$$f'(\xi_1) = \frac{f(x) - f(x_0)}{x - x_0}, \quad f'(\xi_2) = \frac{f(x_1) - f(x)}{x_1 - x}$$

Falls nun $f'$ (streng) monoton wachsend ist, dann ist $f'(\xi_1) \leq f'(\xi_2)$ (respektive $f'(\xi_1) < f'(\xi_2)$) und wir erhalten die (strenge) Konvexität von $f$ aus Lemma 7.39.

Angenommen $f$ ist konvex und $x_0 < x_1$ sind zwei Punkte in $I$. Dann folgt aus Lemma 7.39, dass für alle $h \in (0, \frac{1}{2}(x_1 - x_0))$ gilt

$$\frac{f(x_0 + h) - f(x_0)}{h} \leq \frac{f(x_1 - h) - f(x_0 + h)}{(x_1 - h) - (x_0 + h)} \leq \frac{f(x_1) - f(x_1 - h)}{h},$$

woraus mit $h \searrow 0$

$$f'(x_0) \leq \frac{f(x_1) - f(x_0)}{x_1 - x_0} \leq f'(x_1)$$

folgt. Wir verweisen auf Übung 7.41 für den letzten Beweisschritt (der weniger häufig für Anwendungen von Bedeutung ist). □

**Übung 7.41** (Strenge Konvexität)**.** *Seien $f, I$ wie in Proposition 7.40. Angenommen $f$ ist streng konvex. Zeigen Sie, dass $f'$ streng monoton wachsend ist.*

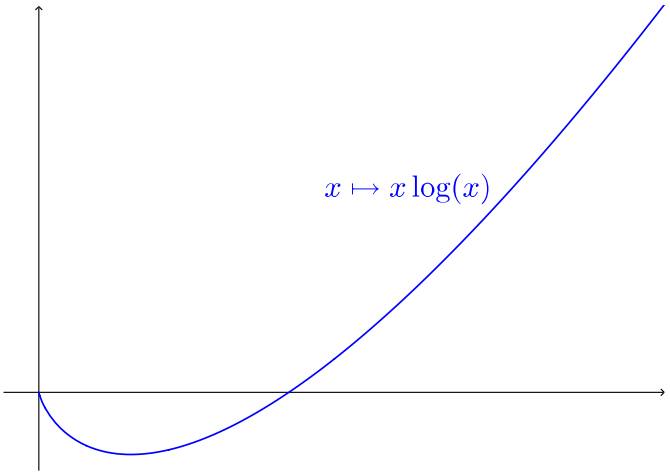
Aus Proposition 7.40 und Korollar 7.35 ergibt sich folgendes Korollar.

**Korollar 7.42** (Konvexität und die zweite Ableitung)**.** *Sei $I \subseteq \mathbb{R}$ ein Intervall mit Endpunkten $a < b$ und $f : I \to \mathbb{R}$ eine zweimal differenzierbare Funktion. Falls $f''(x) \geq 0$ für alle $x \in I$, dann ist $f$ konvex. Falls $f''(x) > 0$ für alle $x \in I$, dann ist $f$ streng konvex.*

**Beispiel 7.43.** *Die Funktion $f : x \in (0, \infty) \mapsto x \log(x)$ ist streng konvex. Dies ergibt sich aus Korollar 7.42, da $f$ glatt ist und*

$$f'(x) = \log(x) + x\frac{1}{x} = \log(x) + 1, \quad f''(x) = \frac{1}{x} > 0.$$

*für alle $x > 0$. Des Weiteren wissen wir bereits aus Beispiel 5.76, dass $\lim_{x\to 0} x \log(x) = 0$. Zuletzt bemerken wir, dass $\lim_{x\to 0} f'(x) = -\infty$, was alles im Graphen von $f$ ersichtlich ist.*




**Lemma 7.44** (Jensensche Ungleichung)**.** *Sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ eine konvexe Funktion. Seien $n \in \mathbb{N}$, $x_1, x_2, \ldots, x_n \in I$ und $t_1, t_2, \ldots, t_n \in [0, 1]$ mit $\sum_{k=1}^n t_k = 1$. Dann gilt*

$$f\left(\sum_{k=1}^n t_k x_k\right) \leq \sum_{k=1}^n t_k f(x_k). \tag{7.7}$$

*Beweis.* Wir verwenden Induktion über $n \geq 2$. Für $n = 1$ haben wir $\sum_{k=1}^1 t_k x_k = x_1$, $\sum_{k=1}^1 t_k f(x_k) = f(x_1)$ und (7.7) ist trivialerweise erfüllt. Für $n = 2$ ist (7.7) gerade (7.6) (mit $t_1 = 1 - t$ und $t_2 = t \in [0, 1]$). Angenommen die Aussage ist für $n \geq 2$ erfüllt. Seien $x_1, \ldots, x_{n+1} \in I$ und $t_1, \ldots, t_{n+1} \in [0, 1]$. Falls $t_{n+1} = 0$ folgt (7.7) direkt aus der Annahme

für $n$. Also angenommen $t_{n+1} > 0$. Dann gilt

$$\begin{aligned}
f\left(\sum_{k=1}^{n+1} t_k x_k\right) &= f\left(\sum_{k=1}^{n-1} t_k x_k + (t_n + t_{n+1})\left(\frac{t_n}{t_n + t_{n+1}} x_n + \frac{t_{n+1}}{t_n + t_{n+1}} x_{n+1}\right)\right) \\
&\leq \sum_{k=1}^{n-1} t_k f(x_k) + (t_n + t_{n+1}) f\left(\frac{t_n}{t_n + t_{n+1}} x_n + \frac{t_{n+1}}{t_n + t_{n+1}} x_{n+1}\right) \\
&\leq \sum_{k=1}^{n-1} t_k f(x_k) + (t_n + t_{n+1})\left(\frac{t_n}{t_n + t_{n+1}} f(x_n) + \frac{t_{n+1}}{t_n + t_{n+1}} f(x_{n+1})\right) \\
&= \sum_{k=1}^{n+1} t_k f(x_k)
\end{aligned}$$

per Induktionsannahme angewendet auf $x_1, \ldots, x_{n-1}, \frac{t_n}{t_n+t_{n+1}} x_n + \frac{t_{n+1}}{t_n+t_{n+1}} x_{n+1}$ und Konvexität von $f$. □

Die Jensenschen Ungleichung hat zahlreiche Anwendungen und kann abhängig von der konvexen Funktion $f$ verschiedene Formen annehmen. Ein Beispiel dafür ist in der nächsten Übung enthalten.

**Übung 7.45** (Harmonisches, geometrisches und arithmetisches Mittel)**.** *Zeigen Sie die Ungleichung*

$$\frac{n}{\frac{1}{x_1} + \ldots + \frac{1}{x_n}} \leq \sqrt[n]{x_1 \cdots x_n} \leq \frac{x_1 + \ldots + x_n}{n}$$

*für das harmonische, das arithmetische und das geometrische Mittel von* $x_1, \ldots, x_n \in \mathbb{R}_{>0}$.

Auch für das Riemann-Integral gilt eine Version der Jensenschen Ungleichung.

**Übung 7.46** (Integralform der Jensenschen Ungleichung)**.** *Sei* $\varphi : [0,1] \to I$ *stetig und sei* $f : I \to \mathbb{R}$ *eine stetige, konvexe Funktion. Dann gilt*

$$f\left(\int_0^1 \varphi(t)\,\mathrm{d}t\right) \leq \int_0^1 f(\varphi(t))\,\mathrm{d}t.$$

Wir schliessen die Diskussion von Konvexität mit folgenden Übungen.

**Übung 7.47** (Minima von konvexen Funktionen)**.** *Sei* $I \subseteq \mathbb{R}$ *ein Intervall und* $f : I \to \mathbb{R}$ *eine konvexe Funktion. Zeigen Sie, dass jedes lokale Minimum von* $f$ *ein (globales) Minimum ist.*

### 7.2.4 Mittelwertsatz nach Cauchy

Wir möchten nun den Mittelwertsatz etwas verallgemeinern.

**Satz 7.48** (Erweiterter Mittelwertsatz)**.** *Seien* $f$ *und* $g$ *stetige Funktionen auf einem Intervall* $[a,b]$ *mit* $a < b$*, so dass* $f$ *und* $g$ *auf* $(a,b)$ *differenzierbar sind. Dann existiert ein* $\xi \in (a,b)$

*mit*

$$g'(\xi)(f(b) - f(a)) = f'(\xi)(g(b) - g(a)). \tag{7.8}$$

*Falls zusätzlich $g'(x) \neq 0$ für alle $x \in (a, b)$ gilt, dann gilt $g(a) \neq g(b)$ und*

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Man beachte, dass der Mittelwertsatz (Theorem 7.29) gerade der Spezialfall $g : x \in [a, b] \mapsto x$ des obigen Satzes ist und man somit in der Tat von einem erweiteren Mittelwertsatz sprechen darf.

*Bemerkung.* Genau wie der Mittelwertsatz hat der Mittelwertsatz von Cauchy eine geometrische Interpretation, nur muss man dieses Mal in der zweidimensionalen Ebene suchen. Dort besagt der Mittelwertsatz von Cauchy unter den getroffenen Annahmen, dass die Kurve $t \mapsto (f(t), g(t))$ eine Tangente besitzt, die parallel zur Gerade durch die Punkte $(f(a), g(a))$, $(f(b), g(b))$ ist.

*Beweis von Satz 7.48.* Wir definieren eine Funktion $F : [a, b] \to \mathbb{R}$ durch

$$F(x) = g(x)\big(f(b) - f(a)\big) - f(x)\big(g(b) - g(a)\big)$$

für alle $x \in \mathbb{R}$. Dann gilt

$$\begin{aligned} F(a) &= g(a)\big(f(b) - f(a)\big) - f(a)\big(g(b) - g(a)\big) = g(a)f(b) - f(a)g(b) \\ F(b) &= g(b)\big(f(b) - f(a)\big) - f(b)\big(g(b) - g(a)\big) = F(a). \end{aligned}$$

Nach dem Satz von Rolle (Satz 7.28) existiert somit ein $\xi \in (a, b)$ mit

$$F'(\xi) = g'(\xi)(f(b) - f(a)) - f'(\xi)(g(b) - g(a)) = 0.$$

Dies beweist die erste Behauptung (7.8) des Satzes.

Falls zusätzlich $g'(x) \neq 0$ für alle $x \in (a, b)$, dann folgt aus dem Satz von Rolle (Satz 7.28), dass $g(b) \neq g(a)$ (wieso?). Nach Division von (7.8) mit $g'(\xi)(g(b) - g(a))$ ergibt sich die zweite Behauptung des Satzes. □

### 7.2.5 Regel von de l'Hôpital

Die folgende Regel stellt eine einfache Anwendung des obigen erweiterten Mittelwertsatzes dar und macht die Berechnung von vielen konkreten Grenzwerten deutlich einfacher.

**Satz 7.49** (Regel von de l'Hôpital)**.** *Seien $a < b$ in $\overline{\mathbb{R}}$ und seien $f, g : (a, b) \to \mathbb{R}$ zwei differenzierbare Funktionen mit $g(x) \neq 0$ und $g'(x) \neq 0$ für alle $x \in (a, b)$. Angenommen der Grenzwert $\lim_{x \searrow a} \frac{f'(x)}{g'(x)}$ existiert in $\overline{\mathbb{R}}$ und eine der beiden folgenden Bedingungen ist erfüllt:*

- *(„$\frac{0}{0}$“)* $\lim_{x \searrow a} f(x) = \lim_{x \searrow a} g(x) = 0$

- *(„$\frac{\infty}{\infty}$“)* $\lim_{x\searrow a} g(x) = +\infty$ *(oder* $\lim_{x\searrow a} g(x) = -\infty$*).*

*Dann existiert auch der Grenzwert* $\lim_{x\searrow a} \frac{f(x)}{g(x)}$ *und es gilt*

$$\lim_{x\searrow a} \frac{f(x)}{g(x)} = \lim_{x\searrow a} \frac{f'(x)}{g'(x)}.$$

*Analoge Aussagen gelten für die Bewegungen* $x \nearrow b$ *oder* $x \to x_0 \in (a,b)$*. Im letzten Fall erlauben wir* $g(x_0) = 0$ *oder auch* $g'(x_0) = 0$*, solange* $g(x) \neq 0$ *und* $g'(x) \neq 0$ *für alle* $x \in (a,b) \setminus \{x_0\}$*.*

Wir nennen „$\frac{0}{0}$“ und „$\frac{\infty}{\infty}$“ unbestimmte Formen, da wir bisher keine allgemeinen Werkzeuge hatten, um den Grenzwert von $\frac{f(x)}{g(x)}$ zu berechnen für den Fall, dass sowohl $f(x)$ als auch $g(x)$ beide gegen 0 oder beide gegen $\infty$ streben. Obige Regel ist in diesen Fällen sehr oft nützlich und viel einfacher anwendbar als unsere bisherigen Berechnungen (siehe unter anderem Beispiel 5.76). Wie bereits erwähnt ist die Regel ein Korollar des erweiterten Mittelwertsatzes.

*Beweis.* Wir betrachten zuerst die Bewegung $x \searrow a$ für ein $a \in \mathbb{R}$ und setzen

$$L = \lim_{x\searrow a} \frac{f'(x)}{g'(x)} \in \overline{\mathbb{R}}.$$

Die unbestimmte Form „$\frac{0}{0}$“: Angenommen es gilt $\lim_{x\searrow a} f(x) = \lim_{x\searrow a} g(x) = 0$. In diesem Fall können wir also $f$ und $g$ stetig auf $[a,b)$ fortsetzen, in dem wir $f(a) = g(a) = 0$ setzen. Da $L = \lim_{x\searrow a} \frac{f'(x)}{g'(x)}$ ist, können wir für eine beliebige Umgebung $U$ von $L$ ein $y \in (a,b)$ wählen mit $\frac{f'(\xi)}{g'(\xi)} \in U$ für alle $\xi \in (a,y)$. Für ein beliebiges $x \in (a,y)$ wenden wir nun den erweiterten Mittelwertsatz (Satz 7.48) auf $[a,x]$ an und erhalten

$$\frac{f(x)}{g(x)} = \frac{f(x)-f(a)}{g(x)-g(a)} = \frac{f'(\xi_x)}{g'(\xi_x)} \in U$$

für ein $\xi_x \in (a,x) \subseteq (a,y)$. Da die Umgebung $U$ von $L$ beliebig war, existiert der Grenzwert $\lim_{x\searrow a} \frac{f(x)}{g(x)}$ und ist gleich $L$. Dies beweist den ersten Fall.

VO: 06.12.17

Die unbestimmte Form „$\frac{\infty}{\infty}$“ mit reellem Grenzwert: Wir nehmen an, dass der Nenner die Aussage $\lim_{x\searrow a} g(x) = +\infty$ (oder $\lim_{x\searrow a} g(x) = -\infty$) erfüllt sowie der Grenzwert $L \in \mathbb{R}$ reell ist.[1]

Sei nun $\varepsilon > 0$. Nach Definition von $L$ und der Annahme dieses Falles gibt es ein $y_\varepsilon \in (a,b)$ mit $\left|\frac{f'(\xi)}{g'(\xi)} - L\right| < \varepsilon$ für alle $\xi \in (a,y_\varepsilon)$. Für ein beliebiges $x \in (a,y_\varepsilon)$ gibt es nach dem erweiterten Mittelwertsatz (Satz 7.48) jeweils ein $\xi_x \in (x,y_\varepsilon) \subseteq (a,y_\varepsilon)$ mit

$$\frac{f(x)-f(y_\varepsilon)}{g(x)-g(y_\varepsilon)} = \frac{f'(\xi_x)}{g'(\xi_x)} \in (L-\varepsilon, L+\varepsilon), \tag{7.9}$$

[1] Man beachte, dass diese Annahme auch gemeinsam mit $L = \lim_{x\searrow a} \frac{f'(x)}{g'(x)}$ keinerlei Information über das asymptotische Verhalten von $f$ für $x \searrow a$ enthält.

wobei $g(x) \neq g(y_\varepsilon)$ auf Grund der zweiten Aussage in Satz 7.48 und unserer Annahmen an $g$. Anders ausgedrückt erhalten wir also

$$\frac{f(x)}{g(x)-g(y_\varepsilon)} = \frac{f'(\xi_x)}{g'(\xi_x)} + \frac{f(y_\varepsilon)}{g(x)-g(y_\varepsilon)}. \tag{7.10}$$

Intuitiv ausgedrückt ist hier die linke Seite etwa $\frac{f(x)}{g(x)}$ für $x$ sehr nahe an $a$, denn $g(y_\varepsilon)$ ist fest gewählt während $|g(x)|$ beliebig gross wird. Des Weiteren ist die rechte Seite nahe an $L$, da $\frac{f(y_\varepsilon)}{g(x)-g(y_\varepsilon)}$ gegen Null geht für $x \searrow a$.

Formaler geht man wie folgt vor. Wir bemerken zuerst, dass (7.10) zu

$$\begin{aligned}\frac{f(x)}{g(x)} &= \frac{f'(\xi_x)}{g'(\xi_x)}\frac{g(x)-g(y_\varepsilon)}{g(x)} + \frac{f(y_\varepsilon)}{g(x)} \\ &= \frac{f'(\xi_x)}{g'(\xi_x)} - \frac{f'(\xi_x)}{g'(\xi_x)}\frac{g(y_\varepsilon)}{g(x)} + \frac{f(y_\varepsilon)}{g(x)}\end{aligned} \tag{7.11}$$

für alle $x \in (a, y_\varepsilon)$ umgeformt werden kann. Da $\lim_{x\searrow a}|g(x)| = \infty$ gilt, existiert weiter ein $\tilde{y}_\varepsilon \in (a, y_\varepsilon)$, so dass

$$\big(|L|+\varepsilon\big)\frac{|g(y_\varepsilon)|}{|g(x)|} < \varepsilon \text{ und } \left|\frac{f(y_\varepsilon)}{g(x)}\right| < \varepsilon$$

für alle $x \in (a, \tilde{y}_\varepsilon)$. Zusammenfassend gilt für ein $x \in (a, \tilde{y}_\varepsilon)$ und eine Wahl $\xi_x \in (a, y_\varepsilon)$ wie in (7.9) die Abschätzung

$$\left|\frac{f'(\xi_x)}{g'(\xi_x)}\right| \leq \left|\frac{f'(\xi_x)}{g'(\xi_x)} - L\right| + |L| \leq |L| + \varepsilon$$

und damit die Ungleichung

$$\begin{aligned}\left|\frac{f(x)}{g(x)} - L\right| &\leq \left|\frac{f'(\xi_x)}{g'(\xi_x)} - L\right| + \left|\frac{f'(\xi_x)}{g'(\xi_x)}\frac{g(y_\varepsilon)}{g(x)}\right| + \left|\frac{f(y_\varepsilon)}{g(x)}\right| \\ &< \varepsilon + (|L|+\varepsilon)\,\frac{|g(y_\varepsilon)|}{|g(x)|} + \varepsilon < 3\varepsilon,\end{aligned}$$

was zu beweisen war.

Die unbestimmte Form „$\frac{\infty}{\infty}$“ mit uneigentlichem Grenzwert: Wir nehmen nun an, dass $L = \infty$. Dann lässt sich die Aussage nach kleinen Anpassungen analog wie oben beweisen. In der Tat existiert nach Annahme zu jedem $\varepsilon > 0$ ein $y_\varepsilon \in (a, b)$, so dass $\frac{f'(\xi)}{g'(\xi)} > \frac{1}{\varepsilon}$ für alle $\xi \in (a, y_\varepsilon)$. Für dieses fest gewählte $y_\varepsilon$ wählen wir nun ein $\tilde{y}_\varepsilon \in (a, y_\varepsilon)$ mit $\left|\frac{f(y_\varepsilon))}{g(x)-g(y_\varepsilon)}\right| < 1$ und $\frac{g(x)-g(y_\varepsilon)}{g(x)} > \frac{1}{2}$ für alle $x \in (a, \tilde{y}_\varepsilon)$. Dann ist nach (7.10)

$$\frac{f(x)}{g(x)} = \frac{g(x)-g(y_\varepsilon)}{g(x)}\frac{f(x)}{g(x)-g(y_\varepsilon)} \geq \frac{1}{2}\frac{f(x)}{g(x)-g(y_\varepsilon)} > \frac{1}{2}\left(-1+\frac{1}{\varepsilon}\right) = -\frac{1}{2} + \frac{1}{2\varepsilon}.$$

Daraus folgt die Aussage, da $\varepsilon > 0$ beliebig war. Der Fall $L = -\infty$ lässt sich analog beweisen oder durch Vorzeichenwechsel von $f$ auf obigen Fall zurückführen.

Restliche Fälle: Wir beschäftigen uns nun mit den übrig bleibenden Fällen im Satz, die

wir jeweils auf einen der obigen Fälle reduzieren können, in dem wir die unabhängige Variable $x$ im Definitionsbereich geeignet ersetzen.

Betrachtet man die Bewegung $x \to b$ für $b = \infty$, so lässt sich dieser Fall auf die vorherigen Fälle zurückführen. In der Tat können wir durch Einschränkung der Funktionen ohne Beschränkung der Allgemeinheit $a > 0$ annehmen und die Funktionen

$$F : x \in (0, \tfrac{1}{a}) \mapsto f(\tfrac{1}{x}), \quad G : x \in (0, \tfrac{1}{a}) \mapsto g(\tfrac{1}{x})$$

definieren und nun stattdessen den Grenzwert $\lim_{x \searrow 0} \frac{F(x)}{G(x)}$ betrachten. Die weiteren Schritte überlassen wir hierbei den Leserinnen und Lesern – siehe Übung 7.50. Die Bewegungen $x \nearrow b$, $x \to x_0$ und $x \to -\infty$ lassen sich ebenso auf die bereits betrachteten Fälle zurückführen. □

**Wichtige Übung 7.50.** *Vervollständigen Sie den obigen Beweis, indem Sie die Reduktionen in allen verbleibenden Fällen komplett ausführen.*

**Übung 7.51** (Rechnen mit der Regel von de l'Hôpital)**.** *Berechnen Sie die Grenzwerte*

$$\lim_{x \to 0} \frac{2\sin(x) - \sin(2x)}{x - \sin(x)}, \quad \lim_{x \to \infty} x^{\frac{1}{x}}, \quad \lim_{x \to 0} \frac{\sin(x^2)}{x \tan(x)}.$$

**Übung 7.52.** *Sei $[a, b]$ ein abgeschlossenes Intervall mit $a < b$ und sei $f : [a, b] \to \mathbb{R}$ stetig. Angenommen $x_0 \in [a, b]$ ist ein Punkt, so dass $f$ auf $[a, b] \setminus \{x_0\}$ differenzierbar ist und angenommen der Grenzwert $\lim_{x \to x_0} f'(x)$ existiert. Zeigen Sie, dass $f$ bei $x_0$ differenzierbar ist und dass $f'$ bei $x_0$ stetig ist.*

Wir möchten kurz anmerken, dass sich höhere Ableitungen mitunter zwar als einen direkten Grenzwert ausdrücken lassen, doch die Existenz dieses Grenzwertes nicht zur mehrmaligen Differenzierbarkeit äquivalent sein muss.

**Übung 7.53** (Zweite Ableitung als Grenzwert)**.** *Sei $I = (a, b) \subseteq \mathbb{R}$ ein Intervall mit $a < b$ und sei $f : I \to \mathbb{R}$ zweimal differenzierbar. Zeigen Sie die Formel*

$$f''(x) = \lim_{h \to 0} \frac{f(x + h) - 2f(x) + f(x - h)}{h^2}$$

*für alle $x \in I$. Verifizieren Sie anhand der Vorzeichenfunktion $x \in \mathbb{R} \mapsto \operatorname{sgn}(x)$, dass die Existenz des obigen Grenzwerts nicht zweimalige Differenzierbarkeit impliziert.*

## 7.3 Trigonometrische Funktionen

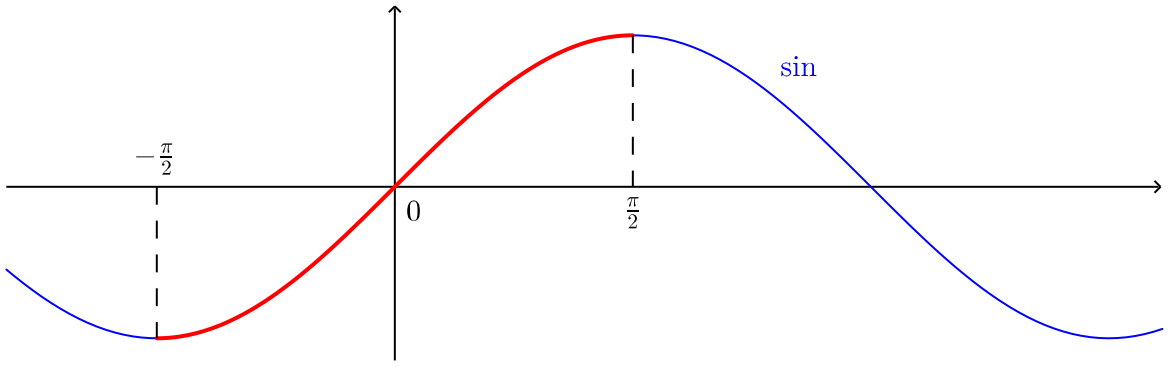
### 7.3.1 Sinus und Arkussinus

Nach Übung 7.4 sind sin und cos glatt und es gelten die Formeln

$$(\sin(x))' = \cos(x), \quad (\cos(x))' = -\sin(x). \tag{7.12}$$

Nach Satz 6.74 (genauer Übung 6.76) sind die Nullstellen von $\sin : \mathbb{R} \to \mathbb{R}$ die Menge $\mathbb{Z}\pi$ der ganzzahligen Vielfachen von $\pi$, womit nach dem Zwischenwertsatz und $\sin\left(\frac{\pi}{2}\right) = 1$ folgt, dass $\sin(x) > 0$ für alle $x \in (0, \pi)$. In diesem Intervall hat des Weiteren der Kosinus eine Nullstelle in $\frac{\pi}{2}$, ist positiv auf dem Intervall $\left[0, \frac{\pi}{2}\right)$ und negativ auf dem Intervall $\left[\frac{\pi}{2}, \pi\right)$. Verbinden wir dies, Gleichung (7.12) und das Kriterium für Monotonie differenzierbarer Funktion aus Korollar 7.35, so ergibt sich, dass sin auf dem Intervall $\left[0, \frac{\pi}{2}\right]$ streng monoton wachsend ist, auf dem Intervall $\left[\frac{\pi}{2}, \pi\right]$ streng monoton fallend ist und auf dem Intervall $[0, \pi]$ konkav (also nach unten gekrümmt) ist.

Da der Sinus eine ungerade Funktion ist, ergibt sich des Weiteren, dass der Sinus auf $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ streng monoton wachsend ist. Wir erkennen diese Eigenschaften direkt wieder im Graphen des Sinus.




Figur 7.4: Der Graph des Sinus mit dem (Teil-)Graphen der Einschränkung $\sin|_{\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]}$ hervorgehoben.

Die Einschränkung

$$\sin|_{\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]} : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \to [-1, 1], \; x \mapsto \sin(x)$$

ist also streng monoton wachsend und bijektiv (nach dem Zwischenwertsatz). Die Umkehrfunktion bezeichnen wir mit

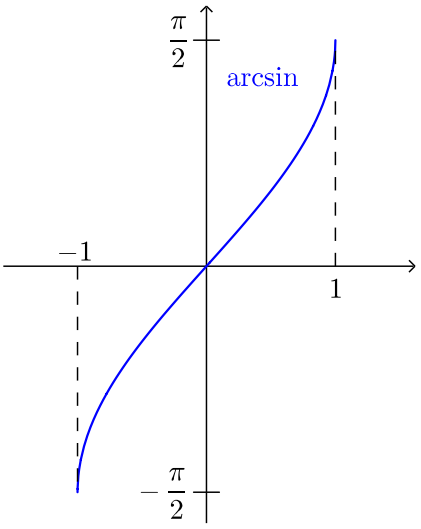
$$\arcsin : [-1, 1] \to \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$$

und nennen wir den **Arkussinus**. Nach dem Satz über die Differenzierbarkeit der inversen Funktion (Satz 7.14) ist der Arkussinus bei $s$ differenzierbar, falls die Ableitung des Sinus bei $x = \arcsin(s)$ nicht Null ist. In der Tat verschwindet die Ableitung des Sinus $\sin' = \cos$ genau an den Randpunkten $-\frac{\pi}{2}, \frac{\pi}{2}$ von $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. Für $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ und $s = \sin(x)$ ergibt sich nach

Satz 7.14

$$\arcsin'(s) = \frac{1}{\cos(x)} = \frac{1}{\sqrt{1-s^2}},$$

da $\cos(x)$ positiv ist für $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ und somit $\cos(x) = \sqrt{1-\sin(x)^2} = \sqrt{1-s^2}$ unter Verwendung von $\sin(x)^2 + \cos(x)^2 = 1$.

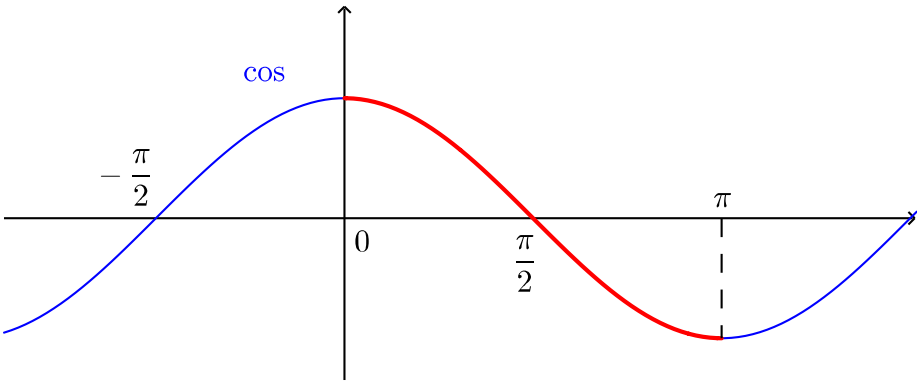



### 7.3.2 Kosinus und Arkuskosinus

Die obige Diskussion über Monotonie des Sinus kann analog durchgeführt werden für den Kosinus. Es ergibt sich, dass der Kosinus bei 0 ein lokales Extremum annimmt, auf dem Intervall $[0, \pi]$ streng monoton fallend ist, auf dem Intervall $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ konkav ist und auf dem Intervall $\left[\frac{\pi}{2}, \pi\right]$ konvex (nach oben gekrümmt) ist. Insbesondere ist die Einschränkung

$$\cos|_{[0,\pi]} : [0, \pi] \to [-1, 1]$$

bijektiv.




Figur 7.5: Der Graph des Kosinus mit hervorgehobenem (Teil-)Graphen der Einschränkung $\cos|_{[0,\pi]}$ hervorgehoben.

Die Umkehrabbildung heisst **Arkuskosinus** und wird als

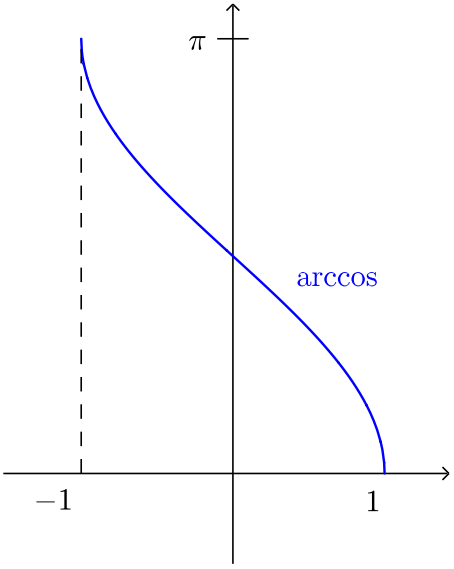
$$\arccos : [-1, 1] \to [0, \pi]$$

geschrieben.

Ebenso können wir die Ableitungsregeln für die Umkehrabbildung anwenden und erhalten bei $s = \cos(x)$ für $x \in (0, \pi)$

$$\arccos'(s) = \frac{1}{-\sin(x)} = -\frac{1}{\sqrt{1-s^2}},$$

da der Sinus auf $(0, \pi)$ positiv ist.




Wir möchten an dieser Stelle daran erinnern, dass sich jede Linearkombination von Sinus und Kosinus als ein Vielfaches des Sinus (oder des Kosinus) schreiben lässt (siehe die entsprechende Übung in Abschnitt 6.9.2).

### 7.3.3 Tangens und Arkustangens

Da $\sin(x + \pi) = -\sin(x)$ und $\cos(x + \pi) = -\cos(x)$ für alle $x \in \mathbb{R}$ gilt, ergibt sich, dass $\tan(x + \pi) = \tan(x)$ für alle $x \in \mathbb{R} \setminus \left(\mathbb{Z}\pi + \frac{\pi}{2}\right)$. Des Weiteren ist

$$\tan'(x) = \left(\frac{\sin(x)}{\cos(x)}\right)' = \frac{\cos(x)\cos(x) - \sin(x)(-\sin(x))}{\cos^2(x)} = \frac{1}{\cos^2(x)}$$

für alle $x \in \mathbb{R} \setminus \left(\mathbb{Z}\pi + \frac{\pi}{2}\right)$ nach der Quotientenregel in Korollar 7.11, Gleichung (7.12) und der trigonometrischen Identität $\sin^2(x) + \cos^2(x) = 1$. Insbesondere ist der Tangens auf jedem Intervall der Form $\left(n\pi - \frac{\pi}{2}, n\pi + \frac{\pi}{2}\right)$ für $n \in \mathbb{Z}$ streng monoton wachsend.

Des Weiteren gilt

$$\lim_{x \nearrow \frac{\pi}{2}} \tan(x) = \lim_{x \nearrow \frac{\pi}{2}} \frac{\sin(x)}{\cos(x)} = +\infty$$

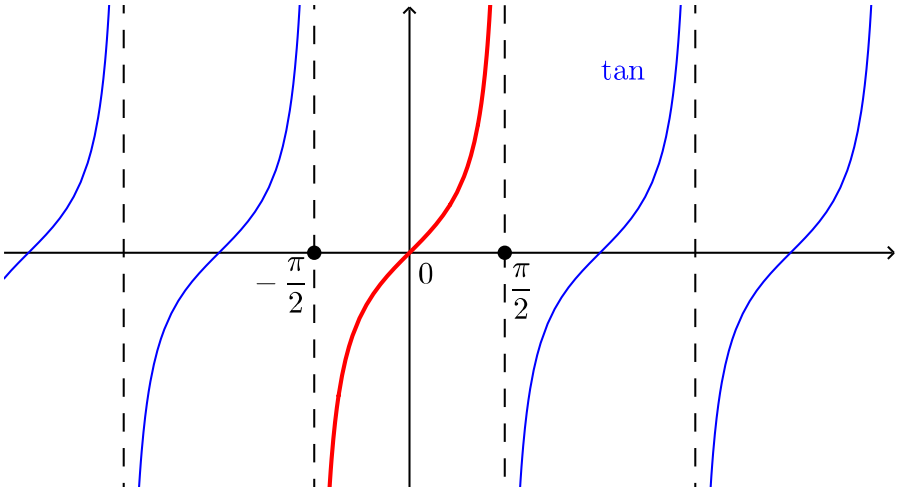
wegen $\sin\left(\frac{\pi}{2}\right) = 1$, $\cos\left(\frac{\pi}{2}\right) = 0$ und $\cos(x) > 0$ für $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, und

$$\lim_{x \searrow -\frac{\pi}{2}} \tan(x) = \lim_{x \searrow -\frac{\pi}{2}} \frac{\sin(x)}{\cos(x)} = -\infty$$

wegen $\sin\left(-\frac{\pi}{2}\right) = -1$. Wie zuvor folgt nun aus dem Zwischenwertsatz, dass die Einschränkung

$$\tan|_{\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)} : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \to \mathbb{R}$$

bijektiv ist.

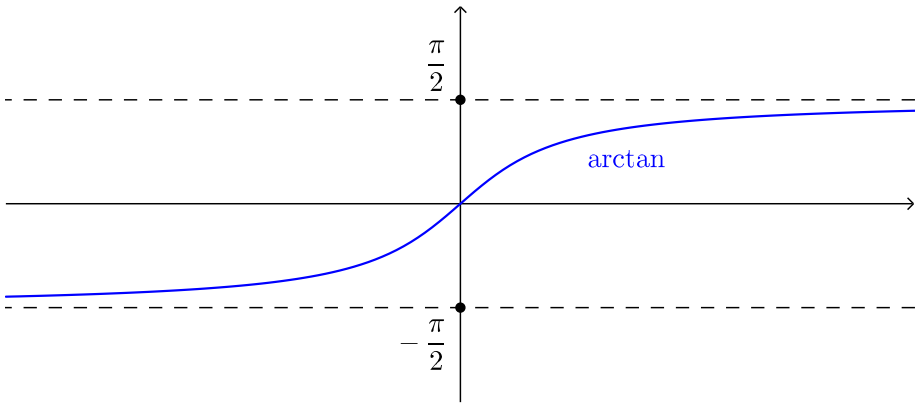



Figur 7.6: Der Graph des Tangens mit hervorgehobener Einschränkung $\tan|_{\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)}$.

Die Umkehrabbildung

$$\arctan : \mathbb{R} \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$$

wird als der **Arkustangens** bezeichnet.




Figur 7.7: Der Graph des Arkustangens.

Nach Satz 7.14 ist der Arkustangens differenzierbar und es gilt bei $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ und $s = \tan(x)$

$$\arctan'(s) = \frac{1}{\frac{1}{\cos^2(x)}} = \cos^2(x).$$

Des Weiteren gilt

$$1 + s^2 = 1 + \frac{\sin^2(x)}{\cos^2(x)} = \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} = \frac{1}{\cos^2(x)}.$$

Daraus folgt schlussendlich, dass für alle $s \in \mathbb{R}$

$$\arctan'(s) = \frac{1}{1 + s^2}$$

gilt.

### 7.3.4 Kotangens und Arkuskotangens

Der Kotangens respektive seine Umkehrfunktion, der Arkuskotangens, zeigen sehr ähnliches Verhalten gegenüber des Tangens respektive gegenüber des Arkustangens. Die Einschränkung $\cot|_{(0,\pi)} : (0, \pi) \to \mathbb{R}$ ist streng monoton fallend und bijektiv. Die Umkehrabbildung

$$\operatorname{arccot} : \mathbb{R} \to (0, \pi)$$

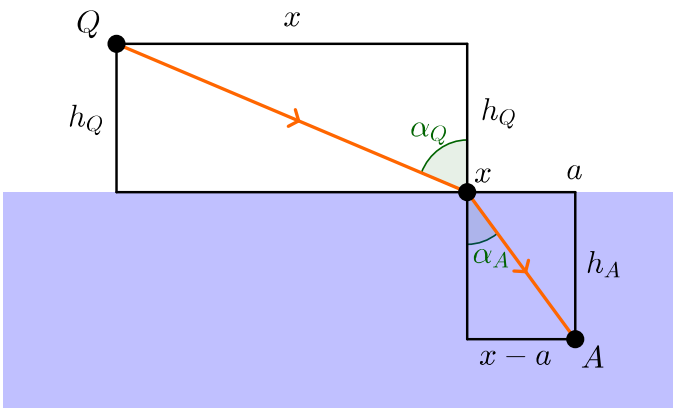
wird **Arkuskotangens** genannt und hat die Ableitung

$$\operatorname{arccot}'(s) = -\frac{1}{1 + s^2}$$

für alle $s \in \mathbb{R}$.

### 7.3.5 Ein physikalisches Beispiel*

**Beispiel 7.54** (Brechungsgesetz von Snellius)**.** *Wir werden hier das Brechungsgesetz der geometrischen Optik von Snellius (1580-1624) aus dem Fermat-Prinzip herleiten. Dabei besagt das Fermat-Prinzip, dass das Licht immer den Weg der kürzesten Reisezeit wählt. Gegeben sei eine geradlinige Grenze zwischen zwei Medien (zum Beispiel Luft und Glas oder Luft und Wasser) und die Lichtgeschwindigkeit $c_1$ und $c_2$ in diesen beiden Medien. (Die universelle Naturkonstante $c$ ist die Lichtgeschwindigkeit im Vakuum.) Wir wählen einen Punkt $Q$ im ersten Medium als Lichtquelle und wollen den Weg, den das Licht zu einem Punkt $A$ im zweiten Medium nimmt, bestimmen.*




*Wir wählen unser Koordinatensystem so, dass die Grenze zwischen den beiden Medien genau die $x$-Achse ist, $Q$ auf der positiven Hälfte der $y$-Achse liegt und die $x$-Koordinate von $A$ gleich $a > 0$ ist. Die Reisezeit des Lichts, das zuerst geradlinig von $Q$ nach $(x, 0)$ (ein Ort des möglichen Grenzübertritts) und dann „nach Brechung" von $(x, 0)$ nach $A$ geht, ist durch*

$$t(x) = \frac{1}{c_1}\sqrt{h_Q^2 + x^2} + \frac{1}{c_2}\sqrt{h_A^2 + (a-x)^2}$$

*beschrieben. Die Funktion $t : \mathbb{R} \to \mathbb{R},\ x \mapsto t(x)$ ist auf ganz $\mathbb{R}$ differenzierbar (wir nehmen $h_Q > 0$ und $h_A > 0$ an) und die Abbildung ist durch*

$$t'(x) = \frac{x}{c_1\sqrt{h_Q^2 + x^2}} - \frac{a-x}{c_2\sqrt{h_A^2 + (a-x)^2}}$$

*für alle $x \in \mathbb{R}$ gegeben. Diese verschwindet bei $x \in \mathbb{R}$ genau dann, wenn*

$$\begin{aligned}
\frac{x}{c_1\sqrt{h_Q^2 + x^2}} &= \frac{a-x}{c_2\sqrt{h_A^2 + (a-x)^2}}, \\
\frac{\sin(\alpha_Q)}{c_1} &= \frac{\sin(\alpha_A)}{c_2} \\
\frac{\sin(\alpha_A)}{\sin(\alpha_Q)} &= \frac{c_2}{c_1},
\end{aligned}$$

*wobei $\alpha_A$ und $\alpha_Q$ die beiden Winkel in obigem Bild bei dem Punkt $(x, 0)$ sind. Wir bemerken noch, dass obige Gleichung in $x$ für genau ein $x \in \mathbb{R}$ erfüllt ist und $t : \mathbb{R} \to \mathbb{R}$ für dieses $x_0 \in (0, a)$ tatsächlich ein globales Minimum annimmt. Die letzte Gleichung, welche die Winkel $\alpha_A, \alpha_Q$ und die Lichtgeschwindigkeiten $c_1, c_2$ in Verbindung bringt, wird das Brechungsgesetz von Snellius genannt.*

**Übung 7.55.** *Beweisen Sie die letzte Aussage in Beispiel 7.54.*

*Hinweis: Kann die Gleichung für $x \leq 0$ oder $x \geq a$ erfüllt sein? Verwenden Sie weiter, dass $\alpha_Q = \arctan\left(\frac{x}{h_Q}\right) \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ streng monoton wachsend von $x \in \mathbb{R}$ abhängt und $\alpha_A = \arctan\left(\frac{a-x}{h_A}\right) \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ streng monoton fallend von $x \in \mathbb{R}$ abhängt, wodurch $\frac{\sin(\alpha_A)}{\sin(\alpha_Q)}$ streng*

*monoton fallend von* $x \in (0, a)$ *abhängt. Wir bemerken noch, dass* $\lim_{x\to\infty} t(x) = +\infty$ *und* $\lim_{x\to-\infty} t(x) = +\infty$.

### 7.3.6 Verwendung der trigonometrischen Funktionen

Ab jetzt werden wir die trigonometrischen Funktionen und all ihre Monotonieeigenschaften, ihre Ableitungen und auch die Ableitungen ihrer Umkehrabbildungen meist ohne Referenz auf diesen Abschnitt verwenden. Deswegen wollen wir die Ableitungsregeln hier nochmals zusammenfassen. Es gilt

$$\begin{aligned}
\sin'(x) &= \cos(x), & \cos'(x) &= -\sin(x),\\
\tan'(x) &= \frac{1}{\cos^2(x)}, & \cot'(x) &= -\frac{1}{\sin^2(x)},\\
\arcsin'(x) &= \frac{1}{\sqrt{1-x^2}}, & \arccos'(x) &= -\frac{1}{\sqrt{1-x^2}},\\
\arctan'(x) &= \frac{1}{1+x^2}, & \operatorname{arccot}'(x) &= -\frac{1}{1+x^2}
\end{aligned}$$

für $x$ im jeweiligen Definitionsbereich der betrachteten Funktion.

## 7.4 Hyperbolische Funktionen

Wir möchten in diesem kurzen Abschnitt die zu Abschnitt 7.3 analoge Diskussion für die in Abschnitt 6.7.1 eingeführten hyperbolischen Funktionen durchführen. Wir erinnern daran, dass

$$\sinh(x) = \frac{\mathrm{e}^x - \mathrm{e}^{-x}}{2}, \quad \cosh(x) = \frac{\mathrm{e}^x + \mathrm{e}^{-x}}{2}, \quad \tanh(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{\mathrm{e}^x - \mathrm{e}^{-x}}{\mathrm{e}^x + \mathrm{e}^{-x}}$$

für alle $x \in \mathbb{R}$.

### 7.4.1 Der Areasinus Hyperbolicus

Nach Übung 7.4 gilt $\sinh'(x) = \cosh(x) > 0$ für alle $x \in \mathbb{R}$. Somit ist also nach Korollar 7.35 der Sinus Hyperbolicus streng monoton wachsend. Da $\lim_{x\to\infty} \sinh(x) = \infty$ und $\lim_{x\to-\infty} \sinh(x) = -\infty$ gilt, erhalten wir nach dem Zwischenwertsatz 3.59, dass

$$\sinh : \mathbb{R} \to \mathbb{R}$$

streng monoton wachsend und bijektiv ist. Die Umkehrabbildung

$$\operatorname{arsinh} : \mathbb{R} \to \mathbb{R}$$

nennen wir den **Areasinus Hyperbolicus**. Nach dem Satz zur Differenzierbarkeit der inversen Funktion ist arsinh differenzierbar und es gilt für $x \in \mathbb{R}$ und $s = \sinh(x)$

$$\operatorname{arsinh}'(s) = \frac{1}{\cosh(x)} = \frac{1}{\sqrt{1 + \sinh^2(x)}} = \frac{1}{\sqrt{1+s^2}}.$$

Der **Areasinus Hyperbolicus** besitzt im Gegensatz zu den Umkehrfunktionen arcsin, arccos und arctan eine geschlossene Form. In der Tat gilt

$$\operatorname{arsinh}(s) = \log\left(s + \sqrt{1+s^2}\right)$$

für alle $s \in \mathbb{R}$, wobei man beachten sollte, dass der Ausdruck rechts für alle $s \in \mathbb{R}$ Sinn ergibt. Kurzes Nachrechnen ergibt für $s \in \mathbb{R}$ und $x = \log(s + \sqrt{1+s^2})$

$$\frac{\mathrm{e}^x - \mathrm{e}^{-x}}{2} = \frac{1}{2}\left(s + \sqrt{1+s^2} - \frac{1}{s + \sqrt{1+s^2}}\right) = \frac{1}{2}\left(s + \sqrt{1+s^2} - \frac{s - \sqrt{1+s^2}}{s^2 - 1 - s^2}\right) = s$$

wie gewünscht.

### 7.4.2 Der Areakosinus Hyperbolicus

Der Kosinus Hyperbolicus erfüllt $\cosh'(x) = \sinh(x)$ und $\cosh''(x) = \cosh(x) > 0$ für alle $x \in \mathbb{R}$ nach Übung 7.4. Insbesondere ist der Kosinus Hyperbolicus streng konvex nach Korollar 7.42 und hat ein globales Minimum bei 0 nach Korollar 7.37 (wieso?). Für $x > 0$

gilt $\cosh'(x) > 0$ und somit ist cosh auf $\mathbb{R}_{\geq 0}$ streng monoton wachsend. Da $\cosh(0) = 1$ und $\lim_{x\to\infty} \cosh(x) = +\infty$, folgt, dass

$$\cosh : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 1}$$

streng monoton wachsend und bijektiv ist. Die Umkehrabbildung

$$\operatorname{arcosh} : \mathbb{R}_{\geq 1} \to \mathbb{R}_{\geq 0}$$

wird der **Areakosinus Hyperbolicus** genannt, ist auf $\mathbb{R}_{>1}$ differenzierbar und erfüllt

$$\operatorname{arcosh}'(s) = \frac{1}{\sinh(x)} = \frac{1}{\sqrt{s^2 - 1}}$$

für $s > 1$ und $s = \cosh(x)$ mit $x > 0$. Des Weiteren gilt

$$\operatorname{arcosh}(s) = \log\left(s + \sqrt{s^2 - 1}\right)$$

für alle $s > 1$. Der Nachweis der obigen Eigenschaften des Areakosinus Hyperbolicus und der noch folgenden Eigenschaften überlassen wir Interessierten.

### 7.4.3 Der Areatangens Hyperbolicus

Der **Areatangens Hyperbolicus** ist die Umkehrfunktion

$$\operatorname{artanh} : (-1, 1) \to \mathbb{R},\ x \mapsto \tfrac{1}{2} \log\left(\frac{1+x}{1-x}\right)$$

der streng monoton wachsenden Bijektion

$$\tanh : \mathbb{R} \to (-1, 1).$$

Des Weiteren ist nach dem Satz zur inversen Funktion (Satz 7.14) artanh differenzierbar und es gilt

$$\operatorname{artanh}'(s) = \frac{1}{1 - s^2}$$

für alle $s \in (-1, 1)$.

**Übung 7.56.** *Verifizieren Sie die oben aufgestellten Behauptungen.*

## 7.5 Erste Differentialgleichungen

Eine **Differentialgleichung** ist eine Gleichung der Form $F(y', y, x) = 0$, die eine reellwertige Funktion $y = y(x)$, ihre Ableitung $y' = \frac{\mathrm{d}y}{\mathrm{d}x}$ und die unabhängige Variable $x$ mittels einer reellwertigen Abbildung $F$ auf einer offenen Teilmenge von $\mathbb{R}^3$ verknüpft. Gesucht sind die Funktionen $y : I \to \mathbb{R}$, die die Gleichung $F(y'(x), y(x), x) = 0$ für alle $x \in I$ erfüllen, wobei der Definitionsbereich $I$ der gesuchten Funktion $y$ ein Intervall $I \subseteq \mathbb{R}$ sein soll.

Da es normalerweise mehrere Lösungen gibt, verlangt man üblicherweise noch etwas mehr Information, nämlich den Wert der Funktion $y(x_0) = y_0$ bei einem fest gewählten Ausgangspunkt $x_0 \in I$. Eine Differentialgleichung $F(y', y, x) = 0$ gemeinsam mit der Bedingung $y(x_0) = y_0$ wird ein **Anfangswertproblem** genannt.

Um genau zu sein, nennt man eine Differentialgleichung $F(y', y, x) = 0$ eine **gewöhnliche Differentialgleichung erster Ordnung**. Hier steht „gewöhnlich“ dafür, dass wir nur eine unabhängige Variable verwenden. „Erster Ordnung“ steht dafür, dass nur die erste (und keine höheren Ableitungen) der gesuchten Lösung in der durch $F$ gegebenen Gleichung erscheint. Wir werden häufig Differentialgleichungen der **expliziten Form** $y' = f(y, x)$ betrachten, wobei dann $f$ eine reellwertige Funktion darstellt, die auf einer offenen Teilmenge von $\mathbb{R}^2$ definiert sein soll.

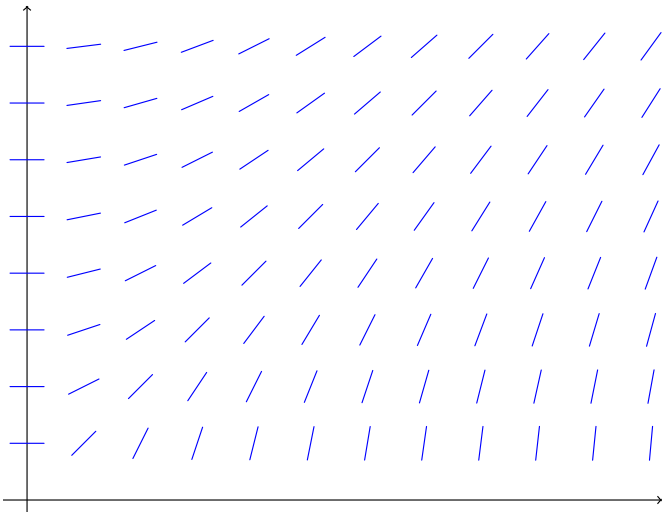
Ein Beispiel einer solchen gewöhnlichen Differentialgleichung erster Ordnung ist die Gleichung $y' = \frac{x}{y}$. Durch Anstarren der Gleichung lässt sich eine Lösung dieser Differentialgleichung erraten, nämlich $y = x$. Wir stellen aber fest, dass es ausgehend von dieser Lösung nicht klar ist, ob es andere Lösungen gibt und wie solche aussehen könnten. Auch unklar ist, wie viele Lösungen vor und nach Angabe eines Anfangswert existieren (und ob überhaupt welche existieren). Wir begnügen uns in diesem Abschnitt mit der Betrachtung von einigen Spezialfällen und werden gegen Ende des zweiten Semesters nochmals ausführlicher auf das Thema der Differentialgleichungen eingehen.

Wir möchten kurz anmerken, dass gewöhnliche Differentialgleichungen (erster Ordnung und höherer Ordnungen) in vielen Gebieten der Mathematik und der Naturwissenschaften auftreten. Beispielsweise hat das Zerfallsgesetz der Physik, mit dem (unter anderem) die Anzahl $N$ Teilchen in einem radioaktiven Stoff beschrieben werden können, eine natürliche Beschreibung in der Differentialgleichung $\dot{N} = \frac{\mathrm{d}N}{\mathrm{d}t} = -\lambda N$ (für einen physikalischen Parameter $\lambda > 0$).[2]

Wir können uns die rechte Seite der Differentialgleichung $y' = f(y, x)$ auch als Richtungsfeld vorstellen, welches bei jedem Punkt $(x, y)$ eine vorgegebene Steigung der gesuchten Lösung der Differentialgleichung angibt. Dieses **Richtungsfeld** können wir mit kleinen Strichen in der richtigen Steigung in einem Bild visualisieren; die gesuchte Lösung sollte dann einen Graphen besitzen, der bei jedem Punkt des Graphen den vorgegebenen Strich bei dem Punkt als Tangente besitzt. Noch umgangssprachlicher formuliert, gibt das Richtungsfeld unendlich viele bereits fertig verlegte Schienen in der Ebene an und die gesuchte Lösung eines Anfangswertproblems zeigt uns, wohin der Zug fährt, wenn er in einem bestimmten Punkt wegfährt und

[2] Das macht exakt gesehen kaum Sinn, denn die Anzahl sollte ja eigentlich eine natürliche Zahl sein und eine differenzierbare Funktion auf einem Intervall mit Werten in $\mathbb{N}$ ist konstant. Da aber die Anzahl sehr gross und die Änderung der Anzahl in einer kleinen Zeitspanne relativ gesehen klein ist, macht es aus praktischen Gründen doch Sinn, dies als Differentialgleichung aufzufassen.

den vorgegebenen Schienen folgt. In dieser Formulierung können wir fragen, ob wir immer angeben können, wohin der Zug fährt oder ob es nicht vielleicht auch „Weichen im Richtungsfeld“ geben könnte.



Figur 7.8: Das Richtungsfeld zur Differentialgleichung $y' = \frac{x}{y}$.

Wir bemerken noch, dass das Lösen der Differentialgleichung, also das Auffinden der Lösung $y = y(x)$, mitunter schwierig ist, doch das Nachprüfen, ob eine Lösung vorliegt, auf Grund unserer Ableitungsregeln aus Abschnitt 7.1 meist sehr einfach ist.

**Applet 7.57** (Einige Richtungsfelder und Anfangswertprobleme)**.** *Wir betrachten einige Funktionen $f$, die verschiedene Richtungsfelder angeben, und die dazugehörigen Anfangswertprobleme für einen bewegbaren Startpunkt $(x_0, y_0)$.*

### 7.5.1 Differenzengleichungen

Ein diskretes Analogon der Differentialgleichung $y' = F(y, x)$ ist die Differenzengleichung

$$y(n+1) - y(n) = \triangle y(n) = F(y(n), n)$$

für eine gesuchte Funktion $y$ auf $\mathbb{N}_0$ und eine gegebene Funktion $F$ auf $\mathbb{R} \times \mathbb{N}_0$. Für diese Gleichung kann man rekursiv eine Lösung $y$ bestimmen, wenn man eine Anfangswertbedingung $y(0) = y_0 \in \mathbb{R}$ gegeben hat (siehe die Besprechung der Rekursion in Abschnitt 2.2). In der Tat können wir

$$y(1) = y(0) + F(y(0), 0), \quad y(2) = y(1) + F(y(1), 1), \quad \ldots$$

setzen und erhalten eine rekursiv bestimmte Lösung. Man beachte dabei, dass die Lösung eindeutig bestimmt ist (wieso?).

Die Differentialgleichung $y' = F(y, x)$ sollte als eine kontinuierliche Version der Differenzengleichung $\triangle y(n) = F(y(n), n)$ aufgefasst werden. Diese Analogie ist wohlgemerkt nicht ausschliesslich oberflächlich, sondern kann sowohl in der Praxis als auch in der Theorie zu wichtigen Ergebnissen führen.

Es kann zum Beispiel sein, dass man eigentlich an einem diskreten Problem interessiert ist, aber die einzelnen Schritte näherungsweise einer kleinen Zeitspanne $\triangle x$ in einem kontinuierlichen Problem entsprechen. In diesem Fall ist es manchmal einfacher, anstelle des diskreten Problems die entsprechende Differentialgleichung zu betrachten.

Umgekehrt kann es sein, dass ein Anfangswertproblem $y' = F(y, x)$, $y(0) = y_0$ auf dem Intervall $I = [0, \infty)$ gegeben ist, aber die Funktion $F$ zu kompliziert ist, um dieses mittels Standardfunktionen zu lösen. Stattdessen kann man in diesem Fall die Differentialgleichung in eine Differenzengleichung verwandeln. Sei also $\triangle x = h > 0$ eine kleine positive Zahl. Dann kann man eine Funktion $\tilde{y}$ auf $\mathbb{N}_0 h$ durch

$$\tilde{y}(0) = y_0, \quad \tilde{y}(h) = \tilde{y}(0) + F(\tilde{y}(0), 0)h, \quad \tilde{y}(2h) = \tilde{y}(h) + F(\tilde{y}(h), h)h, \quad \dots$$

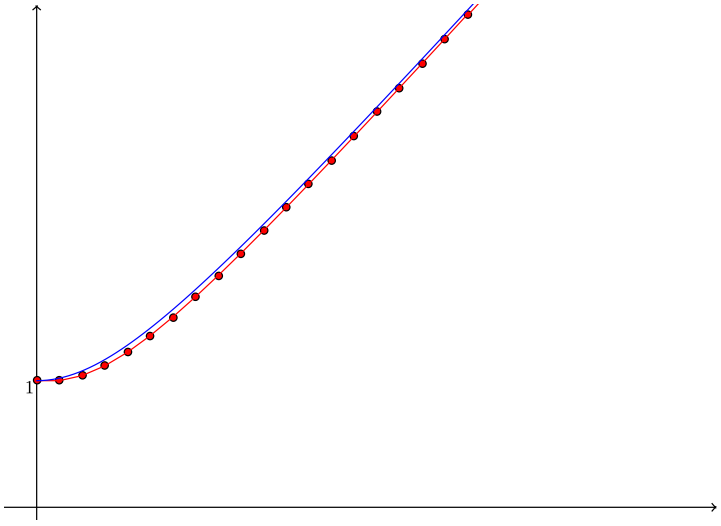
definieren. Dabei hofft man, dass die gesuchte Lösung der Differentialgleichung $y$ und die Lösung obiger Differenzengleichung $\tilde{y}$ einander wegen

$$\begin{aligned}
y(0) &= y_0 = \tilde{y}(0), \\
\frac{y(h) - y(0)}{h} &\approx \frac{\mathrm{d}y}{\mathrm{d}x}(0) = F(y(0), 0) = \frac{\tilde{y}(h) - \tilde{y}(0)}{h} \\
&\implies y(h) \approx \tilde{y}(h) \\
\frac{y(2h) - y(h)}{h} &\approx \frac{\mathrm{d}y}{\mathrm{d}x}(h) = F(y(h), h) \approx F(\tilde{y}(h), h) = \frac{\tilde{y}(2h) - \tilde{y}(h)}{h} \\
&\implies y(2h) \approx \tilde{y}(2h) \\
&\dots
\end{aligned}$$

ähnlich sind.

Diese Beschreibung ist natürlich bloss eine Heuristik (das Zeichen $\approx$ sollte dies ersichtlich machen). Sie kann jedoch in gewissen Fällen zu einer approximativen numerischen Lösung oder gar zu einem Beweis der Existenz einer Lösung führen.

**Übung 7.58** (Rekursive Näherung)**.** *Zeigen Sie, dass obige Heuristik für das Anfangswertproblem $y' = y$, $y(0) = 1$ zu einer Lösung führt.*



Figur 7.9: Vergleich der rekursiv definierten approximativen Lösung $\tilde{y}$ (rote Punkte) mit linearer Interpolation (ebenfalls in Rot) zur Lösung $y$ in Blau. Das betrachtete Anfangswertproblem ist dabei $y' = \frac{x}{y}$, $y(0) = 1$, welches von $y = \sqrt{x^2+1}$ gelöst wird.

### 7.5.2 Stammfunktionen

Eine der einfachsten Differentialgleichungen ist eine Gleichung der Form

$$y' = f(x)$$

für eine gegebene Funktion $f : I \to \mathbb{R}$ auf einem Intervall $I \subseteq \mathbb{R}$. Eine Lösung $F : I \to \mathbb{R}$, das heisst, eine differenzierbare Funktion $F : I \to \mathbb{R}$ mit $F'(x) = f(x)$ für alle $x \in I$, wird eine **Stammfunktion** von $f$ genannt. Wir schreiben auch

$$\int f(x)\,\mathrm{d}x = F(x) + C,$$

falls $F' = f$, wobei $C \in \mathbb{R}$ eine unbestimmte Konstante – die **Integrationskonstante** – ist. Wir bemerken, dass $F(x) + C$ auch $(F(x) + C)' = f(x)$ erfüllt, falls $F$ eine Stammfunktion und $C \in \mathbb{R}$ eine Konstante ist. Wir bezeichnen $\int f(x)\,\mathrm{d}x$ als das **unbestimmte Integral**. Dabei ist es noch nicht klar, inwiefern das unbestimmte Integral von der Wahl einer Stammfunktion von $f$ abhängt, und was die Integraldarstellung mit dem Riemann-Integral zu tun hat. Die Notation wird zum Teil in folgendem Lemma (und vollständig in Kapitel 8) erklärt.

**Lemma 7.59** (Integrationskonstante)**.** *Sei $I \subseteq \mathbb{R}$ ein Intervall, $f : I \to \mathbb{R}$ eine Funktion und $F, F_1 : I \to \mathbb{R}$ Stammfunktionen von $f$. Dann gibt es eine Konstante $C \in \mathbb{R}$ mit $F_1 = F + C$. In anderen Worten, alle Lösungen von $y' = f$ sind gegeben durch die Formel $y = F(x) + C$, wenn wir die Konstante $C \in \mathbb{R}$ variieren.*

*Beweis.* Sei $G = F_1 - F$. Dann ist $G'(x) = F_1'(x) - F'(x) = f(x) - f(x) = 0$ für alle $x \in I$. Aber nach Korollar 7.33 des Mittelwertsatzes muss eine differenzierbare Funktion auf einem Intervall mit Ableitung Null konstant sein. Somit folgt $G(x) = C$ für ein $C \in \mathbb{R}$ und alle $x \in I$, und damit ebenso die Aussage. □

Das unbestimmte Integral bezeichnet per Definition die Umkehroperation zur Differentiation und die Integrationskonstante deutet an, dass wir jede weitere Stammfunktion erhalten können, indem wir zu einer bekannten Stammfunktion eine beliebige Konstante addieren.

Da wir bereits viele Ableitungsregeln (siehe Abschnitte 7.1 und 7.3) kennen, können wir diese rückwärts als Integrationsregeln (zur Bestimmung des unbestimmten Integrals) lesen. Wir werden dies systematisch im nächsten Kapitel besprechen und wollen hier nur einige erste Regeln ansprechen. Zum Beispiel gilt für $s \in \mathbb{R}$ (oder sogar für $s \in \mathbb{C}$)

$$\int x^s\,\mathrm{d}x = \begin{cases} \frac{1}{s+1}x^{s+1} + C & \text{falls } s \neq -1 \\ \log|x| + C & \text{falls } s = -1 \end{cases}$$

nach Beispiel 7.15,

$$\begin{aligned}
\int \exp(x)\,\mathrm{d}x &= \exp(x) + C \\
\int \cos(x)\,\mathrm{d}x &= \sin(x) + C \\
\int \sin(x)\,\mathrm{d}x &= -\cos(x) + C \\
\int \sinh(x)\,\mathrm{d}x &= \cosh(x) + C \\
\int \cosh(x)\,\mathrm{d}x &= \sinh(x) + C \\
\int \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x &= \arcsin(x) + C \\
\int \frac{1}{1+x^2}\,\mathrm{d}x &= \arctan(x) + C \\
\int \frac{1}{\sqrt{1+x^2}}\,\mathrm{d}x &= \operatorname{arsinh}(x) + C = \log\left(x + \sqrt{1+x^2}\right) + C \\
\int \frac{1}{\sqrt{x^2-1}}\,\mathrm{d}x &= \operatorname{arcosh}(x) + C = \log\left(x + \sqrt{x^2-1}\right) + C
\end{aligned}$$

nach Beispiel 7.3(iii), Übung 7.4(ii) und den Abschnitten 7.3–7.4.

Wir begnügen uns vorerst mit dieser Liste und besprechen im nächsten Kapitel weitere Methoden der Berechnung des unbestimmten Integrals, aber erst nachdem wir die Frage „Was hat das so definierte unbestimmte Integral mit dem Riemann-Integral zu tun?“ beantwortet haben und dadurch noch mehr Motivation für die Betrachtung des unbestimmten Integral erhalten haben.

VO: 07.12.17

### 7.5.3 Lineare Differentialgleichungen erster Ordnung

Der nächst einfache Typ einer Differentialgleichung besteht aus den **linearen Differentialgleichungen erster Ordnung**, welche von der Form

$$y' + f(x)y = g(x) \tag{7.13}$$

für zwei gegebene Funktionen $f, g$ besteht. In der Gleichung $y' + f(x)y = g(x)$ wird die Funktion $g$ auch die **Störfunktion** genannt. Falls die Störfunktion Null ist, nennen wir (7.13) **homogen** und sonst **inhomogen**.

Die Bezeichnung „linear“ entstammt der Tatsache, dass sich Gleichung (7.13) in der Tat als ein lineares Gleichungssystem auf geeigneten Vektorräumen auffassen lässt. Informell sieht man schnell, dass die Abbildung $y \mapsto y' + f(x)y$ linear ist. Welche Vektorräume man dabei jedoch betrachten soll, hängt stark von den Eigenschaften der Funktionen $f$ und $g$ ab. Wir werden später etwas genauer auf diese Fragestellung eingehen, wenn wir allgemeiner Lösbarkeit von Differentialgleichungen diskutieren. Nun möchten wir aber eine Lösung von (7.13) finden, wobei wir zuerst den homogenen Fall thematisieren.

**Lemma 7.60** (Homogene lineare Differentialgleichungen erster Ordnung)**.** *Sei $I \subseteq \mathbb{R}$ ein Intervall, $f : I \to \mathbb{R}$ eine Funktion und $F : I \to \mathbb{R}$ eine Stammfunktion von $f$. Die Lösungen $y : I \to \mathbb{R}$ der homogenen linearen Differentialgleichung erster Ordnung $y' + f(x)y = 0$ sind genau die Vielfachen der Funktion $x \in I \mapsto \exp(-F(x))$.*

*Beweis.* Für die Funktion $y : x \in I \mapsto A\exp(-F(x))$ zu $A \in \mathbb{R}$ gilt

$$y'(x) = A\exp(-F(x))(-F'(x)) = -f(x)A\exp(-F(x)) = -f(x)y(x)$$

für alle $x \in I$ und somit $y' + f(x)y = 0$.

Sei nun $y$ eine beliebige Lösung der homogenen Differentialgleichung $y' + f(x)y = 0$. Wir definieren die Funktion $\tilde{y} : x \in I \mapsto \exp(F(x))y(x)$ und berechnen

$$\tilde{y}'(x) = \exp(F(x))f(x)y(x) + \exp(F(x))y'(x) = \exp(F(x))(f(x)y(x) - f(x)y(x)) = 0$$

für alle $x \in I$. Daher ist auf Grund von Korollar 7.33 $\tilde{y} = A$ für eine Konstante $A \in \mathbb{R}$ und somit

$$y(x) = \exp(-F(x))\tilde{y}(x) = A\exp(-F(x))$$

für alle $x \in I$ wie gewünscht. □

**Lemma 7.61** (Inhomogene lineare Differentialgleichungen erster Ordnung)**.** *Sei $I \subseteq \mathbb{R}$ ein Intervall und seien $f, g : I \to \mathbb{R}$ Funktionen. Falls es eine Lösung $y_{\text{part}} : I \to \mathbb{R}$ der Differentialgleichung $y' + fy = g$ gibt (die auch die **partikuläre** Lösung genannt wird), dann ist die allgemeine Lösung $y_{\text{inhom}}$ von der Form $y_{\text{inhom}} = y_{\text{part}} + y_{\text{hom}}$, wobei $y_{\text{hom}}$ die allgemeine Lösung der homogenen Gleichung $y' + f(x)y = 0$ ist.*

**Übung 7.62.** *Beweisen Sie Lemma 7.61. Vergleichen Sie Lemma 7.61 des Weiteren mit folgender Tatsache aus der linearen Algebra. Ist $F : V \to W$ eine lineare Abbildung zwischen Vektorräumen $V, W$ über einem Körper $K$ und $v \in V$ und $w \in W$ erfüllen $F(v) = w$, dann ist jedes $\tilde{v} \in V$ mit $F(\tilde{v}) = w$ von der Form $\tilde{v} = v + v_0$ für $v_0$ im Kern der Abbildung $F$.*

Lemma 7.61 ist natürlich nur dann interessant, wenn eine partikuläre Lösung bekannt ist. Ein nützlicher Trick, um eine solche zu finden, ist die **Variation der Konstanten**. Hierbei nimmt man an, dass in der Lösung $y_{\text{hom}}(x) = A\exp(-F(x))$ der homogenen Gleichung $A = A(x)$ eine differenzierbare Funktion der Variable $x$ statt einer Konstante ist. Das heisst, wir setzen $y(x) = A(x)\exp(-F(x))$ für alle $x \in I$, berechnen

$$\begin{aligned} y'(x) &= A'(x)\exp(-F(x)) - A(x)f(x)\exp(-F(x)) \\ y'(x) + f(x)y(x) &= A'(x)\exp(-F(x)) \end{aligned}$$

und wollen also $A'(x)\exp(-F(x)) = g(x)$ lösen. Dies führt zu $A'(x) = g(x)\exp(F(x))$ und

$$A(x) = \int g(x)\exp(F(x))\,\mathrm{d}x.$$

Zusammenfassend kann man also eine partikuläre Lösung der inhomogenen Differentialgleichung $y' + fy = g$ finden, indem man eine Stammfunktion $A(x)$ von $g(x)\exp(F(x))$ findet und dann

$$y_{\text{part}} : x \in I \mapsto A(x)\exp(-F(x))$$

setzt.

Die allgemeine Lösung $y_{\text{inhom}}$ der inhomogenen Differentialgleichung $y' + fy = g$ enthält nach Lemma 7.61 eine unbekannte Konstante (die in einer Lösung der homogenen Gleichung versteckt ist). Wenn nun zusätzlich ein Anfangswert $y(x_0) = y_0$ für $x_0 \in I$ gegeben ist, dann kann man diesen zur Bestimmung der Konstante verwenden und dadurch das Anfangswertproblem lösen.

Um sich obiges Verfahren für das Anfangswertproblem $y' + fy = g$, $y(x_0) = y_0$ zu merken, kann man auch folgendes „Kochrezept“ durchlaufen, die zum Teil die Leibniz-Notation verwenden und wegen obiger Diskussion zum richtigen Resultat führen.

- **Trennung der Variablen in der homogenen Gleichung:**

$$\begin{aligned}
y'_{\text{hom}} + f(x)y_{\text{hom}} &= 0 \\
\frac{\mathrm{d}y_{\text{hom}}}{\mathrm{d}x} &= -f(x)y_{\text{hom}} \\
\frac{\mathrm{d}y_{\text{hom}}}{y_{\text{hom}}} &= -f(x)\,\mathrm{d}x \\
\int \frac{\mathrm{d}y_{\text{hom}}}{y_{\text{hom}}} &= -\int f(x)\,\mathrm{d}x \\
\ln|y_{\text{hom}}| &= -F(x) + C \\
|y_{\text{hom}}| &= \mathrm{e}^C \exp(-F(x)) \\
y_{\text{hom}}(x) &= A\exp(-F(x))
\end{aligned}$$

- **Variation der Konstanten:** Mit dem Ansatz $y_{\text{part}} = A(x)\exp(-F(x))$ und der Differentialgleichung $y'_{\text{part}} + f(x)y_{\text{part}} = g(x)$ erhält man eine Differentialgleichung für $A(x)$.

- **Bestimmung der Konstanten:** Setze

$$y = y_{\text{part}} + y_{\text{hom}} = A(x)\exp(-F(x)) + A\exp(-F(x))$$

  und $y(x_0) = y_0 = A(x_0)\exp(-F(x_0)) + A\exp(-F(x_0))$, um $A \in \mathbb{R}$ zu bestimmen.

Wir bemerken allerdings, dass der erste Schritt obiges Kochrezepts eigentlich nicht alle Lösungen lieferte. In der Tat haben wir zur Vereinfachung der Notation $A = \pm\mathrm{e}^C \neq 0$ gesetzt, dann aber einfach von der Konstante $A \in \mathbb{R}$ gesprochen, wodurch wir die triviale Lösung der homogenen Differentialgleichung $y = 0$ wiedergewonnen haben. Dieses und auch ähnliche Kochrezepte sind später sehr nützlich um Lösungen zu finden. Einmal gefunden, ist es normalerweise auch ein leichtes zu überprüfen, ob eine Funktion eine Lösung darstellt (und dies

wird üblicherweise auch der Fall sein). Doch sollten wir uns bewusst sein, dass das Kochrezept möglicherweise nicht alle Lösungen liefert.

**Beispiel 7.63** (Trennung der Variablen)**.** *Wir möchten das Anfangswertproblem*

$$\begin{aligned} y' - 2xy &= \mathrm{e}^{x^2} \\ y(0) &= 1 \end{aligned}$$

*auf* $\mathbb{R}$ *lösen.*

- *Nach Lemma 7.60 berechnen wir zu* $f(x) = -2x$

$$\int (-2x)\,\mathrm{d}x = -x^2 + C$$

  *und wählen somit* $F(x) = -x^2$ *als Stammfunktion. Die allgemeine Lösung der homogenen Differentialgleichung* $y' - 2xy = 0$ *ist somit von der Form* $y_{\mathrm{hom}}(x) = A\mathrm{e}^{x^2}$ *für* $A \in \mathbb{R}$.

- *Um eine partikuläre Lösung zu finden, brauchen wir eine Stammfunktion der Funktion* $g(x)\exp(F(x)) = 1$, *wobei* $g(x) = \mathrm{e}^{x^2}$. *Wir setzen somit* $y_{\mathrm{part}}(x) = x\mathrm{e}^{x^2}$.

- *Die allgemeine Lösung der inhomogenen Gleichung ist nach Lemma 7.61 somit von der Form* $y(x) = A\mathrm{e}^{x^2} + x\mathrm{e}^{x^2}$. *Unter Verwendung des Anfangswerts erhalten wir* $y(0) = A = 1$ *und somit ist die eindeutig bestimmte Lösung des obigen Anfangswertproblems durch*

$$y(x) = \mathrm{e}^{x^2} + x\mathrm{e}^{x^2} = (x+1)\mathrm{e}^{x^2}$$

  *gegeben. An dieser Stelle empfiehlt es sich durch Einsetzen zu überprüfen, dass* $y$ *tatsächlich eine Lösung ist.*

**Übung 7.64.** *Finden Sie eine Lösung des Anfangswertproblems*

$$\begin{aligned} y' - \left(\frac{4}{x} + 1\right) y &= x^4 \\ y(1) &= 1 \end{aligned}$$

*auf dem Intervall* $(0, \infty)$.

### 7.5.4 Zweite Ordnung

Wir wollen in diesem Unterabschnitt den einfachsten Typ einer gewöhnlichen Differentialgleichung zweiter Ordnung betrachten, nämlich lineare Differentialgleichungen mit konstanten Koeffizienten. Hierbei wird für vorgegebene $a_0, a_1 \in \mathbb{C}$ die Gleichung

$$y'' + a_1 y' + a_0 y = 0 \tag{7.14}$$

als **homogene lineare gewöhnliche Differentialgleichungen zweiter Ordnung mit konstanten Koeffizienten** bezeichnet. Für eine vorgegebene Funktion $g$ wird

$$y'' + a_1 y' + a_0 y = g$$

als **inhomogene lineare gewöhnliche Differentialgleichungen zweiter Ordnung mit konstanten Koeffizienten und Störfunktion** $g$ bezeichnet.

Bevor wir das allgemeine Verfahren zur Berechnung der Lösungen dieser Differentialgleichungen vorstellen, wollen wir zuerst einige Spezialfälle betrachten, welche auch als Motivation für das allgemeine Verfahren betrachtet werden sollten.

**Beispiel 7.65** (Einige homogene Gleichungen)**.** *Die folgenden Behauptungen sind sehr leicht mit Differentiation überprüfbar. (Die Anmerkungen zu den betrachteten Polynomen soll gewisse Übereinstimmungen andeuten und sollte zumindest nach der Besprechung des allgemeinen Verfahrens für Sie Sinn machen.)*

*(a) Für die Differentialgleichung $y'' = 0$ sind die Funktionen*

$$x \in \mathbb{R} \mapsto y(x) = C_1 + C_2 x$$

*für beliebige Konstanten $C_1, C_2 \in \mathbb{C}$ Lösungen. (Die $0$ ist die einzige Lösung der Gleichung $T^2 = 0$.)*

*(b) Für die Differentialgleichung $y'' - y' = 0$ sind die Funktionen*

$$x \in \mathbb{R} \mapsto y(x) = C_1 + C_2 \exp(x)$$

*für beliebige Konstanten $C_1, C_2 \in \mathbb{C}$ Lösungen. (Die Lösungen der Gleichung $T^2 - T = 0$ sind gegeben durch $0$ und $1$.)*

*(c) Für die Differentialgleichung $y'' + y = 0$ sind die Funktionen*

$$x \in \mathbb{R} \mapsto y(x) = C_1 \sin(x) + C_2 \cos(x)$$

*für beliebige Konstanten $C_1, C_2 \in \mathbb{C}$ Lösungen. Ebenso sind die Funktionen*

$$x \in \mathbb{R} \mapsto y(x) = D_1 \exp(\mathrm{i}x) + D_2 \exp(-\mathrm{i}x)$$

*für beliebige Konstanten $D_1, D_2 \in \mathbb{C}$ Lösungen, und auf Grund des Zusammenhangs zwischen der komplexen Exponentialabbildung und den trigonometrischen Funktionen in Satz 6.72 werden dadurch die gleiche Menge an Lösungen beschrieben. (Die Lösungen der Gleichung $T^2 + 1 = 0$ sind gegeben durch $\mathrm{i}, -\mathrm{i} \in \mathbb{C}$.)*

Obige Beispiele legen den Ansatz $y = \exp(\alpha x)$ für die Lösungen der Differentialgleichung nahe, wobei man $\alpha \in \mathbb{C}$ noch bestimmen muss (und dies nicht unbedingt alle Lösungen liefert). In der Tat das allgemeine Kochrezept für die Lösung der homogenen Differentialgleichung in

(7.14) besteht darin, dass wir die Koeffizienten der Differentialgleichung als Koeffizienten des sogenannten **charakteristischen Polynoms**

$$p(T) = T^2 + a_1 T + a_0$$

verwenden. Anschliessend müssen die Nullstellen der Gleichung $p(T) = 0$ berechnet werden.

- Falls es zwei verschiedene Nullstellen $\alpha_1, \alpha_2 \in \mathbb{C}$ gibt, so sind die Funktionen

$$y(x) = C_1 \exp(\alpha_1 x) + C_2 \exp(\alpha_2 x) \tag{7.15}$$

  für beliebige Konstanten $C_1, C_2 \in \mathbb{C}$ Lösungen der homogenen Differentialgleichung (7.14).

- Falls die Koeffizienten $a_0, a_1$ reell sind und die Nullstellen $\alpha_1 = \alpha$ durch eine komplexe Zahl $\alpha = \beta + \gamma \mathrm{i}$ mit $\beta \in \mathbb{R}$ und $\gamma > 0$ und die Konjugierte $\alpha_2 = \overline{\alpha}$ beschrieben werden, so sind die Funktionen

$$y(x) = C_1 \exp(\beta x) \sin(\gamma x) + C_2 \exp(\beta x) \cos(\gamma x) \tag{7.16}$$

  für beliebige Konstanten $C_1, C_2 \in \mathbb{C}$ Lösungen der homogenen Differentialgleichung (7.14).

- Falls es nur eine Nullstelle $\alpha \in \mathbb{C}$ des charakteristischen Polynoms $p(T) = (T - \alpha)^2$ gibt, so sind die Funktionen

$$y(x) = C_1 \exp(\alpha x) + C_2 x \exp(\alpha x) \tag{7.17}$$

  für beliebige Konstanten $C_1, C_2 \in \mathbb{C}$ Lösungen der homogenen Differentialgleichung (7.14).

Mit Hilfe der Ableitungsregeln lässt sich nun überprüfen, dass dieses Verfahren in der Tat Lösungen liefert. Hierfür ist der wichtigste Schritt die folgende Rechnung. Angenommen $a_0, a_1, \alpha \in \mathbb{C}$ sind Konstanten, welche wir verwenden, um das Polynom $p(T) = T^2 + a_1 T + a_0$ und die Funktion $y : x \in \mathbb{R} \mapsto y(x) = \exp(\alpha x)$ zu definieren. Dann gilt $y' = \alpha y$, $y'' = \alpha^2 y$ und deshalb $y'' + a_1 y' + a_0 y = p(\alpha) y(x)$, was den Zusammenhang zwischen den Nullstellen des charakteristischen Polynom und den Lösungen der homogenen Differentialgleichung erklärt.

**Wichtige Übung 7.66** (Homogene Gleichung)**.**

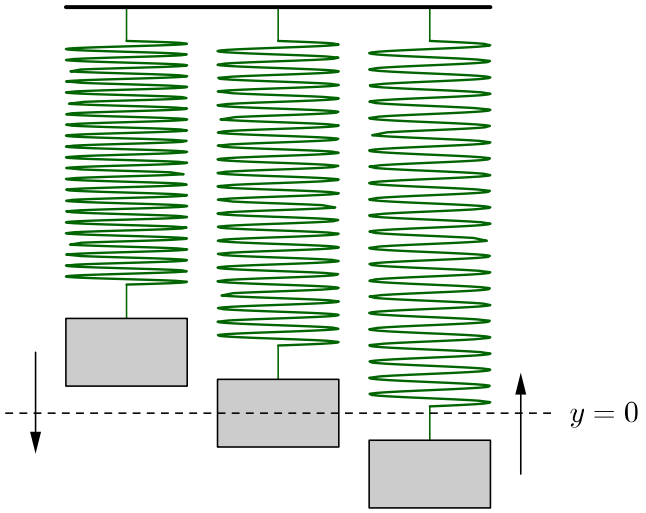
*(i) Zeigen Sie, dass die Menge der Lösungen der homogenen Differentialgleichung* (7.14) *auf* $\mathbb{R}$ *einen Teilraum des Vektorraums aller zweimal differenzierbarer Funktionen auf* $\mathbb{R}$ *bildet.*

*(ii) Zeigen Sie, dass die Funktionen in* (7.15), (7.16) *beziehungsweise* (7.17) *Lösungen von* (7.14) *sind.*

*(iii) Zeigen Sie, dass wir in den Spezialfällen von Beispiel 7.65 tatsächlich alle Lösungen der jeweiligen Differentialgleichung gefunden haben.*

Wir behandeln nun ein wichtiges Beispiel aus der Physik.

**Beispiel 7.67** (Gedämpfte Schwingung)**.** *Wir bringen ein Gewicht an einer elastischen Feder an und wählen das Koordinatensystem, so dass $y = 0$ dem Ruhezustand (wo sich das Gewicht nicht bewegt) entspricht.*




*Wir wollen die Position $y(t)$ des Gewichts als Funktion der Zeit $t$ betrachten. Nach den Newtonschen Grundgesetzen der Bewegung ist die zweite Ableitung $\ddot{y}$ (nach der Zeit) multipliziert mit der Masse $m$ des Gewichts gleich der Kraft, die auf das Gewicht wirkt. Eine Komponente dieser Kraft entsteht durch die Ausdehnung der Feder und orientiert sich in Richtung Ruhezustand. Nach dem Hookeschen Gesetz ist diese Kraft durch $-ky$ gegeben, wobei $k > 0$ die Federkonstante genannt wird. Weiter wirken üblicherweise Reibungkräfte auf die Bewegung. Wir nehmen an, dass die entsprechende Krafteinwirkung durch $-d\dot{y}$ gegeben ist, wobei $d \geq 0$ die Dämpfungskonstante ist. Die Differentialgleichung, die die Bewegung $y(t)$ der Masse beschreibt, ist somit*

$$\begin{aligned} m\ddot{y} &= -d\dot{y} - ky \\ \ddot{y} + \frac{d}{m}\dot{y} + \frac{k}{m}y &= 0, \end{aligned}$$

*was also eine gewöhnliche lineare homogene Differentialgleichung zweiter Ordnung ist. Wir setzen zur Vereinfachung der Notation $m = 1$. Das charakteristische Polynom obiger Differentialgleichung ist*

$$p(T) = T^2 + dT + k$$

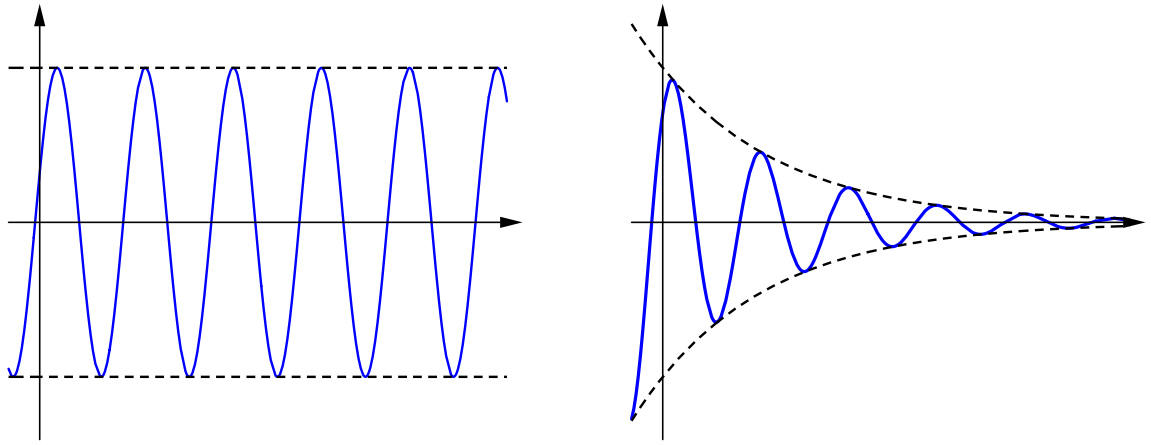
*mit Nullstellen*

$$\alpha = -\frac{d}{2} \pm \sqrt{\frac{d^2}{4} - k}.$$

*Für* $\frac{d^2}{4} - k < 0$ *sind dies komplex konjugierte Nullstellen* $\alpha = \beta \pm \gamma \mathrm{i}$, *wobei* $\beta = -\frac{d}{2}$ *und* $\gamma = \sqrt{\left|\frac{d^2}{4} - k\right|}$. *In diesem Fall erhält man somit Lösungen von der Form*

$$\mathrm{e}^{-\frac{d}{2}t}\big(A\sin(\gamma t) + B\cos(\gamma t)\big),$$

*siehe auch folgendes Bild.*



Figur 7.10: Die ungedämpfte Schwingung links (mit $d = 0$) und die gedämpfte Schwingung (mit $d > 0$) rechts.

*Falls die Reibung stark ist und* $\frac{d^2}{4} - k > 0$ *ist, dann wird die Schwingung zerstört – dies kann mit einem Türschliess-Mechanismus verglichen werden. Die Lösungen sind dann von der Form* $A\mathrm{e}^{\alpha_1 t} + B\mathrm{e}^{\alpha_2 t}$, *wobei* $\alpha_1 = -\frac{d}{2} + \sqrt{\frac{d^2}{4} - k} < 0$ *und* $\alpha_2 = -\frac{d}{2} - \sqrt{\frac{d^2}{4} - k} < 0$. *Es gibt also in diesem Fall zwei potenziell sehr verschiedene Lösungen, die unterschiedlich schnell zum Ruhezustand der Feder (also* $y = 0$*) streben.*

*Der Grenzfall* $\frac{d^2}{4} - k = 0$ *ist nochmals anders, da wir in diesem Fall bis jetzt nur eine eindimensionale Lösungsmenge bestehend aus allen Vielfachen von* $y_1 = \mathrm{e}^{-\frac{d}{2}t}$ *gefunden haben. Mit obiger Anleitung erhält man eine weitere Lösung* $y_2(t) = t\mathrm{e}^{-\frac{d}{2}t}$, *die linear unabhängig zu* $y_1$ *ist (wieso?). In der Tat gilt*

$$\begin{aligned}\dot{y}_2(t) &= \mathrm{e}^{-\frac{d}{2}t} - \tfrac{d}{2}t\mathrm{e}^{-\frac{d}{2}t}\\ \ddot{y}_2(t) &= -d\mathrm{e}^{-\frac{d}{2}t} + t\left(\tfrac{d^2}{4}\right)\mathrm{e}^{-\frac{d}{2}t}\end{aligned}$$

*und somit*

$$\ddot{y}_2(t) + d\dot{y}_2(t) + \tfrac{d^2}{4}y_2(t) = -d\mathrm{e}^{-\frac{d}{2}t} + \tfrac{d^2}{4}t\mathrm{e}^{-\frac{d}{2}t} + d\mathrm{e}^{-\frac{d}{2}t} - \tfrac{d^2}{2}t\mathrm{e}^{-\frac{d}{2}t} + \tfrac{d^2}{4}t\mathrm{e}^{-\frac{d}{2}t} = 0.$$

*für alle* $t \in \mathbb{R}$.

Wir wenden uns nun dem inhomogenen Problem zu, wobei wir annehmen wollen, dass die Störfunktion $g$ sich als Summe von Produkten von Polynomen und Exponentialabbildungen (oder auch dem Sinus und Kosinus) schreiben lässt. In diesen Fällen können wir spezifische

Ansätze verwenden, um eine erste Lösung der inhomogenen Differentialgleichung

$$y'' + a_1 y' + a_0 y = g \tag{7.18}$$

zu finden. Diese Lösung nennen wir auch die **partikuläre Lösung** $y_{\text{part}}$. Die allgemeine Lösung (mit zwei unbekannten Konstanten) der inhomogenen Differentialgleichung lässt sich dann auf Grund der Linearität als Summe der allgemeinen Lösung der homogenen Differentialgleichung (mit zwei unbekannten Konstanten) und der partikulären Lösung schreiben. Wir beschreiben das Rezept zur Berechnung der partikulären Lösung, wobei $p(T)$ wiederum das charakteristische Polynom der homogenen Differentialgleichung darstellt:

- Falls $g(x) = q(x)\mathrm{e}^{\alpha x}$ für ein Polynom $q(T)$ vom Grad $n$ und $\alpha \in \mathbb{C}$ mit $p(\alpha) \neq 0$, dann definiert man $y_{\text{part}} = Q(x)\mathrm{e}^{\alpha x}$, wobei $Q(T)$ ein Polynom vom Grad $n$ mit noch zu bestimmenden $n+1$ Koeffizienten ist. Nun berechnet man die linke Seite von (7.18) setzt dies gleich $g$ und verwendet diese Gleichung, um die Koeffizienten von $Q$ zu bestimmen.
- Falls $g(x) = q(x)\mathrm{e}^{\alpha x}$ für ein Polynom $q(T)$ vom Grad $n$ und $\alpha \in \mathbb{C}$ mit $p(\alpha) = 0$, dann wiederholt man obiges Verfahren, allerdings mit dem Ansatz $y_{\text{part}} = Q(x)x^{\ell}\mathrm{e}^{\alpha x}$, wobei $\ell$ die Vielfachheit der Nullstelle $\alpha$ von $p(T)$ angibt.
- Falls $g$ als eine Linearkombination von Ausdrücken wie oben dargestellt werden kann, dann können wir obiges Verfahren getrennt anwenden und die jeweiligen partikulären Lösungen der vereinfachten Differentialgleichungen addieren.
- Falls $\alpha \in \mathbb{C} \setminus \mathbb{R}$ ist und wir an reellwertigen Lösungen einer Differentialgleichung mit reellen Koeffizienten interessiert sind, dann kann man $\mathrm{e}^{\alpha x}, \mathrm{e}^{\overline{\alpha} x}$ in obigen Diskussionen durch $\mathrm{e}^{\beta x}\cos(\gamma x)$, $\mathrm{e}^{\beta x}\sin(\gamma x)$ für $\beta, \gamma \in \mathbb{R}$ mit $\alpha = \beta + \gamma\mathrm{i}$ ersetzen.

**Beispiel 7.68** (Partikuläre Lösungen)**.** *Wir betrachten einige Spezialfälle für die obigen Ansätze.*

*(a) Zum Beispiel könnten das charakteristische Polynom durch $p(T) = (T+1)(T+2)$ und das Störglied durch $x \in \mathbb{R} \mapsto g(x) = x\exp(x)$ gegeben sein. In diesem Fall ist der Skalar $\alpha = 1$ in der Exponentialfunktion des Störglieds keine Nullstelle des charakteristischen Polynoms und wir können den Ansatz $y_{\text{part}} = (C_1 x + C_2)\exp(x)$ für noch zu bestimmende Konstanten $C_1, C_2$ verwenden.*

*(b) Für $p(T) = (T+1)(T+2)$ und Störglied $x \in \mathbb{R} \mapsto g(x) = x\exp(-x)$ muss man allerdings den Ansatz $y_{\text{part}} = (C_1 x^2 + C_2 x)\exp(-x)$ für noch zu bestimmende Konstanten $C_1, C_2$ verwenden, da $-1$ sehr wohl eine Nullstelle des charakteristischen Polynoms ist.*

*(c) Falls $p(T) = (T+1)^2$ und das Störglied durch $x \in \mathbb{R} \mapsto g(x) = x\exp(-x)$ gegeben ist, so muss man den Ansatz $y_{\text{part}} = (C_1 x^3 + C_2 x^2)\exp(-x)$ für noch zu bestimmende Konstanten $C_1, C_2$ verwenden, da $-1$ eine doppelte Nullstelle des charakteristischen Polynoms ist.*

*(d) Falls $p(T) = T^2 + 1$ und das Störglied durch $x \in \mathbb{R} \mapsto g(x) = x\sin(x)$ gegeben ist, so muss man den Ansatz $y_{\text{part}} = (C_1 x^2 + C_2 x)\sin(x) + (C_3 x^2 + C_4 x)\cos(x)$ für noch zu bestimmende Konstanten $C_1, C_2, C_3, C_4$ verwenden. Dies ergibt sich ebenso aus obigen Fällen,*

*denn* $\sin(x) = \frac{1}{2\mathrm{i}}(\exp(\mathrm{i}x) - \exp(-\mathrm{i}x))$ *und beide Skalaren* $\mathrm{i}, -\mathrm{i}$ *sind einfache Nullstellen des charakteristischen Polynoms. Deswegen müssen wir den Grad der Polynomfaktoren vor* $\exp(\mathrm{i}x)$ *und vor* $\exp(-\mathrm{i}x)$ *jeweils um* 1 *erhöhen, was zu den getrennten Ansätzen* $(D_1x^2 + D_2x)\exp(\mathrm{i}x)$ *für die partikuläre Lösung für das Störglied* $\frac{1}{2\mathrm{i}}x\exp(\mathrm{i}x)$ *beziehungsweise* $(D_3x^2 + D_4x)\exp(-\mathrm{i}x)$ *für das Störglied* $-\frac{1}{2\mathrm{i}}x\exp(-\mathrm{i}x)$ *führt. Wollen wir die Rechnung über die reellen Zahlen ausführen, so betrachten wir stattdessen obigen Ansatz mit Sinus und Kosinus.*

Wie bereits erwähnt ergibt sich die allgemeine Lösung einer inhomogenen Differentialgleichung dann als die Summe einer partikulären Lösung und der allgemeinen Lösung der zugehörigen homogenen Differentialgleichung. Falls zusätzlich Anfangsbedingungen in der Form $y(x_0) = y_0$ und $y'(x_0) = y_1$ bekannt sind, so lässt sich nun die Lösung des Anfangswertproblems mit Hilfe eines Gleichungssystems in den unbekannten Konstanten lösen. Wir erproben dieses Verfahren im folgenden Beispiel.

**Beispiel 7.69** (Gedämpfte Schwingung mit periodischer Krafteinwirkung)**.** *Wir verändern die homogene Differentialgleichung*

$$\ddot{y} + d\dot{y} + ky = 0$$

*der gedämpften Schwingung aus Beispiel 7.67, indem wir eine periodische Krafteinwirkung der Form* $a\sin(t)$ *für ein* $a \in \mathbb{R}$ *dem System hinzufügen, wodurch wir die inhomogene Differentialgleichung* $\ddot{y}+d\dot{y}+ky = a\sin(t)$ *erhalten. Wir wollen nun eine Lösung des Anfangswertproblems*

$$\begin{aligned} \ddot{y} + d\dot{y} + ky &= a\sin(t) \\ y(0) = 0,\ y'(0) &= 0 \end{aligned}$$

*finden, wobei die Anfangsbedingungen der unbewegten Ruhelage am Anfang entspricht.*

*Die allgemeine Lösung der homogenen Differentialgleichung haben wir bereits in Beispiel 7.67 diskutiert. Wir möchten hier jedoch nur der Fall* $\frac{d^2}{4} - k < 0$ *besprechen. In diesem Fall ist die allgemeine Lösung*

$$y_{\mathrm{hom}}(t) = \mathrm{e}^{-\frac{d}{2}t}\big(A\sin(\gamma t) + B\cos(\gamma t)\big),$$

*wobei* $\gamma = \sqrt{\left|\frac{d^2}{4} - k\right|}$.

*Für die partikuläre Lösung machen wir den Ansatz* $y_{\mathrm{part}}(t) = -C\sin(t) + D\cos(t)$ *und erhalten*

$$\begin{aligned} \dot{y}_{\mathrm{part}} &= -C\cos(t) - D\sin(t) \\ \ddot{y}_{\mathrm{part}} &= C\sin(t) - D\cos(t). \end{aligned}$$

*Wir möchten also* $C, D$ *so bestimmen, dass*

$$\ddot{y}_{\mathrm{part}} + d\dot{y}_{\mathrm{part}} + ky_{\mathrm{part}} = (C - Dd - Ck)\sin(t) + (-D - Cd + Dk)\cos(t) = a\sin(t),$$

*oder nach Vergleich der Koeffizienten (wieso geht das?), dass*

$$\begin{pmatrix} 1-k & -d \\ d & 1-k \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} a \\ 0 \end{pmatrix}$$

*in Matrixschreibweise. Falls die Determinante* $(1-k)^2 + d^2$ *nicht Null ist, sehen wir nach Inversion, dass die Lösung des Gleichungssystems durch*

$$\begin{pmatrix} C \\ D \end{pmatrix} = \frac{1}{(1-k)^2 + d^2} \begin{pmatrix} 1-k & d \\ -d & 1-k \end{pmatrix} \begin{pmatrix} a \\ 0 \end{pmatrix} = \frac{a}{(1-k)^2 + d^2} \begin{pmatrix} 1-k \\ -d \end{pmatrix}$$

*gegeben ist, wodurch*

$$y_{\text{part}}(t) = -\frac{a(1-k)}{(1-k)^2 + d^2} \sin(t) - \frac{ad}{(1-k)^2 + d^2} \cos(t).$$

*Wir addieren zu* $y_{\text{part}}$ *die allgemeine Lösung* $y_{\text{hom}}$ *der homogenen Differentialgleichung, um eine Lösung* $y$ *zu erhalten. Nun setzen wir* $t = 0$ *sowohl in der Lösung* $y$ *als auch ihrer Ableitung und erhalten gemeinsam mit den Anfangsbedingungen*

$$\begin{aligned} y(0) &= -\frac{ad}{(1-k)^2 + d^2} \cdot 1 + 1 \cdot B \cdot 1 = 0 \\ y'(0) &= -\frac{a(1-k)}{(1-k)^2 + d^2} \cdot 1 - \frac{d}{2} \cdot 1 \cdot B + 1 \cdot A \cdot \gamma = 0. \end{aligned}$$

*Somit wählen wir die Konstanten* $A, B$ *als*

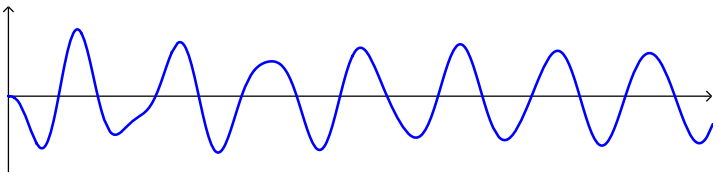
$$\begin{aligned} B &= \frac{ad}{(1-k)^2 + d^2}, \\ A &= \frac{Bd}{2\gamma} + \frac{a(1-k)}{\gamma((1-k)^2 + d^2)} = \frac{a}{\gamma((1-k)^2 + d^2)} \left( \frac{d^2}{2} + (1-k) \right). \end{aligned}$$

*Die Lösung des Anfangswertproblems ist somit gegeben durch*

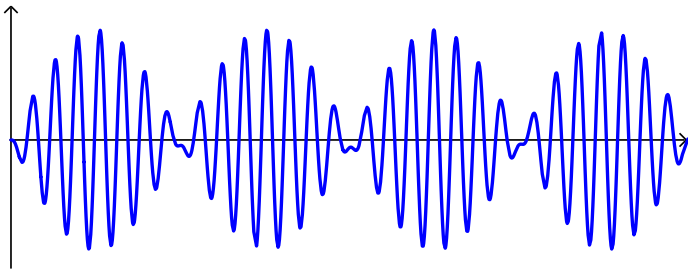
$$\begin{aligned} y(t) &= y_{\text{part}}(t) + y_{\text{hom}}(t) \\ &= -\frac{a(1-k)}{(1-k)^2 + d^2} \sin(t) - \frac{ad}{(1-k)^2 + d^2} \cos(t) \\ &\quad + \mathrm{e}^{-\frac{d}{2}t} \left( \frac{a}{\gamma((1-k)^2 + d^2)} \left( \frac{d^2}{2} + (1-k) \right) \sin(\gamma t) + \frac{ad}{(1-k)^2 + d^2} \cos(\gamma t) \right). \end{aligned}$$

*Man kann in diesem einfachen Modell bereits einige physikalische Beobachtungen wiedererkennen.*

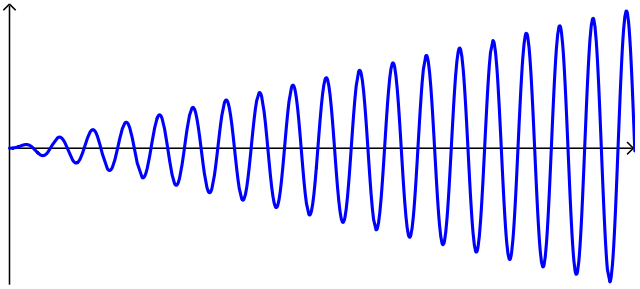
- *Bei positiver Dämpfung* $d > 0$ *strebt die Lösung gegen eine Schwingung mit Frequenz der vorgegebenen Krafteinwirkung. Die Amplitude dieser Schwingung ist durch* $\frac{a}{\sqrt{(1-k)^2+d^2}}$ *gegeben (siehe die entsprechende Übung in Abschnitt 6.9.2) und hängt also linear von der vorgegeben Amplitude der Krafteinwirkung ab. Sie wird grösser, wenn* $k$ *nahe bei* $1$ *und* $d$ *nahe bei* $0$ *ist.*



- *Bei verschwindender Dämpfung $d = 0$ ergibt sich eine Lösung, die eine Überlagerung der „Eigenschwingung“ des Systems und der „vorgegebenen Schwingung“ der Krafteinwirkung ist. Falls $k \neq 1$ nahe bei $1$ ist, dann können sich diese Schwingungen manchmal gleichzeitig verstärken und manchmal fast auslöschen, siehe folgendes Bild.*



- *Bei $d = 0$ und $k = 1$ gilt obige Diskussion nicht mehr. In diesem Fall ist eine partikuläre Lösung durch $y_{\text{part}}(t) = -\frac{a}{2}t\cos(t)$ gegeben (wieso?) und eine Lösung des Anfangswertproblems durch $y(t) = -\frac{a}{2}t\cos(t) + \frac{a}{2}\sin(t)$. Diese Lösung ist also unbeschränkt (unabhängig davon wie klein $a \neq 0$ auch sein mag). Physikalisch begründet sich dies mit der Tatsache, dass die Krafteinwirkung mit dem System im gleichen Takt ist und die Schwingung (bei Abwesenheit von Reibung) aufschaukelt.*



*Die Phänomene in diesem Beispiel werden als Resonanz bezeichnet und können in verschiedenen Situationen beobachtet werden. Zum Beispiel kann sich die Schwingung von einer Saite einer Gitarre auf die gleiche Saite einer anderen Gitarre übertragen, ohne dass diese sich berühren müssen. Denn durch Anzupfen der Saite auf der ersten Gitarre erzeugt man einen Ton, der als Schwingung der Luft in einer gewissen Frequenz definiert werden kann. Diese Schwingung breitet sich aus und übt auf die gleiche Saite der zweiten Gitarre in der richtigen Frequenz eine Kraft aus, die zwar sehr klein ist, aber in der Frequenz mit der Eigenschwingung der Saite der zweiten Gitarre übereinstimmt. Aus diesen Grund beginnt die Saite der zweiten Gitarre sichtbar zu schwingen. (Zur Not kann man dies auch nur mit einer Gitarre beobachten, aber hierzu muss man auf zwei Saiten die gleiche Note greifen.)*

**Applet 7.70** (Harmonischer Oszillator)**.** *Sie können sowohl die Koeffizienten der Differentialgleichung* $\ddot{y} + d\dot{y} + ky = a \sin t$ *als auch die Anfangsbedingungen* $y(0) = y_0$ *und* $\dot{y}(0) = y_1$ *einstellen.*

VO: 13.12.17

## 7.6 Weitere Lernmaterialien

### 7.6.1 Verwendung des Kapitels

Der Begriff der Ableitung und die Ableitungsregel sind fundamentales Grundwissen für alle weitere Untersuchungen in der Analysis I/II-Vorlesung und vielen Anwendungen in Physik und anderen Wissenschaften. Die Berechnung der Ableitung einer vorgegebenen Funktion ist meist ziemlich einfach, da die wenigen Regeln für alle üblichen algebraischen Verknüpfungen von bekannten Funktionen anwendbar sind. Bloss wenn eine Funktion durch eine Fallunterscheidung definiert wird, müssen wir mitunter auf die ursprüngliche Definition zurückgreifen und können damit vielleicht die Ableitung bei den Problempunkten berechnen. Wir empfehlen Ihnen dies zu üben bis Sie Ableitungen mit 100%-iger Sicherheit berechnen können.

Wie wir gesehen haben, ist der Mittelwertsatz der Differentialrechnung *das Hilfsmittel* für den Zusammenhang zwischen der Ableitung und dem Verhalten der ursprünglichen Funktion auf einem Intervall. Für den Beweis des Mittelwertsatzes benötigten wir (über den Umweg des Satzes von Rolle) den Satz über die Existenz des Maximums auf einem kompakten Intervall, wo wiederum das Supremum und damit die Vollständigkeit der reellen Zahlen notwendig waren. Insbesondere bildet der Mittelwertsatz die Grundlage von allen Kurvendiskussionen, welche Monotonieeigenschaften aber auch Krümmungseigenschaften in der Form von Konvexität und Konkavität einer vorgegebenen Funktion beschreiben. Obwohl viele Ableitungsregeln und auch viele weitere Eigenschaften, die eine Funktion mit ihrer Ableitung verknüpft, auch für komplex-wertige Funktionen gelten, gilt der Mittelwertsatz nicht für komplex-wertige Funktionen.

Die Kurvenbesprechung verwendeten wir zum Beispiel für den Sinus, den Kosinus und den Tangens, was zu der Definition der inversen Funktionen Arkussinus, Arkusconsinus und Arkustangens führte. Wir werden diese Funktionen, deren Eigenschaften, Definitionsbereiche und Ableitungen ab nun auch ohne Verweise auf Abschnitt 7.3 verwenden. Wir haben damit die üblichen Funktionen eingeführt und kennen auch die Ableitungen von all diesen Funktionen.[3] Wir werden aber auch noch weiteren Ihnen wahrscheinlich unbekannten, aber in gewissen Anwendungen wichtigen Funktionen begegnen.

Des Weiteren haben wir mit der Regel von de l'Hôpital *das Hilfsmittel* für die Berechnung von Grenzwerten gefunden. Dies ist hilfreich, da die Berechnung von Ableitungen im Vergleich sehr einfach ist. In gewissen Fällen wie zum Beispiel $\lim_{x\searrow 0} x\log(x)$ erfordert die Anwendung der Regel von de l'Hôpital ein wenig Geschick, da man zuerst den Ausdruck $x\log(x)$ als einen Bruch darstellen muss. Doch auch dies ist mit ein wenig Übung nicht schwierig, da man hier einfach der „schwierigeren Funktion $\log(x)$ den Vortritt gibt“ um stattdessen $\frac{\log(x)}{x^{-1}}$ zu betrachten und dies nach einmaliger Anwendung der Regel von de l'Hôpital zum Erfolg führt. Dies ist auch hilfreich zur Berechnung von Grenzwerten von Folgen, falls diese durch eine konkrete Formel $f(n)$ definiert ist und für diese Formel der Grenzwert $\lim_{x\to\infty} f(x)$ berechnet werden kann.

[3] Wir meinen mit „üblichen Funktionen“ solche, die sowohl auf den meisten Taschenrechnern vorhanden sind und auch im Gymnasium unterrichtet werden. Abgesehen davon unterscheiden sich diese Funktionen aber kaum von anderen Funktionen, die wir zum Teil noch kennenlernen werden und sich vielleicht nicht durch die üblichen Funktionen ausdrücken lassen.

Wir haben auch den Begriff der Differentialgleichung eingeführt und einige erste Anfangswertprobleme lösen können. Diese Begriffe sind ebenso von fundamentaler Bedeutung für die Physik und viele weitere Wissenschaften. Denn immer wenn eine Grösse $y$ von einem Zeitparamter $t$ abhängt und gewisse Gesetzmässigkeiten für das Änderungsverhalten $\dot{y}(t)$ in Abhängigkeit von $y(t)$ und dem Zeitparameter $t$ bekannt sind, ergibt sich daraus eine gewöhnliche Differentialgleichung. Wir werden im zweiten Semester nochmals zu diesem Thema zurückkehren, die Frage der eindeutigen Lösbarkeit von Anfangswertproblemen aufwerfen und in grosser Allgemeinheit positiv beantworten können.

### 7.6.2 Übungen

**Übung.** *Finden Sie eine differenzierbare Funktion auf einem offenen Intervall $I$, so dass ein Punkt $x_0 \in I$ mit $f'(x_0) = 0$ existiert, der kein lokales Extremum ist.*

**Übung** (Vielfachheit von Nullstellen)**.** *Sei $f \in \mathbb{R}[T]$ ein Polynom und sei $k \in \mathbb{N}$. Wir erinnern daran, dass eine $k$-fache Nullstelle von $f$ ein $z \in \mathbb{C}$ ist mit der Eigenschaft, dass $(T-z)^k$ das Polynom $f$ teilt, aber $(T-z)^{k+1}$ das Polynom $f$ nicht teilt. Zeigen Sie, dass folgende Eigenschaften für einen Punkt $x_0 \in \mathbb{R}$ äquivalent sind:*

*(i) $x_0$ ist eine $k$-fache Nullstelle von $f$.*

*(ii) $f^{(\ell)}(x_0) = 0$ für alle $\ell \in \mathbb{N}_0$ mit $\ell \leq k-1$ und $f^{(k)}(x_0) \neq 0$.*

*Daraus folgt beispielsweise, dass eine $k$-fache Nullstelle von $f$ eine $k-1$-fache Nullstelle von $f'$ ist (wieso?). Die zweite Eigenschaft (ii) lässt sich direkt auf beliebige, glatte Funktionen $f : \mathbb{R} \to \mathbb{R}$ erweitern. Mit etwas mehr stimmt dies auch für (i) und man erhält wiederum eine äquivalente Charakterisierung.*

*Allerdings muss nicht jede Nullstelle einer von Null verschiedenen, glatten Funktion $f : \mathbb{R} \to \mathbb{R}$ eine endliche Vielfachheit besitzen. Finden Sie ein nicht-triviales Beispiel einer solchen Funktion. Das heisst, finden Sie eine glatte Funktion $f : \mathbb{R} \to \mathbb{R}$ mit einer Nullstelle $x_0 \in \mathbb{R}$, so dass $f^{(k)}(x_0) = 0$ für alle $k \in \mathbb{N}_0$, aber $f|_U \neq 0$ für alle Umgebungen $U$ von $x_0$.*

**Übung** (Satz von Darboux)**.** *Sei $[a,b]$ ein kompaktes Intervall mit Endpunkten $a < b$ und sei $f : [a,b] \to \mathbb{R}$ eine differenzierbare Funktion. Dann nimmt $f'$ zwischen $a$ und $b$ alle Werte zwischen $f'(a)$ und $f'(b)$ an.*

*Gehen Sie wie folgt vor, um einen Beweis zu erhalten.*

*(i) Sei $c$ zwischen $f'(a)$ und $f'(b)$. Betrachten Sie die Funktion $x \in [a,b] \mapsto f(x) - cx \in \mathbb{R}$ um zu argumentieren, dass man ohne Beschränkung der Allgemeinheit $f'(a) \leq 0$, $f'(b) \geq 0$ sowie $c = 0$ annehmen kann.*

*(ii) Verwenden Sie nun den Extremwertsatz (Korollar 3.72) und Proposition 7.17 um zu zeigen, dass ein Punkt $x \in [a,b]$ mit $f'(x) = 0$ existiert.*

**Übung.** *Zeigen Sie, dass die Funktion*

$$f : x \in \mathbb{R} \mapsto \begin{cases} \exp\left(-\frac{1}{1-x^2}\right) & \text{falls } |x| < 1 \\ 0 & \text{falls } |x| \geq 1 \end{cases}$$

*glatt ist.*

**Übung.** *Sei $f : D \to \mathbb{C}$ eine Funktion auf einer Teilmenge $D \subseteq \mathbb{R}$ und sei $a \in D$ Häufungspunkt. Wir erinnern daran, dass $f$ bei $a$ differenzierbar ist, falls der Grenzwert $\lim_{x\to a} \frac{f(x)-f(a)}{x-a}$ existiert. Zeigen Sie, dass $f$ genau dann bei $a$ differenzierbar ist, wenn $\mathrm{Re}(f)$ und $\mathrm{Im}(f)$ bei $a$ differenzierbar sind.*

**Übung** (Challenge)**.** *In dieser Übung möchten wir eine differenzierbare Funktion $f$ auf $[0,1]$ konstruieren, deren Ableitung an allen (und insbesondere überabzählbar vielen) Punkten in der Cantor-Menge nicht stetig ist. Als Ausgangspunkt betrachten wir dazu die Funktion*

$$\varphi : [0,1] \to \mathbb{R}, \quad x \mapsto \begin{cases} x^2(1-x)^2 \sin\left(\frac{1}{x(1-x)}\right) & \text{falls } x \in (0,1) \\ 0 & \text{falls } x \in \{0,1\} \end{cases}.$$

*Zeigen Sie, dass $\varphi$ differenzierbar ist, aber dass $\varphi'$ bei $0$ und $1$ nicht stetig ist.*

*Sei $C$ die Cantor-Menge. Wir verwenden die Notation $x_+, x_-$ zu $x \in [0,1] \setminus C$ aus Abschnitt 6.8.2 und definieren damit $f : [0,1] \to \mathbb{R}$ durch $f(x) = 0$ für $x \in C$ und*

$$f(x) = (x_+ - x_-)^{\frac{3}{2}} \, \varphi\left(\frac{x - x_-}{x_+ - x_-}\right)$$

*für $x \in [0,1] \setminus C$. Zeigen Sie, dass $f$ differenzierbar ist und dass $f'$ bei keinem Punkt in $C$ stetig ist.*

In der folgenden Übung möchten wir eine zahlentheoretische Anwendung des Mittelwertsatzes präsentieren. Wir erinnnern uns daran, dass in Abschnitt 3.2 der Begriff der algebraischen und transzendenten Zahlen eingeführt wurde. Dabei heisst eine komplexe Zahl algebraisch, wenn sie die Nullstelle eines von Null verschiedenen Polynoms mit rationalen Koeffizienten ist, und transzendent sonst. Wir möchten in folgender Übung zeigen, dass die Zahl

$$\sum_{n=0}^{\infty} 10^{-n!}$$

oder Zahlen mit ähnlichen Eigenschaften transzendent sind.

**Übung.** *Eine irrationale reelle Zahl $\alpha \in \mathbb{R}$ ist eine Liouville-Zahl, falls für jedes $n \in \mathbb{N}$ eine rationale Zahl $\frac{p}{q} \in \mathbb{Q}$ existiert mit*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^n}.$$

*In einem gewissen Sinne ist eine Liouville-Zahl also eine irrationale Zahl, die sich sehr gut durch rationale Zahlen approximieren lässt.*

*(i) Zeigen Sie, dass $\sum_{n=0}^{\infty} 10^{-n!}$ eine Liouville-Zahl ist. Verwenden Sie die Ziffernentwicklung zur Basis $10$ und Übung 6.91.*

*Wir wollen nun zeigen, dass jede Liouville-Zahl (und damit auch $\sum_{n=0}^{\infty} 10^{-n!}$) transzendent ist. Dazu behaupten wir, dass es für jede algebraische Zahl $\beta$ eine Konstante $A > 0$ und ein*

*$n \in \mathbb{N}$ gibt, so dass*

$$\left|\beta - \frac{p}{q}\right| \geq \frac{A}{q^n} \tag{7.19}$$

*für alle $\frac{p}{q} \in \mathbb{Q}$.*

*(ii) Sei $f \in \mathbb{Q}[x]$ mit $f(\beta) = 0$ und Grad $n \geq 1$. Sei weiters $\frac{p}{q} \in \mathbb{Q} \cap [\beta - 1, \beta + 1]$ keine Nullstelle von $f$. Zeigen Sie, dass $\left|f\left(\frac{p}{q}\right)\right| \geq \frac{1}{q^n}$.*

*(iii) Argumentieren Sie mit dem Mittelwertsatz, dass $\left|f\left(\frac{p}{q}\right)\right| = |f'(\xi)|\left|\beta - \frac{p}{q}\right|$ für ein $\xi$ zwischen $\beta$ und $\frac{p}{q}$. Schliessen Sie, dass es ein $a > 0$ gibt mit $|f'(x)| \leq a$ für alle $x \in [\beta - 1, \beta + 1]$ und folgern Sie (7.19).*

*(iv) Zeigen Sie, dass jede Liouville-Zahl transzendent ist.*

**Übung** (Konvexität und Mittelpunktseigenschaft)**.** *Sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ eine stetige Funktion. Zeigen Sie, dass $f$ genau dann konvex ist, wenn für alle $x, y \in I$ die Ungleichung*

$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x) + f(y)}{2}$$

*gilt. Begründen Sie auch intuitiv, wieso letztere Ungleichung implizit in (7.6) enthalten ist und somit (zumindest auf den ersten Blick) die schwächere Annahme ist. Sie können dazu mit einem Bild arbeiten. Zeigen Sie unter Verwendung obiger Charakterisierung, dass $x \in \mathbb{R} \mapsto |x|$ eine konvexe Funktion ist.*

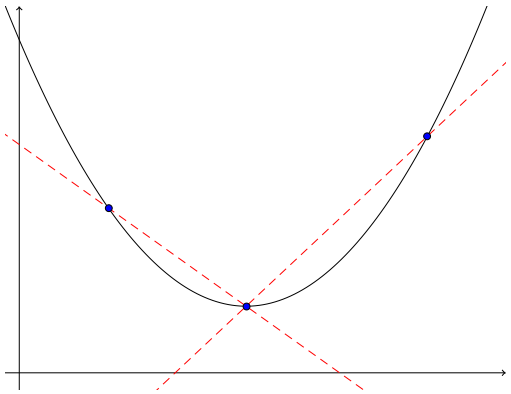
**Übung.** *In dieser Übung möchten wir zeigen, dass konvexe Funktionen fast überall differenzierbar sind. Sei $I \subseteq \mathbb{R}$ ein nicht-leeres Intervall und sei $f : I \to \mathbb{R}$ konvex.*

*(i) Zeigen Sie, dass die links- und rechtsseitigen Ableitungen $f'_-(x) \leq f'_+(x)$ von $f$ bei jedem $x \in I$ ausser vielleicht bei den Endpunkten von $I$ existieren.*

*(ii) Zeigen Sie, dass $f$, möglicherweise abgesehen von den Endpunkten von $I$, stetig ist.*

*(iii) Zeigen Sie, dass es eine höchstens abzählbare Ausnahmemenge $A \subseteq I$ gibt, so dass $f$ bei jedem $x \in I \setminus A$ differenzierbar ist.*

*Hinweis: Für (i) können Sie Lemma 7.39 verwenden, um eine Monotonie der Differenzenquotienten zu zeigen. Für (ii) können Sie folgendes Bild als Hinweis verwenden.*

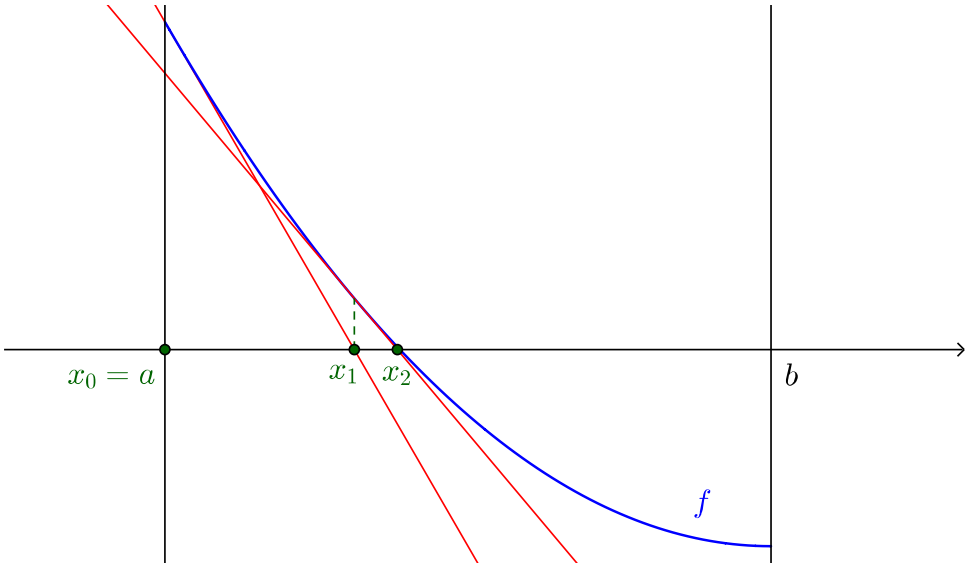


*Für (iii) verwenden Sie Übung 3.58 für $f'_-, f'_+$.*

**Übung** (Newton-Verfahren für konvexe Funktionen)**.** *Seien $a < b$ zwei reelle Zahlen und sei $f : [a, b] \to \mathbb{R}$ eine differenzierbare, konvexe Funktion mit $f(a) > 0$ und $f(b) < 0$.*

*(i) Zeigen Sie, dass $f$ eine eindeutig bestimmte Nullstelle $z$ in $[a, b]$ besitzt.*

*Wir betrachten nun das Newton-Verfahren zur numerischen Bestimmung von Nullstellen. Sei $x_0 = a$ und $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ für $n \in \mathbb{N}_0$. Dabei ist $x_{n+1}$ die Nullstelle der Tangente an $f$ durch $x_n$ für jedes $n \in \mathbb{N}$; siehe dazu folgendes Bild.*




*Nun wollen wir zeigen, dass die Folge $(x_n)_n$ gegen $z$ konvergiert.*

*b) Zeigen Sie, dass die Folge $(x_n)_n$ Grenzwert $z$ hat, falls sie konvergiert.*

*c) Zeigen Sie, dass die Folge $(x_n)_n$ monoton wachsend ist und dass $x_n \leq z$ und somit auch $x_n \in [a, b]$ für alle $n \in \mathbb{N}$ erfüllt ist. Schliessen Sie damit auf die Konvergenz des Newton-Verfahrens mit Startpunkt $a$.*

*Wir werden später mit Hilfe zusätzlicher Werkzeuge das Newton-Verfahren in einem allgemeineren Kontext besprechen können – siehe Beispiel 8.65.*

### 7.6.3 Lernkarten

Sie können wiederum die Lernkarten oder den Graphen für Ihre Wiederholung der Themen des Kapitels verwenden.

# Kapitel 8

# Die Ableitung und das Riemann-Integral

In diesem Kapitel wollen wir die zwei wichtigen Themen „Riemann-Integral“ aus Kapitel 4 und „Ableitung“ aus Kapitel 7 verknüpfen. Wie wir sehen werden, ist dieser Zusammenhang für Anwendungen und für die weitere Theorie von fundamentaler Bedeutung.

## 8.1 Der Fundamentalsatz der Integral- und Differentialrechnung

**Definition 8.1.** Sei $[a, b]$ ein kompaktes Intervall in $\mathbb{R}$ und sei $f \in R([a, b])$ eine auf $[a, b]$ Riemann-integrierbare Funktion. Die Funktion

$$x \in [a, b] \mapsto \int_a^x f(t)\,\mathrm{d}t$$

nennt sich das **Integral mit veränderlichen oberen Grenze** oder **das partikuläre Integral** von $f$.

Man beachte, dass für $f \in R([a, b])$ und $x \in [a, b]$ das Integral $\int_a^x f(t)\,\mathrm{d}t$ wohldefiniert ist (siehe Satz 4.26 über die Intervalladditivität des Riemann-Integrals). Wir erinnern daran, dass wir in Übung 4.28 bereits die Stetigkeit des partikulären Integrals gezeigt haben. Mit etwas stärkeren Annahmen ergibt sich nun folgender Satz von fundamentaler Bedeutung.

**Theorem 8.2** (Ableitung des Integrals)**.** *Sei $[a, b]$ ein kompaktes Intervall mit Endpunkten $a < b$ und sei $f \in R([a, b])$ eine auf $[a, b]$ Riemann-integrierbare Funktion. Falls $f$ bei $x_0 \in [a, b]$ stetig ist, so ist $F : x \in [a, b] \mapsto \int_a^x f(t)\,\mathrm{d}t$ bei $x_0$ differenzierbar und $F'(x_0) = f(x_0)$.*

*Beweis.* Sei $\varepsilon > 0$. Dann existiert ein $\delta > 0$ mit

$$|f(x) - f(x_0)| < \varepsilon$$

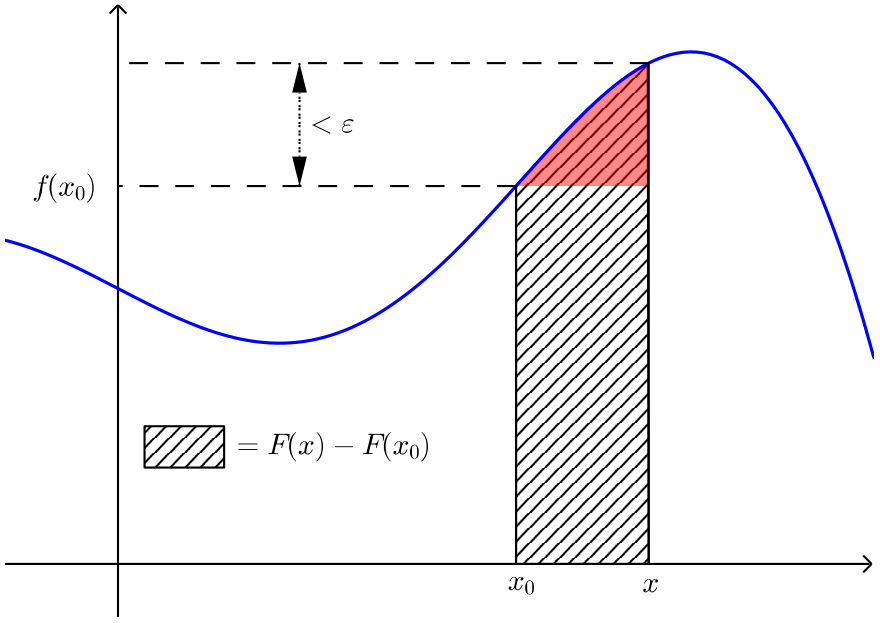
für alle $x \in (x_0 - \delta, x_0 + \delta) \cap [a, b]$. Wir verwenden dies nun in Kombination mit der Dreiecks-Ungleichung für das Riemann-Integral (Satz 4.24) und der Intervalladditivität des Riemann-Integrals (Satz 4.26), um die Aussage zu zeigen. Für $x \in (x_0, x_0 + \delta) \cap [a, b] \setminus \{x_0\}$ gilt

$$\begin{aligned}
\left|\frac{F(x) - F(x_0)}{x - x_0} - f(x_0)\right| &= \left|\frac{1}{x - x_0}\left(\int_a^x f(t)\,\mathrm{d}t - \int_a^{x_0} f(t)\,\mathrm{d}t\right) - f(x_0)\right| \\
&= \left|\frac{1}{x - x_0}\int_{x_0}^x f(t)\,\mathrm{d}t - f(x_0)\right| \\
&= \left|\frac{1}{x - x_0}\int_{x_0}^x f(t)\,\mathrm{d}t - \frac{1}{x - x_0}\int_{x_0}^x f(x_0)\,\mathrm{d}t\right| \\
&= \left|\frac{1}{x - x_0}\int_{x_0}^x (f(t) - f(x_0))\,\mathrm{d}t\right| \\
&\leq \frac{1}{x - x_0}\int_{x_0}^x |f(t) - f(x_0)|\,\mathrm{d}t \leq \frac{1}{x - x_0}\int_{x_0}^x \varepsilon\,\mathrm{d}t \;=\; \varepsilon.
\end{aligned}$$

Analog gilt für $x \in (x_0 - \delta, x_0) \cap [a, b] \setminus \{x_0\}$, dass

$$\begin{aligned}
\left|\frac{F(x) - F(x_0)}{x - x_0} - f(x_0)\right| = \left|\frac{F(x_0) - F(x)}{x_0 - x} - f(x_0)\right| &= \left|\frac{1}{x_0 - x}\int_x^{x_0} (f(t) - f(x_0))\,\mathrm{d}t\right| \\
&\leq \frac{1}{x_0 - x}\int_x^{x_0} |f(t) - f(x_0)|\,\mathrm{d}t \leq \varepsilon.
\end{aligned}$$

Da $\varepsilon > 0$ beliebig war, beweist dies $\lim_{x \to x_0} \frac{F(x) - F(x_0)}{x - x_0} = f(x_0)$ und damit den Satz. □




Figur 8.1: Illustration zum Beweis von Theorem 8.2. Der Wert $F(x) - F(x_0)$ lässt sich schreiben als $f(x_0)(x - x_0)$ plus die Fläche in Rot, die kleiner ist als $\varepsilon(x - x_0)$. Somit ist $\frac{F(x) - F(x_0)}{x - x_0}$, bis auf einen Fehler kleiner als $\varepsilon$, durch $f(x_0)$ gegeben.

**Korollar 8.3** (Ableitung des Integrals)**.** *Sei $[a, b]$ ein kompaktes Intervall mit Endpunkten $a < b$ und sei $f \in C([a, b])$ stetig. Dann ist $x \in [a, b] \mapsto \int_a^x f(t)\,\mathrm{d}t$ eine Stammfunktion von $f$*

*und jede Stammfunktion $F : [a, b] \to \mathbb{R}$ von $f$ hat die Form*

$$F(x) = \int_a^x f(t)\,\mathrm{d}t + C \tag{8.1}$$

*für alle $x \in [a, b]$ und eine Konstante $C \in \mathbb{R}$.*

*Beweis.* Nach Satz 4.42 ist $f$ Riemann-integrierbar. Nach Theorem 8.2 ist

$$x \in [a, b] \mapsto \int_a^x f(t)\,\mathrm{d}t$$

differenzierbar und eine Stammfunktion von $f$. Nach Lemma 7.59 unterscheidet sich jede weitere Stammfunktion nur um eine Konstante von dieser, was die Formel im Korollar beweist. □

**Korollar 8.4** (Berechnung des Integrals)**.** *Sei $[a, b]$ ein kompaktes Intervall mit Endpunkten $a < b$ und sei $f \in C([a, b])$ stetig. Falls $F : [a, b] \to \mathbb{R}$ eine Stammfunktion von $f$ ist, dann gilt*

$$\int_a^b f(t)\,\mathrm{d}t = F(b) - F(a).$$

Wir werden auch öfter die Abkürzung $\big[F(x)\big]_a^b = \big[F(x)\big]_{x=a}^{x=b} = F(b) - F(a)$ verwenden.

*Beweis.* Nach Korollar 8.3 gilt (8.1). Setzen wir $x = a$ erhalten wir für die Konstante $C = F(a)$ und das Korollar folgt nun indem wir $x = b$ in (8.1) einsetzen. □

**Korollar 8.5** (Integral der Ableitung)**.** *Sei $[a, b]$ ein kompaktes Intervall mit Endpunkten $a < b$ und sei $F \in C^1([a, b])$ stetig differenzierbar. Dann gilt*

$$F(x) = F(a) + \int_a^x F'(t)\,\mathrm{d}t$$

*für alle $x \in [a, b]$.*

Insbesondere kann der Wert von $F$ an jeder Stelle $x$ vollständig durch den Wert von $F$ bei der Stelle $a$ und die Werte der Funktion $F'$ beschrieben werden.

*Beweis.* Die Funktion $F_0 : x \in [a, b] \mapsto \int_a^x F'(t)\,\mathrm{d}t$ ist nach Korollar 8.3 ebenso wie $F$ eine Stammfunktion. Sei $C \in \mathbb{R}$ mit $F(x) = F_0(x) + C$ für alle $x \in [a, b]$ nach Korollar 8.3. Setzen wir $x = a$ ein, so erhalten wir

$$F(a) = F_0(a) + C = \int_a^a F'(t)\,\mathrm{d}t + C = C.$$

Somit ist

$$F(x) = F(a) + F_0(x) = F(a) + \int_a^x F'(t)\,\mathrm{d}t$$

für alle $x \in [a, b]$, wie zu zeigen war. □

Die obigen Resultate sind auch als der **Fundamentalsatz der Integral- und Differentialrechnung** bekannt und gehen auf die Arbeiten von Leibniz, Newton und Barrow zurück, die weitgehend die Ausgangspunkte der Analysis darstellen. Zur Vereinfachung der Diskussion haben wir Stetigkeit von $f$ beziehungsweise $F'$ angenommen. Diese Annahme lässt sich auf verschiedene Arten abschwächen.

**Übung 8.6** (Theorem 8.2 für „fast überall“ stetige Funktionen)**.** *Sei $[a,b]$ ein kompaktes Intervall mit Endpunkten $a<b$ und sei $f \in R([a,b])$ Riemann-integrierbar und bei höchstens endlich vielen Punkten unstetig. Zeigen Sie, dass die verallgemeinerte Stammfunktion $F$ mit $F(x) = \int_a^x f(t)\,\mathrm{d}t$ für $x \in [a,b]$ stetig ist und bei allen Punkten $x \in [a,b]$ bis auf endlich viele Ausnahmen differenzierbar ist mit $F'(x) = f(x)$.*

**Wichtige Übung 8.7** (Fundamentalsatz für komplexe Funktionen)**.** *In obigem Satz und dessen Korollaren haben wir für die Aussagen eigentlich nicht wirklich verwendet, dass die betrachteten Funktionen reellwertig sind. Zeigen Sie deswegen, dass alle obigen Resultate für komplexwertige Funktionen zutreffen.*

Zusammenfassend wollen wir noch betonen, dass auf Grund obiger Resultate die Berechnung von Riemann-Integralen und damit insbesondere von Flächeninhalten, Schwerpunkkoordinaten, Arbeitsberechnungen, Bogenlängen und Volumen von Rotationskörpern (siehe Abschnitt 8.3) auf die Berechnung von Stammfunktionen zurückgeführt werden kann.

**Übung 8.8** (Mittelwertsatz der Integralrechnung)**.** *Sei $[a,b]$ ein kompaktes Intervall mit Endpunkten $a<b$ und $f:[a,b]\to\mathbb{R}$ stetig. Zeigen Sie, dass ein $\xi \in (a,b)$ existiert mit*

$$\int_a^b f(x)\,\mathrm{d}x = f(\xi)(b-a).$$

*Können Sie zwei verschiedene Beweise finden?*

**Übung 8.9** (Riemann-integrierbare Ableitung)**.** *Wir möchten hier eine etwas stärkere Version von Korollar 8.5 thematisieren. Sei $[a,b]$ ein kompaktes Intervall mit Endpunkten $a<b$ und $F:[a,b]\to\mathbb{R}$ eine differenzierbare, reellwertige Funktion mit $F' \in R([a,b])$. Zeigen Sie, dass $F(x) = F(a) + \int_a^x F'(t)\,\mathrm{d}t$ für alle $x \in [a,b]$.*

**Applet 8.10** (Fundamentalsatz)**.** *Wir sehen hier nochmals die Idee des Beweises des Fundamentalsatzes der Integral- und Differentialrechnung (Theorem 8.2), wobei unten das partikuläre Integral der Funktion im oberen Fenster dargestellt wird.*

### 8.1.1 Differentiation von Potenzreihen

Wir wenden nun den Fundamentalsatz an, um zu zeigen, dass sich Potenzreihen nicht nur integrieren (siehe Satz 6.85), sondern auch differenzieren lassen.

**Korollar 8.11** (Differentiation von Potenzreihen)**.** *Sei $f(x) = \sum_{n=0}^{\infty} c_n x^n$ eine Potenzreihe mit Konvergenzradius $R>0$. Dann ist $f:(-R,R)\to\mathbb{R}$ differenzierbar und es gilt*

$$f'(x) = \sum_{n=1}^{\infty} n c_n x^{n-1}$$

*für alle $x \in (-R, R)$, wobei die Potenzreihe rechts ebenfalls Konvergenzradius $R$ hat.*

*Beweis.* Sei $g(x)$ die Potenzreihe $\sum_{n=1}^{\infty} nc_n x^{n-1} = \sum_{k=0}^{\infty}(k+1)c_{k+1}x^k$ und sei $S$ der zugehörige Konvergenzradius. Nach Satz 6.56 ist $g$ auf $(-S, S)$ stetig und nach Satz 6.85 darf $g(x)$ gliedweise integriert werden. Genauer gesagt besagt Satz 6.85, dass $G : (-S, S) \to \mathbb{R}$ gegeben durch

$$G(x) = \int_0^x g(t)\,\mathrm{d}t = \sum_{k=0}^{\infty} \frac{(k+1)c_{k+1}}{k+1} x^{k+1} = \sum_{k=0}^{\infty} c_{k+1} x^{k+1}$$

wieder eine Potenzreihe mit Konvergenzradius $S$ darstellt. Nach Korollar 8.3 gilt weiters, dass $G$ eine Stammfunktion von $g = G'$ darstellt. Da dies aber abgesehen vom ersten Glied der Reihe genau die Potenzreihe $f$ ist, stimmen die Konvergenzradien $R = S$ überein, $f$ ist differenzierbar und $f'(x) = G'(x) = g(x) = \sum_{n=1}^{\infty} nc_n x^{n-1}$ für alle $x \in (-R, R)$. □

Korollar 8.11 lässt sich leicht anpassen, um zu zeigen, dass durch Potenzreihen gegebene Funktionen glatt sind.

**Übung 8.12** (Verallgemeinerung von Korollar 8.11)**.** *Sei $f(x) = \sum_{n=0}^{\infty} c_n x^n$ eine Potenzreihe mit Konvergenzradius $R > 0$. Zeigen Sie, dass $f : (-R, R) \to \mathbb{R}$ glatt ist und finden Sie eine Darstellung von $f^{(n)}$ durch eine Potenzreihe für jedes $n \in \mathbb{N}$.*

**Übung 8.13** (Koeffizientenvergleich für Potenzreihen)**.** *Seien $f(x) = \sum_{n=0}^{\infty} c_n x^n$ und $g(x) = \sum_{n=0}^{\infty} d_n x^n$ Potenzreihen mit reellen Koeffizienten und positiven Konvergenzradien $R_f, R_g$. Sei $R = \min\{R_f, R_g\}$ und angenommen $f(x) = g(x)$ für alle $x \in (-R, R)$. Zeigen Sie, dass dann $c_n = d_n$ für alle $n \in \mathbb{N}_0$ und damit $R_f = R_g$ gilt.*

**Übung 8.14** (Potenzreihenentwicklung für Wurzeln)**.** *Sei $\alpha \in \mathbb{C}$. Wir wollen hier zeigen, dass*

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n$$

*für alle $x \in (-1, 1)$, wobei die verallgemeinerten Binomialkoeffizienten für $n \in \mathbb{N}_0$ durch*

$$\binom{\alpha}{n} = \frac{\prod_{j=0}^{n-1}(\alpha - j)}{n!} = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}$$

*definiert sind.*

*(a) Zeigen Sie, dass für $\alpha \notin \mathbb{N}_0$ die Potenzreihe*

$$g(x) = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n$$

*Konvergenzradius 1 hat.*

*(b) Berechnen Sie die Ableitung von $g$ und zeigen Sie, dass $f(x) = (1+x)^\alpha$ und $g(x)$ die Differentialgleichung*

$$y' = \alpha \frac{y}{1+x}$$

*erfüllen.*

*(c) Berechnen Sie die Ableitung von $\frac{g(x)}{f(x)}$ und schliessen Sie die Behauptung.*

### 8.1.2 Die alternierende harmonische Reihe

Wir haben bereits in Beispiel 6.20 gesehen, dass die alternierende harmonische Reihe $\sum_{n=1}^\infty \frac{(-1)^{n+1}}{n}$ konvergiert, was auch aus dem Leibniz-Kriterium (Proposition 6.25) folgt. Mit den Resultaten von Kapitel 6 konnten wir den Wert der Reihe aber nicht bestimmen. Nun können wir mit Hilfe des Fundamentalsatzes der Integral- und Differentialrechnung die Identität

$$\sum_{n=1}^\infty \frac{(-1)^{n+1}}{n} = \log(2)$$

beweisen.

Es ist in diesem Fall (vielleicht überraschenderweise) einfacher, eine allgemeinere Aussage zu zeigen. Wir beginnen hierfür mit

$$(\log(1+x))' = \frac{1}{1+x} = \frac{1}{1-(-x)} = \sum_{n=0}^\infty (-1)^n x^n$$

für alle $x \in (-1,1)$, wobei die Reihe rechts Konvergenzradius 1 hat (und an den Endpunkten divergiert). Nach Korollar 8.5 und Satz 6.85 folgt daraus, dass

$$\log(1+x) = \log(1) + \int_0^x \frac{1}{1+t}\,\mathrm{d}t = \sum_{n=0}^\infty \frac{(-1)^n}{n+1} x^{n+1} = \sum_{k=1}^\infty \frac{(-1)^{k+1}}{k} x^k$$

für alle $x \in (-1,1)$. Da die Potenzreihe rechts auch für $x = 1$ konvergiert, ist die Funktion

$$f : x \in (-1,1] \mapsto \sum_{k=1}^\infty \frac{(-1)^{k+1}}{k} x^k$$

nach dem Abelschen Grenzwertsatz (Satz 6.65) auch bei $x = 1$ stetig. Da die Funktion mit Definitionsbereich $(-1,1]$ definiert durch $\log(1+x)$ für $x \in (-1,1]$ ebenfalls stetig ist und für $x \in (-1,1)$ mit $f(x)$ übereinstimmt, ist

$$\log(2) = \lim_{x \nearrow 1} \log(1+x) = \lim_{x \nearrow 1} f(x) = \sum_{n=1}^\infty \frac{(-1)^{n+1}}{n}.$$

### 8.1.3 Die Leibniz-Reihe

Wir verwenden obige Methode nochmals, um

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = \frac{\pi}{4} \tag{8.2}$$

zu beweisen.

Nach dem Leibniz-Kriterium (Satz 6.25) ist die Reihe $\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}$ konvergent. Wir beginnen die Berechnung ihres Wertes mit

$$\arctan'(x) = \frac{1}{1+x^2} = \sum_{n=0}^{\infty} (-1)^n x^{2n}.$$

für $x \in (-1, 1)$. Nach Korollar 8.5 gilt

$$\arctan(x) = \arctan(0) + \int_0^x \frac{1}{1+t^2}\,\mathrm{d}t = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1}$$

Stetigkeit des Arkustangens bei $x = 1$, der Abelsche Grenzwertsatz (Satz 6.65) und die Identität $\arctan(1) = \frac{\pi}{4}$ beweisen nun (8.2).

## 8.2 Integrationsmethoden

Wir erinnern daran, dass das unbestimmte Integral einer Funktion $f$ in der Variablen $x$ der Ausdruck

$$\int f(x)\,\mathrm{d}x = F(x) + C$$

ist, wobei $F$ eine Stammfunktion von $f$ ist. In den Abschnitten 7.5.2 und 7.4 haben wir bereits einige Regeln zur Berechnung konkreter unbestimmter Integrale kennengelernt: Für $s \in \mathbb{R}$ (oder sogar $s \in \mathbb{C}$) ist

$$\int x^s\,\mathrm{d}x = \begin{cases} \frac{1}{s+1}x^{s+1} + C & \text{falls } s \neq -1 \\ \log|x| + C & \text{falls } s = -1 \end{cases}$$

und

$$\begin{aligned}
\int \exp(x)\,\mathrm{d}x &= \exp(x) + C \\
\int \cos(x)\,\mathrm{d}x &= \sin(x) + C \\
\int \sin(x)\,\mathrm{d}x &= -\cos(x) + C \\
\int \sinh(x)\,\mathrm{d}x &= \cosh(x) + C \\
\int \cosh(x)\,\mathrm{d}x &= \sinh(x) + C \\
\int \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x &= \arcsin(x) + C \\
\int \frac{1}{1+x^2}\,\mathrm{d}x &= \arctan(x) + C \\
\int \frac{1}{\sqrt{1+x^2}}\,\mathrm{d}x &= \operatorname{arsinh}(x) + C \\
\int \frac{1}{\sqrt{x^2-1}}\,\mathrm{d}x &= \operatorname{arcosh}(x) + C.
\end{aligned}$$

Des Weiteren gilt für Funktionen $f_1, f_2$ in der Variable $x$ und Zahlen $\alpha_1, \alpha_2 \in \mathbb{R}$

$$\int \alpha_1 f_1(x) + \alpha_2 f_2(x)\,\mathrm{d}x = \alpha_1 \int f_1(x)\,\mathrm{d}x + \alpha_2 \int f_2(x)\,\mathrm{d}x.$$

Denn falls $F_1$ eine Stammfunktion von $f_1$ ist und $F_2$ eine Stammfunktion von $f_2$ ist, so muss die Funktion $\alpha_1 F_1 + \alpha_2 F_2$ auf Grund der Linearität der Ableitung (Proposition 7.5) eine Stammfunktion von $\alpha_1 f_1 + \alpha_2 f_2$ sein.

Auf ähnliche Weise lassen sich die anderen Regeln der Differentiation als Identitäten für unbestimmte Integrale auffassen, wie wir nun ausführen wollen.

### 8.2.1 Partielle Integration

Die Produktregel in Proposition 7.5

$$(uv)' = u'v + uv'$$

für zwei differenzierbare Funktionen $u, v$ führt ebenso zu einer Integrationsregel, nämlich der **partiellen Integration**

$$\begin{aligned} uv + C &= \int (u'v + uv')\,\mathrm{d}x \\ \int uv'\,\mathrm{d}x &= uv - \int u'v\,\mathrm{d}x + C. \end{aligned} \tag{8.3}$$

In der Leibniz-Notation ist $v' = \frac{\mathrm{d}v}{\mathrm{d}x}$ und $u' = \frac{\mathrm{d}u}{\mathrm{d}x}$. Deswegen schreibt man die partielle Integration oft auch als

$$\int u\,\mathrm{d}v = uv - \int v\,\mathrm{d}u + C,$$

was formal bloss als Kurzform der Formel in (8.3) verstanden werden sollte. Die Regel der partiellen Integration ist bereits ein Beispiel, wo eine einfache Regel des Differenzierens eine komplexere Regel des Integrierens als Entsprechung hat. Die Produktregel erlaubt uns, die Ableitung jedes Produkts mittels der Ableitung dessen Faktoren auszudrücken. Die partielle Integration hingegen erlaubt uns, das unbestimmte Integral eines Produkts mittels dem Integral eines Faktors und eines weiteren Integrals auszudrücken. Mit etwas Glück (und Geschick) ist das zweite Integral einfacher und kann anschliessend berechnet werden. Wir demonstrieren dies anhand zweier Beispiele.

**Beispiel 8.15** (Beispiele partieller Integration)**.**

*(i) Wir berechnen das unbestimmte Integral* $\int x\exp(x)\,\mathrm{d}x$*. Dafür setzen wir* $u(x) = x$ *und* $v'(x) = \exp(x)$*. Eine Stammfunktion von* $v'$ *ist* $v(x) = \exp(x)$*. Damit erhalten wir*

$$\int x\exp(x)\,\mathrm{d}x = x\exp(x) - \int 1\cdot\exp(x)\,\mathrm{d}x + C = x\exp(x) - \exp(x) + C.$$

*Wir bemerken, dass es genügt, in solchen Berechnungen immer bloss eine unbekannte Integrationskonstante* $C$ *zu verwenden, da mehrere solche einfach zusammengefasst werden können. (Kontrollieren Sie diese Rechnung durch Ableiten.) Dieselbe Berechnungsmethode führt auch für unbestimmte Integrale der Form* $\int x^n\exp(x)\,\mathrm{d}x$*,* $\int x^n\sin(x)\,\mathrm{d}x$ *und* $\int x^n\cos(x)\,\mathrm{d}x$ *für* $n \in \mathbb{N}$ *zum Erfolg.*

*(ii) Wir wollen das unbestimmte Integral* $\int \log(x)\,\mathrm{d}x$ *berechnen. Es mag zuerst etwas überraschend sein, dass wir dazu partielle Integration verwenden wollen. Sei* $u(x) = \log(x)$ *und*

*$v' = 1$. Dann ist $v(x) = x$ eine Stammfunktion von $v'$, womit*

$$\int \log(x)\,\mathrm{d}x = \int \log(x)\cdot 1\,\mathrm{d}x = \log(x)\cdot x - \int \frac{1}{x}x\,\mathrm{d}x + C = x\log(x) - \int 1\,\mathrm{d}x + C$$
$$= x\log(x) - x + C.$$

*Dies kann man wiederum durch Ableiten verifizieren (was nicht notwendig ist, aber einen sehr einfachen Test darstellt).*

**Übung 8.16.**

*(i) Berechnen Sie $\int x^2 \sin(x)\,\mathrm{d}x$.*

*(ii) Geben Sie eine rekursive Formel zur Berechnung von $\int x^n \exp(x)\,\mathrm{d}x$, $\int x^n \sin(x)\,\mathrm{d}x$ und $\int x^n \cos(x)\,\mathrm{d}x$ für $n \in \mathbb{N}$ an.*

*(iii) Berechnen Sie $\int x^s \log(x)\,\mathrm{d}x$ für jedes $s \in \mathbb{R}$. Beachten Sie hierbei, dass der Fall $s = -1$ getrennt zu behandeln ist.*

*(iv) Berechnen Sie das unbestimmte Integral $\int e^{ax}\sin(bx)\,\mathrm{d}x$ für $a, b \in \mathbb{R}\setminus\{0\}$.*

### 8.2.2 Substitution

Falls eine Funktion $g$ auf einem Intervall $I_u$ das unbestimmte Integral $\int g(u)\,\mathrm{d}u = G(u)+C$ besitzt und $f : I_x \to I_u$ eine stetig differenzierbare Abbildung auf dem Intervall $I_x$ ist, dann gilt

$$\int (g\circ f)(x) f'(x)\,\mathrm{d}x = (G\circ f)(x) + C$$

auf $I_x$. Dies folgt unmittelbar aus der Kettenregel in Satz 7.8 und wird oft auch geschrieben als

$$\int (g\circ f)(x) f'(x)\,\mathrm{d}x = \int g(u)\,\mathrm{d}u \tag{8.4}$$

für die „neue Variable“ $u = f(x)$. Alternativ werden wir die obige **Substitutionsregel** gemeinsam mit der Leibniz-Notation auch in folgender informellen Schreibweise

$$\int (g\circ f)(x) f'(x)\,\mathrm{d}x = \int (g\circ f)(x)\frac{\mathrm{d}u}{\mathrm{d}x}\,\mathrm{d}x = \int g(u)\,\mathrm{d}u = G(u) + C = G(f(x)) + C,$$

verwenden, wobei $u = f(x)$ und $\mathrm{d}u = f'(x)\,\mathrm{d}x$.

**Beispiel 8.17.**

*(i) Es gilt*

$$\int \frac{x}{1+x^2}\,\mathrm{d}x = \int \underbrace{\frac{1}{1+x^2}}_{=\frac{1}{u}} \underbrace{x\,\mathrm{d}x}_{=\frac{1}{2}\,\mathrm{d}u} = \frac{1}{2}\int \frac{1}{u}\,\mathrm{d}u = \frac{1}{2}\log|u| = \frac{1}{2}\log(1+x^2) + C,$$

*wobei* $u = 1 + x^2$ *gesetzt wurde, womit* $\mathrm{d}u = 2x\,\mathrm{d}x$.

*(ii) Es gilt*

$$\int \frac{1}{\sin(x)}\,\mathrm{d}x = \int \frac{1}{2\sin\left(\frac{x}{2}\right)\cos\left(\frac{x}{2}\right)}\,\mathrm{d}x = \int \frac{1}{\tan(u)\cos^2(u)}\,\mathrm{d}u = \int \frac{1}{v}\,\mathrm{d}v = \log|v| + C$$
$$= \log\left|\tan\left(\frac{x}{2}\right)\right| + C,$$

*wobei* $u = \frac{x}{2}$, $\mathrm{d}u = \frac{1}{2}\,\mathrm{d}x$, *und* $v = \tan(u)$, $\mathrm{d}v = \frac{1}{\cos^2(u)}\,\mathrm{d}u$.

Wie bereits erwähnt, benötigt die Integration mehr Übung und Vorraussicht als die Differentiation. Obige Substitutionen benötigen zum Beispiel den Blick ob gewisse Faktoren vielleicht die gewünschte Ableitung $f'$ einer inneren Funktionen $f$ darstellen könnte. Manchmal ist dies naheliegend wie in Beispiel 8.17(i), doch manchmal erfordert dies Erfahrung und eine längere Suche wie in Beispiel 8.17(ii).

### 8.2.3 Integration rationaler Funktionen

Wir erinnern daran, dass eine rationale Funktion eine Funktion der Form $x \mapsto \frac{p(x)}{q(x)}$ für Polynome $p(t), q(t) \in \mathbb{R}[t]$ und $q(t) \neq 0$ ist, wobei der Definitionsbereich $\mathbb{R}$ ohne die Nullstellen von $q(t)$ ist. Wir wollen hier ein Verfahren zur Berechnung des unbestimmten Integrals einer rationalen Funktion besprechen. Nach Division mit Rest für Polynome (siehe Übung 3.17) können wir als erstes ein Polynom abspalten, so dass die verbleibende rationale Funktion von der Form $\frac{p_1(x)}{q(x)}$ für $\deg(p_1) < \deg(q)$ ist.

Da Polynome mittels der Formel $\int x^n\,\mathrm{d}x = \frac{1}{n+1}x^{n+1} + C$ integriert werden können, nehmen wir nun an, dass der Grad von $p$ kleiner als der Grad von $q$ ist. Wir betrachten zuerst einige Spezialfälle.

**Beispiel 8.18** (Integration von elementaren rationalen Funktionen)**.** *Sei* $a \in \mathbb{R}$ *beliebig und* $n \geq 2$ *eine natürliche Zahl.*

*(i) Es gilt*

$$\int \frac{1}{x-a}\,\mathrm{d}x = \int \frac{1}{u}\,\mathrm{d}u = \log|u| + C = \log|x-a| + C,$$

*wobei* $u = x - a$ *gesetzt wurde und* $\mathrm{d}u = \mathrm{d}x$ *ist.*

*(ii) Für* $n \geq 2$ *gilt*

$$\int \frac{1}{(x-a)^n}\,\mathrm{d}x = \int u^{-n}\,\mathrm{d}u = \frac{1}{-n+1}u^{-n+1} + C = \frac{1}{-n+1}(x-a)^{-n+1} + C,$$

*wobei wieder* $u = x - a$ *gesetzt wurde und* $\mathrm{d}u = \mathrm{d}x$ *ist.*

*(iii) Falls $a \neq 0$ ist, so gilt*

$$\int \frac{1}{a^2+x^2}\,\mathrm{d}x = \frac{1}{a^2}\int \frac{1}{1+\left(\frac{x}{a}\right)^2}\,\mathrm{d}x = \frac{1}{a}\int \frac{1}{1+u^2}\,\mathrm{d}u = \frac{1}{a}\arctan(u)+C$$
$$= \frac{1}{a}\arctan\left(\frac{x}{a}\right)+C,$$

*wobei $u = \frac{x}{a}$ gesetzt wurde und $\mathrm{d}u = \frac{1}{a}\,\mathrm{d}x$ ist.*

*(iv) Es gilt*

$$\int \frac{x}{a^2+x^2}\,\mathrm{d}x = \frac{1}{2}\int \frac{1}{u}\,\mathrm{d}u = \frac{1}{2}\log|u|+C = \frac{1}{2}\log(a^2+x^2)+C,$$

*wobei $u = a^2+x^2$ und $\mathrm{d}u = 2x\,\mathrm{d}x$.*

*(v) Für $n \geq 2$ gilt*

$$\int \frac{x}{(a^2+x^2)^n}\,\mathrm{d}x = \frac{1}{2}\int \frac{1}{u^n}\,\mathrm{d}u = \frac{1}{2(1-n)}u^{1-n}+C = \frac{1}{2(1-n)}(a^2+x^2)^{1-n}+C,$$

*wobei $u = a^2+x^2$ und $\mathrm{d}u = 2x\,\mathrm{d}x$.*

Im Allgemeinen verwenden wir die sogenannte **Partialbruchzerlegung** von $\frac{p(x)}{q(x)}$, um die Integration rationaler Funktionen auf obige Beispiele zurückzuführen. In der Tat lässt sich $\frac{p(x)}{q(x)}$ als Linearkombination von einfacheren rationalen Funktionen darstellen. Diese sind von der Form

$$\frac{1}{(x-a)},\quad \frac{1}{(x-a)^2},\quad \dots\quad ,\quad \frac{1}{(x-a)^k},$$

oder von der Form

$$\frac{A_1x+B_1}{(x-\lambda)(x-\overline{\lambda})},\quad \dots\quad ,\quad \frac{A_\ell x+B_\ell}{\big((x-\lambda)(x-\overline{\lambda})\big)^\ell},$$

wobei $a \in \mathbb{R}$ eine Nullstelle von $q$ mit Vielfachheit $k$ und $\lambda \in \mathbb{C}\setminus\mathbb{R}$ eine Nullstelle von $q$ mit Vielfachheit $\ell \in \mathbb{N}$ ist und $A_1, B_1, \dots, A_\ell, B_\ell \in \mathbb{R}$.

**Beispiel 8.19** (Integration rationaler Funktionen und die Partialbruchzerlegung)**.**

*(i) Wir wollen das unbestimmte Integral $\int \frac{x^4+1}{x^2(x+1)}\,\mathrm{d}x$ bestimmen. Als erstes führen wir Division mit Rest*

$$\begin{array}{l}
(x^4+1):(x^3+x^2) = x-1\\
\underline{-x^4-x^3}\\
\quad -x^3+1\\
\quad \underline{x^3+x^2}\\
\qquad\quad x^2+1 \quad \text{Rest}
\end{array}$$

*durch, womit*

$$\int \frac{x^4+1}{x^2(x+1)}\,\mathrm{d}x = \int \left(x - 1 + \frac{x^2+1}{x^2(x+1)}\right)\mathrm{d}x = \frac{x^2}{2} - x + \int \frac{x^2+1}{x^2(x+1)}\,\mathrm{d}x.$$

*Um die Partialbruchzerlegung von* $\frac{x^2+1}{x^2(x+1)}$ *zu erhalten, setzen wir*

$$\frac{x^2+1}{x^2(x+1)} = \frac{A}{x^2} + \frac{B}{x} + \frac{C}{x+1}$$

*für noch unbekannte Zahlen* $A, B, C \in \mathbb{R}$*, multiplizieren mit* $x^2(x+1)$ *und erhalten*

$$x^2 + 1 = A(x+1) + Bx(x+1) + Cx^2.$$

*Nun setzen wir in diesem* $x = 0$ *um* $A = 1$ *zu erhalten und* $x = -1$*, um* $C = 2$ *zu erhalten. Für* $x = 1$ *ergibt sich nun* $2 = 1 \cdot 2 + B \cdot 2 + 2 \cdot 1$ *und somit* $B = -1$*. (Alternativ kann man auch beide Seiten ausmultiplizieren, die Koeffizienten links und rechts vergleichen, und auf diese Weise drei Gleichungen in den unbekannten Variablen* $A, B, C$ *erhalten.) Daher ist*

$$\begin{aligned}\int \frac{x^2+1}{x^2(x+1)}\,\mathrm{d}x &= \int \frac{1}{x^2}\,\mathrm{d}x - \int \frac{1}{x}\,\mathrm{d}x + 2\int \frac{1}{x+1}\,\mathrm{d}x \\ &= -\frac{1}{x} - \log|x| + 2\log|x+1| + D\end{aligned}$$

*(ii) Wir berechnen das unbestimmte Integral* $\int \frac{1}{x(x^2+2x+2)}\,\mathrm{d}x$*. Man beachte dabei, dass das Polynom* $x^2+2x+2$ *keine reellen Nullstellen hat. Für die Partialbruchzerlegung machen wir den Ansatz*

$$\frac{1}{x(x^2+2x+2)} = \frac{A}{x} + \frac{Bx+C}{x^2+2x+2}.$$

*Nun multiplizieren wir mit* $x(x^2+2x+2)$ *und erhalten*

$$1 = A(x^2+2x+2) + (Bx+C)x.$$

*Für* $x = 0$ *ergibt sich* $A = \frac{1}{2}$*. Daher ist*

$$1 = (\tfrac{1}{2} + B)x^2 + (1+C)x + 1$$

*und* $B = -\frac{1}{2}$ *und* $C = -1$. *Es folgt*

$$\begin{aligned}\int \frac{1}{x(x^2+2x+2)}\,\mathrm{d}x &= \tfrac{1}{2}\int \frac{1}{x}\,\mathrm{d}x - \tfrac{1}{2}\int \frac{x+2}{x^2+2x+2}\,\mathrm{d}x \\ &= \tfrac{1}{2}\log|x| - \tfrac{1}{2}\int \frac{x+2}{(x+1)^2+1}\,\mathrm{d}x \\ &= \tfrac{1}{2}\log|x| - \tfrac{1}{2}\int \frac{u+1}{u^2+1}\,\mathrm{d}x \\ &= \tfrac{1}{2}\log|x| - \tfrac{1}{4}\log|u^2+1| - \tfrac{1}{2}\arctan(u) + D \\ &= \tfrac{1}{2}\log|x| - \tfrac{1}{4}\log((x+1)^2+1) - \tfrac{1}{2}\arctan(x+1) + D,\end{aligned}$$

*wobei wir* $u = x+1$ *gesetzt haben und Beispiele 8.18 (c) und (d) verwendet haben.*

In manchen Fällen kann obiges Verfahren auch auf das Integral $\int \frac{1}{(a^2+x^2)^n}\,\mathrm{d}x$ für ein $a \in \mathbb{R}$ und $n \geq 2$ führen, was wir mit der trigonometrischen Substitution $\tan(u) = \frac{x}{a}$ (siehe unten) behandeln können. Eine andere, allgemeinere Herangehensweise möchten wir in folgender Bemerkung für Interessierte behandeln.

*Bemerkung* (Integration rationaler Funktionen mit mehrfachen komplexen Nullstellen). Wie oben schon bemerkt, kann man nach der Partialbruchzerlegung ein Integral einer rationalen Funktion auf die Integration von Ausdrücken der Form $\frac{1}{(x-a)^k}$ oder von $\frac{Ax+B}{(x^2+bx+c)^k}$ für $k \in \mathbb{N}$ und für Konstanten $a, A, B, b, c \in \mathbb{R}$ zurückführen, wobei die Polynome der Form $x^2+bx+c$ keine rellen Nullstellen haben. Für die Berechnung eines Integrals des zweiten Typs mit $k > 1$ möchten wir hier einen Algorithmus erläutern, wobei wir uns auf den Fall $x^2+bx+c = x^2+1$ beschränken (auf welchen man den allgemeinen Fall mit quadratischem Ergänzen zurückführen kann).

Seien also $k > 1$ und ein Polynom $q$ von Grad kleiner als $2k$ gegeben. Dann ist das unbestimmte Integral $\int \frac{q(x)}{(x^2+1)^k}\,\mathrm{d}x$ immer von der Form

$$\frac{p(x)}{(x^2+1)^{k-1}} + \alpha\arctan(x) + \beta\log(x^2+1) + C. \tag{8.5}$$

für ein Polynom $p$ von Grad kleiner $2k-2$ und Konstanten $\alpha, \beta$. Durch Ableiten, auf den gemeinsamen Nenner bringen und Vergleich der Koeffizienten lässt sich somit die Stammfunktion ermitteln.

**Übung 8.20.** *Wir möchten in dieser Übung den oben erklärten Algorithmus genauer erklären und beginnen mit einem konkreten Beispiel.*

*(i) Berechnen Sie das Integral* $\int \frac{1}{(x^2+1)^2}\,\mathrm{d}x$.

*Sei nun* $k > 1$ *und* $q$ *ein Polynom von Grad kleiner als* $2k$.

*(ii) Zeigen Sie, dass die Ableitung von* $\frac{p(x)}{(x^2+1)^{k-1}}$ *für ein beliebiges Polynom* $p$ *von Grad kleiner als* $2k-2$ *durch*

$$\frac{(x^2+1)p'(x) - 2(k-1)xp(x)}{(x^2+1)^k}$$

*gegeben ist.*

*(iii) Berechnen Sie die Matrixdarstellung $M$ der Abbildung*

$$\phi : p(x) \mapsto (x^2+1)p'(x) - 2(k-1)xp(x)$$

*bezüglich der Basis der Monome.*

*(iv) Schliessen Sie auf die Darstellung in (8.5), indem Sie zeigen, dass das Bild von $\phi$ zusammen mit $(x^2+1)^{k-1}$ und $2x(x^2+1)^{k-1}$ den Vektorraum der Polynome von Grad kleiner gleich $2k-1$ aufspannt.*

VO: 14.12.17

### 8.2.4 Trigonometrische Substitution

In allen bisherigen Beispielen der Substitutionsregel in Abschnitt 8.2.2 hatten wir das Glück, dass das vorhandene Integral (vielleicht nach etwas Arbeit) bereits die richtige Struktur besass. Man verwendet die Substitutionsregel aber oft auch bevor wir noch wissen welches Integral sich eigentlich nach der Substitution ergibt, wobei es gewisse Funktionentypen gibt bei denen eine gewisse Substitution erfahrungsgemäss erfolgreich sein könnte. Wir wenden uns nun einem konkreten Beispiel dessen zu.

**Beispiel 8.21** (Kreisfläche)**.** *Wir möchten für $r > 0$ das unbestimmte Integral $\int \sqrt{r^2 - x^2}\,\mathrm{d}x$ berechnen. Auf Grund der trigonometrischen Identitäten $\sqrt{r^2 - r^2\sin^2(\theta)} = r\cos(\theta)$ bietet es sich nun an, die Funktion*

$$f : I_\theta = \left(-\tfrac{\pi}{2}, \tfrac{\pi}{2}\right) \to I_x = (-r, r),\ \ \theta \mapsto x = r\sin(\theta)$$

*für die Substitution zu verwenden. Denn mit dieser Substitution haben wir die Hoffnung, die Wurzel in einen anderen Ausdruck zu verwandeln.*

*Allerdings ist dies umgekehrt zu der Substitution in Abschnitt 8.2.2, da wir hier die „neue Variable“ $\theta$ verwenden um die „alte Variable“ $x = r\sin(\theta)$ auszudrücken. (Anstatt wie in Abschnitt 8.2.2 wo wir die neue Variable $u = f(x)$ als Funktion der alten Variable $x$ gesehen haben). Da $f$ bijektiv ist, ist dies kein Problem: denn $x = r\sin(\theta) \in (-r, r)$ ist zu $\theta = \arcsin(\frac{x}{r}) \in (-\frac{\pi}{2}, \frac{\pi}{2})$ äquivalent. Weiters ist die Ableitung von $f = \frac{\mathrm{d}x}{\mathrm{d}\theta}$ gleich $r\cos\theta$ und damit auf ganz $(-\frac{\pi}{2}, \frac{\pi}{2})$ ungleich $0$. Gemeinsam mit dem Satz über die Ableitung der inversen Funktion (Satz 7.14) erhalten wir daher*

$$\begin{aligned}
\int \sqrt{r^2 - x^2}\,\mathrm{d}x &= \int r\cos(\theta)\left(\tfrac{\mathrm{d}x}{\mathrm{d}\theta}\right)\left(\tfrac{\mathrm{d}\theta}{\mathrm{d}x}\right)\mathrm{d}x \\
&= \int r^2\cos^2(\theta)\,\mathrm{d}\theta \\
&= r^2 \int \frac{1+\cos(2\theta)}{2}\,\mathrm{d}\theta \\
&= \tfrac{r^2}{2}\left(\theta + \tfrac{1}{2}\sin(2\theta)\right) + C \\
&= \tfrac{r^2}{2}\arcsin\left(\tfrac{x}{r}\right) + \tfrac{1}{2}x\sqrt{r^2 - x^2} + C,
\end{aligned}$$

*wobei wir eben* $x = r\sin(\theta)$, $\sqrt{r^2 - x^2} = r\cos(\theta)$, $\mathrm{d}x = r\cos(\theta)\,\mathrm{d}\theta$ *und die trigonometrischen Identitäten*

$$\begin{aligned}\cos(2\theta) &= \cos^2(\theta) - \sin^2(\theta) = 2\cos^2(\theta) - 1\\ \cos^2(\theta) &= \frac{\cos(2\theta)+1}{2}\\ \sin(2\theta) &= 2\sin(\theta)\cos(\theta)\end{aligned}$$

*für* $\theta \in (-\frac{\pi}{2}, \frac{\pi}{2})$ *verwendet haben.*

*Veranschaulichen Sie sich die Substitution und die wichtigsten der obigen Identitäten in einem rechtwinkeligen Dreieck. Geben Sie weiters eine geometrische Interpretation der beiden Terme des unbestimmten Integrals bei der Berechnung des bestimmten Integrals* $\int_0^b \sqrt{r^2 - x^2}\,\mathrm{d}x$ *für* $0 < b \leq r$ *an.*

*Dies zeigt, dass*

$$G_0 : x \in (-r, r) \mapsto \tfrac{1}{2}r^2 \arcsin\left(\tfrac{x}{r}\right) + \tfrac{1}{2}x\sqrt{r^2 - x^2}$$

*eine Stammfunktion von* $g_0 : x \in (-r, r) \mapsto \sqrt{r^2 - x^2}$ *ist (was wie immer viel einfacher zu überprüfen ist). Da aber sogar die Funktion* $g : x \in [-r, r] \mapsto \sqrt{r^2 - x^2}$ *stetig ist, besitzt* $g$ *nach Korollar 8.3 auch auf ganz* $[-r, r]$ *eine Stammfunktion* $G_1$*, welche auf* $(-r, r)$ *mit* $G_0 + C$ *übereinstimmt. Da aber*

$$G : x \in [-r, r] \mapsto \tfrac{1}{2}r^2 \arcsin\left(\tfrac{x}{r}\right) + \tfrac{1}{2}x\sqrt{r^2 - x^2}$$

*auch eine auf ganz* $[-r, r]$ *stetige Funktion definiert, folgt aus Stetigkeit von* $G$ *und* $G_1$*, dass* $G_1 = G + C$ *und damit ist* $G$ *auf ganz* $[-r, r]$ *eine Stammfunktion von* $g$*. Wir bemerken allerdings, dass* $\arcsin\left(\frac{x}{r}\right)$ *keine Ableitung in den Punkten* $-r$ *und* $r$ *besitzt. (Wieso ist dies kein Widerspruch zu obiger Diskussion?)*

Substitutionen wie obige nennen sich vielfach **trigonometrische Substitutionen**. Wir werden bei diesen Berechnungen nicht immer so sorgfältig argumentieren und vielmehr der Leibniz Notation vertrauen, doch muss immer Invertierbarkeit der Funktion gegeben sein wenn wir die alte Variable durch die neue Variable ausdrücken. Für die folgende Auflistung der trigonometrischen Substitutionen sei $n \in \mathbb{Z}$.

- In Ausdrücken der Form $(a^2 - x^2)^{\frac{n}{2}}$ für $a > 0$ führt wie bereits im obigen Beispiel oft die Substitution $x = a\sin(\theta)$ mit $\theta \in (-\frac{\pi}{2}, \frac{\pi}{2})$ zum Ziel, wobei sich damit $\mathrm{d}x = a\cos(\theta)\,\mathrm{d}\theta$ und $(a^2 - x^2)^{\frac{1}{2}} = a\cos(\theta)$ ergibt.
- In Ausdrücken der Form $(a^2 + x^2)^{\frac{n}{2}}$ für $a > 0$ führt oft die Substitution $x = a\tan(\theta)$ mit $\theta \in (-\frac{\pi}{2}, \frac{\pi}{2})$ zum Ziel, wobei sich damit $\mathrm{d}x = \frac{a}{\cos^2(\theta)}\,\mathrm{d}\theta$ und $(a^2 + x^2)^{\frac{1}{2}} = \frac{a}{\cos(\theta)}$ ergibt.
- Obwohl dies keine trigonometrische Substitution darstellt, bemerken wir noch Folgendes. Falls ein „einzelnes“ $x$ vor dem Ausdruck $(a^2 - x^2)^{\frac{n}{2}}$ oder dem Ausdruck $(a^2 + x^2)^{\frac{n}{2}}$ steht, ist die Substitution $u = a^2 - x^2$ respektive $u = a^2 + x^2$ teilweise viel einfacher.

Als Merkhilfe kann es helfen für die beiden trigonometrischen Substitution ein rechtwinkeliges Dreieck zu skizzieren und abhängig von der Substitution die Seiten mit Hilfe von Pythagoras entsprechend zu beschriften.

**Beispiel 8.22** (Trigonometrische Substitution)**.**

*(i) Es gilt für $a > 0$*

$$\begin{aligned}\int \frac{1}{(a^2+x^2)^{\frac{3}{2}}}\,\mathrm{d}x &= \int \frac{\cos^3(\theta)}{a^3} a \frac{1}{\cos^2(\theta)}\,\mathrm{d}\theta = \tfrac{1}{a^2}\int \cos(\theta)d\theta = \tfrac{1}{a^2}\sin(\theta)+C \\ &= \frac{x}{a^2\sqrt{a^2+x^2}}+C,\end{aligned}$$

*wobei wir $x = a\tan(\theta)$, $\sqrt{a^2+x^2} = a\frac{1}{\cos(\theta)}$, $\mathrm{d}x = a\frac{1}{\cos^2(\theta)}\,\mathrm{d}\theta$ verwendet haben. (Veranschaulichen Sie sich die Substitution und obige Identitäten in einem Bild.)*

*(ii) Es ist*

$$\int x\sqrt{1-x^2}\,\mathrm{d}x = -\tfrac{1}{2}\int u^{\frac{1}{2}}\,\mathrm{d}u = -\tfrac{1}{2}\tfrac{2}{3}u^{\frac{3}{2}}+C = -\tfrac{1}{3}(1-x^2)^{\frac{3}{2}}+C,$$

*wobei $u = 1-x^2$, $\mathrm{d}u = -2x\,\mathrm{d}x$.*

### 8.2.5 Weitere Integrationsmethoden

Es gibt viele weitere Methoden zur Integration; viele davon beruhen auf spezielle Substitutionen.

Beispielsweise lassen sich gewisse unbestimmte Integrale mit hyperbolischen Substitutionen berechnen. Sei $n \geq -1$. In Ausdrücken der Form $(x^2-a^2)^{\frac{n}{2}}$ für $a \in \mathbb{R}$ führt oft die Substitution $x = a\cosh(u)$ zum Ziel, wobei sich damit $\mathrm{d}x = a\sinh(u)\,\mathrm{d}u$ und $(x^2-a^2)^{\frac{1}{2}} = a\sinh(u)$ ergibt.

**Beispiel 8.23.** *Wir berechnen*

$$\begin{aligned}\int \sqrt{x^2-1}\,\mathrm{d}x &= \int \sinh^2(u)\,\mathrm{d}u = \cosh(u)\sinh(u) - \int \cosh^2(u)\,\mathrm{d}u \\ &= \cosh(u)\sinh(u) - \int \sinh^2(u)+1\,\mathrm{d}u + C \\ &= \cosh(u)\sinh(u) - u - \int \sinh^2(u)\,\mathrm{d}u + C,\end{aligned}$$

*wobei $x = \cosh(u)$ und $\mathrm{d}x = \sinh(u)\,\mathrm{d}u$. Nach Auflösen ergibt sich somit*

$$\int \sqrt{x^2-1}\,\mathrm{d}x = \int \sinh^2(u)\,\mathrm{d}u = \frac{\cosh(u)\sinh(u)-u}{2}+C = \frac{x\sqrt{x^2-1}-\operatorname{arcosh}(x)}{2}+C.$$

Eine andere Methode, die wir hier kurz erwähnen möchten, ist die sogenannte Halbwinkelmethode (oder auch Weierstrass-Substitution). Diese ist dann nützlich, wenn man das Integral einer rationalen Funktion in $\cos(x)$ und $\sin(x)$ wie zum Beispiel $\frac{\cos^2(x)}{\sin(x)+2017}$ in die Integration einer rationalen Funktion in $u = \tan\left(\frac{x}{2}\right)$ umwandeln möchte (siehe auch Beispiel 8.17 (b)).

**Übung 8.24** (Halbwinkelmethode)**.** *Wir möchten das unbestimmte Integral* $\int \frac{\cos(x)}{2+\sin(x)}\,\mathrm{d}x$ *mit der Substitution* $u = \tan\left(\frac{x}{2}\right)$ *berechnen. Zeigen Sie dafür zuerst die Identitäten*

$$\sin(x) = \frac{2u}{1+u^2}, \quad \cos(x) = \frac{1-u^2}{1+u^2}.$$

*Zeigen Sie anschliessend, dass das obige Integral nach Substitution zu einem Integral einer rationalen Funktion in* $u$ *wird und berechnen Sie es.*

Manchmal führt man auch die eine oder die andere Substitution durch, weil in der zu integrierenden Funktion eine verschachtelte Funktion vorliegt und man einfach keine andere Methode zur Verfügung hat. Zum Beispiel bei dem Integral $\int \sin(\sqrt{x})\,\mathrm{d}x$ steht keine der erwähnten Methoden zur Verfügung, doch ist man versucht $u = \sqrt{x}$ zu setzen um zu sehen was sich daraus ergibt. Dies führt in der Tat zum Erfolg (wieso?). Ebenso in dem Integral der Form $\int \frac{1}{1+\exp(x)}\,\mathrm{d}x$ führt der Ansatz $u = \exp(x)$ zu einem unbestimmten Integral einer rationalen Funktion (wieso?).

### 8.2.6 Neue Funktionen

Manchmal führen obige Methoden zur Bestimmung eines unbestimmten Integrals einer Funktion zu keinem Ergebnis. Dies kann daran liegen, dass die gesuchte Stammfunktion sich nicht mit den bisher bekannten Funktionen ausdrücken lässt.

**Beispiel 8.25** (Integralsinus)**.** *Der* ***Integralsinus*** *ist die Stammfunktion* $\mathrm{Si} : \mathbb{R} \to \mathbb{R}$ *der stetigen Funktion*

$$x \in \mathbb{R} \mapsto \begin{cases} \frac{\sin(x)}{x} & \textit{falls } x \neq 0 \\ 1 & \textit{falls } x = 0 \end{cases}$$

*mit der Normalisierung* $\mathrm{Si}(0) = 0$*. Er lässt sich als Potenzreihe schreiben, denn nach Satz 6.85 gilt*

$$\mathrm{Si}(x) = \int_0^x \frac{\sin(t)}{t}\,\mathrm{d}t = \int_0^x \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} t^{2n}\,\mathrm{d}t = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!(2n+1)} x^{2n+1}$$

*für alle* $x \in \mathbb{R}$*.*

**Beispiel 8.26** (Integralkosinus)**.** *Der Integralkosinus* $\mathrm{Ci} : (0,\infty) \to \mathbb{R}$ *ist definiert als die Stammfunktion von* $x \in (0,\infty) \mapsto \frac{\cos(x)}{x} \in \mathbb{R}$ *mit der Normalisierung* $\lim_{x\to\infty} \mathrm{Ci}(x) = 0$*.*

Dabei möchten wir auf folgende Übung verweisen, die zeigt, dass der Integralkosinus so wohldefiniert ist.

**Übung 8.27.** *Sei* $F$ *eine Stammfunktion von* $x \in (0,\infty) \mapsto \frac{\cos(x)}{x} \in \mathbb{R}$*. Zeigen Sie, dass der Grenzwert* $\lim_{x\to\infty} F(x)$ *existiert. Drücken Sie* Ci *als Summe einer Konstanten (der sogenannten Euler-Mascheroni Konstanten), der Logarithmusfunktion und einer Potenzreihe aus.*

Unter Verwendung uneigentlicher Integrale werden wir später weitere wichtige Funktionen kennenlernen, die sich nicht in Termen bekannter Funktionen ausdrücken lassen – siehe zum Beispiel 8.45.

### 8.2.7 Das bestimmte Integral

Alle obigen Regeln zur Berechnung des unbestimmten Integrals lassen sich nach dem Fundamentalsatz der Integral- und Differentialrechnung eins zu eins auch für das Riemann-Integral, welches im Gegensatz zum unbestimmten Integral auch das bestimmte Integral genannt wird, anwenden. Dabei haben wir zwei Möglichkeiten.

- Eine erste Möglichkeit ist mit obigen Methoden zuerst das unbestimmte Integral zu berechnen und dann Korollar 8.4 zur Berechnung des Riemann-Integrals zu verwenden.
- Falls wir aber nur an einem einzigen Riemann-Integral interessiert sind, ist es oft einfacher, die Ausdrücke ausserhalb des Integrals so früh wie möglich zu berechnen. Wir erklären dies im Folgenden für die partielle Integration und die Substitution.

Sind $u, v$ zwei stetig differenzierbare Funktionen auf einem kompakten Intervall $[a, b]$ mit Endpunkten $a < b$. Dann gilt

$$\int_a^b uv' \,\mathrm{d}x = [uv]_a^b - \int_a^b u'v \,\mathrm{d}x.$$

Denn falls $F$ eine Stammfunktion von $uv'$ und $G$ eine Stammfunktion von $u'v$ ist, dann gilt für alle $x \in [a, b]$

$$F(x) + C_1 = \int uv' \,\mathrm{d}x = u(x)v(x) - \int u'v \,\mathrm{d}x + C_2 = u(x)v(x) - G(x) + C_3$$

für gewisse Integrationskonstanten $C_1, C_2, C_3$. Somit ist nach Korollar 8.4

$$\begin{aligned}\int_a^b uv' \,\mathrm{d}x = [F(x)]_a^b = F(b) - F(a) &= u(b)v(b) - G(b) - (u(a)v(a) - G(a)) \\ &= [u(x)v(x)]_a^b - [G(x)]_a^b = [u(x)v(x)]_a^b - \int_a^b u'v \,\mathrm{d}x.\end{aligned}$$

Ebenso können wir bei einer Substitution in Abschnitt 8.2.2 die Grenzen für ein Riemann-Integral enstsprechend der Substitution neu berechnen. Sei $I_x$ ein Intervall mit Endpunkten $a_x < b_x$, sei $I_u$ ein weiteres Intervall, sei $g : I_u \to \mathbb{R}$ stetig und $f : I_x \to I_u$ stetig differenzierbar. Für ein kompaktes Intervall $[a, b]$ mit Endpunkten $a < b$ in $I_x$ gilt dann

$$\int_a^b g \circ f(x) f'(x) \,\mathrm{d}x = \int_{f(a)}^{f(b)} g(u) \,\mathrm{d}u.$$

In der Tat, wenn $G$ eine Stammfunktion von $g$ auf $I_u$ ist, dann ist nach der Kettenregel $G \circ f$ eine Stammfunktion von $x \in I_x \mapsto g \circ f(x)f'(x)$. Nach Korollar 8.4 gilt also

$$\int_a^b g \circ f(x)f'(x)\,\mathrm{d}x = [G \circ f]_a^b = G(f(b)) - G(f(a)) = [G]_{f(a)}^{f(b)} = \int_{f(a)}^{f(b)} g(u)\,\mathrm{d}u.$$

Die Annahme der Stetigkeit an $g$ kann abgeschwächt werden – siehe die entsprechende Übung im Abschnitt 8.8.2.

### 8.2.8 Leibniz-Notation

Wir werden die Leibniz-Notation in der Berechnung von unbestimmten und bestimmten Integralen wie bereits oben im Folgenden immer wieder verwenden. Diese Notation verpackt in einem natürlichen Formalismus die partielle Integration

$$\int u\,\mathrm{d}v = uv - \int v\,\mathrm{d}u + C$$

und die Substitutionsregeln

$$\int g(u(x))\frac{\mathrm{d}u}{\mathrm{d}x}\,\mathrm{d}x = \int g(u)\,\mathrm{d}u$$
$$\int g(u(x))\,\mathrm{d}x = \int g(u)\frac{\mathrm{d}x}{\mathrm{d}u}(u)\,\mathrm{d}u,$$

wobei wir in der zweiten Formulierung der Substitution vorraussetzen, dass $u : I_x \to I_u$ bijektiv mit nicht verschwindender Ableitung ist und dadurch im linken Integral mit $1 = \frac{\mathrm{d}x}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x}$ multiplizieren konnten und die erste Formulierung der Substitutionsregel anwenden konnten. Wie wir gesehen haben, sind diese Regeln Umformulierungen der Produktregel für die Ableitung und der Kettenregel für die Ableitung (gemeinsam mit der Ableitungsregel für die inverse Abbildung).

Bei konkreten Integralberechnungen verwenden wir mitunter auch Gleichungen, die $\mathrm{d}x$ und $\mathrm{d}u$ miteinander verbinden. Zum Beispiel bei der trigonometrischen Substitution $x = a\sin\theta$ (für $x \in (-a, a)$ und $\theta \in (-\frac{\pi}{2}, \frac{\pi}{2})$) verwenden wir auch die Formel $\mathrm{d}x = a\cos\theta\,\mathrm{d}\theta$, die formal gesehen keine Bedeutung hat (und deswegen auf keinen Fall in dieser Form in Beweisen auftreten sollte), doch eben im Zuge der Substitution in der Formulierung der Leibniz-Notation einen bequemen Zwischenschritt darstellt.

Informell taucht in Anwendungen das Symbol $\mathrm{d}x$ auch oft in Diskussionen auf, die zu einem Riemann-Integral führen, wobei $\mathrm{d}x$ dann für ein (sehr) kleines $\Delta x$ stehen sollte. In Anwendungen werden häufig die Begriffe der Riemann-Summe oder der additiven Intervallfunktion vermieden, wobei es genau diese Begriffe sind, die diese Verwendung von $\mathrm{d}x$ genau und formal korrekt machen würden (siehe Abschnitte 4.4 und 5.5). Auf jeden Fall hat in diesem Zusammenhang eine Formel der Gestalt $\mathrm{d}x = a\cos\theta\,\mathrm{d}\theta$ auch eine Interpretation: Da $\Delta x$ die Länge eines kleinen Teilintervalls von $(-a, a)$ angibt und $\Delta\theta$ die Länge des entsprechenden Teilintervalls in $(-\frac{\pi}{2}, \frac{\pi}{2})$ so gibt die Ableitung $a\cos\theta$ (bis auf einen kleinen und wie sich herausstellt vernachlässigbaren Fehler) den Grössenunterschied $\frac{\Delta x}{\Delta\theta}$ an, der bei Betrachtung

von etwaigen Riemann-Summen in der Variable $x$ und der Variable $\theta$ als zusätzlicher Faktor auftreten würde. Wir müssen dies nicht genauer ausführen oder die Substitution auf diese Art und Weise beweisen, da wir ja mittels der Kettenregel und dem Fundamentalsatz der Integral- und Differentialrechnung bereits die Substitutionsregel für Riemann-Integrale bewiesen haben und obiger Formalismus diese nur auf eine andere Art präsentiert. Dieser Beweis über den Fundamentalsatz verwendet allerdings etwas stärkere Annahmen als notwendig (siehe folgende Übung für den direkten Beweis mit schwächeren Annahmen).

**Übung 8.28** (Substitution für Riemann-integrierbare Funktionen)**.** *Sei $[a, b]$ ein kompaktes Intervall in $\mathbb{R}$ mit Endpunkten $a < b$ und $f : [a, b] \to [c, d]$ eine stetig differenzierbare Funktion mit $f'(t) \neq 0$ für alle $t \in [a, b]$. Dann ist für jede Riemann-integrierbare Funktion $g : [c, d] \to \mathbb{R}$ auch $t \in [a, b] \mapsto g \circ f(t)f'(t)$ Riemann-integrierbar und*

$$\int_a^b g \circ f(t)f'(t)\,\mathrm{d}t = \int_{f(a)}^{f(b)} g(x)\,\mathrm{d}x.$$

## 8.3 Anwendungen

### 8.3.1 Flächeninhalte

Wir wollen hier nochmals Beispiele für Flächenberechnungen besprechen, welche unter anderem den Namen der Umkehrfunktionen der hyperbolischen Funktionen erklären.

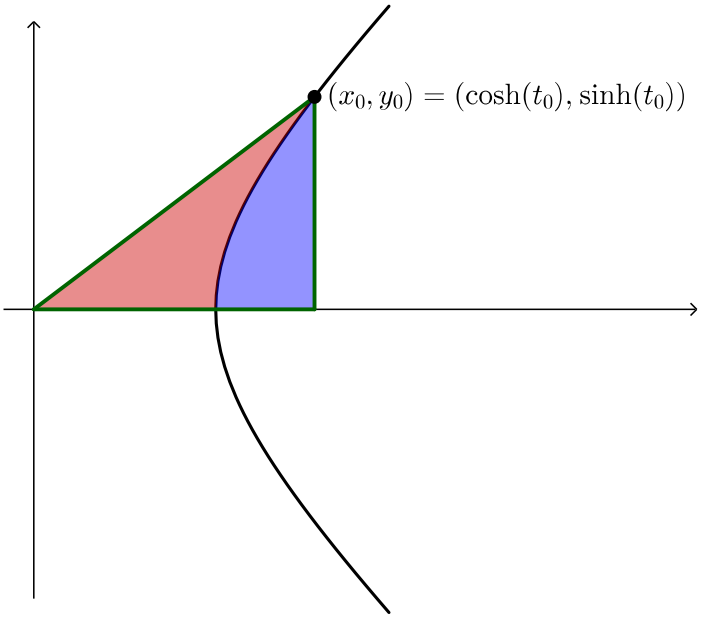
**Beispiel 8.29.** *Wir berechnen den Flächeninhalt des Kreises mit Radius $r > 0$. Dieser ist durch $2\int_{-r}^{r}\sqrt{r^2-x^2}\,\mathrm{d}x$ definiert (wieso?), und gemeinsam mit Bespiel 8.21 ergibt sich daraus*

$$2\int_{-r}^{r}\sqrt{r^2-x^2}\,\mathrm{d}x = 2\left[\tfrac{1}{2}r^2\arcsin\left(\tfrac{x}{r}\right)+\tfrac{1}{2}x\sqrt{r^2-x^2}\right]_{-r}^{r} = r^2\pi$$

**Beispiel 8.30** (Hyperbolische Umkehrfunktionen)**.** *Wir verwenden die Funktion*

$$t\in\mathbb{R}\mapsto(x,y)=(\cosh(t),\sinh(t))\in\mathbb{R}^2 \tag{8.6}$$

*um die „positive Hälfte“ der Hyperbel $\left\{(x,y)\mid x^2-y^2=1\right\}$ zu parametrisieren. Wir stellen uns den Parameter $t\in\mathbb{R}$ vorerst als Zeit vor. In diesem Sinne beschreibt* (8.6) *eine Bewegung im $\mathbb{R}^2$. Wir wollen den Flächeninhalt des folgenden Gebietes in Rosa zwischen dem Ursprung und einem Teil der Hyperbel berechnen.*




*Dieser ist der Flächeninhalt $\frac{1}{2}x_0y_0=\frac{1}{2}\cosh(t_0)\sinh(t_0)$ des eingezeichneten Dreiecks minus dem Flächeninhalt unterhalb der Hyperbel zwischen der 1 und $x_0=\cosh(t_0)$ in Blau. Letztere Fläche ist durch $\int_1^{x_0}\sqrt{x^2-1}\,\mathrm{d}x$ gegeben. Um dieses Integral zu berechnen, verwenden wir die*

*hyperbolische Substitution* $x = \cosh(t)$, $\mathrm{d}x = \sinh(t)\,\mathrm{d}t$ *und erhalten*

$$\int_1^{x_0} \sqrt{x^2-1}\,\mathrm{d}x = \int_0^{t_0} \sinh^2(t)\,\mathrm{d}t = \int_0^{t_0} \frac{\mathrm{e}^{2t} - 2 + \mathrm{e}^{-2t}}{4}\,\mathrm{d}t = \left[\frac{\mathrm{e}^{2t} - \mathrm{e}^{-2t}}{8} - \frac{1}{2}t\right]_0^{t_0}$$
$$= \left[\frac{1}{4}\sinh(2t) - \frac{1}{2}t\right]_0^{t_0} = \frac{1}{2}x_0y_0 - \frac{1}{2}t_0.$$

*Somit ist der Flächeninhalt des gesuchten Gebiets*

$$\frac{1}{2}t_0 = \frac{1}{2}\operatorname{arcosh}(x_0) = \frac{1}{2}\operatorname{arsinh}(y_0).$$

*Dies erklärt die Namen „Areasinus Hyperbolicus“ und „Areakosinus Hyperbolicus“ der Umkehrfunktionen der hyperbolischen Funktionen (wieso?).*

### 8.3.2 Bogenlänge

Im Folgenden möchten wir einen stetige Funktion $\gamma : [a,b] \to \mathbb{R}^d$ für $d \geq 2$, auch **Weg** oder **Kurve** von $\gamma(a)$ nach $\gamma(b)$ genannt, betrachten. Dabei fassen wir $t \in [a,b]$ als Zeitparameter und $\gamma(t)$ als die Position zum Zeitpunkt $t$ auf.

Falls alle Komponenten $\gamma_1, \ldots, \gamma_d$ von $\gamma = (\gamma_1, \ldots, \gamma_d)^t$ stetig differenzierbar sind, interpretieren wir für einen Zeitpunkt $t \in [a,b]$ den Ausdruck $\|\dot\gamma(t)\|_2 = \sqrt{\dot\gamma_1(t)^2 + \ldots + \dot\gamma_d(t)^2}$ als die Geschwindigkeit zum Zeitpunkt $t \in [a,b]$.

Wir möchten nun die Bogenlänge des Weges $\gamma$ als die gesamte Strecke, die zwischen den Zeiten $a$ und $b$ zurückgelegt wurde, definieren. Dabei soll gelten, dass die zurückgelegte Strecke zwischen gleichen Zeiten $\alpha$ und $\alpha$ Null ist und dass sich Strecken additiv verhalten, also dass die zwischen den Zeiten $\alpha < \beta$ zurückgelegte Strecke plus die zwischen den Zeiten $\beta < \gamma$ zurückgelegte Strecke gerade die zwischen den Zeiten $\alpha < \gamma$ zurückgelegte Strecke ist. Im Sinne von Definition 4.29 ist die zurückgelegte Strecke also eine additive Intervallfunktion auf $[a,b]$.

Des Weiteren möchten wir natürlich verlangen, dass die in einem Teilintervall $[\alpha,\beta] \subseteq [a,b]$ mit $\alpha < \beta$ zurückgelegte Strecke zwischen $(\beta - \alpha)$ mal die minimale Geschwindigkeit in $[\alpha,\beta]$ und $(\beta - \alpha)$ mal die maximale Geschwindigkeit in $[\alpha,\beta]$ liegt. Nach Proposition 4.30 ist daher die einzig vernünftige Definition der **Bogenlänge** des Weges $\gamma$ der Ausdruck

$$L(\gamma) = \int_a^b \|\dot\gamma(t)\|_2\,\mathrm{d}t = \int_a^b \sqrt{\dot\gamma_1(t)^2 + \ldots + \dot\gamma_d(t)^2}\,\mathrm{d}t.$$

Anders formuliert ist also die Länge des zurückgelegten Weges das Integral über die Geschwindigkeitsfunktion.

**Beispiel 8.31** (Umfang des Kreises)**.** *Wir betrachten den Weg*

$$\gamma : t \in [0, 2\pi] \mapsto (\cos(t), \sin(t))^t \in \mathbb{R}^2.$$

*Wegen* $\gamma(0) = \gamma(2\pi) = (1,0)^t$ *sind der Start- und der Endpunkt von* $\gamma$ *gleich (wir sagen auch, dass der Weg* $\gamma$ ***geschlossen*** *ist). Auch gilt für die Geschwindigkeit zu jedem Zeitpunkt*

$t \in [0, 2\pi]$

$$\|\dot{\gamma}(t)\|_2 = \sqrt{\dot{\gamma}_1(t)^2 + \dot{\gamma}_2(t)^2} = \sqrt{\sin^2(t) + \cos^2(t)} = 1.$$

*Der Weg (oder die Kurve) $\gamma$ durchläuft (wegen $\cos^2(t) + \sin^2(t) = 1$ für alle $t \in \mathbb{R}$) den Einheitskreis also mit konstanter Geschwindigkeit Eins. Deswegen gilt*

$$L(\gamma) = \int_0^{2\pi} \sqrt{\dot{\gamma}_1(t)^2 + \dot{\gamma}_2(t)^2}\,\mathrm{d}t = \int_0^{2\pi} 1\,\mathrm{d}t = 2\pi.$$

*Des Weiteren besucht $\gamma$ jeden Punkt (bis auf den Endpunkt) genau einmal (siehe auch Abschnitt 6.6.4). Einen solchen Weg nennen wir auch **einfach**. Deswegen lässt sich die Bogenlänge von $\gamma$ auch als den Umfang des Einheitskreises auffassen, der somit $2\pi$ ist. Dies gilt analog für Teilstrecken und definiert den Begriff **Winkel** als Bogenlänge am Einheitskreis.*

Sei $\gamma : [a, b] \to \mathbb{R}^d$ ein stetig differenzierbarer Weg ausgehend von einem Intervall $[a, b]$ mit Endpunkten $a < b$. Eine stetig differenzierbare **Reparametrisierung** von $\gamma$ ist ein Weg der Form $\gamma \circ \psi : [\tilde{a}, \tilde{b}] \to \mathbb{R}^d$, wobei $[\tilde{a}, \tilde{b}]$ ein kompaktes Intervall mit Endpunkten $\tilde{a} < \tilde{b}$ ist und $\psi : [\tilde{a}, \tilde{b}] \to [a, b]$ eine stetig differenzierbare, monotone wachsende, bijektive Funktion ist. Wenn wir uns $\gamma$ als einen „Fahrplan eines Autobusses“ vorstellen, dann entspricht $\psi$ einer „Fahrplanänderung“.

Intuitiv ausgedrückt ist eine Reparametrisierung eines Weges also ein Weg mit denselben Endpunkten (da $\psi(\tilde{a}) = a$ und $\psi(\tilde{b}) = b$) und der immer in dieselbe Richtung geht (wegen Monotonie). Anschaulich kann man deswegen erwarten, dass jede Reparametrisierung eines Weges dieselbe Bogenlänge hat. Auch wollen wir zeigen, dass ein nie anhaltender Weg so reparametrisiert werden kann, dass der neue Weg Einheitsgeschwindigkeit hat. Falls $\gamma : [a, b] \to \mathbb{R}^d$ nie anhält oder genauer falls $\|\dot{\gamma}(t)\|_2 > 0$ für alle $t \in [a, b]$, so nennen wir $\gamma$ **regulär**.

**Lemma 8.32** (Reparametrisierungen eines Weges)**.** *Sei $\gamma : [a, b] \to \mathbb{R}^d$ ein stetig differenzierbarer Weg für $a < b$. Dann hat jede Reparametrisierung von $\gamma$ dieselbe Bogenlänge. Falls $\gamma$ regulär ist, gibt es eine Reparametrisierung von $\gamma$ mit Einheitsgeschwindigkeit, welche auch die **Parametrisierung nach Bogenlänge** genannt wird.*

In Beispiel 8.31 ist der betrachtete Weg bereits nach Bogenlänge parametrisiert.

*Beweis.* Sei $[\tilde{a}, \tilde{b}]$ ein kompaktes Intervall mit Endpunkten $\tilde{a} < \tilde{b}$ und $\psi : [\tilde{a}, \tilde{b}] \to [a, b]$ eine stetig differenzierbare, monotone wachsende, bijektive Funktion. Dann gilt

$$\begin{aligned}
L(\gamma \circ \psi) &= \int_{\tilde{a}}^{\tilde{b}} \sqrt{(\gamma_1 \circ \psi)'(s)^2 + \ldots + (\gamma_d \circ \psi)'(s)^2}\,\mathrm{d}s \\
&= \int_{\tilde{a}}^{\tilde{b}} \sqrt{\dot{\gamma}_1(\psi(s))^2 \psi'(s)^2 + \ldots + \dot{\gamma}_d(\psi(s))^2 \psi'(s)^2}\,\mathrm{d}s \\
&= \int_{\tilde{a}}^{\tilde{b}} \sqrt{\dot{\gamma}_1(\psi(s))^2 + \ldots + \dot{\gamma}_d(\psi(s))^2}\,\psi'(s)\,\mathrm{d}s \\
&= \int_a^b \sqrt{\dot{\gamma}_1(t)^2 + \ldots + \dot{\gamma}_d(t)^2}\,\mathrm{d}t = L(\gamma).
\end{aligned}$$

Für die zweite Aussage konstruieren wir nun eine geeignete Funktion $\psi$ wie oben. Sei

$$\phi : [a,b] \to [0, L(\gamma)],\ t \mapsto \int_a^t \sqrt{\dot\gamma_1(s)^2 + \ldots + \dot\gamma_d(s)^2}\,\mathrm{d}s.$$

Wegen $\dot\phi(t) = \sqrt{\dot\gamma_1(t)^2 + \ldots + \dot\gamma_d(t)^2} > 0$ für alle $t \in [a,b]$ sowie $\phi(a) = 0$ und $\phi(b) = L(\gamma)$ ist $\phi : [a,b] \to [0, L(\gamma)]$ eine streng monoton wachsende, stetig differenzierbare Bijektion. Insbesondere ist $\psi = \phi^{-1} : [0, L(\gamma)] \to [a,b]$ ebenfalls streng monoton wachsend und stetig differenzierbar. Zur Zeit $s \in [0, L(\gamma)]$ berechnen wir nun die Geschwindigkeit von $\gamma \circ \psi$. Ist $t = \psi(s)$, so gilt wegen $\psi'(s) > 0$ auch

$$\begin{aligned}\sqrt{(\gamma_1 \circ \psi)'(s)^2 + \ldots + (\gamma_d \circ \psi)'(s)^2} &= \sqrt{\dot\gamma_1(t)^2\psi'(s)^2 + \ldots + \dot\gamma_d(t)^2\psi'(s)^2} \\ &= \psi'(s)\sqrt{\dot\gamma_1(t)^2 + \ldots + \dot\gamma_d(t)^2} \\ &= \frac{1}{\dot\phi(t)}\sqrt{\dot\gamma_1(t)^2 + \ldots + \dot\gamma_d(t)^2} = 1.\end{aligned}$$

Somit hat die Reparametrisierung $\gamma \circ \psi$ die gewünschte Eigenschaft. □

**Übung 8.33** (Eindeutigkeit der Parametrisierung)**.** *In Lemma 8.32 wird bereits von* der *Parametrisierung nach Bogenlänge gesprochen. Wir wollen dies hier begründen. Sei $\gamma : [a,b] \to \mathbb{R}^d$ ein stetig differenzierbarer, regulärer Weg. Nach Lemma 8.32 dürfen wir annehmen, dass $\gamma$ Einheitsgeschwindigkeit hat. Zeigen Sie, dass es keine weitere Reparametrisierung von $\gamma$ mit Einheitsgeschwindigkeit gibt.*

**Übung 8.34** (Totale Variation des Weges)**.** *In dieser Übung wollen wir noch eine weitere Begründung für die Definition der Bogenlänge eines Weges $\gamma : [a,b] \to \mathbb{R}^d$ geben. Hierfür interpretieren wir $d(v,w) = \|v - w\|_2$ als den Abstand zweier Punkte $v, w \in \mathbb{R}^d$. Die* ***totale Variation*** *von $\gamma : [a,b] \to \mathbb{R}^d$ ist definiert als*

$$V(\gamma) = \sup_{\mathfrak{Z}} \sum_{k=1}^n \|\gamma(x_k) - \gamma(x_{k-1})\|,$$

*wobei das Supremum über alle Zerlegungen $\mathfrak{Z} = \{x_0 = a < x_1 < \cdots < x_n = b\}$ von $[a,b]$ genommen wird. Nehmen Sie nun an, dass $\gamma$ stetig differenzierbar ist und zeigen Sie $V(\gamma) = L(\gamma)$.*

Für einen Weg $\gamma : [a,b] \to \mathbb{R}^d$ und eine stetige Funktion $f : \mathbb{R}^d \to \mathbb{R}$ kann ein Integral der Form

$$\int_a^b f(\gamma(t))\|\dot\gamma(t)\|_2\,\mathrm{d}t$$

auch physikalische Bedeutung haben. Zum Beispiel kann der Weg einen verbogenen Draht (mit konstanter Dichte $1kg/m$) beschreiben. In diesem Fall gibt

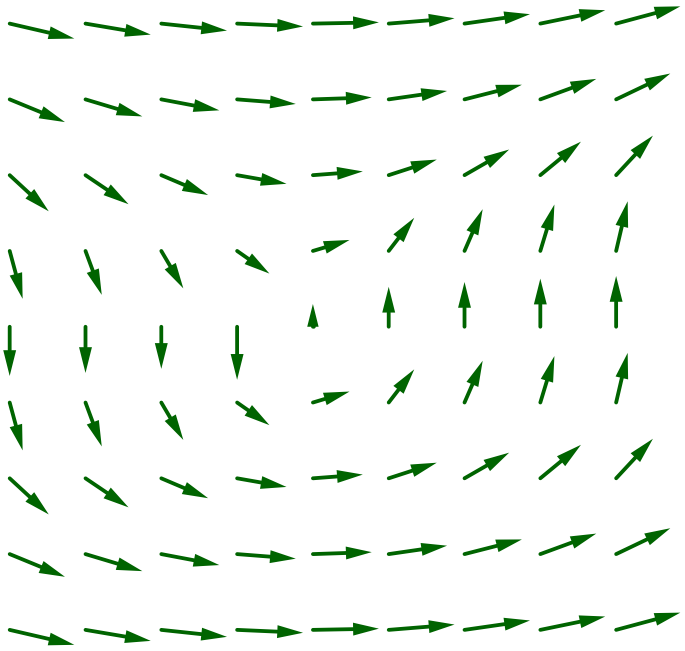
$$\frac{1}{L(\gamma)}\int_a^b \gamma_j(t)\|\dot\gamma(t)\|_2\,\mathrm{d}t$$

die $j$-te Koordinate des Schwerpunktes des Drahtes an, wobei $j \in \{1, \ldots, d\}$.

### 8.3.3 Wegintegrale von Vektorfeldern

Wir kommen nun zu einem weiteren Typ von Wegintegralen, der sowohl für die Physik als auch für die weitere Analysis wichtig sein wird. Hierfür betrachten wir nochmals reelle Zahlen $a < b$ und einen stetig differenzierbaren Weg $\gamma : [a, b] \to \mathbb{R}^d$. Wir interpretierten $\|\dot{\gamma}(t)\|_2$ ja bereits als Geschwindigkeit (in $m/s$) des Weges zum Zeitpunkt $t \in [a, b]$ (in $s$) und wollen analog dazu die Ableitung $\dot{\gamma}(t)$ als den Geschwindigkeitsvektor zum Zeitpunkt $t$ interpretieren (mit jeder Koordinate in $m/s$), der eben nicht nur die augenblickliche Geschwindigkeit als eindimensionale Grösse angibt, sondern auch die Richtung der Bewegung beschreibt.

Sei $\mathbf{f} : \mathbb{R}^d \to \mathbb{R}^d$ eine stetige Funktion (siehe Definition 5.96), welche wir als ein **Kraftfeld** interpretieren und bei jedem Punkt $\mathbf{v} \in \mathbb{R}^d$ die Richtung und Stärke einer Krafteinwirkung zum Beispiel auf Grund von Wind angibt (mit jeder Koordinate in $N$). Wir nennen in diesem Zusammenhang $\mathbf{f}$ auch ein **Vektorfeld** und visualisieren für $d = 2$ (und etwas schwieriger auch für $d = 3$) dieses durch eine Ansammlung von Vektoren bei mehreren Punkten im Definitionsbereich, siehe folgendes Bild.



Das innere Produkt $\langle \mathbf{f}(\gamma(t)), \dot{\gamma}(t) \rangle$ gibt damit die Leistung (in $W = Nm/s$) an, die bei Bewegung mit vorgeschriebener Geschwindigkeit von der Krafteinwirkung zum Zeitpunkt $t$ geleistet wird. Hierbei kann es vorkommen, dass Krafteinwirkung und Geschwindigkeit ähnliche Richtungen haben und das innere Produkt positiv ist. Ebenso kann es aber vorkommen, dass Krafteinwirkung und Geschwindigkeit entgegengesetzt sind und das innere Produkt negativ ist. In diesem Sinne (siehe auch Abschnitt 4.4.4) berechnet das sogenannte **Wegintegral**

$$\int_\gamma \mathbf{f} \cdot \mathrm{d}\mathbf{s} = \int_a^b \langle \mathbf{f}(\gamma(t)), \dot{\gamma}(t) \rangle \, \mathrm{d}t$$

die Arbeit, die von der Krafteinwirkung insgesamt geleistet wurde.

Wir werden im zweiten Semester derartige Integrale nochmals genauer untersuchen und dann zum Beispiel folgende Frage beantworten können: Wie kann man einem Kraftfeld $\mathbf{f}$ ansehen, ob das Wegintegral nur von Anfangspunkt $\gamma(a)$ und Endpunkt $\gamma(b)$ abhängt und nicht von der Wahl des konkreten Weges von $\gamma(a)$ nach $\gamma(b)$?

**Beispiel 8.35** (Abhängigkeit von der Wahl des Weges)**.** *Sei* $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2$ *definiert durch* $\mathbf{f}(x,y) = \binom{y}{x^2}$. *Wir betrachten den Weg* $\gamma : [0,1] \to \mathbb{R}^2$ *definiert durch* $\gamma(t) = \binom{t}{t^2}$ *für* $t \in [0,1]$. *Dann ist das Wegintegral von* $\mathbf{f}$ *über den Weg* $\gamma$ *von* $\gamma(0) = \binom{0}{0}$ *nach* $\gamma(1) = \binom{1}{1}$ *durch*

$$\int_0^1 \left\langle \begin{pmatrix} t^2 \\ t^2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2t \end{pmatrix} \right\rangle \mathrm{d}t = \int_0^1 (t^2 + 2t^3)\,\mathrm{d}t = \frac{1}{3} + \frac{2}{4} = \frac{5}{6}$$

*gegeben. Verwenden wir allerdings den Weg* $\eta : [0,1] \to \mathbb{R}^2$ *definert durch* $\eta(t) = \binom{t^2}{t}$ *für* $t \in [0,1]$, *so sind zwar Anfangs- und Endpunkte unverändert, doch ist das Wegintegral durch*

$$\int_0^1 \left\langle \begin{pmatrix} t \\ t^4 \end{pmatrix}, \begin{pmatrix} 2t \\ 1 \end{pmatrix} \right\rangle \mathrm{d}t = \int_0^1 (2t^2 + t^4)\,\mathrm{d}t = \frac{2}{3} + \frac{1}{5} = \frac{13}{15}$$

*gegeben.*

**Applet 8.36** (Wegintegral)**.** *Wir stellen sowohl das Vektorfeld* $\mathbf{f}$, *einen verschiebbaren Weg* $\gamma$ *mit animiertem Punkt* $\gamma(t)$, *die Ableitung* $\gamma'(t)$ *und darunter den Graph der Funktion* $t \in [0,1] \mapsto \langle \mathbf{f}(\gamma(t)), \gamma'(t) \rangle$ *dar.*

### 8.3.4 Volumen von Rotationskörpern

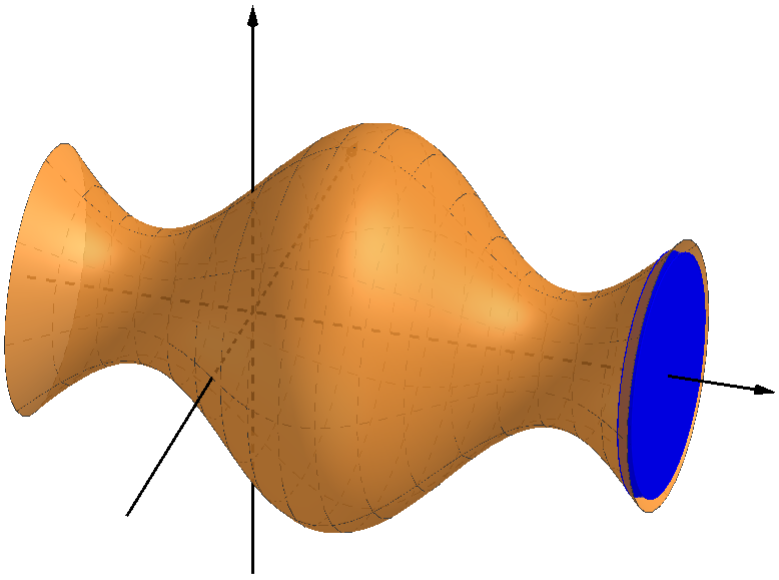
Sei $[a,b] \subseteq \mathbb{R}$ ein kompaktes Intervall mit Endpunkten $a < b$ und $f : [a,b] \to \mathbb{R}_{\geq 0}$ stetig. Wir betrachten das Gebiet

$$G = \left\{ (x,y) \in \mathbb{R}^2 \mid a \leq x \leq b,\ 0 \leq y \leq f(x) \right\}$$

und den zugehörigen Körper

$$K = \left\{ (x,y,z) \in \mathbb{R}^3 \mid a \leq x \leq b,\ 0 \leq \sqrt{y^2 + z^2} \leq f(x) \right\},$$

der sich aus Rotation von $G$ um die $x$-Achse ergibt. Sind die beiden Zylinder $Z_1, Z_2$ mit Radius $\min_{x\in[a,b]} f(x)$ respektive $\max_{x\in[a,b]} f(x)$ um die $x$-Achse gegeben, so will man wegen den Enthaltungen $Z_1 \subseteq K \subseteq Z_2$, dass das Volumen von $K$ zwischen $\pi \left(\min_{x\in[a,b]} f(x)\right)^2 (b-a)$ und $\pi \left(\max_{x\in[a,b]} f(x)\right)^2 (b-a)$ liegt. Wir halten dies in folgendem Bild fest, wo gemeinsam mit dem Rotationskörper eine von vielen „Scheiben“, die zusammen den Körper approximieren, dargestellt werden.



Deswegen (siehe auch Übung 8.38) definieren wir das **Volumen des Rotationskörpers** $K$ durch

$$\pi \int_a^b f(x)^2 \, \mathrm{d}x. \tag{8.7}$$

**Beispiel 8.37** (Volumen der Kugel)**.** *Sei* $K = \left\{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 \leq r^2\right\}$ *für* $r > 0$ *die Kugel mit Radius* $r$*. Die Kugel* $K$ *lässt sich auch als Rotationskörper mittels der Funktion* $f : x \in [-r, r] \mapsto \sqrt{r^2 - x^2}$ *auffassen. Ihr Volumen ist deswegen durch*

$$\begin{aligned}\pi \int_{-r}^{r} \left(\sqrt{r^2 - x^2}\right)^2 \mathrm{d}x = \pi \int_{-r}^{r} (r^2 - x^2) \, \mathrm{d}x = \pi \left[r^2 x - \tfrac{x^3}{3}\right]_{-r}^{r} &= \pi \left(r^3 - \tfrac{r^3}{3} + r^3 - \tfrac{r^3}{3}\right) \\ &= \tfrac{4\pi}{3} r^3\end{aligned}$$

*gegeben.*

**Übung 8.38.** *Motivieren Sie die Definition des Volumen eines Rotationskörpers mit mehr Details in Analogie zu Abschnitt 8.3.2 unter Verwendung von Proposition 4.30.*

### 8.3.5 Oberflächen von Rotationskörpern

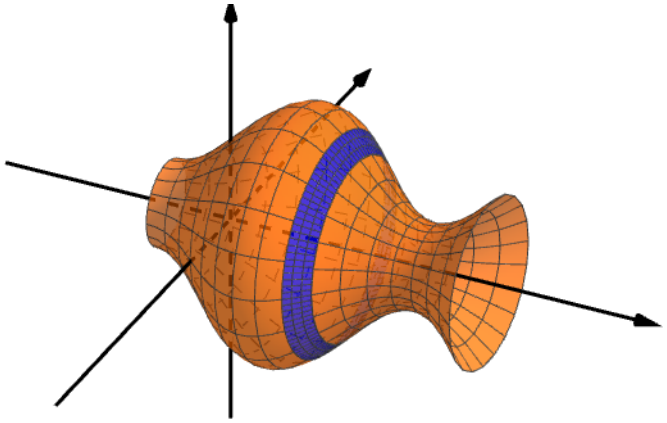
Obwohl Proposition 4.30 oft ein guter Wegweiser für das Auffinden einer geeigneten Definition darstellt, müssen oder können wir diese nicht immer als Grundlage wählen. Manchmal begnügen wir uns mit geometrischer Intuition als Motivation der Definition.[1]

[1] Man kann die Sinnhaftigkeit einer Definition zwar hinterfragen, doch kann man eine Definition ohnehin nicht beweisen.

Wir betrachten $a < b$ in $\mathbb{R}$, eine stetig differenzierbare Funktion $f : [a, b] \to \mathbb{R}$ und den Rotationskörper

$$K = \left\{(x, y, z) \in \mathbb{R}^3 \mid a \leq x \leq b,\ 0 \leq \sqrt{y^2 + z^2} \leq f(x)\right\}$$

wie im letzten Abschnitt.



In einem kleinen Teilintervall $[x_{k-1}, x_k] \subseteq [a, b]$ der Länge $\triangle x_k$ ist die Funktion $y = f(x)$ der Tangente $y - f(\xi) = f'(\xi)(x - \xi)$ für ein $\xi \in [x_{k-1}, x_k]$ sehr nahe. Ausserdem wird die Oberfläche des Anteils von $K$, der dem Intervall $[x_{k-1}, x_k]$ entspricht, sehr gut durch die Aussenoberfläche des Kegelstumpfs beschrieben, der entsteht, wenn man obiges Tangentenstück zwischen $x_{k-1}$ und $x_k$ um die $x$-Achse rotiert. Die Aussenoberfläche eines Kegelstumpfs ist näherungsweise $\ell \cdot U$, wobei $\ell$ die Länge der Aussenkante des Kegelstumpfs und $U$ der Umfang einer der beiden Kreise darstellt (wieso?).

Die Oberfläche sollte also näherungsweise durch

$$\begin{aligned}\sum_{k=1}^{n} \sqrt{(\triangle x_k)^2 + (\triangle y_k)^2} \cdot 2\pi f(x_k) &= 2\pi \sum_{k=1}^{n} \sqrt{1 + \left(\frac{\triangle y_k}{\triangle x_k}\right)^2} f(x_k) \triangle x_k \\ &= 2\pi \sum_{k=1}^{n} \sqrt{1 + f'(\xi_k)^2} f(x_k)(x_k - x_{k-1})\end{aligned}$$

gegeben sein, wobei $\xi_k \in [x_{k-1}, x_k]$ für jedes $k \in \{1, \ldots, n\}$ einen Zwischenpunkt darstellt. Deswegen definieren wir nun die **Oberfläche des Rotationskörpers** $K$ als

$$2\pi \int_a^b \sqrt{1 + f'(x)^2}\ f(x)\, \mathrm{d}x.$$

**Beispiel 8.39** (Kugeloberfläche)**.** *Wie in Beispiel 8.37 betrachten wir zu $r > 0$ die Funktion $x \in [-r, r] \mapsto \sqrt{r^2 - x^2}$, deren Rotationskörper gerade die Kugel von Radius $r$ ist. Für alle $x \in [-r, r]$ ist*

$$f'(x) = -\frac{x}{\sqrt{r^2 - x^2}}.$$

*Damit ist die Kugeloberfläche gleich*

$$2\pi \int_{-r}^{r} \sqrt{1 + \frac{x^2}{r^2 - x^2}} \sqrt{r^2 - x^2}\, \mathrm{d}x = 2\pi \int_{-r}^{r} \sqrt{r^2}\, \mathrm{d}x = 2\pi r [x]_{-r}^{r} = 4\pi r^2.$$

VO: 18.12.17

## 8.4 Das uneigentliche Integral

Wir wollen nun den Begriff des Riemann-Integrals auf mehrere Arten erweitern.

### 8.4.1 Uneigentliche Integrationsgrenzen

Für $a \in \mathbb{R}$ und eine komplexwertige Funktion $f : [a, \infty) \to \mathbb{C}$ mit $f|_{[a,b]} \in R([a, b])$ für alle $b > a$ definieren wir das **uneigentliche Integral**

$$\int_a^\infty f(x)\,\mathrm{d}x = \lim_{b\to\infty} \int_a^b f(x)\,\mathrm{d}x,$$

falls der Grenzwert existiert. Weiter sagen wir, dass das uneigentliche Integral **konvergiert**, falls der obige Grenzwert in $\mathbb{C}$ existiert. Ansonsten nennen wir das uneigentliche Integral $\int_a^\infty f(x)\,\mathrm{d}x$ **divergent**.

**Beispiel 8.40.** *Es gilt*

$$\int_0^\infty \frac{1}{1+x^2}\,\mathrm{d}x = \lim_{b\to\infty} \int_0^b \frac{1}{1+x^2}\,\mathrm{d}x = \lim_{b\to\infty} \arctan(b) = \frac{\pi}{2}.$$

**Beispiel 8.41.** *Es gilt für $\alpha \in \mathbb{R}$*

$$\int_1^\infty x^{-\alpha}\,\mathrm{d}x = \begin{cases} \frac{1}{\alpha-1} & \textit{falls } \alpha > 1 \\ +\infty & \textit{falls } \alpha \leq 1. \end{cases}$$

*Insbesondere ist das obige uneigentliche Integral genau dann konvergent, wenn $\alpha > 1$.*

*In der Tat ist*

$$\int_1^b x^{-\alpha}\,\mathrm{d}x = \begin{cases} \left[\frac{1}{-\alpha+1}x^{-\alpha+1}\right]_1^b = \frac{1}{-\alpha+1}b^{-\alpha+1} - \frac{1}{-\alpha+1} & \textit{falls } \alpha \neq 1 \\ [\log(x)]_1^b = \log(b) & \textit{falls } \alpha = 1 \end{cases}$$

*und*

$$\lim_{b\to\infty} \frac{1}{-\alpha+1} b^{-\alpha+1} = \begin{cases} +\infty & \textit{falls } \alpha < 1 \\ 0 & \textit{falls } \alpha > 1 \end{cases}$$

$$\lim_{b\to\infty} \log(b) = +\infty.$$

**Übung 8.42.** *Berechnen Sie $\int_1^\infty x^{-\alpha}\,\mathrm{d}x$ auch für $\alpha \in \mathbb{C}$.*

Uneigentliche Integrale der Form $\int_{-\infty}^b f(x)\,\mathrm{d}x$ sind ähnlich definiert. Ebenso definieren wir für eine komplexwertige Funktion $f : \mathbb{R} \to \mathbb{C}$ mit $f|_{[a,b]} \in R([a, b])$ für alle $a < b$ das uneigentliche Integral

$$\int_{-\infty}^\infty f(x)\,\mathrm{d}x = \int_{-\infty}^0 f(x)\,\mathrm{d}x + \int_0^\infty f(x)\,\mathrm{d}x = \lim_{a\to-\infty} \int_a^0 f(x)\,\mathrm{d}x + \lim_{b\to\infty} \int_0^b f(x)\,\mathrm{d}x,$$

falls beide Grenzwerte existieren. Wir möchten dazu anmerken, dass man sich bewusst dazu entscheidet, die Bewegungen gegen $-\infty$ respektive $+\infty$ komplett getrennt zu behandeln. Alles andere würde zu komischen Phänomenen führen, wie folgendes Beispiel zeigt.

**Beispiel 8.43.** *Das uneigentliche Integral* $\int_{-\infty}^{\infty} x \,\mathrm{d}x$ *existiert nicht, da* $\int_{-\infty}^{0} x \,\mathrm{d}x$ *sowie* $\int_{0}^{\infty} x \,\mathrm{d}x$ *nicht existieren. Wir bemerken aber, dass der Grenzwert* $\lim_{c\to\infty} \int_{-c}^{c} x \,\mathrm{d}x = \lim_{c\to\infty} 0 = 0$ *existieren würde aber* $\lim_{c\to\infty} \int_{-c}^{c+1} x \,\mathrm{d}x = \lim_{c\to\infty} \frac{1}{2}((c+1)^2 - c^2) = \infty$ *wäre.*

Wie wir nun besprechen wollen, haben uneigentlichen Integrale oft sehr enge Beziehungen zu Reihen. Genau wie bei Folgen und Reihen (siehe Satz 5.34 und Proposition 6.11) ist es bei uneigentlichen Integralen nicht-negativer Funktionen einfacher über Konvergenz zu entscheiden.

**Lemma 8.44.** *Sei* $a \in \mathbb{R}$ *und* $f : [a, \infty) \to \mathbb{R}_{\geq 0}$ *eine nicht-negative Funktion mit* $f \in R([a, b])$ *für alle* $b > a$*. Entweder konvergiert das uneigentliche Integral über* $f$ *oder es divergiert gegen Unendlich. In beiden Fällen gilt*

$$\int_a^\infty f(x) \,\mathrm{d}x = \sup \left\{ \int_a^b f(x) \,\mathrm{d}x \mid b > a \right\}.$$

*Beweis.* Die Funktion $b \in [a, \infty) \mapsto \int_a^b f(x) \,\mathrm{d}x$ ist monoton wachsend. Wenn das Supremum $S = \sup \left\{ \int_a^b f(x) \,\mathrm{d}x \mid b > a \right\}$ Unendlich ist, dann divergiert das uneigentliche Integral auf Grund der Monotonie gegen Unendlich. Wenn $S < \infty$ ist, dann gibt es zu $\varepsilon > 0$ ein $B > a$ mit

$$S - \varepsilon \leq \int_a^B f(x) \,\mathrm{d}x \leq S$$

Insbesondere gilt für $b > B$ auf Grund der Monotonie und der Definition von $S$ dieselbe Ungleichung auch für $\int_a^b f(x) \,\mathrm{d}x$. Dies beweist die Konvergenz des uneigentlichen Integrals. □

**Beispiel 8.45** (Gaussche Glockenkurve)**.** *Wir wollen das uneigentliche Integral*

$$\int_{-\infty}^{\infty} \mathrm{e}^{-x^2} \,\mathrm{d}x = \int_{-\infty}^{-1} \mathrm{e}^{-x^2} \,\mathrm{d}x + \int_{-1}^{1} \mathrm{e}^{-x^2} \,\mathrm{d}x + \int_{1}^{\infty} \mathrm{e}^{-x^2} \,\mathrm{d}x$$

*besprechen, wobei die Funktion* $x \in \mathbb{R} \mapsto \mathrm{e}^{-x^2}$ *die* Gaussche Glockenkurve *genannt wird. Auf Grund von Lemma 8.44 reicht es aus eine „Majorantenfunktion“ zu finden, die ein konvergentes uneigentliches Integral definiert. Für* $x \in [1, \infty)$ *gilt zum Beispiel* $x^2 \geq x$ *und daher* $\mathrm{e}^{-x^2} \leq \mathrm{e}^{-x}$*, woraus*

$$\int_1^\infty \mathrm{e}^{-x^2} \,\mathrm{d}x \leq \int_1^\infty \mathrm{e}^{-x} \,\mathrm{d}x < \infty$$

*folgt. Dies zeigt die Konvergenz des zweiten uneigentlichen Integrals, auf Grund der Symmetrie der Funktion ist daher auch* $\int_{-\infty}^{\infty} \mathrm{e}^{-x^2} \,\mathrm{d}x$ *konvergent. Wir werden den Wert* $I$ *dieses Integral erst im zweiten Semester berechnen können. Doch wollen wir noch erwähnen, dass die streng*

*monoton wachsende Funktion*

$$\Phi : x \in \mathbb{R} \mapsto I^{-1} \int_{-\infty}^{x} \mathrm{e}^{-t^2} \, \mathrm{d}t$$

*die Verteilungsfunktion der Normalverteilung mit Erwartungswert* 0 *und Standardabweichung* $\frac{1}{\sqrt{2}}$ *genannt wird. Diese Funktion lässt sich ebenso wie die Funktionen aus Abschnitt 8.2.6 nicht durch die sonst üblichen Funktionen ausdrücken und ist in der Wahrscheinlichkeitsrechnung, der Statstik und in vielen Anwendungen von fundamentaler Bedeutung.*

Der folgende Satz charakterisiert nun Konvergenz uneigentlicher Integrale wie in obigem Lemma durch Konvergenz von Reihen (auf hinreichende und notwendige Weise).

**Satz 8.46** (Integraltest für Reihen)**.** *Sei* $f : [1, \infty) \to \mathbb{R}_{\geq 0}$ *eine monoton fallende Funktion. Dann gilt*

$$\sum_{n=2}^{\infty} f(n) \leq \int_{1}^{\infty} f(x) \, \mathrm{d}x \leq \sum_{n=1}^{\infty} f(n).$$

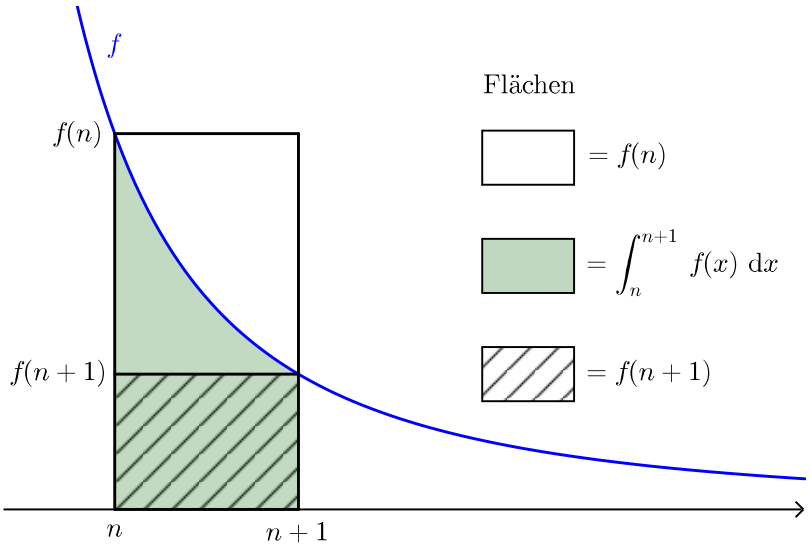
*Insbesondere konvergiert die Reihe* $\sum_{n=1}^{\infty} f(n)$ *genau dann, wenn das uneigentliche Integral* $\int_1^\infty f(x) \, \mathrm{d}x$ *konvergiert. Dies gilt analog für Integrale der Form* $\int_N^\infty f(x) \, \mathrm{d}x$ *für* $N \in \mathbb{N}$.

Wir bemerken, dass auf Grund der Monotonieannahme an $f$ in obigem Satz die Eigenschaft $f|_{[1,b]} \in R([1, b])$ für alle $b > 1$ erfüllt ist nach Satz 4.31.

*Beweis.* Für $n \in \mathbb{N}$ und einen beliebigen Zwischenpunkt $x \in [n, n+1]$ gilt nach Monotonie von $f$ die Ungleichung $f(n+1) \leq f(x) \leq f(n)$ und somit

$$f(n+1) \leq \int_{n}^{n+1} f(x) \, \mathrm{d}x \leq f(n),$$

was auch in folgendem Bild ersichtlich ist.



Nach Summation von 1 bis $n$ erhält man mit Intervalladditivität des Riemann-Integrals

$$\sum_{\ell=2}^{n+1} f(\ell) = \sum_{k=1}^{n} f(k+1) \leq \int_1^{n+1} f(x)\,\mathrm{d}x \leq \sum_{k=1}^{n} f(k).$$

Falls das uneigentliche Integral $\int_1^\infty f(x)\,\mathrm{d}x$ existiert, dann folgt

$$\sum_{\ell=2}^{n+1} f(\ell) \leq \int_1^{n+1} f(x)\,\mathrm{d}x \leq \int_1^\infty f(x)\,\mathrm{d}x.$$

Daher ist die monoton wachsende Folge $\left(\sum_{\ell=2}^{n+1} f(\ell)\right)_n$ nach oben beschränkt und konvergiert somit nach Satz 5.34. Insbesondere gilt auch $\sum_{\ell=2}^{\infty} f(\ell) \leq \int_1^\infty f(x)\,\mathrm{d}x$.

Falls $\sum_{k=1}^{\infty} f(k)$ konvergiert, dann ist für $b > 1$ und $n = \lfloor b \rfloor$

$$\int_1^b f(x)\,\mathrm{d}x \leq \int_1^{n+1} f(x)\,\mathrm{d}x \leq \sum_{k=1}^{\infty} f(k).$$

Nach Lemma 8.44 ist somit das uneigentliche Integral $\int_1^\infty f(x)\,\mathrm{d}x$ konvergent und durch die Zahl $\sum_{k=1}^{\infty} f(k)$ beschränkt. □

**Übung 8.47** (Divergenzrate der harmonischen Reihe)**.** *Verwenden Sie obigen Satz, um den p-Test in Beispiel 6.17 zu erhalten. Imitieren Sie des Weiteren die Methodik im obigen Beweis von Satz 8.46, um die Divergenzrate*

$$\sum_{n=1}^{N} \frac{1}{n} = \log(N) + O(1)$$

*für $N \to \infty$ für die harmonische Reihe zu beweisen.*

**Übung 8.48** (Eine lange Nadel)**.** *Berechnen Sie das Volumen und die Oberfläche des „uneigentlichen Rotationskörpers“, der entsteht, wenn man das Gebiet unter dem Graphen der Funktion $x \in [1, \infty) \mapsto \frac{1}{x}$ um die $x$-Achse rotiert.*

**Übung 8.49** (Ein oszillierendes Integral)**.** *Entscheiden Sie für welche $p \in \mathbb{R}_{\geq 0}$ das uneigentliche Integral $\int_0^\infty x \sin(x^p)\,\mathrm{d}x$ konvergiert.*

### 8.4.2 Das Integral über unbeschränkte Funktionen

Für $a < B$ in $\mathbb{R}$ und eine Funktion $f : [a, B) \to \mathbb{C}$ mit $f|_{[a,b]} \in R([a, b])$ für alle $b \in (a, B)$ definieren wir das **uneigentliche Integral**

$$\int_a^B f(x)\,\mathrm{d}x = \lim_{b \nearrow B} \int_a^b f(x)\,\mathrm{d}x,$$

falls der Grenzwert existiert.

Wie folgende Übung zeigt, steht diese Notation nicht im Widerspruch zum Riemann-Integral.

**Übung 8.50** (Kompatibilität)**.** *Sei $f : [a, B) \to \mathbb{C}$ wie oben. Angenommen $f$ ist beschränkt. Zeigen Sie, dass das uneigentliche Integral $\int_a^B f(x)\,\mathrm{d}x$ existiert und gleich dem Riemann-Integral $\int_a^B f(x)\,\mathrm{d}x$ ist, wobei man $f$ auf beliebige Weise auf den Punkt $B$ erweitert.*

Für $A < b$ in $\mathbb{R}$ und $f : (A, b] \to \mathbb{C}$ eine Funktion mit $f|_{[a,b]} \in R([a, b])$ für alle $a \in (A, b]$ definieren wir analog das uneigentliche Integral

$$\int_A^b f(x)\,\mathrm{d}x = \lim_{a \searrow A} \int_a^b f(x)\,\mathrm{d}x.$$

**Beispiel 8.51.** *Wir berechnen $\int_0^1 \log(x)\,\mathrm{d}x$ mittels*

$$\begin{aligned}\int_0^1 \log(x)\,\mathrm{d}x = \lim_{a \searrow 0} \int_a^1 \log(x)\,\mathrm{d}x &= \lim_{a \searrow 0} [x \log(x) - x]_a^1 \\ &= \lim_{a \searrow 0} (\log(1) - 1 - a \log(a) + a) = -1\end{aligned}$$

*nach Beispiel 8.15 und Beispiel 5.76.*

Weitere uneigentliche Integrale führen wir mittels Intervalladditivität auf obige uneigentliche Integrale zurück. Wir überlassen es Interessierten, sich hier einige Möglichkeiten auszudenken, und führen stattdessen ein Beispiel vor.

**Beispiel 8.52.** *Wir betrachten das Integral $\int_{-1}^1 \frac{1}{x}\,\mathrm{d}x$. Dieses ist uneigentlich, da $x \mapsto \frac{1}{x}$ auf jeder Umgebung von $0$ unbeschränkt ist. Es gilt daher auf Grund der Definition in diesem Fall, dass*

$$\int_{-1}^1 \frac{1}{x}\,\mathrm{d}x = \int_{-1}^0 \frac{1}{x}\,\mathrm{d}x + \int_0^1 \frac{1}{x}\,\mathrm{d}x,$$

*wobei beide Integrale rechts uneigentlich sind und das uneigentliche Integral $\int_{-1}^1 \frac{1}{x}\,\mathrm{d}x$ per Definition genau dann existiert, wenn die beiden Integrale rechts existieren. Des Weiteren gilt*

$$\begin{aligned}\int_{-1}^0 \frac{1}{x}\,\mathrm{d}x &= \lim_{b \nearrow 0} \int_{-1}^b \frac{1}{x}\,\mathrm{d}x = \lim_{b \nearrow 0} \log|b| - \log|-1| = \lim_{b \nearrow 0} \log|b| = -\infty \\ \int_0^1 \frac{1}{x}\,\mathrm{d}x &= \lim_{a \searrow 0} \int_a^1 \frac{1}{x}\,\mathrm{d}x = \lim_{a \searrow 0} \log(1) - \log(a) = +\infty,\end{aligned}$$

*wodurch $\int_{-1}^1 \frac{1}{x}\,\mathrm{d}x$ nicht existiert (und wir diesem auch nicht das Symbol $\infty$ oder $-\infty$ zuweisen).*

**Beispiel 8.53** (Bogenlänge des Kreises)**.** *Wir wollen nochmals die Bogenlänge des Kreises berechnen. Doch verwenden wir diesmal die Gleichung $y = f(x) = \sqrt{1 - x^2}$ als Definition des*

*oberen Halbkreises. Die Bogenlänge des Kreises ist demnach gegeben durch das Integral*

$$\begin{aligned} 2\int_{-1}^{1}\sqrt{1+f'(x)^2}\,\mathrm{d}x &= 4\int_0^1\sqrt{1+\frac{x^2}{1-x^2}}\,\mathrm{d}x \\ &= 4\lim_{b\nearrow 1}\int_0^b\frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x \\ &= 4\lim_{b\nearrow 1}\arcsin(b)-0 = 4\tfrac{\pi}{2} = 2\pi. \end{aligned}$$

**Übung 8.54.** *Berechnen Sie* $\int_0^1 \frac{1}{\sqrt{x}}\,\mathrm{d}x$ *und* $\int_0^{\frac{\pi}{2}}\tan(x)\,\mathrm{d}x$.

**Übung 8.55** (Absolute Konvergenz)**.** *Sei* $a\in\mathbb{R}$ *und* $f:[a,\infty)\to\mathbb{C}$ *eine komplexwertige Funktion mit* $f\in R([a,b])$ *für alle* $b>a$. *Wir nennen das uneigentliche Integral* $\int_a^\infty f(x)\,\mathrm{d}x$ ***absolut konvergent****, falls* $\int_a^\infty |f(x)|\,\mathrm{d}x$ *konvergent ist. Zeigen Sie, dass absolute Konvergenz des uneigentlichen Integrals* $\int_a^\infty f(x)\,\mathrm{d}x$ *auch die Konvergenz dieses Integrals impliziert.*

*Bemerkung.* Zusammenfassend haben wir bei den Definitionen in diesem Abschnitt bei jedem Problempunkt eines möglichen Riemann-Integrals einen Grenzwert verwendet, um den Integralbegriff zu erweiteren. Dies wirft nochmals die Frage auf, ob es nicht vielleicht einen Integralbegriff gibt, der diese und auch andere bereits erwähnte Probleme des Riemann-Integrals auf natürliche Art und Weise löst. Diese Frage wird im zweiten Studienjahr des Mathematikstudiums mit der Theorie des Lebesgue-Integrals in der Vorlesung „Mass und Integral“ positiv beantwortet.

### 8.4.3 Die Gamma-Funktion*

Die **Gamma-Funktion** $\Gamma$ ist bei $s\in(0,\infty)$ durch das konvergente uneigentliche Integral

$$\Gamma(s)=\int_0^\infty x^{s-1}\mathrm{e}^{-x}\,\mathrm{d}x \tag{8.8}$$

definiert. Für $s\in(0,1)$ ist dies aus zwei Gründen ein uneigentliches Integral und wir müssen die Integrationsgrenzen $A=0$ und $B=\infty$ getrennt untersuchen. Für $a>0$ und $b>a$ gilt jedoch

$$\int_a^b x^{s-1}\mathrm{e}^{-x}\,\mathrm{d}x = \tfrac{1}{s}\left[x^s\mathrm{e}^{-x}\right]_a^b+\tfrac{1}{s}\int_a^b x^s\mathrm{e}^{-x}\,\mathrm{d}x.$$

Wir setzen $b=1$ und erhalten

$$\int_0^1 x^{s-1}\mathrm{e}^{-x}\,\mathrm{d}x=\lim_{a\searrow 0}\left(\tfrac{1}{s}\left[x^s\mathrm{e}^{-x}\right]_a^1+\tfrac{1}{s}\int_a^1 x^s\mathrm{e}^{-x}\,\mathrm{d}x\right)=\tfrac{1}{s\mathrm{e}}+\tfrac{1}{s}\int_0^1 x^s\mathrm{e}^{-x}\,\mathrm{d}x,$$

wobei das Integral rechts (für alle $s\in(0,\infty)$) ein eigentliches Riemann-Integral darstellt. Für $a=1$ erhalten wir

$$\int_1^\infty x^{s-1}\mathrm{e}^{-x}\,\mathrm{d}x=\lim_{b\to\infty}\left(\tfrac{1}{s}\left[x^s\mathrm{e}^{-x}\right]_1^b+\tfrac{1}{s}\int_1^b x^s\mathrm{e}^{-x}\,\mathrm{d}x\right)=-\tfrac{1}{s\mathrm{e}}+\tfrac{1}{s}\int_1^\infty x^s\mathrm{e}^{-x}\,\mathrm{d}x.$$

Um die Konvergenz von $\int_1^\infty x^s \mathrm{e}^{-x}\,\mathrm{d}x$ zu zeigen, wollen wir den Integraltest für Reihen in Satz 8.46 verwenden. Die erste Vorraussetzung des Integraltests ist erfüllt, da die Funktion $f(x) = x^s \mathrm{e}^{-x}$ nichtnegativ ist. Die zweite Vorraussetzung ist, dass $f(x)$ monoton abnehmend sein soll. Wir berechnen daher die Ableitung und sehen, dass

$$f'(x) = sx^{s-1}\mathrm{e}^{-x} - x^s \mathrm{e}^{-x}.$$

Da $f'(x) < 0$ für alle $x > s$, sehen wir, dass diese Vorraussetzung zumindest für $x \geq N = \lfloor s \rfloor + 1$ erfüllt ist. Es folgt daher, dass das Integral $\int_N^\infty x^s \mathrm{e}^{-x}\,\mathrm{d}x$ genau dann konvergiert wenn

$$\sum_{n=N}^{\infty} n^s \mathrm{e}^{-n} < \infty$$

konvergiert, was aber nach dem Quotientenkriterium in Korollar 6.32 (oder dem Wurzelkriterium in Korollar 6.30) für Reihen in der Tat zutrifft.

Addieren wir die beiden Integrale wieder und verwenden wir die Definition in (8.8) so erhalten wir, dass $\Gamma(s)$ für alle $s \in (0, \infty)$ wohldefiniert ist und

$$\Gamma(s+1) = s\Gamma(s) \tag{8.9}$$

erfüllt.

Oft wird die Gamma-Funktion als eine Erweiterung der Fakultätsfunktion auf $\mathbb{N}$ bezeichnet. In der Tat gilt für alle $n \in \mathbb{N}$

$$\Gamma(n+1) = n!\ ,$$

was wir hier noch beweisen möchten. Für $n = 0$ gilt

$$\Gamma(1) = \int_0^\infty \mathrm{e}^{-x}\,\mathrm{d}x = \lim_{b\to\infty} - \left[\mathrm{e}^{-x}\right]_0^b = 1$$

und somit folgt nach (8.9) und Induktion

$$\Gamma(n+1) = n\Gamma(n) = n(n-1)\Gamma(n-1) = \ldots = n!\ \Gamma(1) = n!\ .$$

Wir werden in der Fortsetzung dieser Vorlesung weitere Eigenschaften der Gamma-Funktion nachweisen können. Beispielsweise stellt sich heraus, dass die Gamma-Funktion glatt ist. Wir können dies hier aber nicht zeigen, da $\Gamma$ in (8.8) durch ein sogenanntes Parameterintegral definiert ist. Auch können wir den Wert

$$\begin{aligned}\Gamma(\tfrac{1}{2}) &= \int_0^\infty \frac{1}{\sqrt{x}}\mathrm{e}^{-x}\,\mathrm{d}x = \lim_{\varepsilon\to 0}\lim_{b\to\infty}\int_\varepsilon^b \frac{1}{\sqrt{x}}\mathrm{e}^{-x}\,\mathrm{d}x = \lim_{\varepsilon\to 0}\lim_{b\to\infty} 2\int_{\sqrt{\varepsilon}}^{\sqrt{b}} \mathrm{e}^{-u^2}\,\mathrm{d}u \\ &= 2\int_0^\infty \mathrm{e}^{-u^2}\,\mathrm{d}u = \int_{-\infty}^\infty \mathrm{e}^{-u^2}\,\mathrm{d}u\end{aligned}$$

mit den uns bis jetzt bekannten Integrationsmethoden nicht berechnen, werden jedoch später mittels einem zweidimensionalen Integral sehen, dass dieser $\sqrt{\pi}$ ist.

Die Gamma-Funktion enthüllt ihre wahre Schönheit erst wenn man komplexe Parameter $z \in \mathbb{C} \setminus \{0, -1, -2, \ldots\}$ erlaubt. Wir laden Interessierte ein, diese Funktion in folgender Übung zu konstruieren.

**Übung 8.56** (Challenge)**.** *Für $z \in \mathbb{C}$ mit $\Re(z) > 0$ definiert man*

$$\Gamma(z) = \int_0^\infty x^{z-1} \mathrm{e}^{-x} \, \mathrm{d}x. \tag{8.10}$$

*(a) Zeigen Sie, dass $\int_0^1 x^{z-1} \mathrm{e}^{-x} \, \mathrm{d}x$ für alle $z \in \mathbb{C}$ mit $\Re(z) > 0$ konvergiert.*
*(b) Zeigen Sie, dass $\int_1^\infty x^{z-1} \mathrm{e}^{-x} \, \mathrm{d}x$ für alle $z \in \mathbb{C}$ mit $\Re(z) > 0$ konvergiert.*
*(c) Zeigen Sie* (8.9) *für alle $z \in \mathbb{C}$ mit $\Re(z) > 0$.*
*(d) Verwenden Sie* (8.9) *um rekursiv $\Gamma(z)$ für $z \in \mathbb{C} \setminus \{0, -1, -2, \ldots\}$ mit $\Re(z) > -1$, oder $\Re(z) > -2$, …, zu definieren, so dass anschliessend $\Gamma(z)$ für alle $z \in \mathbb{C} \setminus \{0, -1, -2, \ldots\}$ definiert ist und* (8.9) *auf dem ganzen Definitionsbereich erfüllt.*

*Hinweis: Sie können in (a) und (b) Übung 8.55 verwenden.*

Hilbert (1862–1943) verwendete in seinem Artikel [Hil93] von 1893 uneigentliche Integrale im Stile der Gamma-Funktion, um zu beweisen, dass e (wie erstmals von Hermite in 1873 bewiesen) und $\pi$ (wie erstmals von Lindemann 1882 bewiesen) transzendent sind. Wir bemerken dabei, dass sich die blosse Irrationalität dieser Zahlen deutlich einfacher beweisen lässt – für e gibt es hierzu eine Übung in Abschnitt 6.9.2 und für $\pi$ eine Übung in Abschnitt 8.8.2. Transzendenzbeweise sind jedoch im Allgemeinen deutlich schwieriger. Wie schwierige derartige Aussagen tatsächlich sind, illustriert vielleicht die Tatsache, dass immer noch nicht bekannt ist, ob $\mathrm{e} + \pi$ eine transzendente Zahl ist oder nicht. Hilbert's Beweis der Transzendenz von e und $\pi$ ist mit den uns bisher bekannten Hilfsmitteln allerdings gut lesbar, weswegen wir Ihnen einen Blick auf diese Lektüre und die damit verbundene Zeitreise empfehlen möchten.

## 8.5 Taylor Approximation

Wir erinnern daran, dass die Ableitung $f'(x_0)$ einer reellwertigen differenzierbaren Funktion $f$ auf einem Intervall die Steigung der Tangente

$$y(x) = f(x_0) + f'(x_0)(x - x_0)$$

des Graphen von $f$ bei $x_0$ angibt, wobei die Tangente den Graphen gut approximiert (im Sinne von $f(x) = y(x) + o(x - x_0)$ für $x \to x_0$). Wir wollen hier die Güte der Approximation erhöhen indem wir statt linearen Approximationen auch noch höhere polynomiale Approximationen erlauben.

Sei also $f : (a, b) \to \mathbb{C}$ eine $n$-mal differenzierbare Funktion auf einem offenen, nicht-leeren (möglicherweise unbeschränkten) Intervall $(a, b) \subseteq \mathbb{R}$. Die **n-te Taylor-Approximation** von $f$ um einen Punkt $x_0 \in (a, b)$ ist das Polynom

$$P^f_{x_0,n}(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k.$$

Die Koeffizienten wurden dabei so gewählt, dass $P^f_{x_0,n}(x_0) = f(x_0)$ und allgemeiner

$$(P^f_{x_0,n})^{(k)}(x_0) = f^{(k)}(x_0)$$

für $k \in \{0, \ldots, n\}$ gilt (wieso?). Die vielleicht naive Hoffnung ist hierbei, dass die Taylor-Approximation, wie der Name sagt, die Funktion $f$ approximiert. Falls $f$ glatt ist, dann ist die **Taylorreihe** von $f$ um $x_0 \in (a, b)$ definiert als die Potenzreihe

$$T^f_{x_0}(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k.$$

Damit wir hier von einer Potenzreihe sprechen dürfen, erweitern wir die Definition 6.54 wie folgt. Eine **Potenzreihe um einen Punkt** $z_0 \in \mathbb{C}$ in der Variable $z \in \mathbb{C}$ ist ein formaler Ausdruck der Form $\sum_{n=0}^{\infty} a_n(z - z_0)^n$, wobei $a_n \in \mathbb{C}$ für alle $n \in \mathbb{N}$ die Koeffizienten der Potenzreihe sind. Der Konvergenzradius

$$R = \left( \limsup_{n\to\infty} \sqrt[n]{|a_n|} \right)^{-1}$$

ist wie in Abschnitt 6.4.1 definiert und die Potenzreihe $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ konvergiert für alle $z \in \mathbb{C}$ mit Abstand kleiner $R$ von $z_0$ und divergiert für alle $z \in \mathbb{C}$ mit $|z - z_0| > R$ (was aus Satz 6.56 folgt, wie?).

Soweit ist nicht klar, was der Konvergenzradius der Taylor-Reihe einer glatten Funktion $f$ um einen Punkt $x_0$ im Definitionsbereich ist und ob er positiv ist. Die naive Hoffnung ist, dass die Taylor-Reihe, wo definiert, gleich der Funktion $f$ ist, da die Taylorreihe bei $x_0$ die gleichen Ableitungen wie $f$ hat. Dies ist in der Tat für viele Funktionen der Fall, doch, wie folgendes Beispiel zeigt, nicht immer.

**Beispiel 8.57** (Verschwindende Taylorreihe)**.** *Wir betrachten die Funktion*

$$\psi : x \in \mathbb{R} \mapsto \begin{cases} \exp\left(-\frac{1}{x}\right) & \text{falls } x > 0 \\ 0 & \text{falls } x \leq 0 \end{cases},$$

*die nach Beispiel 7.23 glatt ist und $\psi^{(n)}(0) = 0$ für alle $n \in \mathbb{N}_0$ erfüllt. Wir zeigen, dass sich $\psi$ nicht durch eine Potenzreihe um 0 darstellen lässt.*

*Wir nehmen also indirekt an, dass es eine Potenzreihe $f(x) = \sum_{n=0}^{\infty} c_n x^n$ mit reellen Koeffizienten und Konvergenzradius $R > 0$ gibt, so dass $f(x) = \psi(x)$ für alle $x \in (-R, R)$. Nach Korollar 8.11 (oder genauer Übung 8.12) ist $f^{(n)}(0) = c_n n!$. Aber da $f(x) = \psi(x)$ für alle $x \in (-R, R)$ angenommen wurde und da Ableiten eine lokale Operation ist, schliessen wir, dass $f^{(n)}(0) = 0$ und somit $c_n = 0$ für alle $n \in \mathbb{N}_0$. Dies widerspricht jedoch $\psi(x) > 0$ für alle $x > 0$. Also kann $\psi$ in keiner Umgebung von 0 durch eine Potenzreihe dargestellt werden.*

Folgender Satz liefert nun einen direkten Vergleich zwischen einer Funktion und ihrer Taylor-Approximationen. Er impliziert für viele Funktionen, dass sie mit ihrer Taylorreihe übereinstimmen.

**Theorem 8.58** (Taylor-Approximation)**.** *Sei $(a, b) \subseteq \mathbb{R}$ ein offenes, nicht-leeres Intervall und sei $f : (a, b) \to \mathbb{C}$ eine $(n + 1)$-mal stetig differenzierbare Funktion. Sei $x_0 \in (a, b)$. Dann gilt für alle $x \in (a, b)$*

$$f(x) = P^f_{x_0,n}(x) + R^f_{x_0,n}(x),$$

*wobei $P^f_{x_0,n}$ die $n$-te Taylor-Approximation ist und wir den Fehlerterm $R^f_{x_0,n}$ durch das sogenannte **Integral-Restglied***

$$x \in (a, b) \mapsto R^f_{x_0,n}(x) = \int_{x_0}^{x} f^{(n+1)}(t) \frac{(x - t)^n}{n!} \, \mathrm{d}t$$

*darstellen können. Dies gilt auch für Funktionen auf $[x_0, b)$ und Punkte $x \in [x_0, b)$ (beziehungsweise $(a, x_0]$ und $x \in (a, x_0]$).*

Die Annahme im Theorem, dass $f$ eine $(n + 1)$-mal stetig differenzierbare Funktion ist, ist essentiell für obige Formulierung. In der Tat ist damit $f^{(n+1)}$ eine stetige Funktion und das Integral der stetigen Funktion $t \mapsto f^{(n+1)}(t) \frac{(x-t)^n}{n!}$ im Integral-Restglied existiert.

*Beweis von Satz 8.58.* Das Theorem ergibt sich mit Induktion über $n$ und partieller Integration (die ebenso für komplexwertige Funktionen auf $(a, b)$ gilt, wieso?). Ist $n = 0$ und $f : (a, b) \to \mathbb{R}$ stetig differenzierbar, so gilt nach dem Fundamentalsatz der Differential- und Integralrechnung (genauer Korollar 8.5)

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(t) \, \mathrm{d}t = P^f_{x_0,0}(x) + R^f_{x_0,0}(x).$$

Ist $f$ nun zweimal stetig differenzierbar (also $n = 1$), so können wir auf obiges Integral partielle Integration mit $u(t) = f'(t)$ und $v(t) = t - x$ anwenden und erhalten

$$\begin{aligned} f(x) &= f(x_0) + \big[f'(t)(t-x)\big]_{t=x_0}^{t=x} - \int_{x_0}^{x} f''(t)(t-x)\,\mathrm{d}t \\ &= f(x_0) + f'(x_0)(x-x_0) + \int_{x_0}^{x} f''(t)\frac{(x-t)^1}{1!}\,\mathrm{d}t \\ &= P^f_{x_0,1}(x) + R^f_{x_0,1}(x). \end{aligned}$$

Wir sehen also wie sich ausgehend vom Fundamentalsatz mittels partieller Integration der nächste Fall des Satzes ergibt.

Angenommen die Aussage des Satzes stimmt für $n - 1 \geq 0$ und sei $f : (a, b) \to \mathbb{R}$ eine $(n+1)$-mal stetig differenzierbare Funktion. Dann gilt auf Grund der Induktionsvorraussetzung

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \int_{x_0}^{x} f^{(n)}(t)\frac{(x-t)^{n-1}}{(n-1)!}\,\mathrm{d}t$$

für alle $x \in (a, b)$.

Wir setzen $u(t) = f^{(n)}(t)$ und $v(t) = -\frac{(x-t)^n}{n!}$, bemerken $v'(t) = \frac{(x-t)^{n-1}}{(n-1)!}$ und wenden partielle Integration an, um

$$\begin{aligned} f(x) &= \sum_{k=0}^{n-1} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k - \left[f^{(n)}(t)\frac{(x-t)^n}{n!}\right]_{t=x_0}^{t=x} + \int_{x_0}^{x} f^{(n+1)}(t)\frac{(x-t)^n}{n!}\,\mathrm{d}t \\ &= \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \int_{x_0}^{x} f^{(n+1)}(t)\frac{(x-t)^n}{n!}\,\mathrm{d}t \end{aligned}$$

zu erhalten. Dies beweist den Induktionsschritt und damit den Satz. □

VO: 20.12.17

Oft werden wir das Theorem von Taylor (Theorem 8.58) in folgender Form verwenden.

**Korollar 8.59** (Taylor-Abschätzung)**.** *Sei $(a, b) \subseteq \mathbb{R}$ ein offenes, nicht-leeres Intervall, sei $f : (a, b) \to \mathbb{C}$ eine $(n + 1)$-mal stetig differenzierbare Funktion und seien $x_0, x \in (a, b)$ zwei Punkte. Wir setzen $M_{n+1} = \max\big\{|f^{(n+1)}(t)| \mid t$ zwischen $x_0$ und $x\big\}$. Dann gilt*

$$|f(x) - P^f_{x_0,n}(x)| = |R^f_{x_0,n}(x)| \leq \frac{M_{n+1}|x-x_0|^{n+1}}{(n+1)!}.$$

*Insbesondere ist $f(x) = P^f_{x_0,n}(x) + O((x-x_0)^{n+1})$ für $x \to x_0$.*

*Beweis.* Angenommen $x \geq x_0$. Dann gilt

$$\begin{aligned} |R^f_{x_0,n}(x)| \leq \int_{x_0}^{x} |f^{(n+1)}(t)|\frac{(x-t)^n}{n!}\,\mathrm{d}t &\leq \frac{M_{n+1}}{n!}\int_{x_0}^{x}(x-t)^n\,\mathrm{d}t \\ &= \frac{M_{n+1}}{n!}\left[-\frac{1}{n+1}(x-t)^{n+1}\right]_{x_0}^{x} = \frac{M_{n+1}(x-x_0)^{n+1}}{(n+1)!}. \end{aligned}$$

Der Beweis für $x < x_0$ ist analog ausgehend von

$$|R^f_{x_0,n}(x)| = \left| \int_{x_0}^{x} f^{(n+1)}(t) \frac{(x-t)^n}{n!} \, \mathrm{d}t \right| \leq \int_{x}^{x_0} |f^{(n+1)}(t)| \frac{(t-x)^n}{n!} \, \mathrm{d}t. \qquad \square$$

*Bemerkung.* Wir hätten andere Versionen des obigen Satz zu Taylor-Approximationen bereits in Abschnitt 7.2 für reellwertige Funktionen beweisen können. Dies sogar unter der etwas schwächeren Annahme an die $(n+1)$-te Ableitung, dass diese bloss zwischen $x_0$ und $x$ existieren soll und nicht unbedingt stetig sein muss. Unter dieser Voraussetzung gibt es ein $\xi_C$ zwischen $x_0$ und $x$, so dass das sogenannte **Restglied nach Cauchy** durch

$$R^f_{x_0,n}(x) = \frac{1}{n!} f^{(n+1)}(\xi_C)(x-\xi_C)^n(x-x_0)$$

gegeben ist. Es gibt unter denselben Voraussetzungen auch ein $\xi_L$ zwischen $x_0$ und $x$, so dass das **Restglied nach Lagrange** durch

$$R^f_{x_0,n}(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi_L)(x-x_0)^{n+1}$$

gegeben ist. Wir verweisen dafür auf folgende Übung.

**Übung 8.60** (Andere Formeln für das Restglied)**.** *Gegeben sei ein nicht-leeres Intervall $(a,b) \subseteq \mathbb{R}$, Zahlen $x, x_0 \in (a,b)$ und eine $(n+1)$-mal differenzierbare Funktion $f : (a,b) \to \mathbb{R}$.*

*(a) Zeigen Sie, dass die Ableitung der Funktion*

$$F : t \in (a,b) \mapsto \sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(x-t)^k,$$

*durch $t \in (a,b) \mapsto \frac{f^{(n+1)}(t)}{n!}(x-t)^n$ gegeben ist.*

*(b) Wenden Sie den Mittelwertsatz (Theorem 7.29) auf die obige Funktion $F$ an, um die Formel $R^f_{x_0,n}(x) = \frac{1}{n!} f^{(n+1)}(\xi_C)(x-\xi_C)^n(x-x_0)$ für das Restglied nach Cauchy zu beweisen.*

*(c) Verwenden Sie den verallgemeinerten Mittelwertsatz (Satz 7.48) für obiges $F$ und die Funktion $g : t \mapsto (x-t)^{n+1}$, um die Formel $R^f_{x_0,n}(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi_L)(x-x_0)^{n+1}$ für das Restglied nach Lagrange zu beweisen.*

**Beispiel 8.61** (Extremwerte)**.** *Wir wollen die Taylor-Approximation verwenden um die Diskussion in Abschnitt 7.1.3 und Korollar 7.37 zu verfeinern. Sei $(a,b) \subseteq \mathbb{R}$ ein offenes, nicht-leeres Intervall und sei $f : (a,b) \to \mathbb{C}$ eine $n$-mal stetig differenzierbare Funktion. Angenommen $x_0 \in (a,b)$ erfüllt*

$$f'(x_0) = \ldots = f^{(n-1)}(x_0) = 0.$$

*Dann gelten folgende Implikationen.*

- *Falls $f^{(n)}(x_0) < 0$ ist und $n$ gerade ist, so nimmt $f$ in $x_0$ ein lokales Maximum an.*

- *Falls $f^{(n)}(x_0) > 0$ ist und $n$ gerade ist, so nimmt $f$ in $x_0$ ein lokales Minimum an.*
- *Falls $f^{(n)}(x_0) \neq 0$ und $n$ ungerade ist, so ist $x_0$ kein lokales Extremum von $f$.*

*Alle drei Aussagen folgen aus der Taylor-Approximation in Theorem 8.58, die in diesem Fall für $x \in (a, b)$ die Form*

$$f(x) = f(x_0) + \int_{x_0}^{x} f^{(n)}(t) \frac{(x-t)^{n-1}}{(n-1)!} \, \mathrm{d}t$$

*annimmt. Falls $f^{(n)}(x_0) > 0$ ist, existiert ein $\delta > 0$ mit $f^{(n)}(t) > 0$ für alle $t \in (x_0 - \delta, x_0 + \delta)$. Wir betrachten nun mehrere Fälle. Ist $x \in (x_0, x_0 + \delta)$, so ist obiges Integral positiv und damit $f(x) > f(x_0)$. Ist $x \in (x_0 - \delta, x_0)$ und $n - 1$ gerade, so ist obiges Integral (auf Grund der umgekehrten Reihenfolge der Integrationsgrenzen) negativ, womit $f(x) < f(x_0)$ gilt und $x_0$ kein lokales Extremum von $f$ ist. Ist hingegen $x \in (x_0 - \delta, x_0)$ und $n - 1$ ungerade, so ergibt sich auf dieselbe Weise $f(x) > f(x_0)$, womit $f$ ein lokales Minimum in $x_0$ annimmt. Für $f^{(n)}(x_0) < 0$ können wir obige Diskussion auf $-f$ anwenden.*

### 8.5.1 Analytische Funktionen

**Applet 8.62** (Taylor-Approximationen). *Wir stellen einige Taylor-Approximationen bei verschiebbaren Fusspunkten zu bekannten Funktionen dar.*

Betrachtet man eine glatte Funktion $f$ in Theorem 8.58, für die der Wert der Ableitung von $f$ nicht zu wild wird für hohe Ableitungen, dann geht das Restglied zu einem fest gewählten $x$ für $n \to \infty$ gegen Null und die Funktion $f$ wird durch ihre Taylorreihe beschrieben. Wie bereits gesehen, muss dies jedoch nicht unbedingt der Fall sein. Eine mögliche Abschätzung, die ausreichen würde, ist

$$\max_{t \in [x_0, x]} |f^{(n+1)}(t)| \leq cA^n$$

für alle $n \in \mathbb{N}$ und zwei Konstanten $c, A \geq 1$. Eine Abschätzung dieser Art findet man beispielsweise für $\exp, \sin, \cos$ und Kombinationen dieser Funktionen.

**Übung 8.63.** *Sei $[a, b] \subseteq \mathbb{R}$ ein kompaktes Intervall mit Endpunkten $a < b$.*

*(a) Sei $f : [a, b] \to \mathbb{R}$ ein Polynom oder eine der Funktionen $\exp, \sin, \cos$. Zeigen Sie, dass Konstanten $c, A \geq 1$ mit*

$$\max_{t \in [a,b]} |f^{(n+1)}(t)| \leq cA^n \tag{8.11}$$

*existieren.*

*(b) Zeigen Sie, dass falls $f, g : [a, b] \to \mathbb{R}$ Funktionen mit der Eigenschaft (8.11) auch $f + g$ und $f \cdot g$ diese Eigenschaften besitzen.*

Eine Funktion, die sich um jeden Punkt im Definitionsbereich als Potenzreihe darstellen lässt, nennen wir analytisch. Genauer sei $(a, b) \subseteq \mathbb{R}$ ein offenes, nicht-leeres Intervall und $f$ :

$(a,b) \to \mathbb{C}$ eine glatte Funktion. Die Funktion $f$ heisst **analytisch**, falls die Taylorreihe von $f$ um jeden Punkt in $x_0 \in (a,b)$ positiven Konvergenzradius $R$ hat und auf $(x_0-R, x_0+R) \cap (a,b)$ mit $f$ übereinstimmt.

Wir wollen ein weiteres Beispiel, wo sich eine glatte Funktion mit ihrer Taylorreihe vergleichen lässt, Interessierten als Übung überlassen (vergleichen Sie dies mit Übung 8.14).

**Übung 8.64.** *Sei $\alpha \in \mathbb{C}$. Zeigen Sie, dass*

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n$$

*für alle $x \in (-1,1)$, wobei für $n \in \mathbb{N}_0$*

$$\binom{\alpha}{n} = \frac{1}{n!} \prod_{k=0}^{n-1} (\alpha - k) = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}.$$

*Bemerkung.* Analytische Funktionen haben im Zusammenhang mit „holomorphen Funktionen“ auf offenen Teilmengen von $\mathbb{C}$ nach $\mathbb{C}$ eine schöne, geschlossene Behandlung, die in der Vorlesung „Funktionentheorie“ im zweiten Studienjahr thematisiert wird. Der komplexe Blickwinkel erklärt zum Beispiel, warum die analytische Funktion $x \in \mathbb{R} \to \frac{1}{1+x^2}$ bei $x_0 = 0$ eine Taylorreihe mit Konvergenzradius 1 besitzt, obwohl die Funktion auf ganz $\mathbb{R}$ definiert ist.

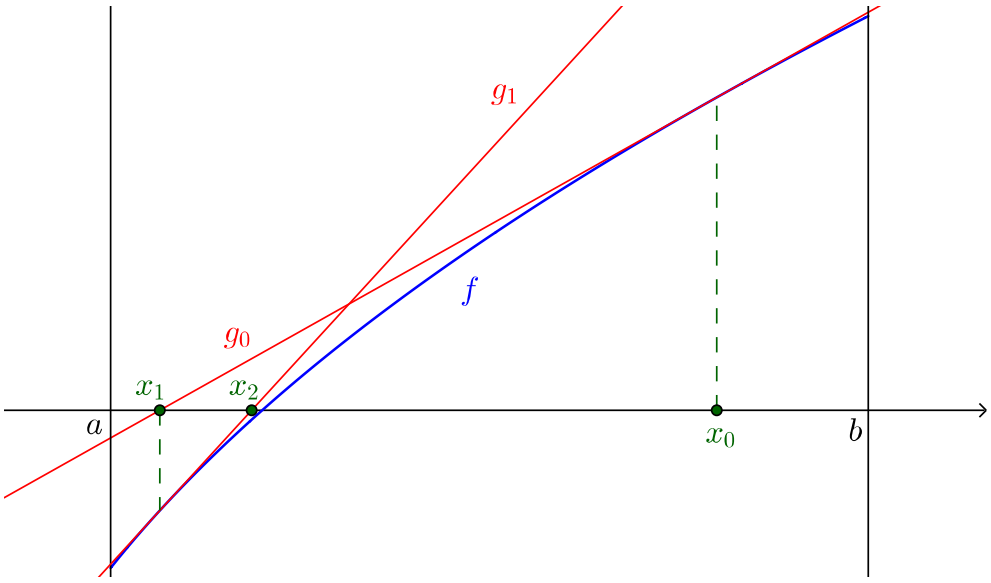
### 8.5.2 Konvergenzgeschwindigkeit des Newton-Verfahrens*

Als Anwendung des Satzes von Taylor (Theorem 8.58) möchten wir in diesem Teilabschnitt das Newton-Verfahren zur approximativen Berechnung einer Nullstelle einer gegebenen Funktion diskutieren. Weitere Anwendung werden in den nächsten Abschnitten dieses Kapitels folgen.

**Beispiel 8.65.** *Sei $[a,b] \subseteq \mathbb{R}$ ein kompaktes Intervall mit Endpunkten $a < b$ und $f : [a,b] \to \mathbb{R}$ eine stetig differenzierbare Funktion mit $f'(x) > 0$ für alle $x \in [a,b]$. (Wir erinnern daran, dass dies im Allgemeinen nicht dasselbe wie strenge Monotonie von $f$ ist, aber strenge Monotonie impliziert.) Angenommen es gilt $f(a) < 0$ und $f(b) > 0$. Dann gibt es auf Grund des Zwischenwertsatzes (Satz 3.59) ein $z \in [a,b]$ mit $f(z) = 0$, das wegen der strengen Monotonie von $f$ eindeutig bestimmt ist. Beginnend mit einem gewählten Punkt $x_0 \in [a,b]$ definieren wir rekursiv*

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \tag{8.12}$$

*für alle $n \in \mathbb{N}_0$ (falls $x_n \in [a,b]$). Die Idee hinter diesem Verfahren möchten wir in folgendem Bild erklären.*




Figur 8.2: In einem ersten Schritt approximiert man $f$ mit der Tangenten $g_0 : t \mapsto f(x_0) + f'(x_0)(t-x_0)$ bei $x_0$. Von dieser berechnet man nun die eindeutige Nullstelle $x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$. In einem zweiten Schritt verwendet man nun diesen neuen Punkt $x_1$ und berechnet die eindeutige Nullstelle $x_2$ der Tangenten $g_1$ bei $x_1$. Dies führt man iterativ so fort.

*Falls die Folge $(x_n)_n$ definiert ist (das heisst, $x_n \in [a,b]$ für alle $n \in \mathbb{N}$) und konvergiert, dann ist ihr Grenzwert eine Nullstelle von $f$ und somit unter unseren Annahmen gleich der Nullstelle $z$. In der Tat gilt für $z' = \lim_{n\to\infty} x_n$*

$$z' = \lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} \left( x_n - \frac{f(x_n)}{f'(x_n)} \right) = z' - \frac{f(z')}{f'(z')}$$

*und somit $f(z') = 0$ und $z' = z$.*

*Das Konvergenzverhalten der Folge $(x_n)_n$ ist im Allgemeinen „sehr chaotisch“. Unter etwas stärkeren Annahmen möchten wir nun aber zeigen, dass die Folge $(x_n)_n$ konvergiert und die Konvergenzgeschwindigkeit untersuchen.*

*Wir nehmen zusätzlich an, dass $f$ zweimal stetig differenzierbar ist und definieren*

$$M = \max\left\{|f''(t)| \mid t \in [a,b]\right\}$$
$$m = \min\left\{|f'(t)| \mid t \in [a,b]\right\}.$$

*Nun wenden wir Korollar 8.59 an, um den Fehlerterm $R^f_{x_0,1}$ bei der Approximation von $f$ durch die Tangente um $x_0$ abzuschätzen. Bei der Nullstelle $z$ ergibt dies*

$$|R^f_{x_0,1}(z)| \leq \frac{M|z-x_0|^2}{2} \tag{8.13}$$

*Des Weiteren gilt per Definition*

$$0 = f(z) = f(x_0) + f'(x_0)(z-x_0) + R^f_{x_0,1}(z). \tag{8.14}$$

*Wir dividieren Gleichung* (8.14) *durch* $f'(x_0)$ *und erhalten*

$$0 = z - \underbrace{\left(x_0 - \frac{f(x_0)}{f'(x_0)}\right)}_{=x_1} + \frac{R^f_{x_0,1}(z)}{f'(x_0)}.$$

*Mit der Abschätzung* (8.13) *ergibt dies*

$$|x_1 - z| = \left|\frac{R^f_{x_0,1}(z)}{f'(x_0)}\right| \leq \tfrac{M}{2m}\,|x_0 - z|^2 \tag{8.15}$$

*Falls nun* $\varepsilon \in (0, \frac{m}{M})$ *klein genug ist, so dass* $B_\varepsilon(z) \subseteq [a,b]$ *gilt, und falls* $x_0 \in B_\varepsilon(z)$ *ist, dann folgt aus* (8.15)*, dass*

$$|x_1 - z| \leq \tfrac{M}{2m}\,|x_0 - z|^2 \leq \tfrac{M}{2m}\,|x_0 - z|\varepsilon \leq \tfrac{1}{2}|x_0 - z| < \frac{\varepsilon}{2}.$$

*Daher liegt* $x_1$ *auch in* $[a,b]$ *und ist bereits näher an* $z$ *als* $x_0$*. Komplett analog beweist man, dass für alle* $n \in \mathbb{N}_0$

$$|x_{n+1} - z| \leq \tfrac{1}{2}\,|x_n - z|$$

*gilt, was mit vollständiger Induktion auch*

$$|x_n - z| \leq 2^{-n}|x_0 - z|$$

*für alle* $n \in \mathbb{N}_0$ *zeigt. Insbesondere konvergiert die Folge* $(x_n)_n$ *und wir können das Newton-Verfahren (unter geeigneter Wahl eines Startpunktes* $x_0$*) verwenden, um die Nullstelle mit hoher Genauigkeit zu approximieren. Die Konvergenz ist noch schneller als diese Abschätzung beweist; man spricht von quadratischer Konvergenz.*

**Übung 8.66** (Quadratische Konvergenz)**.** *Wir möchten die Konvergenzgeschwindigkeit des Newton-Verfahrens hier genauer analysieren. In obigem Beispiel wurde gezeigt, dass sich der Fehler in jedem Schritt mindestens halbiert. Wir betrachten nun das Argument etwas genauer (und behalten dementsprechend die Notation). Sei* $\beta = \frac{M}{2m}$*. Zeigen Sie für den mit* $\beta$ *gewichteten Abstand* $\beta|x_n - z|$ *die Abschätzung*

$$\beta|x_n - z| \leq (\beta|x_0 - z|)^{2^n}$$

*für jedes* $n \in \mathbb{N}$ *gilt.*

**Übung 8.67.** *Verwenden Sie das Newton-Verfahren, um* $\sqrt{2}$ *(oder alternativ eine andere irrationale, reelle algebraische Zahl) auf drei Dezimalstellen genau zu berechnen. Für die genaue, numerische Berechnung des Resultats dürfen Sie einen Rechner oder ein Computer-Programm benutzen.*

VO: 21.12.17
+ Christmas
Special

## 8.6 Numerische Integration

Wie bereits erwähnt wurde, gibt es Integrale, die sich nicht in Ausdrücken der üblichen (den uns bisher bekannten) Funktionen darstellen lassen. Besonderes in diesen Fällen sind folgende Abschätzungen zur approximativen Berechnung von Integralen sehr nützlich. Im nächsten Abschnitt werden wir weitere Beispiele sehen, die die zugrundeliegende Idee des folgenden Satzes in anderen Überlegungen gewinnbringend einsetzen. VO: 2. Semester

**Satz 8.68.** *Seien $a < b$ reelle Zahlen, $f : [a, b] \to \mathbb{R}$ eine Funktion, $n \in \mathbb{N}$, $h = \frac{b-a}{n}$ die Schrittweite und $x_\ell = a + \ell h$ für $\ell \in \{0, \dots, n\}$.*

*(a) (Rechtecksregel*) Falls $f$ stetig differenzierbar ist, dann gilt*

$$\int_a^b f(x)\,\mathrm{d}x = h\big(f(x_0) + \ldots + f(x_{n-1})\big) + F_1,$$

*wobei der Fehler $F_1$ durch $|F_1| \leq \frac{(b-a)^2}{2n} \max_{x\in[a,b]} |f'(x)|$ beschränkt ist.*

*(b) (Sehnentrapezregel) Falls $f$ zweimal stetig differenzierbar ist, dann gilt*

$$\begin{aligned}\int_a^b f(x)\,\mathrm{d}x &= h\left(\frac{f(x_0)+f(x_1)}{2} + \ldots + \frac{f(x_{n-1})+f(x_n)}{2}\right) + F_2\\ &= \frac{h}{2}\big(f(x_0) + 2f(x_1) + \ldots + 2f(x_{n-1}) + f(x_n)\big) + F_2,\end{aligned}$$

*wobei der Fehler $F_2$ durch $|F_2| \leq \frac{(b-a)^3}{6n^2} \max_{x\in[a,b]} |f''(x)|$ beschränkt ist.*

*(c) (Simpson-Regel*) Falls $f$ viermal stetig differenzierbar ist und $n$ gerade ist, dann gilt*

$$\begin{aligned}\int_a^b f(x)\,\mathrm{d}x = \frac{h}{3}\big(&f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + 2f(x_4) + \ldots\\ &+ 2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)\big) + F_3,\end{aligned}$$

*wobei der Fehler $F_3$ durch $|F_3| \leq \frac{(b-a)^5}{45n^4} \max_{x\in[a,b]} |f^{(4)}(x)|$ beschränkt ist.*

* *Wir haben zwar die Simpson-Regel ausführlicher besprochen. Doch in Hinblick auf eine mögliche Testfrage eignet sich die Sehnentrapezregel, wo bereits alle Ideen im Beweis auftauchen, besser.*

Insbesondere verhält sich der Fehler für das Rechtecksverfahren wie $O_f(n^{-1})$ für $n \to \infty$, für das Sehnentrapezverfahren wie $O_f(n^{-2})$ für $n \to \infty$ und für das Simpson-Verfahren wie $O_f(n^{-4})$ für $n \to \infty$.

Wir möchten anmerken, dass die Konstanten in obigen Abschätzungen nicht optimal sind. Alle drei obigen Approximationsverfahren sind sogenannte Newton-Cotes-Verfahren. Die wesentliche Idee eines solchen Verfahrens ist die folgende: zuerst schreibt man nach Intervalladditivität

$$\int_a^b f(x) = \sum_{\ell=0}^{n-1} \int_{x_\ell}^{x_{\ell+1}} f(x)\,\mathrm{d}x.$$

Nun approximiert man jedes obige Integralstück $\int_{x_\ell}^{x_{\ell+1}} f(x)\,\mathrm{d}x$ durch einen Ausdruck der Form $\sum_{k=1}^{K} w_k f(z_k)$ für Gewichte $w_1, \dots, w_K \in (0,1)$ mit $\sum_{k=1}^{K} w_k = 1$ und Stützpunkte $z_1, \dots, z_K \in [x_\ell, x_{\ell+1}]$. Beispielsweise nimmt man für die Sehnentrapezregel zwei Stützpunkte ($K = 2$), nämlich die beiden Endpunkte des Intervalls $[x_\ell, x_{\ell+1}]$, mit den Gewichten $w_1 = w_2 = \frac{1}{2}$.

Die Summe der Fehler, die auf den Stücken $\int_{x_\ell}^{x_{\ell+1}} f(x)\,\mathrm{d}x$ zustandekommen, ergeben dann den Gesamtfehler der Approximation. Obiger Satz gibt demnach an, wie dieser Fehler kontrolliert werden kann.

Vor dem Beweis des obigen Satzes möchten wir kurz erklären, wie sich die Simpson-Regel als Newton-Cotes-Verfahren auffassen lässt. Für die äquidistante Zerlegung des Intervalles $[a,b]$ mit den Punkten $y_\ell = a + \ell\frac{b-a}{m}$ für $\ell = 0, \dots, m$ betrachtet man auf $[y_\ell, y_{\ell+1}]$ die Gewichte $w_1 = \frac{1}{6}$, $w_2 = \frac{4}{6}$ und $w_3 = \frac{1}{6}$ und die Stützpunkte $y_\ell$, $\frac{y_\ell + y_{\ell+1}}{2}$ und $y_{\ell+1}$. Das dazugehörige Newton-Cotes-Verfahren ist genau das Simpson-Verfahren (wieso?). Wir wenden uns nun dem Beweis des obigen Satzes zu.

*Beweis.* Für (a) verwenden wir den Mittelwertsatz, wonach es zu $\ell \in \{0, \dots, n-1\}$ und $x \in [x_\ell, x_{\ell+1}]$ ein $\xi_x \in (x_\ell, x_{\ell+1})$ gibt mit $f(x) - f(x_\ell) = f'(\xi_x)(x - x_\ell)$. Insbesondere gilt

$$|f(x) - f(x_\ell)| = |f'(\xi_x)|(x - x_\ell) \le (x - x_\ell) \max_{t\in[a,b]} |f'(t)|.$$

Damit erhalten wir

$$\begin{aligned}\left|\int_{x_\ell}^{x_{\ell+1}} f(x)\,\mathrm{d}x - f(x_\ell)h\right| &\le \int_{x_\ell}^{x_{\ell+1}} |f(x) - f(x_\ell)|\,\mathrm{d}x \le \max_{t\in[a,b]} |f'(t)| \int_{x_\ell}^{x_{\ell+1}} (x - x_\ell)\,\mathrm{d}x \\ &\le \frac{h^2}{2} \max_{t\in[a,b]} |f'(t)|.\end{aligned}$$

Durch Summation, Intervalladditivität des Integrals und die Dreiecksungleichung erhalten wir

$$\left|\int_a^b f(x)\,\mathrm{d}x - \sum_{\ell=0}^{n-1} f(x_\ell)h\right| \le \frac{nh^2}{2} \max_{t\in[a,b]} |f'(t)| = \frac{(b-a)^2}{2n} \max_{t\in[a,b]} |f'(t)|.$$

Für (b) betrachten wir zuerst zu $\ell \in \{0, \dots, n-1\}$ die Endpunkte $x_- = x_\ell$ und $x_+ = x_{\ell+1}$ und den Mittelpunkt $\tilde{x} = \frac{x_- + x_+}{2}$ des Intervalls $[x_-, x_+]$. Des Weiteren definieren wir den Wert $M_2 = \max_{x\in[a,b]} |f''(x)|$. Nach Korollar 8.59 gilt für die Approximation durch das erste Taylor-Polynom um $\tilde{x}$

$$\big|f(t) - \big(f(\tilde{x}) + f'(\tilde{x})(t - \tilde{x})\big)\big| \le \frac{M_2}{2!}\,|t - \tilde{x}|^2 \le \frac{M_2}{2}\left(\frac{h}{2}\right)^2 \le \frac{M_2}{8}h^2$$

für alle $t \in [x_-, x_+]$. Wir verwenden dies für die Endpunkte $t = x_-$ und $t = x_+$ des Intervalls $[x_-, x_+]$ und erhalten aus der Dreiecksungleichung

$$\begin{aligned}&\left|\frac{f(x_-)+f(x_+)}{2} - f(\tilde{x})\right| \\ &\qquad = \tfrac{1}{2}\big|\big(f(x_-) - f(\tilde{x}) - f'(\tilde{x})(x_- - \tilde{x})\big) + \big(f(x_+) - f(\tilde{x}) - f'(\tilde{x})(x_+ - \tilde{x})\big)\big| \le \frac{M_2 h^2}{8},\end{aligned}$$

da sich der lineare Term wegen $x_- - \tilde{x} = -(x_+ - \tilde{x})$ aufhebt.

Aus demselben Grund erhalten wir

$$\begin{aligned}\left|\int_{x_-}^{x_+} f(t)\,\mathrm{d}t - f(\tilde{x})h\right| &= \left|\int_{x_-}^{x_+} \big(f(t) - f(\tilde{x})\big)\,\mathrm{d}t\right| = \left|\int_{x_-}^{x_+} \big(f(t) - (f(\tilde{x}) + f'(\tilde{x})(t-\tilde{x}))\big)\,\mathrm{d}t\right| \\ &\le \tfrac{M_2}{2}\int_{x_-}^{x_+} (t-\tilde{x})^2\,\mathrm{d}t = \tfrac{M_2}{2}\int_{-\frac{h}{2}}^{\frac{h}{2}} s^2\,\mathrm{d}s = \frac{M_2 h^3}{3\cdot 8}\end{aligned}$$

Zusammenfassend gilt also für $\ell \in \{0, \dots, n-1\}$ und $\tilde{x}_\ell = \frac{x_\ell + x_{\ell+1}}{2}$

$$\left|f(\tilde{x}_\ell) - \frac{f(x_{\ell-1}) + f(x_\ell)}{2}\right| \le \frac{M_2 h^2}{8} \tag{8.16}$$

$$\left|\int_{x_{\ell-1}}^{x_\ell} f(t)\,\mathrm{d}t - f(\tilde{x}_\ell)h\right| \le \frac{M_2 h^3}{3\cdot 8} \tag{8.17}$$

Wir multiplizieren (8.16) mit $h$ und summieren sowohl (8.16) als auch (8.17) über $\ell$ in $\{1, \dots, n\}$. Daraus folgt mit Intervalladditivität des Riemann-Integrals

$$\left|\int_a^b f(x)\,\mathrm{d}x - h\sum_{\ell=1}^{n} \frac{f(x_{\ell-1}) + f(x_\ell)}{2}\right| \le n\left(\frac{M_2 h^3}{8} + \frac{M_2 h^3}{3\cdot 8}\right) = \frac{(b-a)^3 M_2}{6n^2}.$$

Für (c) betrachten wir wieder zuerst zu $k \in \left\{0, \dots, \frac{n}{2} - 1\right\}$ die Endpunkte $x_- = x_{2k}$ und $x_+ = x_{2k+2}$ und den Mittelpunkt $\tilde{x} = x_{2k+1}$ des Intervalls $[x_-, x_+]$ und verwenden Korollar 8.59 bei $\tilde{x}$. Dies ergibt für $t \in [x_-, x_+]$

VO: Wir haben Teil (c) in der VO bewiesen, doch reicht es den etwas einfacheren Beweis von Teil (b) zu lernen.

$$\begin{aligned}&\left|f(t) - \left(f(\tilde{x}) + f'(\tilde{x})(t-\tilde{x}) + \frac{f''(\tilde{x})}{2}(t-\tilde{x})^2 + \frac{f'''(\tilde{x})}{6}(t-\tilde{x})^3\right)\right| \\ &\qquad \le \frac{M_4(t-\tilde{x})^4}{4!} \le \frac{M_4 h^4}{4!},\end{aligned} \tag{8.18}$$

wobei $M_4 = \max_{x\in[a,b]} |f^{(4)}(x)|$. Durch Integration über $[x_-, x_+]$ erhalten wir

$$\left|\int_{x_-}^{x_+} f(t)\,\mathrm{d}t - \left(2hf(\tilde{x}) + \frac{f''(\tilde{x})}{2}\int_{-h}^{h} s^2\,\mathrm{d}s\right)\right| \le \frac{M_4}{4!}\int_{-h}^{h} s^4\,\mathrm{d}s$$

oder auch

$$\left|\int_{x_-}^{x_+} f(t)\,\mathrm{d}t - \left(2hf(\tilde{x}) + \frac{f''(\tilde{x})}{3}h^3\right)\right| \le \frac{M_4 h^5}{60}$$

Setzen wir $t = x_-$ und $t = x_+$ in Gleichung (8.18), so erhalten wir

$$\frac{h}{3}\left|\big(f(x_-) + f(x_+)\big) - \big(2f(\tilde{x}) + f''(\tilde{x})h^2\big)\right| \leq \frac{h \cdot 2 \cdot M_4 h^4}{3 \cdot 24} = \frac{M_4 h^5}{36},$$

da sich die linearen und kubischen Terme gegenseitig aufheben. Daher gilt

$$\begin{aligned}
&\left|\int_{x_-}^{x_+} f(t)\,\mathrm{d}t - \tfrac{h}{3}\big(f(x_-) + 4f(\tilde{x}) + f(x_+)\big)\right| \\
&\quad = \left|\int_{x_-}^{x_+} f(t)\,\mathrm{d}t - \tfrac{h}{3}\big(6f(\tilde{x}) + f''(\tilde{x})h^2\big) + \tfrac{h}{3}\big(2f(\tilde{x}) + f''(\tilde{x})h^2\big) - \tfrac{h}{3}\big(f(x_-) + f(x_+)\big)\right| \\
&\qquad\qquad \leq \tfrac{1}{60}M_4h^5 + \tfrac{1}{36}M_4h^5 = \tfrac{2}{45}M_4h^5.
\end{aligned}$$

Nach Summation über $k \in \left\{0, \ldots, \frac{n}{2} - 1\right\}$ ergibt sich die Folge $1, 4, 2, 4, 2, \ldots, 2, 4, 1$ der Gewichte für die Funktionswerte in der Simpson-Regel wie im Satz und die Abschätzung genau wie im Beweis von (b) oben. □

**Übung 8.69.** *Erklären Sie unter Verwendung der Simpson-Regel, wie man* $\pi = 4\int_0^1 \frac{1}{1+x^2}\,\mathrm{d}x$ *auf beliebig viele Dezimalstellen genau bestimmen kann.*

### 8.6.1 Landau-Notation II

Wie in obigem Beweis der Simpson-Regel ersichtlich wurde, ist das genaue Buchführen der Konstanten relativ anstrengend und der eigentliche Wert, den man zum Schluss erhält, steckt nicht so sehr in der konkreten Konstante sondern in den anderen Ausdrücken. Wir möchten nun deswegen ein weiteres Stück Notation einführen, welche uns erlaubt bei Abschätzungen unsere Denkleistung auf das Wesentliche zu fokussieren.

Sei $X$ eine Menge und seien $f, g : X \to \mathbb{C}$ zwei Funktion. Wir schreiben

$$f(x) = O(g(x)) \quad \text{für } x \in X,$$

falls eine Konstante $C > 0$ existiert mit $|f(x)| \leq C|g(x)|$ für alle $x \in X$.

Die obige Notation kann leicht mit der Landau-Notation aus Abschnitt 5.6 verwirrt werden, weswegen wir uns Mühe geben werden, den Zusatz „für $x \in X$“ auch immer anzugeben. In einem gewissen Sinne ist die Aussage $f(x) = O(g(x))$ für $x \in [a, \infty)$ eine „globale Aussage“, da alle Zahlen $x$ in $[a, \infty)$ betroffen sind. Hingegen ist $f(x) = O(g(x))$ für $x \to \infty$ eine „lokale Aussage“, da nur Zahlen $x \in [a, \infty)$ betroffen sind, die gross genug sind (das heisst nahe genug bei unendlich sind, siehe Abschnitt 5.6).

In gewissen Situation bedeutet die Notation aber dasselbe, wie folgende Übung zeigt.

**Übung 8.70.** *Sei* $a \in \mathbb{R}$ *und seien* $f, g : [a, \infty) \to \mathbb{R}$ *zwei stetige Funktionen mit* $g(x) > 0$ *für alle* $x \in [a, \infty)$. *Zeigen Sie, dass die Aussagen*

$$\begin{aligned}
f(x) &= O(g(x)) \quad \textit{für } x \in [a, \infty) \\
f(x) &= O(g(x)) \quad \textit{für } x \to \infty
\end{aligned}$$

*äquivalent sind.*

Die Gross-O-Notation ist per Definition also insbesondere dann nützlich, wenn wir Konstanten „verstecken“ wollen. Beispielsweise ist $100! = O(1)$. Damit die Notation auch in Rechnungen nützlich ist, erlauben wir auch arithmetische Operationen mit ihr (vergleiche auch Abschnitt 5.6): Ist $X$ eine Menge und sind $f_1, f_2, g : X \to \mathbb{C}$ drei Funktionen, dann bedeutet

$$f_1(x) = f_2(x) + O(g(x)) \quad \text{für } x \in X,$$

dass $f_1(x) - f_2(x) = O(g(x))$ für $x \in X$ oder intuitiv ausgedrückt, dass die Differenz von $f_1$ und $f_2$ durch $g$ kontrolliert ist. Des Weiteren folgt aus $f_1 = O(g)$ für $x \in X$, auch $f_1 f_2 = O(g f_2)$ für $x \in X$.

Die Gross-O-Notation wird auch dazu verwendet, Unabhängigkeiten von gewissen Parametern zum Ausdruck zu bringen. Wir möchten dies an einem Beispiel illustrieren.

**Beispiel 8.71** (Parameterabhängigkeit bei Landau-Notation)**.** *Sei $k \in \mathbb{N}$ und $\alpha \in (-\infty, 2]$.*

*(a) Nach Taylorapproximation im Sinne von Korollar 8.59 gilt*

$$\left| x^{\frac{3}{2}} - \left( k^{\frac{3}{2}} + \tfrac{3}{2} k^{\frac{1}{2}} (x-k) \right) \right| \leq \tfrac{1}{2} \max_{t \in [1,\infty)} \left| \tfrac{3}{4} t^{-\frac{1}{2}} \right| (x-k)^2 = \tfrac{3}{8} (x-k)^2$$

*für $x \in [1, \infty)$ . Insbesondere gilt*

$$x^{\frac{3}{2}} = k^{\frac{3}{2}} + \tfrac{3}{2} k^{\frac{1}{2}} (x-k) + O((x-k)^2) \quad \textit{für } x \in [1, \infty),$$

*wobei die in $O(\cdot)$ versteckte implizite Konstante also nicht vom Parameter $k$ abhängt.*

*(b) Hängt die implizite Konstante, nicht wie in (a) oben, vom Parameter ab, so indizieren wir den Parameter bei $O(\cdot)$. Ein konkretes Beispiel: Wenn $\alpha \in (-\infty, 2]$, dann gilt*

$$x^\alpha = 1 + \alpha(x-1) + O_\alpha((x-1)^2)$$

*für $x \in [\frac{1}{2}, \infty)$ auf Grund der Taylorapproximation in Korollar 8.59.*

*(c) Auch eine Abhängigkeit des Definitionsbereichs vom Parameter ist denkbar und oft nützlich. Für alle Zahlen $t \in \left[k - \frac{1}{2}, k + \frac{1}{2}\right]$ gilt $(\log)''(t) = -\frac{1}{t^2} = O\left(\frac{1}{k^2}\right)$ und daher*

$$\log(t) = \log(k) + \tfrac{1}{k}(t-k) + O\left(\tfrac{(t-k)^2}{k^2}\right) = \log(k) + \tfrac{1}{k}(t-k) + O\left(\tfrac{1}{k^2}\right)$$

*auf Grund der Taylorapproximation in Korollar 8.59.*

*Im Allgemeinen sollte man die Landau-Notation sorgfältig verwenden, da sich hier oft Fehler einschleichen insbesondere bei der Frage der Abhängigkeit der impliziten Konstanten.*

## 8.7 Drei asymptotische Formeln

### 8.7.1 Das Wallissche Produkt

Die Formel

$$\pi = \lim_{n\to\infty} \frac{1}{n} \frac{(2^n n!)^4}{((2n)!)^2} \tag{8.19}$$

nach John Wallis (1616-1703) wird das **Wallissche Produkt** genannt.

Um Gleichung (8.19) zu beweisen, wollen wir zuerst für jedes $k \in \mathbb{N}_0$ das Integral

$$I_k = \int_0^\pi \sin^k(x)\,\mathrm{d}x$$

mit vollständiger Induktion berechnen.

Für $k = 0$ gilt $I_0 = \pi$ und für $k = 1$ gilt $I_1 = \int_0^\pi \sin(x)\,\mathrm{d}x = 2$. Mit partieller Integration erhalten wir für alle $k \geq 2$

$$\begin{aligned} I_k &= \int_0^\pi \sin^k(x)\,\mathrm{d}x = \int_0^\pi \sin^{k-1}(x)\sin(x)\,\mathrm{d}x \\ &= \Big[\sin^{k-1}(x)\big(-\cos(x)\big)\Big]_0^\pi + \int_0^\pi (k-1)\sin^{k-2}(x)\cos^2(x)\,\mathrm{d}x \\ &= (k-1)\int_0^\pi \sin^{k-2}(x)(1-\sin^2(x))\,\mathrm{d}x \\ &= (k-1)(I_{k-2} - I_k), \end{aligned}$$

woraus sich die Rekursionsformel

$$I_k = \frac{k-1}{k} I_{k-2} \tag{8.20}$$

für jedes ganze $k \geq 2$ ergibt.

Somit ist für $n \in \mathbb{N}$

$$\begin{aligned} I_{2n} = \frac{2n-1}{2n} I_{2n-2} &= \frac{2n-1}{2n} \cdot \frac{2n-3}{2n-2} \cdots \frac{1}{2} I_0 \\ &= \frac{(2n)(2n-1)(2n-2)\cdots 1}{\big((2n)(2n-2)\cdots 2\big)^2} I_0 = \frac{(2n)!}{(2^n n!)^2} \cdot \pi, \end{aligned}$$

wobei wir den Bruch mit $2n(2n-2)\cdots 2$ erweitert haben um im Zähler $(2n)!$ zu erhalten. Auf dieselbe Weise ergibt sich

$$I_{2n+1} = \frac{2n}{2n+1} \cdot \frac{2n-2}{2n-1} \cdots \frac{2}{3} I_1 = \frac{\big((2n)(2n-2)\cdots 2\big)^2}{(2n+1)(2n)\cdots 1} \cdot 2 = \frac{(2^n n!)^2}{(2n+1)!} \cdot 2.$$

Die Folge $(I_n)_n$ erfüllt noch eine weitere Eigenschaft: Denn für $x \in [0, \pi]$ ist $\sin(x) \in [0, 1]$ und somit

$$\sin^{2n+2}(x) \leq \sin^{2n+1}(x) \leq \sin^{2n}(x)$$

für alle $x \in [0, \pi]$ and $n \in \mathbb{N}$. Durch Integration folgt die Monotonieungleichung

$$I_{2n+2} = \int_0^\pi \sin^{2n+2}(x)\,\mathrm{d}x \leq I_{2n+1} = \int_0^\pi \sin^{2n+1}(x)\,\mathrm{d}x \leq I_{2n} = \int_0^\pi \sin^{2n}(x)\,\mathrm{d}x$$

und mittels (8.20) auch

$$I_{2n+2} = \frac{2n+1}{2n+2} \underbrace{\frac{(2n)!}{(2^n n!)^2}\pi}_{I_{2n}} \leq \frac{(2^n n!)^2}{(2n+1)!} 2 \leq I_{2n} = \frac{(2n)!}{(2^n n!)^2}\pi.$$

Wir multiplizieren diese Ungleichungskette mit $\frac{(2^n n!)^2}{(2n)!}\frac{2n+1}{2n}$ und erhalten

$$\frac{(2n+1)^2}{2n(2n+2)}\pi \leq \frac{1}{n}\frac{(2^n n!)^4}{((2n)!)^2} \leq \frac{2n+1}{2n}\pi.$$

Dies beweist die Formel (8.19) nach Wallis unter Anwendung des Sandwich-Lemmas (Lemma 5.31).

VO: 19.2.18

### 8.7.2 Stirling-Formel

Wir verwenden nun das Wallissche Produkt und Taylor-Approximationen, um das asymptotische Verhalten von $n!$ zu bestimmen. Genauer formuliert wollen wir die **Stirling-Formel**

$$\lim_{n\to\infty} \frac{n!}{\sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^n} = 1 \tag{8.21}$$

beweisen. Diese ist nach James Stirling (1692-1770) benannt und wird auch als $n! \approx \sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^n$ geschrieben.

Der Grundgedanke für den Beweis von (8.21) ist die Gleichung

$$\log(n!) = \log(1) + \log(2) + \ldots + \log(n)$$

für $n \in \mathbb{N}$ genauer zu betrachten, wobei wir die Summe aber durch das Integral $\int_{\frac{1}{2}}^{n+\frac{1}{2}} \log(x)\,\mathrm{d}x$ ersetzen wollen, da dieses einfacher zu berechnen ist.

Für $k \in \mathbb{N}$ und $t \in \left[k - \frac{1}{2}, k + \frac{1}{2}\right]$ gilt nach Beispiel 8.71

$$\log(t) = \log(k) + \tfrac{1}{k}(t-k) + O\left(\tfrac{1}{k^2}\right).$$

Wir erhalten

$$\int_{k-\frac{1}{2}}^{k+\frac{1}{2}} \log(t)\,\mathrm{d}t = \log(k) + O\left(\tfrac{1}{k^2}\right)$$

für alle $k \in \mathbb{N}$, da sich der lineare Term bei der Integration weglöscht.

Für ein festes $n \in \mathbb{N}$ und die Summe über $k \in \{1, \dots, n\}$ folgt daraus

$$\begin{aligned}\log(n!) &= \sum_{k=1}^{n} \log(k) = \sum_{k=1}^{n} \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} \log(t)\,\mathrm{d}t + \sum_{k=1}^{n} O\left(\tfrac{1}{k^2}\right) = \int_{\frac{1}{2}}^{n+\frac{1}{2}} \log(t)\,\mathrm{d}t + \sum_{k=1}^{n} O\left(\tfrac{1}{k^2}\right) \\ &= (n+\tfrac{1}{2})\log(n+\tfrac{1}{2}) - (n+\tfrac{1}{2}) - c + \sum_{k=1}^{n} O\left(\tfrac{1}{k^2}\right),\end{aligned}$$

wobei sich die Konstante $c \in \mathbb{R}$ durch Auswerten der Stammfunktion bei der unteren Grenze $\frac{1}{2}$ ergibt. Wiederum nach Taylorapproximation gilt $\log(n+\frac{1}{2}) = \log(n) + \frac{1}{n} \cdot \frac{1}{2} + O(\frac{1}{n^2})$. Wir multiplizieren diese Approximation mit $(n+\frac{1}{2})$ und erhalten (mit $(n+\frac{1}{2})O(\frac{1}{n^2}) = O(\frac{1}{n})$) somit

$$\begin{aligned}\log(n!) &= (n+\tfrac{1}{2})\left(\log(n) + \tfrac{1}{2n} + O(\tfrac{1}{n^2})\right) - (n+\tfrac{1}{2}) - c + \sum_{k=1}^{n} O\left(\tfrac{1}{k^2}\right) \\ &= (n+\tfrac{1}{2})\log(n) + \tfrac{1}{2} + O\left(\tfrac{1}{n}\right) - n - \tfrac{1}{2} - c + \sum_{k=1}^{n} O\left(\tfrac{1}{k^2}\right) \\ &= \frac{1}{2}\log(n) + n\log(n) - n + a_n\end{aligned}$$

für eine konvergente Folge $(a_n)_n$ in $\mathbb{R}$ (wobei wir Proposition 6.28 verwendet haben um die absolut konvergenten Fehlerterme $\varepsilon_n = O(\frac{1}{n^2})$ aufzusummieren). Wenden wir die Exponentialabbildung an, so erhalten wir für alle $n \in \mathbb{N}$

$$n! = \sqrt{n} n^n \mathrm{e}^{-n} b_n, \tag{8.22}$$

wobei $b_n = \exp(a_n)$ ist. Da $(a_n)_n$ konvergent ist, ist auch $(b_n)_n$ konvergent und der Grenzwert $B = \lim_{n\to\infty} \exp(a_n) = \exp(\lim_{n\to\infty} a_n)$ ist positiv.

Um $B$ zu berechnen, verwenden wir (8.22) gemeinsam mit dem Wallisschen Produkt (8.19). Es gilt

$$\pi = \lim_{n\to\infty} \frac{1}{n} \frac{(2^n n!)^4}{((2n)!)^2} = \lim_{n\to\infty} \frac{1}{n} \frac{\left(2^n \sqrt{n} n^n \mathrm{e}^{-n} b_n\right)^4}{\left(\sqrt{2n}(2n)^{2n} e^{-2n} b_{2n}\right)^2} = \lim_{n\to\infty} \frac{b_n^4}{2b_{2n}^2},$$

da die Potenzen von 2, von $n$ und von $e$ sich jeweils direkt kürzen, und die Wurzelterme gemeinsam mit dem $\frac{1}{n}$ vor dem Bruch sich zu einem $\frac{1}{2}$ vereinfachen. Daraus folgt

$$\pi = \frac{B^4}{2B^2} = \tfrac{1}{2}B^2$$

und somit muss $B = \sqrt{2\pi}$ gelten. Aus Gleichung (8.22) erhalten wir

$$\lim_{n\to\infty} \frac{n!}{\sqrt{n}\left(\frac{n}{\mathrm{e}}\right)^n} = \lim_{n\to\infty} b_n = \sqrt{2\pi},$$

was zur Stirling-Formel (8.21) äquivalent ist.

Auf Grund der Bedeutung von der Fakultätsfunktion, zum Beispiel in kombinatorischen Überlegungen, ist es nicht verwunderlich, dass die Sterling-Formel vielfache Anwendungen

besitzt. In den folgenden Übungen wollen wir einige wenige solcher Anwendungen präsentieren.

**Übung 8.72** (Besuche bei der Irrfahrt auf $\mathbb{Z}$)**.** *Wir betrachten eine (diskrete) zufällige Bewegung auf den ganzen Zahlen ausgehend von Null. Also beginnen wir zur Zeit $t = 0$ im Ursprung, hüpfen dann mit Wahrscheinlichkeit $\frac{1}{2}$ entweder nach links auf die Position $-1$ oder nach rechts auf die Position $1$. Bezeichnet im Allgemeinen $x_t \in \mathbb{Z}$ die Position zur Zeit $t \in \mathbb{N}$, so bewegen wir uns auf die Zeit $t+1$ hin wieder entweder links oder rechts mit gleicher Wahrscheinlichkeit von $\frac{1}{2}$.*

*Berechnen Sie die Wahrscheinlichkeit, dass man sich zur Zeit $t \in \mathbb{N}$ bei Null befindet und approximieren Sie diese mit der Stirling-Formel.*

**Übung 8.73.** *Eine Münze wird $n$-mal geworfen. Sei $\varepsilon > 0$. Zeigen Sie, dass die Wahrscheinlichkeit, dass die Münze höchstens $(\frac{1}{2} - \varepsilon)n$-mal auf Kopf fällt, für $n \to \infty$ gegen Null geht.*

*Hinweis: Sie dürfen dabei annehmen, dass $2^{-n}\binom{n}{k}$ die Wahrscheinlichkeit ist, dass bei $n$-maligem Wurf die Münze $k$-mal auf Kopf fällt. Zeigen Sie auch, dass die Funktion $x \in (0,1) \mapsto x\log(x) + (1-x)\log(1-x)$ ein striktes lokales Maximum bei $x = \frac{1}{2}$ hat.*

**Übung 8.74** (Ein Ausblick auf Algorithmik)**.** *In dieser Übung wollen wir die worst case Geschwindigkeit eines Sortieralgorithmus abschätzen. Sei also gegeben ein Algorithmus, der zu einem Datensatz $a_1, \ldots, a_n \in \mathbb{Z}$ diesen umordnet, das heisst, eine Abbildung $j : \{1, \ldots, n\} \to \{1, \ldots, n\}$ findet, so dass $a_{j(1)} \leq \ldots \leq a_{j(n)}$. Wir interessieren uns nun für die Zahl $W(n)$ der Anzahl Vergleiche zweier Zahlen des Datensatzes, die der Algorithmus schlimmstenfalls durchführen muss, um den Datensatz zu sortieren. Zeigen Sie, dass eine von $n$ unabhängige Konstante $C > 0$ existiert, so dass $W(n) \geq Cn\log(n)$.*

### 8.7.3 Asymptotik der harmonischen Reihe*

Für den Beweis der Stirling-Formel wurde eine Kombination von Taylor-Approximation und Integration verwendet, um das asymptotische Verhalten der Reihe $\log(n!) = \sum_{k=1}^{n} \log(k)$ zu bestimmen. Diese Methode kann auch allgemeiner zur Untersuchung des Divergenzverhaltens von Reihen eingesetzt werden. Zum Beispiel gilt die „asymptotische Entwicklung“

$$\sum_{k=1}^{n} \frac{1}{k} = \log(n) + \gamma + \frac{1}{2n} + O(\tfrac{1}{n^2}) \tag{8.23}$$

für $n \to \infty$, wobei

$$\gamma = \int_1^\infty \left( \frac{1}{\lfloor x \rfloor} - \frac{1}{x} \right) \mathrm{d}x = 0.577215664901532\ldots$$

die **Euler-Mascheroni-Konstante** genannt wird.

**Übung 8.75.** *Wir wollen die Asymptotik* (8.23) *in dieser Übung beweisen.*

*(a) Beweisen Sie, dass das Integral in der Definition von $\gamma$ konvergiert und dass die Gleichung*

$$\sum_{k=1}^{n} \frac{1}{k} = \log(n) + \gamma + \frac{1}{n} - \int_n^\infty \left( \frac{1}{\lfloor x \rfloor} - \frac{1}{x} \right) \mathrm{d}x$$

*für alle $n \in \mathbb{N}$ gilt.*

(b) *Verwenden Sie Taylor-Approximation, um*

- $\log(m+1) = \log(m) + \frac{1}{m} - \frac{1}{2m^2} + O(\frac{1}{m^3})$
- $\frac{1}{m} - \int_m^{m+1} \frac{1}{t}\,\mathrm{d}t = \frac{1}{2m^2} + O(\frac{1}{m^3}) = \frac{1}{2}\int_m^{m+1} t^{-2}\,\mathrm{d}t + O(\frac{1}{m^3})$

*für $m \in \mathbb{N}$ zu beweisen.*

(c) *Verwenden Sie das Integralkriterium (Satz 8.46) um zu zeigen, dass*

$$\sum_{m=n}^{\infty} \frac{1}{m^3} = O(\tfrac{1}{n^2})$$

*für $n \to \infty$.*

(d) *Verbinden Sie die obigen Aussagen zu einem Beweis von* (8.23).

(e) *Wiederholen Sie Beispiel 6.5 und approximieren Sie mit der Formel* (8.23) *die notwendige Anzahl Bausteine $n$ für einen Sprungturm, der $10m$ in den See hinreicht. Vergleichen Sie dies dann zum exakten Resultat $n = 12367$.*

## 8.8 Weitere Lernmaterialien

### 8.8.1 Verwendung des Kapitels

Dieses Kapitel verbindet mit Hilfe des Fundamentalsatz der Analysis die zentralen Begriffe der Ableitung und des Riemann-Integrals. Damit haben wir das vollständige Arsenal an Ableitungsregeln auch für die Berechnung von Integralen verwenden können, womit die Berechnung von Integralen mitunter deutlich einfacher wurde – auch wenn dies zusätzliche Übung erfordert. Umgekehrt haben wir aber auch mit dem Integralrestglied im Satz zur Taylor-Approximation gesehen, dass das Riemann-Integral nützlich sein kann um den Zusammenhang zwischen den Ableitungen und der ursprünglichen Funktion besser zu verstehen. Falls die symbolische Integration sich als nicht machbar erweist, so ist wiederum die Taylor-Approximation nützlich um das Riemann-Integral numerisch – zum Beispiel mit der Simpson-Methode – mit überraschend hoher Genauigkeit zu berechnen. Ebenso ist aber der Satz zur Taylor-Approximation auch von theoretischer Wichtigkeit, da wir mit diesen asymptotische Formeln wie zum Beispiel die Sterling-Formel beweisen können.

Zusammenfassend können wir also sagen, dass dieses Kapitel den Aufbau der eindimensionalen Analysis vollendet. Die Inhalte dieses Kapitels bilden einen zentralen Bestandteil der Analysis I/II Vorlesung und ihrer Anwendungen.

### 8.8.2 Übungen

Da die Schwierigkeit beim Integrieren vor allem in der Auswahl der richtigen Methode liegt, wollen wir in folgender Übung noch einige weitere Aufgaben ohne Angabe der richtigen Technik auflisten. Dazu wollen wir noch erwähnen, dass es oft auch mehr als eine Methode gibt, die zum Erfolg führen kann.

**Übung.** *Berechnen Sie die unbestimmten Integrale*

$$\begin{array}{ll}
\int \frac{1}{\sqrt{x^2+2017^2}}\,\mathrm{d}x, & \int \frac{1}{x\sqrt{1+x^2}}\,\mathrm{d}x, \\
\int \exp(\sqrt[3]{x})\,\mathrm{d}x, & \int \frac{1}{1+\exp(-x)}\,\mathrm{d}x, \\
\int \frac{x}{1+\sqrt{x}}\,\mathrm{d}x, & \int \cos^5(x)\sin^3(x)\,\mathrm{d}x, \\
\int \cos^2(x)\sin^2(x)\,\mathrm{d}x, & \int \frac{1}{\cos^2(x)\sin^2(x)}\,\mathrm{d}x, \\
\int \frac{\log(x)+\sin(\log(x))x^2}{x^3}\,\mathrm{d}x, & \int \frac{1}{\cos(x)}\,\mathrm{d}x
\end{array}$$

**Übung.** *Zeigen Sie, dass jede stetig differenzierbare Funktion $f : [a,b] \to \mathbb{R}$ auf einem kompakten Intervall $[a,b]$ mit Endpunkten $a < b$ beschränkte Variation $V(f)$ (siehe die entsprechende Übung in Abschnitt 4.8.2) hat und dass $V(f)(x) = \int_a^x |f'(t)|\,\mathrm{d}t$ gilt für alle $x \in [a,b]$.*

**Übung** (Abel-Summation und Partielle Integration)**.** *Wir wollen in dieser Übung den Zusammenhang zwischen der Abel-Summation und der partiellen Integration erklären. Sei also $a < b$ und seien $u, v : [a,b] \to \mathbb{C}$ zwei stetig differenzierbare Funktionen. Verwenden Sie die Abel-Summation von Übung 3.3, um die partielle Integration $\int_a^b uv'\,\mathrm{d}x = [uv]_a^b - \int_a^b u'v\,\mathrm{d}x$ zu beweisen.*

**Übung** (Ein alternativer Beweis von Proposition 4.30)**.** *Wir möchten hier einen Beweis von Proposition 4.30 unter der Annahme, dass $f$ stetig ist, durchführen. Gehen Sie wie folgt vor:*

*(i) Zeigen Sie, dass die Abbildung $F : x \in [a,b] \mapsto \mathcal{I}(a,x)$ differenzierbar ist und dass $F' = f$.*

*(ii) Verwenden Sie nun den Fundamentalsatz der Integral- und Differentialrechnung, um auf Proposition 4.30 zu schliessen.*

**Übung** (Volumen einer Vase)**.** *Berechnen Sie das Volumen der Vase*

$$\left\{(x,y,z) \in \mathbb{R}^3 \mid x \in [-\pi, 2\pi],\ 0 \leq \sqrt{y^2+z^2} \leq \sin(x)+2\right\}.$$

**Übung** (Gleichmässige Konvergenz der Ableitungen)**.** *Sei $(f_n)_n$ eine Folge stetig differenzierbarer reellwertiger Funktionen auf einem kompakten Intervall $[a,b]$ mit Endpunkten $a < b$. Angenommen die Folge $(f_n)_n$ konvergiert gleichmässig gegen eine Funktion $f : [a,b] \to \mathbb{R}$ und die Folge der Ableitungen $(f_n')_n$ konvergiert gleichmässig gegen eine Funktion $g : [a,b] \to \mathbb{R}$. Zeigen Sie, dass $f$ stetig differenzierbar ist mit $f' = g$.*

**Übung.** *Wir verwenden obige Übung, um eine glatte Funktion zu konstruieren, deren Taylorreihe um einen Punkt Konvergenzradius Null hat.*

*(i) Zeigen Sie, dass die Reihe $\sum_{n=0}^{\infty} \mathrm{e}^{-n}\cos(n^2x)$ für alle $x \in \mathbb{R}$ absolut konvergiert.*

*(ii) Nach (i) definieren wir die Funktion*

$$f : x \in \mathbb{R} \mapsto \sum_{n=0}^{\infty} \mathrm{e}^{-n}\cos(n^2x).$$

*Zeigen Sie, dass $f$ glatt ist.*

*(iii) Berechnen Sie die Taylorreihe von $f$ um Null und deren Konvergenzradius.*

**Übung** (Krümmung ebener Kurven)**.** *In dieser Übung möchten wir die Krümmung ebener Kurven betrachten. Alle Kurven, die wir dabei betrachten wollen, sollen zweimal differenzierbar, regulär und einfach sein, hier der Einfachheit vorerst inklusive den Endpunkten. Für eine solche Kurve $\gamma : [a,b] \to \mathbb{R}^2$ und ein $p \in \mathbb{R}^2$ auf der Kurve sei $t \in [a,b]$ der eindeutige Zeitpunkt mit $\gamma(t) = p$. Dann ist die Krümmung von $\gamma$ bei $p$ definiert als*

$$\kappa_\gamma(p) = \frac{\langle \dot\gamma(t), R\ddot\gamma(t)\rangle}{\|\dot\gamma(t)\|^3},$$

*wobei $R = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Ist $\gamma$ einfach, aber erfüllt $\gamma(a) = \gamma(b)$, so reicht es anzunehmen, dass $\dot\gamma(a) = \dot\gamma(b)$ und $\ddot\gamma(a) = \ddot\gamma(b)$ gilt, damit obiger Ausdruck für die Krümmung Sinn ergibt.*

*(i) Berechnen Sie die Krümmung der Kurven $t \mapsto (\cos(t), \sin(t))$ und $t \mapsto (\cosh(t), \sinh(t))$ (definiert auf geeigneten Intervallen).*

*(ii) Zeigen Sie, dass die Krümmung unabhängig ist von der Parametrisierung, das heisst, dass für jede Reparametrisierung $\gamma \circ \psi$ einer Kurve $\gamma$ wie oben gilt $\kappa_{\gamma\circ\psi}(p) = \kappa_\gamma(p)$ für alle Punkte $p$ auf der Kurve.*

**Übung** (Existenz von Kurven vorgegebener Krümmung)**.** *Wie in vorheriger Übung möchten wir hier die Krümmung ebener Kurven betrachten, aber dabei zulassen, dass die betrachteten Kurven nicht einfach sind, womit für eine reguläre Kurve $\gamma : [a, b] \to \mathbb{R}^2$ die Krümmung $\kappa_\gamma$ definiert ist als Funktion auf $[a, b]$ via $\kappa_\gamma(t) = \frac{\langle \dot\gamma(t), R\ddot\gamma(t)\rangle}{\|\dot\gamma(t)\|^3}$.*

*Sei nun $\kappa : [a, b] \to \mathbb{R}$ eine beliebige stetige Funktion auf einem Intervall $[a, b]$ mit Endpunkten $a < b$. Wir möchten hier zeigen, dass eine zweimal stetig differenzierbare Kurve $\gamma$ mit Krümmungsfunktion $\kappa$ existiert. Dazu betrachten wir die Funktion*

$$\theta : [a, b] \to \mathbb{R}^2,\ s \mapsto \int_a^s \kappa(r)\,\mathrm{d}r$$

*und setzen*

$$\gamma : [a, b] \to \mathbb{R}^2,\ t \mapsto \int_a^t (\cos(\theta(s)), \sin(\theta(s))\,\mathrm{d}s.$$

*Zeigen Sie, dass die Komponenten von $\gamma$ jeweils zweimal stetig differenzierbar ist und dass $\gamma$ eine reguläre, nach Bogenlänge parametrisierte Kurve ist. Verifizieren Sie anschliessend, dass $\kappa_\gamma = \kappa$ gilt.*

**Übung** (Irrationalität der Kreiszahl)**.** *In dieser Übung möchten wir zeigen, dass $\pi$ irrational ist, wobei wir dem Beweis von Niven [Niv47] folgen werden.*

*Per Widerspruch wollen wir annehmen, dass $\pi = \frac{a}{b}$ ist für $a, b \in \mathbb{N}$. Nun betrachten wir die Polynome*

$$\begin{aligned} f(x) &= \frac{x^n(a - bx)^n}{n!}, \\ F(x) &= f(x) - f^{(2)}(x) + f^{(4)}(x) - \ldots + (-1)^n f^{(2n)}(x) \end{aligned}$$

*für ein $n \in \mathbb{N}$, welches wir später wählen werden.*

*(i) Begründen Sie, wieso $f$, alle Ableitungen von $f$ und $F$ bei $0$ und bei $\pi$ ganzzahlige Werte annehmen. Verifizieren Sie des Weiteren, dass $f|_{[0,\pi]}$ eine nicht-negative Funktion ist, welche genau bei $0$ und $\pi$ verschwindet.*

*(ii) Zeigen Sie, dass*

$$(F'(x)\sin(x) - F(x)\cos(x))' = f(x)\sin(x)$$

*und $\int_0^\pi f(x)\sin(x)\,\mathrm{d}x = F(\pi) + F(0)$.*

*(iii) Schliessen Sie auf einen Widerspruch.*

**Übung** (Summe der Reziproken der Primzahlen)**.** *In dieser Übung möchten wir für natürliche Zahlen $n \in \mathbb{N}$ die Summe $\sum_{p\in\mathbb{P}:p\leq n} \frac{1}{p}$ betrachten, wobei $\mathbb{P} \subseteq \mathbb{N}$ die Menge der Primzahlen*

*bezeichnet. Dabei möchten wir zeigen, dass*

$$\sum_{p\in\mathbb{P}:p\leq n}\frac{1}{p}\geq\log(\log(n))-\log(C)$$

*für alle $n\geq 3$ und eine Konstante $C>1$. Insbesondere gibt es unendlich viele Primzahlen (wieso?). Die obige Ungleichung stellt eine (stark) abgeschwächte Version des zweiten Theorems von Mertens (und damit einen Vorreiter des Primzahlsatzes) dar.*

*(i) Verifizieren Sie für alle $n\in\mathbb{N}$ die Ungleichung*

$$\exp\left(\sum_{p\in\mathbb{P}:p\leq n}\frac{1}{p}\right)\geq\prod_{p\in\mathbb{P}:p\leq n}\left(1+\frac{1}{p}\right).$$

*(ii) Zeigen Sie für alle $n\in\mathbb{N}$*

$$\sum_{k=1}^{n}\frac{1}{k}\leq\left(\sum_{\ell=1}^{n}\frac{1}{\ell^2}\right)\prod_{p\in\mathbb{P}:p\leq n}\left(1+\frac{1}{p}\right)$$

*(iii) Schliessen Sie auf die Aussage.*

### 8.8.3 Lernkarten

Sie können wiederum die Lernkarten oder den Graphen für Ihre Wiederholung der Themen des Kapitels verwenden.

# Kapitel 9

# Metrische Räume

Dieses Kapitel am Anfang des zweiten Semesters stellt einen wichtigen Schritt der Verallgemeinerung dar. In der Tat möchten wir uns von der fast ausschliesslichen Betrachtung von Intervallen als „Grundräume“ wegbewegen und allgemeinere Räume, wo Distanzen gemessen werden können, einführen. In diesem allgemeineren Rahmen werden wir einige wichtige Sätze rund um Konvergenz und Stetigkeit, die uns mitunter bereits für Intervalle bekannt sind, beweisen können. Dies stellt zum einen eine willkommene Wiederholung dar und verschafft uns zum anderen den notwendigen, breiteren Blickwinkel für die weitere Theorie wie zum Beispiel für den Beweis des Fundamentalsatzes der Algebra oder für die Fortsetzung der Diskussion von Differentialgleichungen.

## 9.1 Konvergenz in metrischen Räumen

### 9.1.1 Definition und erste Beispiele

**Definition 9.1** (Metrik)**.** Ein **metrischer Raum** $(X, \mathrm{d})$ ist eine Menge $X$ gemeinsam mit einer Abbildung $\mathrm{d} : X \times X \to \mathbb{R}_{\geq 0}$, die die **Metrik** auf $X$ genannt wird und die folgenden drei Eigenschaften erfüllt:

- (Definitheit) Für alle $x_1, x_2 \in X$ gilt $\mathrm{d}(x_1, x_2) = 0 \iff x_1 = x_2$.
- (Symmetrie) Für alle $x_1, x_2 \in X$ gilt $\mathrm{d}(x_1, x_2) = \mathrm{d}(x_2, x_1)$.
- (Dreiecksungleichung) Für alle $x_1, x_2, x_3 \in X$ gilt $\mathrm{d}(x_1, x_3) \leq \mathrm{d}(x_1, x_2) + \mathrm{d}(x_2, x_3)$.

Intuitiv ausgedrückt weist eine Metrik d auf einer Menge $X$ je zwei Punkten ihre **Distanz** (ihren **Abstand**) zu. In dieser Auffassung besagt die Definitheit der Metrik, dass der einzige Punkt, der Abstand Null zu einem gegebenen Punkt $x_1 \in X$ hat, $x_1$ selbst ist. Symmetrie der Metrik besagt, dass der Abstand von $x_1 \in X$ zu $x_2 \in X$ der gleiche ist wie von $x_2$ zu $x_1$. Fasst man die Distanz zwischen zwei Punkten als die Länge eines kürzesten Weges vom einen zum anderen Punkt auf (was nicht immer möglich ist), dann besagt die Dreiecksungleichung, dass die Länge eines kürzesten Weges von $x_1$ nach $x_3$ höchstens so gross ist wie die Länge eines Weges, den man abläuft, wenn man zuerst den Umweg nach $x_2$ und von dort aus nach $x_3$ geht.

Folgende Beispiele von Metriken sind uns eigentlich bereits bekannt – siehe Lemma 9.2 unten:

- $X = \mathbb{R}$ mit der Standardmetrik d definiert durch $\mathrm{d}(x_1, x_2) = |x_1 - x_2|$ für $x_1, x_2 \in \mathbb{R}$.
- $X = \mathbb{C}$ mit der Standardmetrik d definiert durch $\mathrm{d}(z_1, z_2) = |z_1 - z_2|$ für $z_1, z_2 \in \mathbb{C}$.
- $X = \mathbb{R}^d$ mit der Einsmetrik $\mathrm{d}_1$ definiert durch $\mathrm{d}_1(v_1, v_2) = \|v_1 - v_2\|_1$ für $v_1, v_2 \in \mathbb{R}^d$.
- $X = \mathbb{R}^d$ mit der euklidischen Metrik $\mathrm{d}_2$ definiert durch $\mathrm{d}_2(v_1, v_2) = \|v_1 - v_2\|_2$ für $v_1, v_2 \in \mathbb{R}^d$.
- $X = \mathbb{R}^d$ mit der Maximumsmetrik $\mathrm{d}_\infty$ definiert durch $\mathrm{d}_\infty(v_1, v_2) = \|v_1 - v_2\|_\infty$ für $v_1, v_2 \in \mathbb{R}^d$.

Für $X = \mathbb{R}^d$ werden wir im Normalfall die euklidische Metrik $d_2$ benützen und diese auch einfach mit $\mathrm{d} = \mathrm{d}_2$ bezeichnen. Da wir ab nun vermehrt Vektoren anstatt nur reelle oder komplexe Zahlen betrachten werden, werden wir diese einfach durch $v \in \mathbb{R}^d$ (anstatt $\mathbf{v} \in \mathbb{R}^d$) bezeichnen.

Wie wir auch sehen werden, gibt es viele weitere, interessante Beispiele von metrischen Räumen. Manche dieser erhalten wir mittels Normen $\|\cdot\|$ auf Vektorräumen (siehe Definition 5.84).

**Lemma 9.2** (Eine Norm definiert eine Metrik)**.** *Sei $V$ ein Vektorraum über*[1] $\mathbb{R}$ *und $\|\cdot\|$ eine Norm auf $V$. Dann definiert*

$$\mathrm{d}(v_1, v_2) = \mathrm{d}_{\|\cdot\|}(v_1, v_2) = \|v_1 - v_2\|$$

*für $v_1, v_2 \in V$ eine Metrik* d *auf $V$, die man auch die von der Norm $\|\cdot\|$* ***induzierte Metrik*** *auf $V$ nennt.*

*Beweis.* Es gilt für $v_1, v_2 \in V$

$$\begin{aligned} \mathrm{d}_{\|\cdot\|}(v_1, v_2) = 0 &\iff \|v_1 - v_2\| = 0 \\ &\iff v_1 - v_2 = 0 \\ &\iff v_1 = v_2 \end{aligned}$$

nach Definitheit der Norm $\|\cdot\|$. Nach Homogenität der Norm für $\alpha = -1$ gilt für $v_1, v_2 \in V$

$$\begin{aligned} \mathrm{d}_{\|\cdot\|}(v_1, v_2) &= \|v_1 - v_2\| = \|(-1)(v_2 - v_1)\| \\ &= \|v_2 - v_1\| = \mathrm{d}_{\|\cdot\|}(v_2, v_1) \end{aligned}$$

[1] Hier und auch im Folgenden können wir ebenso Vektorräume über $\mathbb{C}$ betrachten, doch inkludiert der Fall der reellen Vektorräume auch den Fall von komplexen Vektorräumen, weshalb wir Vektorräume über $\mathbb{C}$ hier und im Folgenden nicht mehr getrennt erwähnen werden.

und somit erhalten wir die Symmetrie von $\mathrm{d}_{\|\cdot\|}$. Zuletzt verwenden wir die Dreiecksungleichung der Norm und erhalten

$$\begin{aligned}\mathrm{d}_{\|\cdot\|}(v_1, v_3) &= \|v_1 - v_3\| = \|(v_1 - v_2) + (v_2 - v_3)\| \\ &\leq \|v_1 - v_2\| + \|v_2 - v_3\| = \mathrm{d}_{\|\cdot\|}(v_1, v_2) + \mathrm{d}_{\|\cdot\|}(v_2, v_3)\end{aligned}$$

für alle $v_1, v_2, v_3 \in V$. Dies zeigt die Dreiecksungleichung für $\mathrm{d}_{\|\cdot\|}$, womit also $\mathrm{d}_{\|\cdot\|}$ eine Metrik auf $V$ ist. □

Nicht jede Metrik auf einem Vektorraum muss durch eine Norm gegeben sein. Des Weiteren ist das Messen von Distanzen nicht nur auf Vektorräumen von Interesse. Interessante Beispiele dieser Art möchten wir nun besprechen.

**Beispiel 9.3** (Weitere metrische Räume)**.**

*(i) (Diskrete Metriken) Sei $X$ eine Menge und $\mathrm{d}_{\text{diskret}} : X \times X \to \mathbb{R}_{\geq 0}$ definiert durch*

$$\mathrm{d}_{\text{diskret}}(x_1, x_2) = \begin{cases} 1 & \text{falls } x_1 \neq x_2 \\ 0 & \text{falls } x_1 = x_2 \end{cases}$$

*für $x_1, x_2 \in X$. Dann ist $(X, \mathrm{d}_{\text{diskret}})$ ein metrischer Raum. In der Tat ist $\mathrm{d}_{\text{diskret}}$ definit und symmetrisch per Definition. Des Weiteren erfüllt $\mathrm{d}$ die Dreiecksungleichung: Seien $x_1, x_2, x_3$ Punkte in $X$. Falls $\mathrm{d}(x_1, x_3) = 0$ gilt, dann ist $\mathrm{d}(x_1, x_3) \leq \mathrm{d}(x_1, x_2) + \mathrm{d}(x_2, x_3)$ trivialerweise erfüllt. Falls $\mathrm{d}(x_1, x_3) = 1$ gilt, dann ist $x_1 \neq x_3$ und $x_2$ ist mindestens von einem Punkt in $\{x_1, x_3\}$ verschieden und die Dreiecksungleichung gilt ebenso.*

*Man beachte, dass die diskrete Metrik auf $\mathbb{R}^d$ für $d \geq 2$ nicht durch eine Norm gegeben ist. In der Tat würde eine Norm $\|\cdot\|$ mit $\|v_2 - v_1\| = \mathrm{d}_{\text{diskret}}(v_1, v_2)$ für alle $v_1, v_2 \in \mathbb{R}^d$ widersprüchlicherweise die Homogenitätseigenschaft in Definition 5.84 nicht erfüllen können.*

*(ii) (Manhattanmetrik) Wir setzen $X = [0, 1]^2$ und*

$$\mathrm{d}_{\mathrm{NY}}((x_1, y_1), (x_2, y_2)) = |x_1 - x_2| + |y_1 - y_2|$$

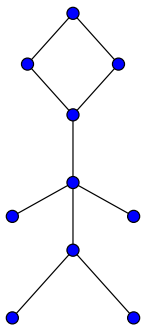
*für $(x_1, y_1), (x_2, y_2) \in [0, 1]^2$. In der Tat erfüllt $\mathrm{d}_{\mathrm{NY}}$ alle Axiome einer Metrik auf $[0, 1]^2$, da $\mathrm{d}_{\mathrm{NY}} = \mathrm{d}_1|_{X \times X}$ die Einschränkung der Einsmetrik $\mathrm{d}_1$ von $\mathbb{R}^2$ auf $X = [0, 1]^2$ ist. Die Metrik $\mathrm{d}_{\mathrm{NY}}$ wird oft auch Manhattan-Metrik genannt. Grund dafür ist, dass man in schachbrettartig angelegten Orten wie zum Beispiel Manhattan auf folgende Weise von $(x_1, y_1)$ nach $(x_2, y_2)$ gelangt: Man geht zuerst bei gleichbleibender $y$-Koordinate von $(x_1, y_1)$ nach $(x_2, y_1)$ und dann bei gleichbleibender $x$-Koordinate von $(x_2, y_1)$ nach $(x_2, y_2)$, oder umgekehrt von $(x_1, y_1)$ nach $(x_1, y_2)$ und dann von $(x_1, y_2)$ nach $(x_2, y_2)$. Es gäbe zwar noch andere Möglichkeiten, aber wenn alle Strassen in Manhattan von West-Ost oder Nord-Süd verlaufen, dann misst $\mathrm{d}_{\mathrm{NY}}$ den relevanten Abstand zwischen zwei Punkten.*

*(iii)* *(Metrik der französischen Eisenbahn) Wir setzen $X = \mathbb{C}$ und definieren die SNCF-Metrik* $\mathrm{d}_{\mathrm{SNCF}}$ *auf $X$ durch*

$$\mathrm{d}_{\mathrm{SNCF}}(z_1, z_2) = \begin{cases} |z_1 - z_2| & \text{falls } z_1, z_2 \text{ linear abhängig über } \mathbb{R} \text{ sind} \\ |z_1| + |z_2| & \text{falls } z_1, z_2 \text{ linear unabhängig über } \mathbb{R} \text{ sind} \end{cases}$$

*für alle $z_1, z_2 \in \mathbb{C}$. Der Grund für den Namen dieser Metrik (siehe Übung 9.4) ist, dass eine Bahnreise von einer französischen Stadt bei $z_1$ zu einer anderen bei $z_2$ meist über den Ursprung $z = 0$ (auch Paris genannt) führt, ausser wenn $z_1$ und $z_2$ auf derselben – von Paris ausgehenden geraden Strecke liegen. Gewissermassen besteht $\mathbb{C}$ in dieser Metrik also aus unendlich vielen Halbgeraden, die sich nur im Ursprung treffen.*

*(iv)* *Ein kombinatorischer Graph ist eine endliche Menge von Punkten, die sogenannten Ecken, von welchen einige mit sogenannten Kanten verbunden sind. Diese lassen sich auf natürliche Weise mit mehreren Metriken ausstatten; der Konkretheit halber betrachten wir einen spezifischen Graphen, doch muss ein Graph nicht unbedingt als Teilmenge von $\mathbb{R}^d$ für $d \geq 2$ gegeben sein.*



*Man kann nun eine Metrik auf den Ecken (durch • gekennzeichnet) dadurch definieren, dass man benachbarten Ecken die Distanz 1 zuweist und dies iteriert. Beispielsweise definiert man die Distanz zweier Ecken, die man über zwei aber nicht weniger Kanten erreichen kann, als 2. Dazu notwendig ist, dass man von einer Ecke zu jeder anderen Ecke über Ablaufen von Kanten gelangen kann (wie bei obigem Graphen) – diese Eigenschaft nennt sich auch Zusammenhang des Graphen.*

*Des Weiteren ist es auch möglich, eine Metrik auf dem kompletten (kontinuierlichen) Graphen zu definieren, indem man die obige Definition auf folgende Weise erweitert. Fasst man eine Kante als Kopie des Intervalles $[0, 1]$ auf, wobei 0 und 1 die zwei Ecken der Kante sind, so kann man eine Distanz auf den Kanten über die Distanz $|x - y|$ auf $[0, 1]$ definieren. Ähnlich wie oben kann man nun damit eine Metrik auf dem gesamten Graphen (inklusive den Kanten) definieren.*

Weitere interessante Beispiele von Metriken finden Sie in Abschnitt 9.9.2.

**Übung 9.4.** *Zeigen Sie, dass die in Beispiel 9.3(iii) definierten Metriken tatsächlich Metriken sind. Führen Sie des Weiteren die Konstruktion der Metriken in (iv) vollständig und formal durch.*

Wir bemerken, dass gegeben eine Teilmenge $Y$ eines metrischen Raumes $(X, \mathrm{d})$ die Einschränkung $\mathrm{d}|_{Y\times Y}$ eine Metrik auf $Y$ definiert. Wenn $Y$ mit dieser Metrik versehen ist, nennen wir $Y$ einen **Teilraum** des metrischen Raumes $(X, \mathrm{d})$ und die Metrik $\mathrm{d}|_{Y\times Y}$ die **induzierte Metrik**. Wenn nicht anders spezifiziert, statten wir Teilmengen eines metrischen Raumes implizit mit der induzierten Metrik aus.

**Wichtige Übung 9.5** (Umgekehrte Dreiecksungleichung)**.** *Sei $(X, \mathrm{d})$ ein metrischer Raum. Zeigen Sie, dass für alle $x_1, x_2, y \in X$ gilt*

$$|\mathrm{d}(x_1, y) - \mathrm{d}(x_2, y)| \leq \mathrm{d}(x_1, x_2).$$

**Übung 9.6** (Deformation der Metrik)**.** *Sei $(X, \mathrm{d})$ ein metrischer Raum. Zeigen Sie, dass*

$$\mathrm{d}_{1/2}(x, y) = \sqrt{\mathrm{d}(x, y)} \text{ und } \tilde{\mathrm{d}}(x, y) = \frac{\mathrm{d}(x, y)}{1 + \mathrm{d}(x, y)}$$

*für $x, y \in X$ zwei Metriken $d_{\frac{1}{2}}$, $\tilde{d}$ auf $X$ definieren.*

### 9.1.2 Konvergenz von Folgen

Zu Beginn von Kapitel 5 haben wir bereits Folgen in beliebigen Mengen definiert. Wir möchten nun Folgen in metrischen Räumen betrachten und für diese genauso wie für Folgen in $\mathbb{R}$ oder $\mathbb{C}$ einen Konvergenzbegriff einführen.

**Definition 9.7** (Konvergenz)**.** Sei $(X, \mathrm{d})$ ein metrischer Raum und $(x_n)_n$ eine Folge in $X$. Wir sagen, dass $(x_n)_n$ gegen einen Punkt $x_0 \in X$ **konvergiert**, falls es für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass $\mathrm{d}(x_n, x_0) < \varepsilon$ für alle $n \geq N$. In diesem Fall schreiben wir $\lim_{n\to\infty} x_n = x_0$ oder $x_n \to x_0$ für $n \to \infty$. Der Punkt $x_0$ nennt sich dabei der **Grenzwert** der Folge $(x_n)_n$. Wir sagen, dass eine Folge $(x_n)_n$ mit $x_n \in X$ für alle $n \in \mathbb{N}$ **in $X$ konvergent ist** falls es ein $x_0 \in X$ mit $\lim_{n\to\infty} x_n = x_0$ gibt.

Wir bemerken zuerst, dass $x_n \to x_0$ für $n \to \infty$ für eine Folge $(x_n)_n$ und einen Punkt $x_0 \in X$ in einem metrischen Raum $(X, \mathrm{d})$ per Definition zur Konvergenz $\mathrm{d}(x_n, x_0) \to 0$ in $\mathbb{R}$ für $n \to \infty$ äquivalent ist (wieso?).

Des Weiteren ist der Grenzwert einer Folge eindeutig bestimmt, wenn er existiert, was mit dem gleichen Argument wie für reelle Folgen gezeigt wird. (Wie?) Man darf somit in der Tat von *dem* Grenzwert sprechen.

Wir bemerken auch, dass Definition 9.7 kompatibel ist mit dem Konvergenzbegriff für normierte Räume (Definition 5.91) in folgendem Sinne. Sei $V$ ein Vektorraum über $\mathbb{R}$ und $\|\cdot\|$ eine Norm auf $V$. Ist $(v_n)_n$ eine Folge in $V$, dann konvergiert $(v_n)_n$ gegen ein $v_0 \in V$ genau dann bezüglich der Norm $\|\cdot\|$, wenn $(v_n)_n$ bezüglich der induzierten Metrik gegen $v_0$ konvergiert.

**Definition 9.8** (Offene Bälle und Umgebungen)**.** Sei $(X, \mathrm{d})$ ein metrischer Raum. Für ein $\varepsilon > 0$ und einen Punkt $x_0 \in X$ nennt man

$$B_\varepsilon(x_0) = \{x \in X \mid \mathrm{d}(x, x_0) < \varepsilon\}$$

den **offenen Ball** mit Radius $\varepsilon$ um $x_0$ oder auch die $\varepsilon$-**Umgebung** von $x_0$. Eine **Umgebung** von $x_0 \in X$ ist eine Teilmenge $V \subseteq X$, die eine $\varepsilon$-Umgebung von $x_0$ für ein $\varepsilon > 0$ enthält.

**Wichtige Übung 9.9** (Offene Bälle)**.** *Beschreiben Sie die offenen Bälle in folgenden metrischen Räumen.*

- $(X, \mathrm{d}_{\text{diskret}})$ *für eine Menge* $X$ *und die diskrete Metrik* $\mathrm{d}_{\text{diskret}}$ *auf* $X$*.*
- $[0,1]^2$ *mit der Manhattanmetrik.*
- $\mathbb{C}$ *mit der französischen Eisenbahnmetrik.*
- $\mathbb{R}^2$ *mit der Metrik* $\mathrm{d}_\infty$ *gegeben durch* $\mathrm{d}_\infty((x_1, y_1), (x_2, y_2)) = \max\{|x_1 - x_2|, |y_1 - y_2|\}$ *für* $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$*.*

Die obige Definition von Umgebungen erlaubt nun eine alternative Formulierung von Konvergenz: Eine Folge $(x_n)_n$ in einem metrischen Raum $(X, \mathrm{d})$ konvergiert genau dann gegen $x_0 \in X$, wenn für jede Umgebung $V \subseteq X$ von $x_0$ **fast alle** (das heisst, alle bis auf endlich viele) Folgenglieder von $(x_n)_n$ in $V$ liegen (wieso?).

Folgende Übung zeigt, dass der Konvergenzbegriff von der Wahl der Metrik abhängt.

**Übung 9.10** (Manhattan und SNCF sind sehr verschieden)**.** *Sei* $X = [0,1]^2$*. Finden Sie eine Folge in* $X$*, die zwar bezüglich der Manhattanmetrik, aber nicht bezüglich der französischen Eisenbahnmetrik konvergiert (wobei wir* $\mathbb{R}^2$ *mit* $\mathbb{C}$ *und damit* $X$ *mit einer Teilmenge von* $\mathbb{C}$ *identifizieren).*

### 9.1.3 Der Raum der stetigen Funktionen

Wir kennen bereits einige Normen auf endlich-dimensionalen Vektorräumen und werden im Verlaufe dieses Kapitels weitere kennenlernen. Für die Analysis sind allerdings nicht nur endlich-dimensionale normierte Vektorräume interessant, sondern oft auch unendlich-dimensionale. Häufig (zum Beispiel bei der Diskussion von Differentialgleichungen) werden dabei sogenannte Funktionenräume untersucht.

**Beispiel 9.11** (Supremumsnorm und Gleichmässige Konvergenz)**.** *Sei* $K = [a, b]$ *ein kompaktes Intervall mit* $a < b$ *in* $\mathbb{R}$*. Wir betrachten den Vektorraum* $V = C(K)$ *der stetigen reellwertigen Funktionen auf* $K$ *und definieren*

$$\|f\|_\infty = \sup_{x \in K} |f(x)| = \max_{x \in K} |f(x)|$$

*für* $f \in C(K)$*. Dann ist* $\|\cdot\|_\infty$ *eine Norm auf* $C(K)$*. Es gilt Definitheit, denn für alle* $f \in C(K)$ *ist* $\|f\|_\infty \geq 0$ *per Definition von* $\|\cdot\|_\infty$ *und* $\|f\|_\infty = 0$ *genau dann, wenn* $f = 0$*. Für* $\alpha \in \mathbb{C}$

*und $f \in C(K)$ gilt des Weiteren*

$$\|\alpha f\|_\infty = \max_{x\in K} |\alpha||f(x)| = |\alpha| \max_{x\in K} |f(x)| = |\alpha|\|f\|_\infty$$

*und somit Homogenität von $\|\cdot\|_\infty$. Schlussendlich gilt für $f_1, f_2 \in C(K)$ auch*

$$\begin{aligned}\|f_1 + f_2\|_\infty &= \max_{x\in K} |f_1(x) + f_2(x)| \\ &\leq \max_{x\in K} \big(|f_1(x)| + |f_2(x)|\big) \\ &\leq \max_{x\in K} |f_1(x)| + \max_{x\in K} |f_2(x)| = \|f_1\|_\infty + \|f_2\|_\infty,\end{aligned}$$

*womit wir die Dreiecksungleichung bewiesen haben und gezeigt haben, dass $\|\cdot\|_\infty$ eine Norm auf $V$ ist.*

*Wir beschreiben nun Konvergenz in $C(K)$ mit der von $\|\cdot\|_\infty$ induzierten Metrik. Sei $(f_n)_n$ eine Folge in $C(K)$ und $f_0 \in C(K)$. Die Funktionenfolge $(f_n)_n$ konvergiert bezüglich der von $\|\cdot\|_\infty$ induzierten Metrik genau dann gegen $f_0$, wenn $(f_n)_n$ gleichmässig gegen $f_0$ konvergiert. In der Tat, falls $\|f_n - f_0\|_\infty \to 0$ für $n \to \infty$ und $\varepsilon > 0$ ist, dann existiert ein $N$ mit $\|f_n - f_0\|_\infty < \varepsilon$ für alle $n \geq N$, was wiederum $|f_n(x) - f(x)| < \varepsilon$ für alle $x \in K$ und $n \geq N$ impliziert. Diese Implikationen lassen sich auch umkehren, was die Behauptung beweist.*

**Übung 9.12** (1-Norm und Konvergenz im Mittel)**.** *Sei $K = [a, b]$ ein kompaktes Intervall mit $a < b$ in $\mathbb{R}$. Wir betrachten den Vektorraum $V = C([a, b])$ und definieren*

$$\|f\|_1 = \int_a^b |f(x)|\,\mathrm{d}x$$

*für $f \in V$.*

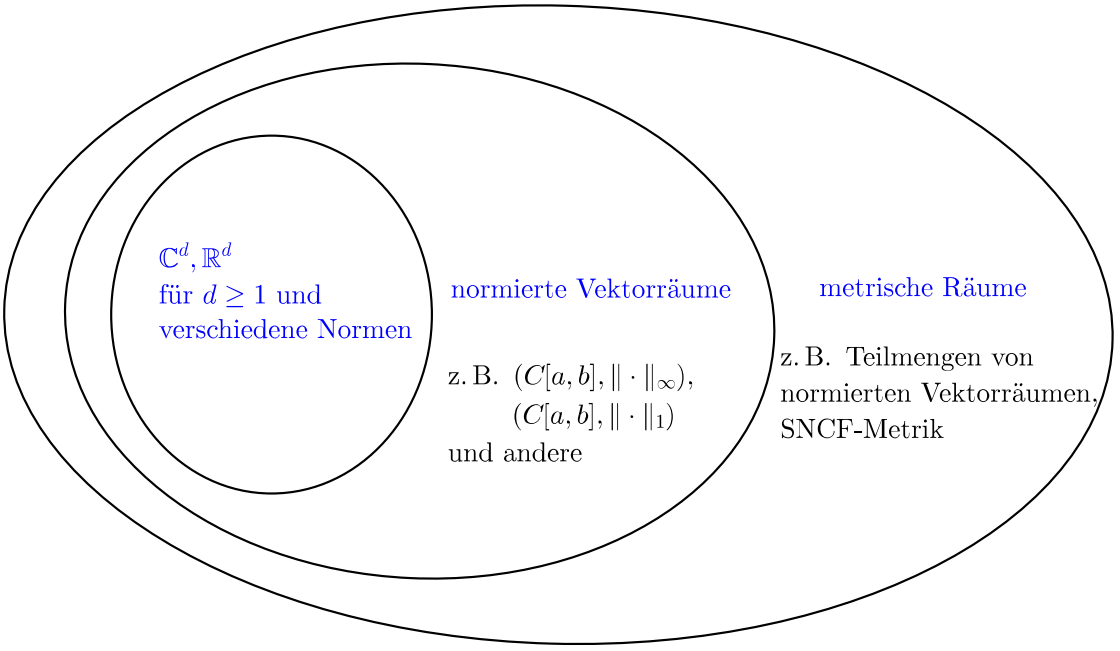
*(i) Zeigen Sie, dass $\|\cdot\|_1$ eine Norm auf $V$ ist. Konvergenz in $V$ bezüglich $\|\cdot\|_1$ nennt man Konvergenz im Mittel.*

*(ii) Sei $(f_n)_n$ eine Folge in $V$. Zeigen Sie, dass Konvergenz $f_n \to f$ für $n \to \infty$ bezüglich $\|\cdot\|_\infty$ für ein $f \in V$ auch die Konvergenz $f_n \to f$ für $n \to \infty$ bezüglich $\|\cdot\|_1$ impliziert.*

*(iii) Finden Sie eine Folge $(f_n)_n$ in $V$ mit $\|f_n\|_1 \to 0$ für $n \to \infty$ und $\|f_n\|_\infty = 1$ für alle $n \in \mathbb{N}$.*

### 9.1.4 Ein kurzer Überblick

Wir fassen die behandelten Begriffe nochmals in einem Diagram zusammen.




Wir werden uns in diesem zweiten Semester vor allem mit $\mathbb{R}^d$ für $d \geq 2$ (mehrdimensionale Analysis) beschäftigen; allerdings werden wir auch $C([a,b])$ mit $\|\cdot\|_\infty$ wie in Beispiel 9.11 oder gewisse Teilmengen von $\mathbb{R}^d$ verwenden. Insbesondere bieten metrische Räume den für uns geeigneten allgemeinen Rahmen.

### 9.1.5 Euklidische Norm

Bei Betrachtung von $\mathbb{R}^d$ (und Teilmengen von $\mathbb{R}^d$) werden wir meist die Norm $\|\cdot\|_2$ (und die von der Norm $\|\cdot\|_2$ auf einer Teilmenge induzierte Metrik) verwenden. Der Grund dafür ist naheliegend, denn diese Norm hat eine klare geometrische Bedeutung, die im Laufe des Semesters zunehmend an Bedeutung gewinnen wird. In der Tat stellt die Definition

$$\|v\|_2 = \sqrt{v_1^2 + \cdots + v_d^2}$$

für $v \in \mathbb{R}^d$ eine mehrdimensionale Verallgemeinerung des Satzes von Pythagoras (siehe die entsprechende Übung in Abschnitt 1.7.6) dar. Des Weiteren können wir für zwei nicht-verschwindende Vektoren $v, w \in \mathbb{R}^d \setminus \{0\}$ die Cauchy-Schwarz-Ungleichung in Proposition 5.86 als die Aussage interpretieren, dass sich das innere Produkt in der Form

$$\langle v, w \rangle = \|v\|\|w\| \cos\varphi$$

schreiben lässt, wobei wir $\varphi = \arccos \frac{\langle v,w \rangle}{\|v\|\|w\|} \in [0, \pi]$ als den **Winkel zwischen** den Vektoren $v$ und $w$ bezeichnen. Wir empfehlen Ihnen, diese Definition an einigen expliziten Beispielen durchzuspielen und sich davon zu überzeugen, dass diese Definition des Winkels den Ihnen intuitiv bekannten Eigenschaften genügt.

Schlussendlich werden wir uns in diesem Semester auch intensiv mit Volumen von Teilmengen von $\mathbb{R}^d$ beschäftigen. Auch hier spielt die euklidischen Norm und das innere Produkt

eine fundamentale Rolle, da wir das Volumen des „$d$-dimensionalen Einheitswürfel“, der von $d$ orthogonalen Vektoren der Länge 1 aufgespannt wird, als 1 definieren werden.

### 9.1.6 Wie sehen metrische Räume aus?

Da metrische Räume sozusagen geometrische Objekte darstellen, drängt sich vielleicht die Frage im Titel des Unterabschnittes auf. Doch ist diese Frage genauso wenig sinnvoll wie zum Beispiel die Frage „Welche Eigenschaften haben chemische Elemente?“. In beiden Fällen hängt die Antwort stark vom betrachteten Spezialfall ab. Zum Beispiel inkludiert die Frage „Wie sehen metrische Räume aus?“ auch die Frage „Wie sehen Teilmengen von $\mathbb{R}^d$ aus?“, denn jede Teilmenge von $\mathbb{R}^d$ kann als eigenständiger metrischer Raum (mit der induzierten Metrik) betrachtet werden. Insbesondere hat auf Grund ihrer Allgemeinenheit diese Frage kaum eine vernünftige Antwort.

Damit Sie sich von der Vielfalt der Antwortmöglichkeiten ein besseres Bild machen können, betrachten wir in den folgenden Applets offene Bälle in einigen wenigen metrische Räumen.

**Applet 9.13** (Bälle in einigen metrischen Räumen)**.** *Die folgenden Apps sollten helfen die Vielfalt der Möglichkeiten für die Gestalt von Bällen in metrischen Räumen visualisieren. Dabei betrachten wir das Einheitsquadrat, die SNCF-Metrik, die Einheitssphäre, den Funktionenraum mit Werten in* $\mathbb{R}$ *und den Funktionenraum mit Werten in* $\mathbb{R}^2$*.*

VO: 21.2.18

## 9.2 Topologische Grundbegriffe

### 9.2.1 Offene und abgeschlossene Teilmengen

Aus Abschnitt 2.4 kennen wir bereits die offenen Teilmengen von $\mathbb{R}$. Dies sind im Wesentlichen jene Teilmengen $U \subseteq \mathbb{R}$, so dass für jedes $x \in U$ Punkte „nahe genug an $x$“ wieder in $U$ liegen. Wir führen nun das analoge Konzept für metrische Räume ein.

**Definition 9.14** (Offene oder abgeschlossene Teilmengen)**.** Sei $(X, \mathrm{d})$ ein metrischer Raum. Eine Teilmenge $O \subseteq X$ heisst **offen** (in $X$), falls es zu jedem Punkt in $O$ einen offenen Ball um diesen Punkt gibt, der in $O$ liegt. Eine Teilmenge $A \subseteq X$ heisst **abgeschlossen** (in $X$), falls ihr Komplement $X \setminus A$ offen ist.

Die leere Menge und die ganze Menge $X$ sind stets sowohl offene als auch abgeschlossene Teilmengen eines metrischen Raumes $(X, \mathrm{d})$. Es kann aber auch weitere (sogenannte abgeschloffene) Teilmengen von $X$ geben, die sowohl offen als auch abgeschlossen sind (siehe Abschnitt 9.2.3).

**Beispiel 9.15** (Offene oder abgeschlossene Teilmengen)**.**

*(i) Sei $(X, \mathrm{d})$ ein metrischer Raum. Sei $x_0 \in X$ ein Punkt und $r > 0$ ein Radius. Der Ball $B_r(x_0) = \{x \in X \mid \mathrm{d}(x, x_0) < r\}$ mit Radius $r$ um $x_0$ ist in der Tat eine offene Teilmenge von $X$. Sei $x \in B_r(x_0)$. Dann gilt $\mathrm{d}(x, x_0) < r$ und somit $\varepsilon := r - \mathrm{d}(x, x_0) > 0$. Der Ball mit Radius $\varepsilon$ um $x$ ist damit in $B_r(x_0)$ enthalten, denn es gilt für $y \in B_\varepsilon(x)$ per Definition $\mathrm{d}(y, x) < \varepsilon$ und somit $\mathrm{d}(y, x_0) \leq \mathrm{d}(y, x) + \mathrm{d}(x, x_0) < \varepsilon + \mathrm{d}(x, x_0) = r$ und $y \in B_r(x_0)$.*

*(ii) Sei $A = \{x_1, \ldots, x_n\} \subseteq X$ eine endliche Teilmenge eines metrischen Raumes $(X, \mathrm{d})$. Dann ist $A$ abgeschlossen. Sei $x \in X \setminus A$ und sei $r = \min\{\mathrm{d}(x, x_1), \ldots, \mathrm{d}(x, x_n)\} > 0$. Dann ist der Ball $B_r(x)$ in $X \setminus A$ enthalten, denn für $k \in \{1, \ldots, n\}$ gilt $\mathrm{d}(x, x_k) \geq r$ und somit $x_k \notin B_r(x)$. Also ist $X \setminus A$ offen und $A$ abgeschlossen.*

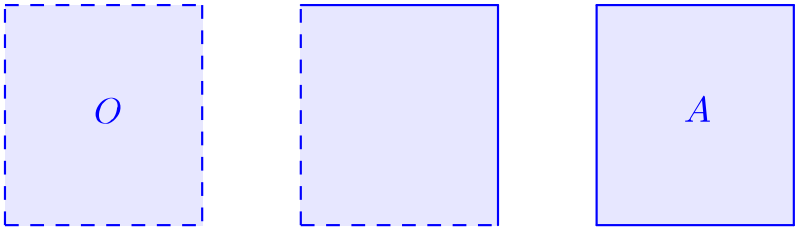
*(iii) Wir betrachten $X = [-1, 1]$ als Teilraum von $\mathbb{R}$. Die Teilmenge $(0, 1]$ von $X$ ist offen, da $(0, 1]$ aus den Punkten in $[-1, 1]$ mit Abstand kleiner $1$ von $1$ besteht. Anders ausgedrückt ist $(0, 1]$ der Ball von Radius $1$ um $1$ im metrischen Raum $X = [-1, 1]$ und somit nach (i) offen. Betrachtet man stattdessen die Teilmenge $[-1, 0) \cup (0, 1] = [-1, 1] \setminus \{0\}$, so ist diese nach (ii) offen, da $\{0\}$ eine abgeschlossene Teilmenge von $[-1, 1]$ ist und ihr Komplement somit offen ist. Wir bemerken, dass für diese Diskussion die Definition des Grundraum $X = [-1, 1]$ mehrmals verwendet wurde (siehe auch Lemma 9.20 und die anschliessende Diskussion).*

*(iv) Wir betrachten den Raum $X = \left\{\frac{1}{n} \mid n \in \mathbb{N}\right\} \cup \{0\}$ als Teilraum von $\mathbb{R}$.*

- *Die Teilmenge $Y_1 = \left\{\frac{1}{m} \mid m \geq n\right\} \cup \{0\}$ von $X$ ist für jedes feste $n \in \mathbb{N}$ offen, da $X \setminus Y_1$ endlich ist. Sie ist auch abgeschlossen, da es zu jedem $y \in X \setminus Y_1$ ein $\varepsilon > 0$ mit $\frac{1}{n} < y - \varepsilon$ gibt und dies $B_\varepsilon(y) \subseteq X \setminus Y_1$ impliziert.*

- *Die Teilmenge* $Y_2 = \left\{\frac{1}{m} \mid m \geq n\right\}$ *von* $X$ *ist nicht abgeschlossen. Tatsächlich ist* $X \setminus Y_2$ *nicht offen, da zu jedem* $\varepsilon > 0$ *ein* $m \in \mathbb{N}$ *mit* $|\frac{1}{m} - 0| < \varepsilon$ *existiert und somit* $B_\varepsilon(0)$ *nicht in* $X \setminus Y_2$ *liegt. Sie ist aber offen, da* $X \setminus Y_2$ *endlich und somit abgeschlossen ist.*

**Übung 9.16** (Abgeschlossener Ball)**.** *Sei* $(X, \mathrm{d})$ *ein metrischer Raum. Zeigen Sie, dass der abgeschlossene Ball* $\{x \in X \mid \mathrm{d}(x, x_0) \leq r\}$ *mit Radius* $r > 0$ *um* $x_0 \in X$ *tatsächlich eine abgeschlossene Teilmenge von* $X$ *ist.*




Figur 9.1: Offene Teilmengen kann man sich als jene Teilmengen vorstellen, die keinen Punkt ihres Randes enthalten und gewissermassen dick sind (wobei wir auf die Definition des Rand noch verzichten wollen). Dies ist im linken Bild illustriert, wo $O$ eine offene Menge (ein Quadrat) im $\mathbb{R}^2$ darstellt und die gestrichelte Linie die Abwesenheit der Randpunkte andeutet. Fügt man Punkte des Randes hinzu, so ist die Menge im Allgemeinen nicht mehr offen (mittleres Bild). Fügt man den kompletten Rand hinzu, so erhält man eine abgeschlossene Menge (rechtes Bild). Abgeschlossene Mengen brauchen im Allgemeinen nicht „dick“ zu sein; beispielsweise sind Geraden im $\mathbb{R}^2$ abgeschlossen.

**Lemma 9.17** (Topologie)**.** *Sei* $(X, \mathrm{d})$ *ein metrischer Raum. Dann gilt*

- *Der Durchschnitt endlich vieler offener Mengen ist offen.*
- *Die Vereinigung beliebig vieler offener Mengen ist offen.*
- *Der Durchschnitt beliebig vieler abgeschlossener Mengen ist abgeschlossen.*
- *Die Vereinigung endlich vieler abgeschlossener Mengen ist abgeschlossen.*

Die Familie der offenen Teilmengen eines metrischen Raumes $(X, \mathrm{d})$ wird die von der Metrik **induzierte Topologie** auf $X$ genannt.

Nach Lemma 9.17 ist also beispielsweise

$$\bigcup_{n \in \mathbb{Z}} (2n, 2n+1) = \ldots \cup (-2, -1) \cup (0, 1) \cup (2, 3) \cup (4, 5) \cup \ldots$$

eine offene Teilmenge von $\mathbb{R}$ mit der Euklidschen Metrik, da für jedes $n \in \mathbb{Z}$ das Intervall $(2n, 2n+1)$ der Ball um $2n + \frac{1}{2}$ ist und somit offen ist.

*Beweis.* Seien $O_1, \ldots, O_n \subseteq X$ offen und sei $x \in \bigcap_{k=1}^n O_k$. Dann gibt es für jedes $k \in \{1, \ldots, n\}$ einen Radius $\varepsilon_k > 0$ mit $B_{\varepsilon_k}(x) \subseteq O_k$, da $x$ in $O_k$ liegt und $O_k$ offen ist per Annahme. Für den Radius $\varepsilon = \min_{k=1,\ldots,n} \varepsilon_k > 0$ gilt nun $B_\varepsilon(x) \subseteq B_{\varepsilon_k}(x) \subseteq O_k$ für jedes $k \in \{1, \ldots, n\}$ und somit $B_\varepsilon(x) \subseteq \bigcap_{k=1}^n O_k$. Da $x \in \bigcap_{k=1}^n O_k$ beliebig ist, schliessen wir, dass $\bigcap_{k=1}^n O_k$ offen ist.

Sei nun $\mathcal{I}$ eine beliebige Menge und für jedes $\alpha \in \mathcal{I}$ eine offene Teilmenge $O_\alpha$ von $X$ gegeben. Für jedes $x \in \bigcup_{\alpha\in\mathcal{I}} O_\alpha$ existiert ein $\beta \in \mathcal{I}$ mit $x \in O_\beta$. Da $O_\beta \subseteq X$ per Annahme offen ist, existiert ein Radius $\varepsilon > 0$ mit $B_\varepsilon(x) \subseteq O_\beta \subseteq \bigcup_{\alpha\in\mathcal{I}} O_\alpha$. Dies beweist, dass $\bigcup_{\alpha\in\mathcal{I}} O_\alpha$ offen ist.

Die beiden Aussagen über abgeschlossene Teilmengen ergeben sich nun aus den ersten zwei Punkten und den Gesetzen von De Morgan. Sind beispielsweise $A_1, \ldots, A_n \subseteq X$ abgeschlossen, dann ist für jedes $k \in \{1, \ldots, n\}$ die Teilmenge $O_k = X \setminus A_k$ offen und somit ist nach dem ersten Punkt $\bigcap_{k=1}^n O_k$ offen. Also ist $X \setminus \bigcap_{k=1}^n O_k = \bigcup_{k=1}^n A_k$ abgeschlossen. □

Betrachten wir $X = \mathbb{R}^d$ für $d \geq 1$ und verschiedene, von (uns bekannten) Normen induzierte Metriken auf $\mathbb{R}^d$, so erhalten wir auf Grund der Normäquivalenz jeweils dieselbe induzierte Topologie. Diese nennt man auch die **Standardtopologie** auf $\mathbb{R}^d$.

**Übung 9.18** (Äquivalente Normen)**.** *Sei* $\mathrm{d}_\infty$ *die von der Maximumsnorm auf* $\mathbb{R}^2$ *induzierte Metrik,* $\mathrm{d}_1$ *die von der* 1*-Norm auf* $\mathbb{R}^2$ *induzierte Metrik und* $\mathrm{d}_2$ *die Euklidsche Metrik. Zeigen Sie, dass die von* $\mathrm{d}_\infty, \mathrm{d}_1$ *respektive* $\mathrm{d}_2$ *induzierten Topologien gleich sind. Formulieren und beweisen Sie die analoge, verallgemeinerte Aussage für äquivalente Normen auf einem Vektorraum.*

**Lemma 9.19** (Charakterisierung offener und abgeschlossener Mengen durch Konvergenz)**.** *Sei* $(X, \mathrm{d})$ *ein metrischer Raum.*

- *Eine Teilmenge* $O \subseteq X$ *ist genau dann offen, wenn für jede konvergente Folge in* $X$ *mit Grenzwert in* $O$ *fast alle Folgenglieder in* $O$ *liegen.*
- *Eine Teilmenge* $A \subseteq X$ *ist genau dann abgeschlossen, wenn für jede konvergente Folge* $(x_n)_n$ *in* $X$ *mit* $x_n \in A$ *für alle* $n \in \mathbb{N}$ *auch der Grenzwert in* $A$ *liegt.*

*Beweis.* Sei $O \subseteq X$ eine offene Teilmenge von $X$ und $(x_n)_n$ eine Folge in $X$ mit Grenzwert $x_0$ in $O$. Dann ist $O$ eine Umgebung von $x_0$, da $O$ offen ist und somit liegen fast alle Folgenglieder von $(x_n)_n$ in $O$.

Sei nun $O \subseteq X$ eine nicht offene Teilmenge. Dann gibt es einen Punkt $x_0 \in O$ mit $B_\varepsilon(x_0) \setminus O \neq \emptyset$ für jedes $\varepsilon > 0$. (Wieso? ) Für jedes $n \in \mathbb{N}$ und $\varepsilon = \frac{1}{n}$ finden wir somit ein $x_n \in B_{1/n}(x_0) \setminus O$. Die Folge $(x_n)_n$ in $X \setminus O$ ist nun konvergent mit Grenzwert $x_0 \in O$, da $\mathrm{d}(x_n, x_0) < \frac{1}{n}$ für alle $n \in \mathbb{N}$, und erfüllt $x_n \notin O$ für jedes $n \in \mathbb{N}$. Dies schliesst den Beweis der ersten Aussage (durch Kontraposition).

Sei $A \subseteq X$ abgeschlossen und sei $(x_n)_n$ eine konvergente Folge in $X$ mit $x_n \in A$ für alle $n \in \mathbb{N}$. Sei $x_0$ der Grenzwert der Folge $(x_n)_n$. Dann ist $O = X \setminus A$ offen und kann nicht den Grenzwert $x_0$ von $(x_n)_n$ enthalten, da sonst fast alle Folgenglieder der Folge $(x_n)_n$ in $O$ liegen müssten. Also ist der Grenzwert $x_0$ in $A$.

Sei schlussendlich $A \subseteq X$ nicht abgeschlossen. Dann ist $O = X \setminus A$ nicht offen und es existiert nach obigem Argument eine Folge $(x_n)_n$ in $A = X \setminus O$ mit Grenzwert $x_0 \in O$. □

Für einen metrischen Raum $(X, \mathrm{d})$ und eine Teilmenge $Y \subseteq X$ haben wir bereits erwähnt, dass wir $Y$ oft auch als einen metrischen Raum betrachten indem wir die Metrik von $X$ einfach

auf $Y$ einschränken. Insbesondere erhalten wir damit eine Definition von offenen Teilmengen von $Y$. Ebenso könnten wir aber auch zuerst offene Teilmengen $O \subseteq X$ definieren und anschliessend dies verwenden um offene Teilmenge von $Y$ als Durchschnitte $O \cap Y$ zu definieren. Dies führt allerdings zu demselben Begriff von offenen Teilmengen von $Y$.

**Lemma 9.20** (Relativ offene oder abgeschlossene Teilmengen)**.** *Sei $(X, \mathrm{d})$ ein metrischer Raum und $Y \subseteq X$ ein Teilraum. Eine Teilmenge von $Y$ ist offen (bezüglich der induzierten Metrik) genau dann wenn sie die Form $O \cap Y$ hat, wobei $O \subseteq X$ eine offene Teilmenge in $X$ ist. Ebenso ist eine Teilmenge von $Y$ abgeschlossen genau dann wenn sie die Form $A \cap Y$ hat, wobei $A \subseteq X$ eine abgeschlossene Teilmenge ist.*

Wir sagen in diesem Zusammenhang auch, dass $O \cap Y$ **in** $Y$ **offen** oder **relativ offen** und $A \cap Y$ **in** $Y$ **abgeschlossen** oder **relativ abgeschlossen** ist. Wenn $Y$ eine offene Teilmenge von $X$ ist, so ist eine relativ offene Teilmenge von $Y$ auch in $X$ offen. Dies gilt analog ebenso für abgeschlossene und relativ abgeschlossene Teilmengen.

Man glaubt vielleicht zuerst, dass obiges ein rein formales Wortspiel ist und eigentlich wenig Zweck erfüllt. Das dem nicht so ist, sollte klar sein, wenn Sie Teilmengen von der Kugeloberfläche $\mathbb{S}^2 = \{v \in \mathbb{R}^3 \mid \|v\| = 1\}$ betrachten und entscheiden wollen, ob der Ball $\{v \in \mathbb{S}^2 \mid \|v - e_1\| < 1\}$ um den ersten Basisvektor $e_1 \in \mathbb{S}^2$ offen ist oder nicht: Als Teilmenge von $\mathbb{R}^3$ ist er nicht offen, doch ist er eine relativ offene Teilmenge von $\mathbb{S}^2$.

*Beweis von Lemma 9.20.* Sei $O \subseteq X$ offen und $y_0 \in O \cap Y$. Dann gibt es ein $\varepsilon > 0$ mit

$$B_\varepsilon^X(y_0) = \{x \in X \mid \mathrm{d}(x, y_0) < \varepsilon\} \subseteq O.$$

Dies impliziert aber

$$B_\varepsilon^Y(y_0) = \{y \in Y \mid \mathrm{d}(y, y_0) < \varepsilon\} \subseteq O \cap Y,$$

womit $O \cap Y$ eine offene Teilmenge von $Y$ ist.

Sei nun umgekehrt $O_Y \subseteq Y$ eine offene Teilmenge. Dann existiert für jedes $y \in O_Y$ ein $\varepsilon_y > 0$ mit $B_{\varepsilon_y}^Y(y) \subseteq O_Y$. Wir definieren $O = \bigcup_{y \in O_Y} B_{\varepsilon_y}^X(y)$. Da $O$ eine Vereinigung von offenen Bällen ist, folgt, dass $O$ eine offene Teilmenge von $X$ ist. Wir zeigen nun, dass $O_Y = O \cap Y$ erfüllt ist. Für $y \in O_Y$ gilt $y \in B_{\varepsilon_y}^X(y) \subseteq O$ und damit $y \in O \cap Y$. Falls umgekehrt $y \in O \cap Y$ ist, dann existiert ein $y' \in O_Y$ mit $y \in B_{\varepsilon_{y'}}^X(y')$. Daraus folgt aber $y \in B_{\varepsilon_{y'}}^Y(y') \subseteq O_Y$. Dies zeigt, dass $O_Y = O \cap Y$ ist.

Die Aussagen über abgeschlossene Teilmengen folgen aus obigem durch Komplementbildung. □

**Übung 9.21** (Beispiele relativ offener Teilmengen)**.**

*(i) Beschreiben Sie die in $Y = \{(x, y) \in \mathbb{R}^2 \mid y \geq 0\}$ offenen Teilmengen, welche Punkte der Achse $\{(x, y) \mid y = 0\}$ enthalten.*

*(ii) Sei $X = \{(x, y) \in \mathbb{R}^2 \mid xy = 0\}$. Charakterisieren Sie die in $X$ offenen Teilmengen.*

Wir werden nun die Begriffe der offenen und abgeschlossenen Teilmengen mit dem Supremum von Teilmengen von $\mathbb{R}$ in Verbindung bringen.

**Lemma 9.22** (Supremum von abgeschlossenen und offenen Teilmengen in $\mathbb{R}$)**.** *Für eine abgeschlossene nicht-leere von oben beschränkte Teilmenge $A \subseteq \mathbb{R}$ gilt $\sup A = \max A \in A$. Für eine offene nicht-leere von oben beschränkte Teilmenge $O \subseteq \mathbb{R}$ gilt $\sup O \notin O$. Dies gilt analog für von unten beschränkte Teilmengen und das Infimum.*

*Beweis.* Für jedes $n \in \mathbb{N}$ existiert nach Satz 2.59 ein $x_n \in A$ mit $\sup A - \frac{1}{n} < x_n \leq \sup A$. Daraus folgt $\lim_{n\to\infty} x_n = \sup A$ und mit Lemma 9.19, dass $\sup A = \max A \in A$.

Falls $\sup O \in O$ wäre, so gäbe es ein $\varepsilon > 0$ mit $B_\varepsilon(\sup O) \subseteq O$. Insbesondere gilt aber damit $\sup O + \varepsilon/2 \in O$ und dies widerspricht der Definition des Supremums in Satz 2.59. Daher muss $\sup O \notin O$ gelten. □

VO: 23.2.18

**Definition 9.23** (Inneres, Rand und Abschluss)**.** Sei $(X, \mathrm{d})$ ein metrischer Raum und $Y \subseteq X$ eine Teilmenge. Ein Punkt $x \in Y$ heisst **innerer Punkt von** $Y$, falls es ein $\varepsilon > 0$ mit $B_\varepsilon(x) \subseteq Y$ gibt. Die Menge aller inneren Punkte

$$Y^\circ = \{x \in Y \mid \exists \varepsilon > 0 : B_\varepsilon(x) \subseteq Y\}$$

wird das **Innere** von $Y$ genannt. Ein Punkt $x \in X$ ist ein **Randpunkt von** $Y$, falls zu jedem $\varepsilon > 0$ beide Durchschnitte $B_\varepsilon(x) \cap Y$ und $B_\varepsilon(x) \cap (X \setminus Y)$ nichtleer sind. Die Menge der Randpunkte

$$\partial Y = \{x \in X \mid B_\varepsilon(x) \cap Y \neq \emptyset \neq B_\varepsilon(x) \cap (X \setminus Y) \text{ für alle } \varepsilon > 0\}$$

wird als der **Rand** von $Y$ bezeichnet. Der **Abschluss** einer Menge wird durch $\overline{Y} = Y \cup \partial Y$ definiert.

Ein Punkt $x$ ist also ein Randpunkt einer Teilmenge $Y \subseteq X$, falls er sowohl von Punkten in der Menge $Y$ als auch von Punkten ausserhalb von $Y$ „approximiert“ werden kann. Beispielsweise ist der Rand des Einheitsballes $B_1(0) \subseteq \mathbb{R}^2$ gerade der Einheitskreis und der Abschluss ist der abgeschlossene Einheitsball $\left\{v \in \mathbb{R}^2 \mid \|v\| \leq 1\right\}$.

**Wichtige Übung 9.24** (Eigenschaften des Abschluss und des Inneren)**.** *Sei $(X, \mathrm{d})$ ein metrischer Raum und sei $Y \subseteq X$ eine Teilmenge.*

*(i) Zeigen Sie, dass $Y^\circ$ eine offene Teilmenge von $X$ ist und jede offene Teilmenge $U \subseteq Y$ enthält.*

*(ii) Zeigen Sie, dass $\overline{Y}$ eine abgeschlossene Teilmenge von $X$ ist und in jeder abgeschlossenen Teilmenge $A \supseteq Y$ enthalten ist.*

### 9.2.2 Häufungspunkte und Dichtheit

Wir übertragen einige weitere Begriffe auf beliebige metrische Räume.

**Definition 9.25** (Häufungspunkte und Dichtheit)**.** Sei $(X, \mathrm{d})$ ein metrischer Raum. Ein Punkt $x_0 \in X$ ist ein **Häufungspunkt einer Folge** $(x_n)_n$ in $X$, falls es eine konvergente Teilfolge $(x_{n_k})$ gibt mit $\lim_{k\to\infty} x_{n_k} = x_0$. Ein Punkt $x_0 \in X$ ist ein **Häufungspunkt einer Teilmenge** $D \subseteq X$, falls für jedes $\varepsilon > 0$ der Durchschnitt $D \cap (B_\varepsilon(x_0) \setminus \{x_0\})$ nicht-leer ist. Eine Teilmenge $D \subseteq X$ heisst **dicht**, falls für jedes $x_0 \in X$ und jedes $\varepsilon > 0$ der Durchschnitt $D \cap B_\varepsilon(x_0)$ nicht-leer ist.

Beispielsweise wissen wir bereits, dass 1 ein Häufungspunkt der Folge $((-1)^n)_n$ in $\mathbb{R}$ ist, 0 ein Häufungspunkt der Teilmenge $(0, 1]$ ist und $\mathbb{Q} \subseteq \mathbb{R}$ dicht ist.

**Wichtige Übung 9.26.** *Sei $(X, \mathrm{d})$ ein metrischer Raum.*

*(i) Zeigen Sie, dass ein Punkt $x_0 \in X$ genau dann ein Häufungspunkt einer Teilmenge $D \subseteq X$ ist, wenn es eine Folge $(x_n)_n$ in $D$ gibt mit $x_n \neq x_0$ für alle $n \in \mathbb{N}$ und Grenzwert $\lim_{n\to\infty} x_n = x_0$.*

*(ii) Zeigen Sie, dass eine Teilmenge $D \subseteq X$ genau dann dicht ist, wenn es für jeden Punkt $x_0 \in X$ eine Folge in $D$ gibt mit Grenzwert $x_0$.*

*(iii) Zeigen Sie, dass eine konvergente Folge $(x_n)_n$ in $X$ den Grenzwert $\lim_{n\to\infty} x_n$ als einzigen Häufungspunkt hat. Insbesondere hat die konstante Folge $(x_0)_n$ den Punkt $x_0$ als Häufungspunkt.*

*(iv) Zeigen Sie, dass eine endliche Teilmenge von $X$ keinen einzigen Häufungspunkt besitzt. Insbesondere sind die Häufungspunkte einer Folge $(x_n)_n$ in $X$ im Allgemeinen nicht dieselben wie jene der Teilmenge $\{x_n \mid n \in \mathbb{N}\}$.*

**Wichtige Übung 9.27** (Charakterisierung der Häufungspunkte einer Folge)**.** *Zeigen Sie in Analogie zu Proposition 5.23, dass für einen metrischen Raum $X$, eine Folge $(x_n)_n$ und einen Punkt $x_0 \in X$ folgende Bedingungen äquivalent sind.*

*(i) Der Punkt $x_0$ ist ein Häufungspunkt der Folge $(x_n)_n$.*

*(ii) Für jedes $\varepsilon > 0$ und jedes $N \in \mathbb{N}$ gibt es ein $n \geq N$ mit $x_n \in B_\varepsilon(x_0)$.*

**Übung 9.28** (Häufungspunkte einer Teilmenge und der Abschluss)**.** *Sei $(X, \mathrm{d})$ ein metrischer Raum und $Y \subseteq X$ eine Teilmenge.*

*(i) Zeigen Sie, dass $x \in \overline{Y}$ genau dann wenn es eine Folge in $Y$ mit Grenzwert gleich $x$ gibt.*

*(ii) Sei $x \in X \setminus Y$. Zeigen Sie, dass $x$ genau dann ein Häufungspunkt von $Y$ ist, wenn $x \in \overline{Y}$.*

*(iii) Zeigen Sie, dass $Y$ genau dann dicht in $X$ ist, wenn $\overline{Y} = X$ gilt.*

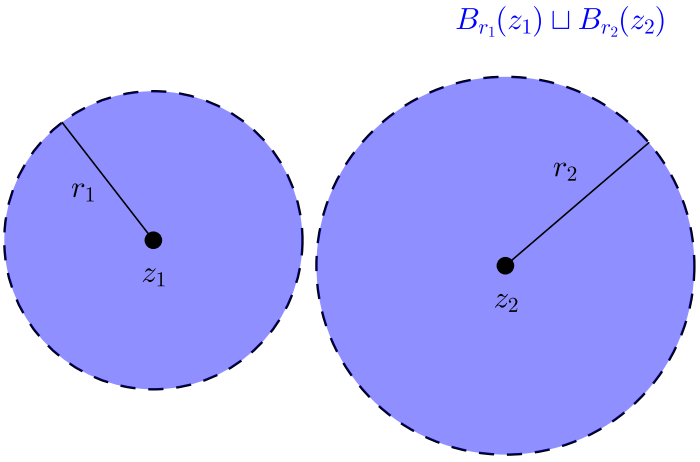
### 9.2.3 Zusammenhang

Wir möchten eine weitere, diesmal neue Eigenschaft von metrischen Räumen einführen, die am Rande schon in Übungen 3.63 und 3.64 angesprochen wurde. Einen metrischen Raum wollen wir zusammenhängend nennen, falls er sich nicht mit offenen Mengen „auseinanderreissen“ lässt. Präzis ausgedrückt:

**Definition 9.29** (Zusammenhang)**.** Sei $(X, \mathrm{d})$ ein nicht-leerer metrischer Raum. Wir nennen $(X, \mathrm{d})$ **zusammenhängend**, falls es keine zwei offene nicht-leere Teilmengen $O_1, O_2 \subseteq X$ gibt mit $X = O_1 \sqcup O_2$. Wir sagen, dass eine Teilmenge $G \subseteq X$ **abgeschloffen** ist, falls $G$ zugleich offen und abgeschlossen ist.

Alternativ ausgedrückt ist ein metrischer Raum $(X, \mathrm{d})$ genau dann zusammenhängend, wenn es ausser der leeren Menge und $X$ keine weiteren abgeschloffenen Teilmengen von $X$ gibt. Man beachte auch, dass der Begriff des Zusammenhangs nur von der Topologie und nicht von der Wahl der Metrik abhängt. Man sagt auch, dass Zusammenhang eine **topologische Eigenschaft** ist.

Ein unzusammenhängender metrischer Raum lässt sich schnell als disjunkte Vereinigung von offenen Teilmengen eines grösseren Raumes konstruieren. Beispielsweise ist für $z_1, z_2 \in \mathbb{C}$ und $r_1, r_2 > 0$ die Menge $X = B_{r_1}(z_1) \cup B_{r_2}(z_2)$ mit der von der Standardmetrik induzierten Metrik unzusammenhängend, falls $|z_1 - z_2| \geq r_1 + r_2$ oder anders ausgedrückt falls die Bälle $B_{r_1}(z_1)$ und $B_{r_2}(z_2)$ disjunkt sind.




Allgemeiner nennen wir eine Teilmenge $Y \subseteq X$ eines metrischen Raumes $(X, \mathrm{d})$ **zusammenhängend**, falls der Teilraum $Y$ mit der von $X$ induzierten Metrik (oder Topologie) zusammenhängend ist. Wir charakterisieren nun die zusammenhängenden Teilmengen von $\mathbb{R}$.

**Proposition 9.30** (Zusammenhang in den reellen Zahlen)**.** *Eine nicht-leere Teilmenge $X \subseteq \mathbb{R}$ ist genau dann zusammenhängend, wenn $X$ ein Intervall ist.*

*Beweis.* Angenommen $X \subseteq \mathbb{R}$ ist kein Intervall und seien $a = \inf(X)$ und $b = \sup(X)$. Falls $(a, b) \subseteq X$ wäre, dann wäre $X$ doch ein Intervall (wieso?). Also gibt es ein $y \in (a, b) \setminus X$ und nach Definition von $a < b$ Punkte

$$x_1 \in O_1 = (-\infty, y) \cap X$$
$$x_2 \in O_2 = (y, \infty) \cap X.$$

Da $y$ nicht in $X$ liegt, gilt $X = O_1 \sqcup O_2$. Des Weiteren sind $O_1$ und $O_2$ offene, nicht-leere Teilmengen von $X$ und somit ist $X$ nicht zusammenhängend.

Sei nun $X = I \subseteq \mathbb{R}$ ein Intervall. Angenommen es existiert eine nicht-leere abgeschloffene Teilmenge $Y \subsetneq I$. Wir wählen $a \in Y$ und $b \in I \setminus Y$. Ohne Beschränkung der Allgemeinheit können wir annehmen, dass $a < b$ (da wir sonst $Y$ durch $I \setminus Y$ ersetzen). Da $Y \subseteq I$ offen und $a \in Y$ ist, ist $Y \cap (a, b)$ nicht-leer. Wir definieren

$$s = \sup(Y \cap [a, b]) = \sup(Y \cap (a, b)) \in [a, b] \subseteq I.$$

Des Weiteren ist $Y \cap (a, b)$ eine offene Teilmenge von $\mathbb{R}$. (Wieso?) Daher zeigt Lemma 9.22, dass $s \notin Y$. Da $Y \subseteq I$ abgeschlossen ist, folgt zum Beispiel aus Lemma 9.19 dass $Y \cap [a, b]$ eine abgeschlossene Teilmenge von $\mathbb{R}$ ist. Daher zeigt Lemma 9.22, dass $s \in Y$ ist. Dieser Widerspruch zeigt also, dass es keine abgeschloffene Teilmengen von $I$ geben kann. □

Nun wollen wir eine mögliche Operation auf zusammenhängenden Teilmengen eines metrischen Raumes diskutieren. Eine weitere findet sich in einer Übung in Abschnitt 9.9.3.

**Lemma 9.31** (Vereinigungen von zusammenhängenden Teilmengen)**.** *Sei $(X, \mathrm{d})$ ein metrischer Raum und seien $Y_1, Y_2$ zwei zusammenhängende Teilräume. Falls der Schnitt $Y_1 \cap Y_2$ nicht-leer ist, dann ist die Vereingung $Y_1 \cup Y_2$ zusammenhängend.*

*Beweis.* Sei $A$ eine nicht-leere, abgeschlossene und offene Teilmenge von $Y_1 \cup Y_2$. Dann ist $A \cap Y_j$ eine abgeschlossene und offene Teilmenge von $Y_j$ für $j \in \{1, 2\}$. Da $A$ nicht-leer ist, ist einer dieser beiden Schnitte nicht-leer – sagen wir $A \cap Y_1$ ist nicht-leer. Da $Y_1$ zusammenhängend ist, folgt $A \cap Y_1 = Y_1$ oder äquivalent $Y_1 \subseteq A$. Da aber $Y_1 \cap Y_2 \neq \emptyset$ folgt $A \cap Y_2 \neq \emptyset$ und somit genauso $A \cap Y_2 = Y_2$ und $Y_2 \subseteq A$. Zusammenfassend ergibt sich $A = Y_1 \cup Y_2$. Dies beweist, dass $Y_1 \cup Y_2$ zusammenhängend ist. □

**Wichtige Übung 9.32.** *Verallgemeinern Sie Lemma 9.31 für beliebige Vereinigungen.*

Lemma 9.31 und Übung 9.32 können beispielsweise angewendet werden, um zu zeigen, dass (endliche) Produkte von zusammenhängenden Räumen zusammenhängend sind (wobei wir noch keine Metriken auf Produkträumen definiert haben). Wir wollen einen Spezialfall davon in folgender Übung diskutieren.

**Übung 9.33** (Zusammenhang der Ebene)**.** *Wir möchten zeigen, dass $\mathbb{R}^2$ zusammenhängend ist (wobei wir $\mathbb{R}^2$ mit der Standardtopologie betrachten). Gehen Sie dazu wie folgt vor:*

*(i) Betrachten Sie $\mathbb{R}^2$ mit der Maximumsnorm. Zeigen Sie, dass die offenen Teilmengen des Teilraumes $\{x\} \times \mathbb{R}$ für $x \in \mathbb{R}$ (respektive $\mathbb{R} \times \{y\}$ für $y \in \mathbb{R}$) von der Form $\{x\} \times U$ für $U \subseteq \mathbb{R}$ offen (respektive $U \times \{y\}$ für $U \subseteq \mathbb{R}$ offen) sind.*

*(ii) Zeigen Sie, dass die Teilräume $\{x\} \times \mathbb{R}$ für $x \in \mathbb{R}$ und $\mathbb{R} \times \{y\}$ für $y \in \mathbb{R}$ zusammenhängend sind.*

*(iii) Verwenden Sie Lemma 9.31 und Übung 9.32 um zu zeigen, dass $\mathbb{R}^2$ zusammenhängend ist.*

## 9.3 Stetigkeit

### 9.3.1 Definition und Charakterisierungen

Wir möchten nun den Stetigkeitsbegriff auf metrische Räume verallgemeinern. Wie wir sehen werden, lässt sich dieser ausschliesslich mittels dem Begriff der Topologie (also mittels der Kollektion offener Mengen) charakterisieren. Wir möchten Stetigkeit allerdings zuerst in Analogie zu Definition 3.46 (respektive der Charakterisierung in Proposition 5.17) einführen.

**Definition 9.34** (Stetigkeit bei einem Punkt)**.** Seien $(X, \mathrm{d}_X), (Y, \mathrm{d}_Y)$ zwei metrische Räume und sei $f : X \to Y$ eine Funktion. Wir sagen, dass $f$ **bei** $x_0 \in X$ **$\varepsilon$-$\delta$-stetig** ist, falls für alle $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass für alle $x \in B_\delta(x_0)$ auch $f(x) \in B_\varepsilon(f(x_0))$ gilt. Wir sagen, dass $f$ **bei** $x_0 \in X$ **folgenstetig** ist, falls für jede konvergente Folge $(x_n)_n$ in $X$ mit Grenzwert $\lim_{n\to\infty} x_n = x_0$ die Folge $(f(x_n))_n$ konvergiert und Grenzwert $\lim_{n\to\infty} f(x_n) = f(x_0)$ hat.

Wie in obiger Definition und vielleicht schon in den letzten zwei Abschnitten ersichtlich wurde, ist eine explizite Bezeichnung d der Metrik in $(X, \mathrm{d})$ nicht immer notwendig (da man zum Beispiel oft stattdessen nur Bälle betrachtet) und führt nur zu zusätzlicher Notation. Wir werden deswegen in Zukunft schlicht $X$ als metrischen Raum bezeichnen, wobei die Metrik also implizit ist, keinen „Namen“ hat und wenn nötig einfach mit $\mathrm{d}(\cdot, \cdot)$ bezeichnet wird.

**Lemma 9.35** (Stetigkeit bei einem Punkt)**.** *Seien $X$ und $Y$ zwei metrische Räume, $f : X \to Y$ eine Funktion und $x_0 \in X$ ein Punkt. Dann ist $f$ genau dann bei $x_0$ $\varepsilon$-$\delta$-stetig, wenn $f$ bei $x_0$ folgenstetig ist.*

Auf Grund der Aussage des obigen Lemma sagen wir auch kurz, dass $f$ **bei** $x_0$ **stetig** ist, falls $f$ bei $x_0$ $\varepsilon$-$\delta$-stetig ist. Des Weiteren ist $f$ **stetig**, wenn $f$ bei jedem Punkt in $X$ stetig ist. In Proposition 9.37 werden weitere wichtige Charakterisierungen von Stetigkeit folgen.

*Beweis.* Angenommen $f$ ist bei $x_0$ $\varepsilon$-$\delta$-stetig. Sei $(x_n)_n$ eine Folge in $X$ mit Grenzwert $x_0 \in X$ und sei $\varepsilon > 0$. Dann existiert ein $\delta > 0$, so dass $f(x) \in B_\varepsilon(f(x_0))$ für alle $x \in B_\delta(x_0)$. Da $(x_n)_n$ gegen $x_0$ konvergiert, existiert nun ein $N \in \mathbb{N}$ mit $x_n \in B_\delta(x_0)$ für alle $n \geq N$. Insbesondere gilt für $n \geq N$ also $f(x_n) \in B_\varepsilon(f(x_0))$. Da $\varepsilon > 0$ beliebig war, gilt somit $\lim_{n\to\infty} f(x_n) = f(x_0)$ und $f$ ist bei $x_0$ folgenstetig wie gewünscht.

Angenommen $f$ ist bei $x_0$ nicht $\varepsilon$-$\delta$-stetig. Dann existiert ein $\varepsilon > 0$, so dass es für jedes $\delta > 0$ ein $x \in B_\delta(x_0)$ gibt mit $f(x) \notin B_\varepsilon(f(x_0))$. Wir wählen nun für jedes $n \in \mathbb{N}$ und $\delta = \frac{1}{n}$ ein solches $x_n \in B_{\frac{1}{n}}(x_0)$. Die Folge $(x_n)_n$ konvergiert somit gegen $x_0$ und es gilt $f(x_n) \notin B_\varepsilon(f(x_0))$ für alle $n \in \mathbb{N}$. Insbesondere konvergiert $(f(x_n))_n$ nicht gegen $f(x_0)$ und $f$ ist nicht folgenstetig. □

Genauso wie für Funktionen auf $\mathbb{R}$ können wir auch in diesem neuen, allgemeinen Rahmen Grenzwerte von Funktionen definieren.

**Definition 9.36.** Sei $X$ ein metrischer Raum, $D \subseteq X$ eine Teilmenge, $x_0 \in X$ ein Häufungspunkt von $D$ und $f : D \to Y$ eine Abbildung in einen weiteren metrischen Raum $Y$. Dann ist

$y_0 \in Y$ der **Grenzwert** von $f(x)$ für $x \to x_0$, geschrieben $y_0 = \lim_{x \to x_0} f(x)$, falls gilt

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x \in D \cap B_\delta(x_0) \setminus \{x_0\} : f(x) \in B_\varepsilon(y_0).$$

Ist $x_0 \in D$, so ist wiederum $f$ genau dann bei $x_0$ stetig, falls der Grenzwert $\lim_{x \to x_0} f(x)$ existiert und gleich $f(x_0)$ ist.

Wie schon erwähnt wurde, wollen wir nun Stetigkeit mit Hilfe der Topologie ausdrücken.

**Proposition 9.37** (Charakterisierungen der Stetigkeit)**.** *Seien $X, Y$ zwei metrische Räume und $f : X \to Y$ eine Funktion. Dann sind folgende Bedingungen äquivalent:*

*(i) Die Funktion $f$ ist stetig.*

*(ii) Für jedes $x \in X$ ist $f$ bei $x$ $\varepsilon$-$\delta$-stetig.*

*(iii) Für jedes $x \in X$ ist $f$ bei $x$ folgenstetig.*

*(iv) Für jede offene Teilmenge $O \subseteq Y$ ist $f^{-1}(O)$ eine offene Teilmenge von $X$.*

*(v) Für jede abgeschlossene Teilmenge $A \subseteq Y$ ist $f^{-1}(A)$ eine abgeschlossene Teilmenge von $X$.*

*Beweis.* Per Definition der Stetigkeit ist (i) äquivalent zu (ii). Nach Lemma 9.35 sind (ii) und (iii) äquivalent. Wir zeigen als nächsten Schritt die Äquivalenz von (i) und (iv).

Sei $f$ stetig und $O \subseteq Y$ offen. Für $x \in f^{-1}(O)$ ist $f(x) \in O$, wodurch es ein $\varepsilon > 0$ gibt mit $B_\varepsilon(f(x)) \subseteq O$, da $O$ offen ist. Da $f$ bei $x$ $\varepsilon$-$\delta$-stetig ist, gibt es ein $\delta > 0$ mit $f(B_\delta(x)) \subseteq B_\varepsilon(f(x)) \subseteq O$ und somit $B_\delta(x) \subseteq f^{-1}(O)$. Da aber $x \in f^{-1}(O)$ beliebig war, ist $f^{-1}(O)$ also offen. Dies beweist (i) $\Longrightarrow$ (iv).

VO: 26.2.18

Angenommen $f$ erfüllt die Bedingung in (iv). Sei $x_0 \in X$ und $\varepsilon > 0$. Wir wissen, dass das Urbild $f^{-1}(B_\varepsilon(f(x_0))) \subseteq X$ den Punkt $x_0$ enthält und offen ist, da der Ball $B_\varepsilon(f(x_0))$ in $Y$ offen ist. Also existiert ein $\delta > 0$ mit $B_\delta(x_0) \subseteq f^{-1}(B_\varepsilon(f(x_0)))$ oder äquivalenterweise $f(B_\delta(x_0)) \subseteq B_\varepsilon(f(x_0))$. Also ist $f$ bei $x_0$ $\varepsilon$-$\delta$-stetig, was die Implikation (iv) $\Longrightarrow$ (i) beweist.

Die Äquivalenz von (iv) und (v) ergibt sich aus den Eigenschaften von Urbildern und der Dualität von offenen und abgeschlossenen Mengen (via der Komplementoperation). Falls $f^{-1}(O)$ für jede offene Menge $O \subseteq Y$ offen ist und $A \subseteq Y$ abgeschlossen ist, dann ist $f^{-1}(A) = X \setminus f^{-1}(Y \setminus A)$ abgeschlossen. Falls $f^{-1}(A)$ für jede abgeschlossene Menge $A \subseteq Y$ abgeschlossen ist und $O \subseteq Y$ offen ist, dann ist $f^{-1}(O) = X \setminus f^{-1}(Y \setminus O)$ offen. □

Mit den Kriterien (iv) und (v) in Proposition 9.37 verfügt man auch über ein nützliches Werkzeug, um Teilmengen eines Raumes auf Offenheit und Abgeschlossenheit zu untersuchen. Zum Beispiel gelten für eine stetige Funktion $f : X \to \mathbb{R}$, dass die Teilmenge $\{x \in X \mid f(x) > 0\}$ offen und die Teilmenge $\{x \in X \mid f(x) \geq 0\}$ abgeschlossen ist. Da es leicht ist (siehe Abschnitt 9.3.3), stetige Funktionen zu konstruieren, kann man dadurch für viele Mengen Offenheit oder Abgeschlossenheit zeigen.

*Bemerkung.* In der Vorlesung „Topologie“ im vierten Semester des Mathematikstudiums werden allgemeiner über die Eigenschaften in Lemma 9.17 Kollektionen offener Teilmengen einer

Menge betrachtet. Klebt man beispielsweise die Ränder des Quadrats $[0,1]^2$ nach vorgegebenen Regeln aneinander, so lässt sich auf dem erhaltenen Quotientenraum eine natürliche Topologie mittels der Topologie auf $[0,1]^2$ definieren. Stetigkeit wird dann wie in Proposition 9.37 dadurch definiert, dass Urbilder offener Mengen offen sein sollen.

### 9.3.2 Zwei stärkere Stetigkeitsbegriffe

Wie wir bereits gesehen haben, sind manchmal folgende Stetigkeitseigenschaften nützlich.

**Definition 9.38.** Seien $(X, \mathrm{d}_X)$ und $(Y, \mathrm{d}_Y)$ zwei metrische Räume und $f : X \to Y$ eine Funktion. Dann heisst $f$ **gleichmässig stetig**, falls es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $x_1, x_2 \in X$ mit $\mathrm{d}_X(x_1, x_2) < \delta$ auch $\mathrm{d}_Y(f(x_1), f(x_2)) < \varepsilon$ gilt. Des Weiteren heisst $f$ **Lipschitz-stetig**, falls es eine sogenannte **Lipschitz-Konstante** $L \geq 0$ gibt mit

$$\mathrm{d}_Y(f(x_1), f(x_2)) \leq L\mathrm{d}_X(x_1, x_2)$$

für alle $x_1, x_2 \in X$.

Per Definition sind gleichmässig stetige Funktionen stetig und wie wir schon gesehen haben, sind stetige Funktionen nicht zwingend gleichmässig stetig. Des Weiteren sind Lipschitz stetige Funktionen gleichmässig stetig. (Wieso?). In Proposition 9.77 werden wir in Analogie zu Satz 3.78 zeigen können, dass unter gewissen, wichtigen Annahmen an $X$ eine stetige Funktion auch gleichmässig stetig ist. Des Weiteren werden wir später sehen, dass eine reellwertige Funktionen auf einer offenen Teilmengen von $\mathbb{R}^d$, die mehrdimensional differenzierbar ist und eine beschränkte Ableitung besitzt, Lipschitz-stetig ist falls der Definitionsbereich noch weitere Annahmen erfüllt.

### 9.3.3 Konstruktion von stetigen Funktionen

Wie schon in Abschnitt 3.5 lassen sich für stetigen Funktionen verschiedene Operationen durchführen, die wiederum zu weiteren stetigen Funktionen führen. Die folgende Proposition fasst viele Themen aus dem ersten Semester zusammen und erweitert diese auf den allgemeineren Rahmen dieses Kapitels.

**Proposition 9.39** (Stetige Funktionen)**.** *Seien $X, Y, Z$ metrische Räume.*

*(i) Falls $f : X \to Y$ und $g : Y \to Z$ stetig sind, dann ist auch $g \circ f : X \to Z$ stetig.*

*(ii) Eine Funktion $f : X \to \mathbb{C}^d$ ist genau dann stetig, wenn die Komponenten*

$$f_j = \pi_j \circ f : X \to \mathbb{C}$$

*für $j = 1, \ldots, d$ stetig sind, wobei $\pi_j : (z_1, \ldots, z_d)^t \in \mathbb{C}^d \mapsto z_j \in \mathbb{C}$ für $j \in \{1, \ldots, d\}$ die Projektion auf die $j$-te Koordinate bezeichnet.*

*(iii) Die Addition $+ : \mathbb{C}^2 \to \mathbb{C}$ ist stetig.*

*(iv) Die Multiplikation $\cdot : \mathbb{C}^2 \to \mathbb{C}$ ist stetig.*

(v) *Die Kehrwertfunktion* $(\ )^{-1} : z \in \mathbb{C}^\times \mapsto z^{-1} \in \mathbb{C}$ *ist stetig.*

(vi) *Polynome in* $\mathbb{C}[z]$*, die komplexe Exponentialabbildung, der reelle Logarithmus, Wurzelfunktionen, die komplexen trigonometrischen Funktionen* sin, cos, tan, cot *die komplexen hyperbolischen Funktionen* sinh, cosh, tanh, *die reellen inversen trigonometrischen Funktionen* arcsin, arccos, arctan *und die reellen inversen hyperbolischen Funktionen* arsinh, arcosh, artanh *sind auf ihrem jeweiligen Definitionsbereich stetig. Des Weiteren definieren Potenzreihen im Inneren des Konvergenzkreises stetige Funktionen.*

Insbesondere können wir mit obiger Proposition die Stetigkeit von beliebig komplizierten Ausdrücken wie zum Beispiel $x \in (0, \infty) \mapsto \sin(\log(x) + x^2 - \exp(x + \mathrm{i}\sqrt[3]{x}))$ zeigen. (Wie?)

*Beweis.* Für (i) sei $O \subseteq Z$ offen, dann ist nach Proposition 9.37(iv) $g^{-1}O \subseteq Y$ offen und damit auch $(g \circ f)^{-1}O = f^{-1}(g^{-1}O) \subseteq X$ offen. Da $O \subseteq Z$ eine beliebige offene Teilmenge ist, folgt die Stetigkeit von $g \circ f$ wiederum nach Proposition 9.37(iv).

Die Eigenschaft in (ii) folgt aus der Charakterisierung der Stetigkeit mittels Folgen in Proposition 9.37(iii) und der Charakterisierung von Folgenkonvergenz in $\mathbb{R}^d$ in Lemma 5.93. Die Aussage in (iii), (iv), (v) folgen ebenso aus Proposition 9.37(iii) und Proposition 5.7. Obwohl logisch unnötig geben wir noch direkte Argumente dafür an.

Für (ii) bemerken wir zuerst, dass die Projektion $\pi_j : \mathbb{C}^d \to \mathbb{C}$ wegen $|\pi_j(v-w)| \leq \|v-w\|$ für $v, w \in \mathbb{C}^d$ Lipschitz-stetig mit Lipschitz-Konstante 1 ist. Daher folgt aus (i) und Stetigkeit von $f : X \to \mathbb{C}^d$, dass auch $\pi_j \circ f : X \to C$ stetig ist für jedes $j \in \{1, \ldots, d\}$. Sei nun umgekehrt $\pi_j \circ f : X \to C$ für jedes $j \in \{1, \ldots, d\}$ stetig. Sei $x_0 \in X$ und $\varepsilon > 0$, dann existiert für jedes $j \in \{1, \ldots, d\}$ ein $\delta_j > 0$ mit $|\pi_j \circ f(x) - \pi_j \circ f(x_0)| < \varepsilon/\sqrt{d}$ für alle $x \in B_{\delta_j}(x_0)$. Wir setzen $\delta = \min(\delta_1, \ldots, \delta_d)$ und erhalten $\|f(x) - f(x_0)\| < \varepsilon$ (wegen $\|v\|_2 \leq \sqrt{d}\|v\|_\infty$ für alle $v \in \mathbb{R}^d$). Dies zeigt Stetigkeit von $f$ bei $x_0$.

Für $(z_0, w_0)^t, (z, w)^t \in \mathbb{C}^2$ gilt

$$|(z + w) - (z_0 + w_0)| \leq |z - z_0| + |w - w_0| \leq 2\|(z, w)^t - (z_0, w_0)^t\|.$$

Daher ist die Addition Lipschitz-stetig mit Lipschitz-Konstante 2, und (iii) folgt.

Für $(z_0, w_0)^t, (z, w)^t \in \mathbb{C}^2$ gilt

$$|zw - z_0w_0| \leq |zw - z_0w| + |z_0w - z_0w_0| = |w||z - z_0| + |z_0||w - w_0|.$$

Sei nun $\varepsilon > 0$ und

$$\|(z, w)^t - (z_0, w_0)^t\| < \delta = \min\left(1, \tfrac{1}{2}\varepsilon \max(|z_0|, 1 + |w_0|)^{-1}\right).$$

Dann gilt $|w| \leq |w - w_0| + |w_0| \leq 1 + |w_0|$ und daher

$$|zw - z_0w_0| \leq (1 + |w_0|)|z - z_0| + |z_0||w - w_0| < \varepsilon.$$

Dies zeigt Stetigkeit der Multiplikation in (iv).

Für $z_0 \in \mathbb{C}^\times$, $\varepsilon > 0$ und $z \in \mathbb{C}$ mit $|z - z_0| < \frac{1}{2}\min(|z_0|, \varepsilon|z_0|^2)$ gilt

$$|z| \geq |z_0| - |z - z_0| > \tfrac{1}{2}|z_0|$$

und daher

$$|z^{-1} - z_0^{-1}| = |zz_0|^{-1}|z_0 - z| \leq 2|z_0|^{-2}\tfrac{1}{2}\varepsilon|z_0|^2 = \varepsilon.$$

Da $\varepsilon > 0$ beliebig war, zeigt dies die Stetigkeit der Kehrwertabbildung in (v).

Für die Stetigkeit der Abbildungen in (vi) siehe Korollar 3.52, Satz 6.68, Korollar 5.58, Satz 6.72, Abschnitt 6.7.1, Abschnitt 7.3, Abschnitt 7.4, und Satz 6.56. □

Eine wichtige Methode der Definition von Funktion, die nicht von Proposition 9.39 abgedeckt wird, ist die Definition mittels Fallunterscheidungen. Dies führt im Allgemeinen nicht zu stetigen Funktionen (charakteristische Funktionen bilden ein schnelles Gegenbeispiel), doch gibt es auch Situationen, wo eine mittels Fallunterscheidungen definierte Funktion stetig ist.

**Übung 9.40** (Stetige Funktionen durch Fallunterscheidung)**.** *Seien $X, Y$ zwei metrische Räume und seien $A_1, A_2 \subseteq X$ zwei abgeschlossene Teilmengen von $X$ mit $X = A_1 \cup A_2$. Angenommen $f_1 : A_1 \to Y$ und $f_2 : A_2 \to Y$ sind zwei stetige Funktionen mit $f_1(x) = f_2(x)$ für alle $x \in A_1 \cap A_2$. Zeigen Sie, dass die damit wohldefinierte Funktion*

$$f : X \to Y,\ x \mapsto \begin{cases} f_1(x) & \textit{falls } x \in A_1 \\ f_2(x) & \textit{falls } x \in A_2 \end{cases}$$

*stetig ist.*

**Wichtige Übung 9.41** (Distanzfunktionen)**.** *Sei $(X, \mathrm{d})$ ein metrischer Raum. In dieser Übung möchten wir den Abstand von Teilmengen von $X$ zu Punkten diskutieren. Zu $x \in X$ und $A \subseteq X$ nicht-leer definieren wir*

$$\mathrm{d}(x, A) = \inf_{a \in A} \mathrm{d}(x, a).$$

*(i) Zeigen Sie, dass eine nicht-leere Teilmenge $A \subseteq X$ genau dann abgeschlossen ist, wenn für alle $x \in X$ die Eigenschaft $\mathrm{d}(x, A) = 0 \iff x \in A$ gilt.*

*(ii) Zeigen Sie, dass die Funktion $x \in X \mapsto \mathrm{d}(x, A) \in \mathbb{R}$ Lipschitz-stetig mit Lipschitz-Konstante 1 ist.*

### 9.3.4 Eine Bemerkung zur Stetigkeit

Für $d \geq 2$ gibt es für die Bewegung $x \to x_0$ im $\mathbb{R}^d$ mehr Möglichkeiten als im eindimensionalen $\mathbb{R}$. In der Definition des Grenzwerts $\lim_{x \to x_0} f(x)$ werden alle mögliche Bewegungen (entlang Geraden, entlang beliebigen Kurven oder auch ohne Formeln) auf $x_0$ zu erlaubt. Wir wollen die Bedeutung dieser Bemerkung mit einigen Beispielen belegen.

**Beispiel 9.42.** *(a) Sei $f : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch*

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{falls } (x,y)^t \neq (0,0)^t \\ 0 & \text{falls } (x,y)^t = (0,0)^t \end{cases}$$

*für $(x,y)^t \in \mathbb{R}^2$. Dann ist $f$ auf der $x$-Achse und ebenso auf der $y$-Achse identisch Null. Hingegen ist $f$ auf einer Geraden $y = ax$ zu $a \neq 0$ gleich $\frac{a}{1+a^2}$. Daher ist $f$ bei $(0,0)^t$ nicht stetig.*

*(b) Sei $f : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch*

$$f(x,y) = \begin{cases} \frac{x^2y}{x^4+y^2} & \text{falls } (x,y)^t \neq (0,0)^t \\ 0 & \text{falls } (x,y)^t = (0,0)^t \end{cases}$$

*für $(x,y)^t \in \mathbb{R}^2$. Dann ist $f$ auf der $x$-Achse und ebenso auf der $y$-Achse identisch Null. Des Weiteren ist $f$ auf einer Geraden $y = ax$ zu $a \neq 0$ gleich $\frac{ax^3}{x^4+a^2x^2} = \frac{ax}{a^2+x^2}$ und damit entlang dieser Gerade stetig. Dennoch ist $f$ bei $(0,0)^t$ nicht stetig, da für $y = x^2$ und $x \neq 0$*

$$f(x,y) = f(x,x^2) = \frac{x^4}{2x^4} = \frac{1}{2}$$

*ist.*

Bei der Definition der Stetigkeit einer Funktion $f : U \to \mathbb{R}$ auf einer offenen Menge $U \subseteq \mathbb{R}^d$ bei einem Punkt $x_0 \in U$ oder auch der Definition des Grenzwerts $\lim_{x\to x_0} f(x)$ darf man also nebst der Annahme, dass $x$ hinreichend nahe an $x_0$ liegt, keine weiteren Annahmen über die genauere Lage von $x$ zu $x_0$ treffen.

### 9.3.5 Stetigkeit und Zusammenhang

Stetige Abbildungen verhalten sich gut bezüglich zusammenhängenden Mengen.

**Proposition 9.43** (Bilder zusammenhängender Mengen unter stetigen Abbildungen)**.** *Seien $X, Y$ metrische Räume und $f : X \to Y$ stetig. Falls $X$ zusammenhängend ist, dann ist das Bild $f(X)$ ein zusammenhängender Teilraum von $Y$.*

Beispielsweise ist also $\mathbb{S}^1 = \{z \in \mathbb{C} \mid |z| = 1\} \subseteq \mathbb{C}$ zusammenhängend, da die stetige Abbildung $x \in \mathbb{R} \mapsto \mathrm{e}^{2\pi \mathrm{i} x} \in \mathbb{S}^1$ surjektiv ist.

*Beweis.* Wir können ohne Beschränkung der Allgemeinheit annehmen, dass $f : X \to Y$ surjektiv ist, indem wir $f : X \to Y$ durch die stetige, surjektive Abbildung $X \to f(X),\ x \mapsto f(x)$ ersetzen. Angenommen $Y$ ist nicht zusammenhängend. Dann existiert eine offene und abgeschlossene Teilmenge $A \subseteq Y$ mit $A \neq \emptyset$ und $A \neq Y$. Das Urbild $f^{-1}(A)$ von $A$ unter $f$ ist somit nach Proposition 9.37 ebenfalls offen und abgeschlossen. Da aber $X$ zusammenhängend ist, muss entweder $f^{-1}(A) = \emptyset$ oder $f^{-1}(A) = X$. Dann gilt aber $A = \emptyset$ respektive $A = Y$ wegen Surjektivität von $f$, was der Annahme an $A$ widerspricht. □

Eine direkte Konsequenz von Proposition 9.43 kennen wir bereits aus Kapitel 3.

**Korollar 9.44** (Zwischenwertsatz)**.** *Sei $I \subseteq \mathbb{R}$ ein Intervall, $f : I \to \mathbb{R}$ eine stetige Funktion und $a, b \in I$. Für jedes $c \in \mathbb{R}$ zwischen $f(a)$ und $f(b)$ gibt es ein $x \in I$ zwischen $a$ und $b$, so dass $f(x) = c$ gilt.*

*Beweis.* Ohne Beschränkung der Allgemeinheit können wir $a < b$ annehmen. Wir wenden Proposition 9.43 auf die stetige Abbildung $f|_{[a,b]} : [a, b] \to \mathbb{R}$ an. Demnach ist $f([a, b])$ zusammenhängend, da nach Proposition 9.30 das Intervall $[a, b]$ in $\mathbb{R}$ zusammenhängend ist. Wieder nach Proposition 9.30 muss $f([a, b]) \subseteq \mathbb{R}$ ein Intervall sein. Da $f(a), f(b) \in f([a, b])$ gilt, liegen alle Werte zwischen $f(a)$ und $f(b)$ im Bild von $f|_{[a,b]}$. □

**Übung 9.45.** *Zeigen Sie folgende Verallgemeinerung des Zwischenwertsatzes: Sei $X$ ein zusammenhängender metrischer Raum und $f : X \to \mathbb{R}$ eine stetige Abbildung. Seien $a, b \in X$. Dann existiert für jedes $c \in \mathbb{R}$ zwischen $f(a)$ und $f(b)$ ein $x \in X$ mit $f(x) = c$.*

Nebst einem besser strukturierten Beweis des Zwischenwertsatzes liefert Proposition 9.43 eine alternative Auffassung von „Zusammenhang“.

**Definition 9.46** (Wege und Wegzusammenhang)**.** Sei $X$ ein nicht-leerer metrischer Raum.

- Ein **Weg** in $X$ ist eine stetige Abbildung $\gamma : [a, b] \to X$ auf einem nicht-leeren Intervall $[a, b] \subseteq \mathbb{R}$ mit **Startpunkt** $\gamma(a)$ und **Endpunkt** $\gamma(b)$. Dabei sagen wir auch, dass $\gamma$ ein Weg von $\gamma(a)$ nach $\gamma(b)$ ist.
- Wir nennen $X$ **wegzusammenhängend**, falls für je zwei Punkte $x, y \in X$ zwei reelle Zahlen $a < b$ und ein Weg $\gamma : [a, b] \to X$ von $x = \gamma(a)$ nach $y = \gamma(b)$ existiert.

Wir können ohne Beschränkung der Allgemeinheit in obiger Definition auch als Definitionsbereich der Wege immer das Intervall $[0, 1]$ wählen. Denn für beliebige reelle Zahlen $a < b$ und einen Weg $\gamma : [a, b] \to X$ können wir stattdessen den Weg $t \in [0, 1] \mapsto \gamma(a + t(b - a))$ betrachten, welcher dieselben Start- und Endpunkte besitzt.

**Lemma 9.47.** *Jeder wegzusammenhängende metrische Raum ist zusammenhängend.*

*Beweis.* Sei $X$ ein nicht-leerer wegzusammenhängender metrischer Raum und sei $x \in X$ fix. Da $X$ wegzusammenhängend ist, existiert für jedes $y \in X$ ein Weg $\gamma_y : [0, 1] \to X$ von $x$ nach $y$. Nach Proposition 9.43 ist $\gamma_y([0, 1])$ für jedes $y \in X$ zusammenhängend. Nach Wahl unserer Wege haben wir

$$X = \bigcup_{y \in X} \gamma_y([0, 1]).$$

Da aber $x = \gamma_y(0) \in \gamma_y([0, 1])$ liegt für jedes $y \in X$, ist der Schnitt $\bigcap_{y \in X} \gamma_y([0, 1])$ nicht-leer. Nach der Verallgemeinerung von Lemma 9.31 in Übung 9.32 ist also $X = \bigcup_{y \in X} \gamma_y([0, 1])$ zusammenhängend. In der Tat, ist $A \subseteq X$ offen und abgeschlossen und ohne Beschränkung der Allgemeinheit $x \in A$, dann ist $A \cap \gamma_y([0, 1])$ offen und abgeschlossen in $\gamma_y([0, 1])$ und somit $\gamma_y([0, 1]) \subseteq A$ für alle $y \in X$. □

Umgekehrt muss ein zusammenhängender Raum nicht wegzusammenhängend sein, wie folgende Übung zeigt.

**Übung 9.48** (Zusammenhang v.s. Wegzusammenhang)**.** *Zeigen Sie, dass der Teilraum*

$$X = \{0\} \times [-1,1] \sqcup \left\{(t, \sin(\tfrac{1}{t})) \in \mathbb{R}^2 \mid t > 0\right\} \subseteq \mathbb{R}^2$$

*zusammenhängend, aber nicht wegzusammenhängend ist.*

Für gewisse Räume ist die Implikation „Zusammenhang $\implies$ Wegzusammenhang“ trotzdem richtig.

**Proposition 9.49.** *Sei $O \subseteq \mathbb{R}^d$ für $d \geq 1$ eine nicht-leere, offene Teilmenge. Dann ist $O$ genau dann wegzusammenhängend, wenn $O$ zusammenhängend ist.*

*Beweis.* Ist $O$ wegzusammenhängend, so ist $O$ auch zusammenhängend nach Lemma 9.47.

Sei nun $O$ zusammenhängend und $x_0 \in O$ ein fester Punkt. Wir definieren die Teilmenge

$$G = \{x \in O \mid \text{es gibt einen Weg in O von } x_0 \text{ nach } x\}$$

und wollen zeigen, dass $G$ sowohl offen als auch abgeschlossen ist.

Sei $x \in G$ und $\gamma : [0,1] \to O$ ein Weg von $x_0$ nach $x$. Da $O$ offen ist, gibt es ein $\varepsilon > 0$ mit $B_\varepsilon(x) \subseteq O$. Für ein $y \in B_\varepsilon(x)$ liegt der gerade Weg $t \in [0,1] \mapsto (1-t)x + ty$, der $x$ mit $y$ verbindet, in $O$. „Klebt“ man nun beide Wege zusammen, so erhält man den Weg

$$t \in [0,2] \mapsto \begin{cases} \gamma(t) & \text{falls } 0 \leq t \leq 1 \\ (2-t)x + (t-1)y & \text{falls } 1 < t \leq 2 \end{cases}$$

von $x_0$ nach $y$ (Stetigkeit folgt direkt oder aus Übung 9.40). Also liegt $y \in G$ und, da $y$ beliebig war, $B_\varepsilon(x) \subseteq G$. Wir haben somit gezeigt, dass $G$ offen ist.

Mit einem ähnlichen Argument zeigt man, dass $O \setminus G$ offen ist. Ist $x \notin G$ und $\varepsilon > 0$ mit $B_\varepsilon(x) \subseteq O$, so liegen alle Punkte in $B_\varepsilon(x)$ auch nicht in $G$. Denn wäre $y \in G \cap B_\varepsilon(x)$, so könnte man durch eine Verkettung von Wegen wie oben auch $x$ mit $x_0$ durch einen Weg verbinden. Also ist $B_\varepsilon(x) \subseteq O \setminus G$ und $O \setminus G$ ist offen.

Da $O$ aber zusammenhängend ist, muss entweder $G$ oder $O \setminus G$ leer sein. Da $x_0 \in G$ liegt, ist also $O \setminus G$ leer und $G = O$. Per Definition von $G$ lässt sich also jeder Punkt in $O$ durch einen Weg mit $x_0$ verbinden. Sind $y, y' \in O$ zwei beliebige Punkte, so ist die Verkettung eines Weges von $y$ nach $x_0$ mit einem Weg von $x_0$ nach $y'$ ein Weg von $y$ nach $y'$. Also ist $O$ wegzusammenhängend. □

VO: 28.02.18

**Übung 9.50** (Verkettung von Wegen)**.** *Wir möchten hier die oben verwendete Verkettung nochmals im Allgemeinen behandeln. Sei $X$ ein metrischer Raum und sei $\gamma : [0,1] \to X$ ein Weg von $x_0 \in X$ nach $x_1 \in X$.*

*(i) (Umkehren von Wegen) Zeigen Sie, dass $t \in [0,1] \mapsto \gamma(1-t)$ ein Weg von $x_1$ nach $x_0$ ist.*

*(ii) (Verketten von Wegen) Sei* $\tilde{\gamma} : [0,1] \to X$ *ein Weg von* $x_1$ *nach* $x_2 \in X$*. Zeigen Sie, dass*

$$t \in [0,1] \mapsto \begin{cases} \gamma(2t) & \textit{falls } 0 \leq t \leq \frac{1}{2} \\ \tilde{\gamma}(2t-1) & \textit{falls } \frac{1}{2} < t \leq 1 \end{cases}$$

*einen Weg von* $x_0$ *nach* $x_2$ *definiert.*

*(iii) Wiederholen Sie damit den Beweis von Proposition 9.49.*

## 9.4 Vollständigkeit

**Definition 9.51** (Cauchy-Folgen und Vollständigkeit)**.** Sei $(X, \mathrm{d})$ ein metrischer Raum. Eine Folge $(x_n)_n$ in $X$ heisst **Cauchy-Folge**, falls es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass für alle $m, n \geq N$ die Abschätzung $\mathrm{d}(x_n, x_m) < \varepsilon$ gilt. Der metrische Raum heisst **vollständig**, falls jede Cauchy-Folge in $X$ konvergiert.

Genauso wie für den Konvergenzbegriff ist der Begriff einer Cauchy-Folge in einem normierten Raum (siehe Abschnitt 5.7.4) kompatibel mit obiger Definition.

Wie dies schon für die reellen Zahlen erwähnt wurde, ist Vollständigkeit gewissermassen jene Eigenschaft, die analytische Methoden erlaubt. Dies ist beispielsweise auch für $\mathbb{R}^d$ mit $d \geq 1$ wahr.

**Proposition 9.52** (Vollständigkeit in endlich dimensionalen Vektorräumen)**.** *Sei $d \geq 1$. Jede abgeschlossene Teilmenge von $\mathbb{R}^d$ ist vollständig.*

Man beachte, dass beliebige Teilmengen von $\mathbb{R}^d$ nicht zwingend vollständig sind; beispielsweise ist $(\frac{1}{n})_n$ eine Cauchy-Folge in $X = (0, 1)$, konvergiert aber gegen 0 in $\mathbb{R}$ und hat somit keinen Grenzwert in $X$.

*Beweis.* Nach Satz 5.95 ist $\mathbb{R}^d$ vollständig. Sei nun $A \subseteq \mathbb{R}^d$ eine abgeschlossene Teilmenge und sei $(a_n)_n$ eine Cauchy-Folge in $A$. Dann besitzt die Folge $(a_n)_n$ einen Grenzwert $b \in \mathbb{R}^d$. Nach Lemma 9.19 muss aber $b$ in $A$ liegen, womit die Cauchy-Folge $(a_n)_n$ also in $A$ konvergiert. Somit ist $A$ vollständig. □

**Übung 9.53** (Vollständigkeit und Abgeschlossenheit)**.** *Sei $(X, \mathrm{d})$ ein metrischer Raum und $A \subseteq X$ eine Teilmenge.*

*(i) Angenommen $X$ ist vollständig und $A$ ist abgeschlossen. Zeigen Sie, dass der Teilraum $A$ auch vollständig ist.*

*(ii) Angenommen $A$ ist vollständig. Zeigen Sie, dass $A \subseteq X$ abgeschlossen ist.*

Genauso wie bei der Diskussion von Konvergenz wollen wir einen Vektorraum $V$ über $\mathbb{R}$ schlicht vollständig bezüglich einer Norm $\|\cdot\|$ nennen, falls er bezüglich der induzierten Metrik vollständig ist. Wie in Abschnitt 5.7.4 ist Vollständigkeit stabil bezüglich Normäquivalenz: Ist $V$ ein Vektorraum über $\mathbb{R}$ und sind $\|\cdot\|_1$ und $\|\cdot\|_2$ zwei äquivalente Normen auf $V$, so ist $V$ genau dann bezüglich $\|\cdot\|_1$ vollständig, wenn $V$ bezüglich $\|\cdot\|_2$ vollständig ist.

*Bemerkung.* Ein Banachraum ist per Definition ein vollständiger normierter Vektorraum. Interessant sind Banachräume unter anderem deswegen, weil sie (und Abbildungen auf ihnen) in der Praxis häufig vorkommen und oft relativ gute Eigenschaften besitzen. Für eine detaillierte Behandlung von Banachräumen verweisen wir auf die Vorlesung der Funktionalanalysis im dritten Studienjahr. Wird die Norm in einem Banachraum noch zusätzlich von einem inneren Produkt induziert, so spricht man von einem Hilbertraum. Hilberträume spielen unter anderem in der vollständigen Behandlung von Fourier-Reihen (siehe Appendix B) eine fundamentale Rolle.

### 9.4.1 Der Banachsche Fixpunktsatz

Sogenannte Fixpunktsätze beschreiben, wann es Punkte gibt, die unter einer Abbildung oder einer anderen Operation festgehalten werden. Sie sind insbesondere oft dann nützlich, wenn es darum geht, Existenzsätze zu beweisen. Beispielsweise werden wir den folgenden Fixpunktsatz für den Beweis der Existenz von Lösungen zu Differentialgleichungen verwenden.

**Satz 9.54** (Banachscher Fixpunktsatz)**.** *Sei* $(X, \mathrm{d})$ *ein nicht-leerer, vollständiger metrischer Raum. Sei* $T : X \to X$ *eine* ***Lipschitz-Kontraktion****, das heisst, eine Abbildung mit der Eigenschaft*

$$\mathrm{d}(T(x_1), T(x_2)) \leq \lambda \mathrm{d}(x_1, x_2)$$

*für alle* $x_1, x_2 \in X$ *und für eine fixe Lipschitz-Konstante* $\lambda < 1$*. Dann existiert ein eindeutig bestimmtes* $x_0 \in X$ *mit* $T(x_0) = x_0$ *(ein* ***Fixpunkt*** *von* $T$*).*

Wir möchten an dieser Stelle anmerken, dass die getroffenen Annahmen in Satz 9.54 notwendig sind – siehe Übung 9.55. Der Banachsche Fixpunktsatz besitzt auch durchaus alltägliche Anwendungen. Beispielsweise zeigt er, dass wenn ein Tourist an der ETH eine Stadtkarte von Zürich zu Boden fallen lässt, ein Punkt auf sich selbst zu liegen kommt (wieso?).

*Beweis.* Wir zeigen zuerst die behauptete Eindeutigkeit. Seien also $x_0, x_0' \in X$ zwei Fixpunkte von $T$. Dann gilt

$$\mathrm{d}(x_0, x_0') = \mathrm{d}(T(x_0), T(x_0')) \leq \lambda \mathrm{d}(x_0, x_0'),$$

was wegen $\lambda < 1$ also $\mathrm{d}(x_0, x_0') = 0$ und somit $x_0 = x_0'$ impliziert.

Für die Existenz wählen wir ein beliebiges $x_1 \in X$ und definieren rekursiv $x_2 = T(x_1)$, $x_3 = T(x_2)$ und allgemein $x_{n+1} = T(x_n)$ für alle $n \in \mathbb{N}$ (die Folge $(T^{\circ n}(x_1))_n$ nennt sich auch die Bahn von $x_1$ unter $T$). Wir möchten nun zeigen, dass die Folge $(x_n)_n$ konvergiert. In der Tat ist der Grenzwert $x_0$ dann ein Fixpunkt, denn es gilt

$$x_0 = \lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} T(x_n) = T(\lim_{n\to\infty} x_n) = T(x_0),$$

da $T : X \to X$ Lipschitz-stetig und somit auch stetig ist.

Wir bemerken zuerst, dass $\mathrm{d}(x_2, x_3) = \mathrm{d}(T(x_1), T(x_2)) \leq \lambda \mathrm{d}(x_1, x_2)$ und allgemeiner

$$\mathrm{d}(x_n, x_{n+1}) \leq \lambda^{n-1} \mathrm{d}(x_1, x_2) \tag{9.1}$$

für alle $n \in \mathbb{N}$. In der Tat folgt Ungleichung (9.1) mittels Induktion $n$. Für $n = 1$ ist die Ungleichung tautologisch erfüllt. Falls (9.1) bereits für $n$ gilt, dann folgt

$$\mathrm{d}(x_{n+1}, x_{n+2}) = \mathrm{d}(T(x_n), T(x_{n+1})) \leq \lambda \mathrm{d}(x_n, x_{n+1}) \leq \lambda^n \mathrm{d}(x_1, x_2)$$

nach Konstruktion, der Voraussetzung im Satz und der Induktionsverankerung.

Wir behaupten nun, dass Ungleichung (9.1) impliziert, dass $(x_n)_n$ eine Cauchy-Folge in $X$ bildet. Daraus folgt mit der vorausgesetzten Vollständigkeit, dass ein Grenzwert $x_0$ der Folge $(x_n)_n$ existiert, was wie oben erklärt dann den Beweis abschliesst. Sei $\varepsilon > 0$. Dann existiert ein $N \in \mathbb{N}$ mit $\frac{\lambda^{N-1}}{1-\lambda}\mathrm{d}(x_1, x_2) < \varepsilon$. Für $n > m \geq N$ folgt damit

$$\begin{aligned}
\mathrm{d}(x_m, x_n) &\leq \mathrm{d}(x_m, x_{m+1}) + \mathrm{d}(x_{m+1}, x_{m+2}) + \ldots + \mathrm{d}(x_{n-1}, x_n) \\
&\leq \lambda^{m-1}\mathrm{d}(x_1, x_2) + \lambda^m \mathrm{d}(x_1, x_2) + \ldots + \lambda^{n-2}\mathrm{d}(x_1, x_2) \\
&\leq \sum_{k=N}^{\infty} \lambda^{k-1}\mathrm{d}(x_1, x_2) = \frac{\lambda^{N-1}}{1-\lambda}\mathrm{d}(x_1, x_2) < \varepsilon,
\end{aligned}$$

was den Beweis abschliesst. □

Wir möchten kurz anmerken, dass der obige Ansatz einen Fixpunkt zu finden, rein geometrisch begründet ist. Ist ein Fixpunkt $x_0 \in X$ gegeben, so zieht die Abbildung $T$ alle Punkte näher an $x_0$ heran, denn es gilt

$$\mathrm{d}(T(x), x_0) = \mathrm{d}(T(x), T(x_0)) \leq \lambda \mathrm{d}(x, x_0).$$

Da also unter Iterieren von $T$ eine Folge entsteht, die $x_0$ immer näher kommt, ist das Vorgehen, den Fixpunkt $x_0$ als Grenzwert einer Bahn zu konstruieren, sehr sinnvoll.

**Übung 9.55** (Notwendigkeit der Annahmen in Satz 9.54)**.**

- *Finden Sie einen nicht-vollständigen metrischen Raum $X$ und eine Lipschitz-Kontraktion $T : X \to X$, die keinen Fixpunkt besitzt.*
- *Finden Sie einen vollständigen metrischen Raum $(X, \mathrm{d})$ und eine Isometrie (das heisst, eine Abbildung $T : X \to X$ mit $\mathrm{d}(T(x_1), T(x_2)) = \mathrm{d}(x_1, x_2)$), die keinen Fixpunkt besitzt.*

### 9.4.2 Wahrscheinlichkeitsmatrizen*

Wir wollen als erste Anwendung des Banachschen Fixpunktsatzes die Existenz eines „stationären Wahrscheinlichkeitsvektors“ zu einer „positiven Wahrscheinlichkeitsmatrix“ zeigen.

**Definition 9.56.** Sei $d \geq 1$. Eine $d \times d$-Matrix $A \in \mathrm{Mat}_{d,d}(\mathbb{R})$ heisst **nicht-negativ** (**positiv**), falls $A_{k\ell} \geq 0$ ($A_{k\ell} > 0$) für alle $k, \ell \in \{1, \ldots, d\}$ gilt. Eine nicht-negative Matrix $P \in \mathrm{Mat}_{d,d}(\mathbb{R})$ ist eine **Wahrscheinlichkeitsmatrix**, falls $\sum_{k=1}^{d} P_{k\ell} = 1$ für alle $\ell \in \{1, \ldots, d\}$ gilt. Ein Vektor $a \in \mathbb{R}^d$ ist ein **Wahrscheinlichkeitsvektor**, falls $a_j \geq 0$ für alle $j \in \{1, \ldots, d\}$ und $\sum_{j=1}^{d} a_j = 1$. Ein Wahrscheinlichkeitsvektor $a \in \mathbb{R}^d$ heisst **stationär** bezüglich einer Wahrscheinlichkeitsmatrix $P \in \mathrm{Mat}_{d,d}(\mathbb{R})$, falls $Pa = a$ gilt (also falls $a$ ein Eigenvektor von $P$ zum Eigenwert 1 ist).

In Anwendungen von obigen Begriffen und dem Resultat unten ist $\{1, \ldots, d\}$ eine Abzählung von verschiedenen Zuständen eines gegebenen Systems. Die Wahrscheinlichkeitsmatrix $P$ gibt zu Zuständen $k, \ell \in \{1, \ldots, d\}$ die Wahrscheinlichkeit $P_{k\ell} \in [0, 1]$ eines Übergangs vom Zustand $\ell$ zum $k$ an. Des Weiteren gibt ein stationärer Wahrscheinlichkeitsvektor (einen

solchen wollen wir finden) eine „zeitunabhängige Wahrscheinlichkeitsverteilung“ über die verschiedenen Zustände an.

Zum Beispiel könnten für $d = 5$ die Zustände das Winterwetter in Zürich angeben: Sonne, Nebel, Wolken, Regen, Schneefall. Die Matrix $P$ beschreibt wie gesagt die Wahrscheinlichkeiten der Übergänge. Beispielsweise gibt $P_{11}$ die Wahrscheinlichkeit an, dass ausgehend von einem sonnigen Tag der nächste Tag wieder sonnig ist; $P_{21}$ gibt die Wahrscheinlichkeit an, dass ausgehend von einem sonnigen Tag der nächste Tag neblig ist und so weiter. Unser Beweis des Resultats weiter unten gemeinsam mit dem Banachschen Fixpunktsatz hat in diesem Zusammenhang die folgende Interpretation. Ausgehend von einem sonnigen Tag gibt also die erste Spalte von $P$ den Wahrscheinlichkeitsvektor an, der das Wetter am nächsten Tag beschreibt. Wendet man nochmals $P$ an, erhält man den Wahrscheinlichkeitsvektor, der das Wetter am nächsten Tag beschreibt. Diese Vektoren stabilisieren sich zu einem stationären Wahrscheinlichkeitsvektor. Dieser Vektor ist derselbe, wenn wir stattdessen von einem Tag mit Nebel, Wolken, Regen oder Schneefall ausgehen.

**Korollar 9.57.** *Sei $d \geq 1$ und $P \in \mathrm{Mat}_{d,d}(\mathbb{R})$ eine positive Wahrscheinlichkeitsmatrix. Dann existiert ein eindeutig bestimmter stationärer Wahrscheinlichkeitsvektor zu $P$, der auch $a_j > 0$ für alle $j \in \{1, \ldots, d\}$ erfüllt.*

Dieser Satz ist ein Spezialfall des Satzes von Perron-Frobenius, welcher ähnliche Existenzaussagen für allgemeinere, nicht-negative Matrizen trifft (siehe Übung 9.58).

*Beweis.* Wir betrachten den Teilraum

$$X = \Big\{ a \in \mathbb{R}^d \mid a \text{ ist ein Wahrscheinlichkeitsvektor} \Big\}.$$

Dieser ist abgeschlossen: Falls $(a_n)_n$ eine Folge in $X$ ist und $a_n \to a \in \mathbb{R}^d$ gilt, dann ist wegen $\pi_j(a_n) \geq 0$ für $j \in \{1, \ldots, d\}$ auch $\pi_j(a) \geq 0$ auf Grund der Stetigkeit von $\pi_j$. Des Weiteren ist $\sum_{j=1}^d \pi_j(a_n) = 1$ und somit auch $\sum_{j=1}^d \pi_j(a) = 1$ nach Stetigkeit der Abbildung $b \mapsto \sum_{j=1}^d \pi_j(b)$. Also ist $X$ abgeschlossen nach der Charakterisierung in Lemma 9.19.

Nach Proposition 9.52 ist $X$ also auch vollständig, wobei wir nach Normäquivalenz wahlweise die Euklidsche Norm $\|\cdot\|_2$ oder die 1-Norm $\|\cdot\|_1$ verwenden dürfen. Wir werden die Norm $\|\cdot\|_1$ und ihre induzierte Metrik $\mathrm{d}_1$ verwenden.

Wir identifizieren die Matrix $P$ mit der zugehörigen linearen Abbildung $P : \mathbb{R}^d \to \mathbb{R}^d$ und definieren $T = P|_X : X \to X$. Dies ist in der Tat wohldefiniert: Für $a \in X$ gilt

$$(Pa)_k = \sum_{\ell=1}^d P_{k\ell} a_\ell \geq 0$$

für alle $k \in \{1, \ldots, d\}$ sowie

$$\sum_{k=1}^d (Pa)_k = \sum_{k=1}^d \sum_{\ell=1}^d P_{k\ell} a_\ell = \sum_{\ell=1}^d a_\ell \Big( \sum_{k=1}^d P_{k\ell} \Big) = \sum_{\ell=1}^d a_\ell = 1$$

und somit $Pa = T(a) \in X$.

Wir möchten nun zeigen, dass $T$ eine Lipschitz-Kontraktion von $X$ ist. Sind $a, a' \in X$, dann gilt

$$\mathrm{d}_1(Pa, Pa') = \|Pa - Pa'\|_1 = \|Py\|_1$$

für $y = a - a' \in Y = \left\{y \in \mathbb{R}^d \mid \sum_{\ell=1}^d y_\ell = 0\right\}$. Man beachte dabei, dass $P(Y) \subseteq Y$ gilt, da für alle $y \in Y$ die Identität $\sum_{k=1}^d (Py)_k = \sum_{k=1}^d \sum_{\ell=1}^d P_{k\ell} y_\ell = \sum_{\ell=1}^d y_\ell \sum_{k=1}^d P_{k\ell} = \sum_{\ell=1}^d y_\ell = 0$ erfüllt ist. Wir möchten nun zeigen, dass $\|Py\|_1 \leq (1-\eta)\|y\|_1$ für alle $y \in Y$ und $\eta = \min_{k,\ell} P_{k\ell} > 0$ gilt, womit nach obigem $T$ eine Lipschitz-Kontraktion mit Lipschitz-Konstante $\lambda = 1 - \eta < 1$ ist.

Sei nun $y \in Y$. Wir definieren die Indexmengen

$$L^+ = \big\{\ell \in \{1, \dots, d\} \mid y_\ell > 0\big\}, \quad L^- = \big\{\ell \in \{1, \dots, d\} \mid y_\ell < 0\big\}$$

und $K^+ = \big\{\ell \in \{1, \dots, d\} \mid (Py)_\ell > 0\big\}$. Dann gilt

$$\|y\|_1 = \sum_{\ell=1}^d |y_\ell| = \sum_{\ell \in L^+} y_\ell - \sum_{\ell \in L^-} y_\ell = 2 \sum_{\ell \in L^+} y_\ell = 2 \sum_{\ell \in L^-} |y_\ell|$$

auf Grund von $\sum_{\ell=1}^d y_\ell = 0$ und aus demselben Grund

$$\|Py\|_1 = 2 \sum_{k \in K^+} (Py)_k.$$

Daher ist

$$\begin{aligned}
\|Py\|_1 &= 2 \sum_{k \in K^+} \sum_{\ell=1}^d P_{k\ell} y_\ell = 2 \sum_{k \in K^+} \left( \sum_{\ell \in L^+} P_{k\ell} y_\ell + \sum_{\ell \in L^-} P_{k\ell} y_\ell \right) \\
&\leq 2 \sum_{\ell \in L^+} y_\ell \sum_{k \in K^+} P_{k\ell} - 2 \sum_{k \in K^+} \sum_{\ell \in L^-} \eta |y_\ell| \\
&\leq 2 \sum_{\ell \in L^+} y_\ell - |K^+| \, \eta \|y\|_1 \leq (1-\eta)\|y\|_1,
\end{aligned}$$

was wir zeigen wollten. Somit ist $T$ eine Lipschitz-Kontraktion. Nach dem Banachschen Fixpunktsatz (Satz 9.54) hat $T$ also einen eindeutigen Fixpunkt $a \in X$.

Wir müssen noch zeigen, dass $a_k > 0$ für alle $k \in \{1, \dots, d\}$ gilt. In der Tat ist

$$a_k = (Pa)_k = \sum_{\ell=1}^d P_{k\ell} a_\ell$$

genau dann gleich null, wenn $P_{k\ell} a_\ell = 0$ und somit, da $P$ positiv ist, $a_\ell = 0$ für alle $\ell$ in $\{1, \dots, d\}$ gilt. Dies ist aber nicht möglich, da $a$ ein Wahrscheinlichkeitsvektor ist. □

**Übung 9.58** (Verallgemeinerungen)**.**

(a) *Sei* $P \in \mathrm{Mat}_{d,d}(\mathbb{R})$ *eine Wahrscheinlichkeitsmatrix mit einer positiven Potenz, zum Beispiel* $P = \begin{pmatrix} \frac{1}{2} & 1 \\ \frac{1}{2} & 0 \end{pmatrix}$. *Zeigen Sie, dass* $P$ *ebenso die Konklusionen in Korollar 9.57 erfüllt und beweisen Sie, dass alle von* $1$ *verschiedenen Eigenwerte von* $P$ *Absolutbetrag kleiner* $1$ *haben.*

(b) *Sei* $P \in \mathrm{Mat}_{d,d}(\mathbb{R})$ *eine Wahrscheinlichkeitsmatrix mit der Eigenschaft, dass es für jedes Indexpaar* $k, \ell \in \{1, \ldots, d\}$ *ein* $n \in \mathbb{N}$ *gibt mit* $(P^n)_{k,\ell} > 0$, *zum Beispiel* $P = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. *Zeigen Sie, dass* $P$ *ebenso die Konklusionen in Korollar 9.57 erfüllt und beweisen Sie, dass alle von* $1$ *verschiedenen Eigenwerte von* $P$ *Absolutbetrag kleiner* $1$ *haben oder eine Wurzel aus* $1$ *sind.*

(c) *Finden Sie weitere Beispiele von Wahrscheinlichkeitsmatrizen, wo andere Phänomene eintreten und obige Aussagen nicht immer gelten.*

## 9.5 Kompaktheit

Wir haben in Abschnitt 3.8 die fundamentalen Eigenschaften Beschränktheit, Existenz eines Maximums (und Minimums) und gleichmässige Stetigkeit von stetigen Funktionen auf kompakten Intervallen bewiesen. Diese Eigenschaften waren dann grundlegend für die Riemann-Integrierbarkeit von stetigen Funktionen (siehe Satz 4.42), den Satz von Rolle und den Mittelwertsatz (Satz 7.28 und Theorem 7.29), und damit für die gesamte Differential- und Integralrechnung des ersten Semesters. Des Weiteren hatten wir in Satz 5.44 gesehen, dass jede Folge in einem kompakten Intervall eine konvergente Teilfolge besitzt. Wir wollen diese und auch noch weitere Eigenschaften im grösseren Rahmen der metrischen Räume untersuchen und werden feststellen, dass viele dieser Eigenschaften äquivalent sind.

### 9.5.1 Definitionen

**Definition 9.59** (Folgenkompaktheit und Beschränktheit)**.** Sei $X$ ein metrischer Raum. Wir sagen, dass $X$ **folgenkompakt** ist, falls jede Folge in $X$ eine konvergente Teilfolge (mit Grenzwert in $X$) besitzt. Ein Teilraum $K \subseteq X$ ist folgenkompakt, falls er als eigenständiger metrischer Raum folgenkompakt ist. Eine Teilmenge $B \subseteq X$ ist **beschränkt**, falls es ein Punkt $x_0 \in X$ und einen Radius $R > 0$ gibt, so dass $B$ im Ball $B_R(x_0)$ enthalten ist.

Da die Definition der Folgenkonvergenz nur von der Topologie (der Familie der offenen Teilmengen) abhängt, ist Folgenkompaktheit ein topologischer Begriff. Ganz im Gegensatz dazu ist Beschränktheit stark von der Wahl der Metrik abhängig. Betrachtet man beispielsweise $\mathbb{R}$ mit den Metriken $\mathrm{d}$ und $\mathrm{d}'$, wobei $\mathrm{d}$ die Euklidsche Metrik ist und $\mathrm{d}'$ die durch $\mathrm{d}'(x,y) = \min\{\mathrm{d}(x,y), 1\}$ für $x, y \in \mathbb{R}$ definierte Metrik ist, so sieht man, dass $\mathbb{R}$ bezüglich $\mathrm{d}'$ beschränkt ist. Allerdings ist die durch $\mathrm{d}'$ gegebene Topologie die Standardtopologie (die durch $\mathrm{d}$ gegebene Topologie) auf $\mathbb{R}$. (Wieso?)

Wir betrachten nun die Beziehung zwischen den beiden Begriffen „Folgenkompaktheit“ und „Beschränktheit“.

**Lemma 9.60** (Notwendige Eigenschaften)**.** *Eine folgenkompakte Teilmenge eines nicht-leeren metrischen Raumes ist abgeschlossen und beschränkt.*

Man beachte, dass die Umkehrung des Lemmas falsch ist (bezüglich der Metrik $\mathrm{d}'$ ist $\mathbb{R}$ abgeschlossen und beschränkt, aber nicht folgenkompakt – wieso?).

*Beweis.* Sei $X$ ein nicht-leerer metrischer Raum und $K \subseteq X$ folgenkompakt. Sei $(x_n)_n$ eine Folge in $K$ mit Grenzwert $x \in X$. Nach Definition 9.59 existiert eine in $K$ konvergente Teilfolge von $(x_n)_n$. Da eine Teilfolge einer konvergenten Folge denselben Grenzwert wie die ursprüngliche Folge besitzt, muss also $x$ in $K$ liegen. Nach der Charakterisierung abgeschlossener Teilmengen aus Lemma 9.19 ist $K$ abgeschlossen.

Sei nun $x_0 \in X$ ein beliebiger Punkt und angenommen $K$ ist nicht beschränkt. Dann gibt es für jedes $n \in \mathbb{N}$ ein $x_n \in K$ mit $\mathrm{d}(x_n, x_0) \geq n$. Da $K$ folgenkompakt ist, existiert nun eine in $K$ konvergente Teilfolge $(x_{n_\ell})_\ell$ der Folge $(x_n)_n$. Sei $x \in K$ ihr Grenzwert. Insbesondere gibt

es zu $\varepsilon = 1$ ein $L \in \mathbb{N}$, so dass $\mathrm{d}(x_{n_\ell}, x) < 1$ für alle $\ell \geq L$ gilt. Damit ist aber für $\ell \geq L$

$$n_\ell \leq \mathrm{d}(x_{n_\ell}, x_0) \leq \mathrm{d}(x_{n_\ell}, x) + \mathrm{d}(x, x_0) < 1 + \mathrm{d}(x, x_0).$$

Dies ist ein Widerspruch, da $n_\ell \to \infty$ für $\ell \to \infty$, und somit ist $K$ beschränkt. □

Wir möchten uns nun einem fundamentalen, weiteren Kompaktheitsbegriff zuwenden.

**Definition 9.61** (Überdeckungskompaktheit)**.** Sei $X$ ein metrischer Raum.

- Eine **offene Überdeckung** $\mathcal{O}$ von $X$ ist eine Kollektion offener Teilmengen von $X$, für die $X = \bigcup_{O \in \mathcal{O}} O$ gilt. Eine (**endliche**) **Teilüberdeckung** einer offenen Überdeckung $\mathcal{O}$ ist eine (endliche) Teilmenge $\mathcal{O}' \subseteq \mathcal{O}$, welche selbst eine offene Überdeckung bildet.
- Der Raum $X$ ist (**abzählbar**) **überdeckungskompakt**, falls für jede (abzählbare) offene Überdeckung eine endliche Teilüberdeckung existiert.
- Wir sagen, dass $X$ das (**abzählbare**) **Schachtelungsprinzip** erfüllt, falls für jede (abzählbare) Kollektion $\mathcal{A}$ abgeschlossener Teilmengen von $X$ mit $A_1 \cap \ldots \cap A_n \neq \emptyset$ für alle $n \in \mathbb{N}$ und $A_1, \ldots, A_n \in \mathcal{A}$ (auch **endliche Schnitteigenschaft** genannt) auch $\bigcap_{A \in \mathcal{A}} A \neq \emptyset$ gilt.

**Beispiel 9.62** (Offene Überdeckungen und Überdeckungskompaktheit)**.**

*(i) Jeder endliche metrische Raum $X$ ist überdeckungskompakt. In der Tat ist die Potenzmenge $\mathcal{P}(X)$ endlich und somit jede offene Überdeckung bereits endlich.*

*(ii) Der Teilraum $X = (0, 1] \subseteq \mathbb{R}$ ist nicht überdeckungskompakt. Tatsächlich besitzt die offene Überdeckung $\mathcal{O} = \left\{ \left(\frac{1}{n}, 1\right] \mid n \geq 2 \right\}$ keine endliche Teilüberdeckung. (Wieso?).*

*(iii) Falls wir stattdessen den Teilraum $X = [0, 1] \subseteq \mathbb{R}$ betrachten, dann bildet die Kollektion $\mathcal{O} = \left\{ \left(\frac{1}{n}, 1\right] \cap X \mid n \in \mathbb{N} \right\}$ keine offene Überdeckung mehr, da $0 \notin \left(\frac{1}{n}, 1\right]$ für alle $n \in \mathbb{N}$. Geben wir zu dieser Kollektion eine weitere offene Menge, die $0$ enthält (wie zum Beispiel $[0, \varepsilon) = (-\varepsilon, \varepsilon) \cap X$ für $\varepsilon > 0$), hinzu, dann erhalten wir eine offene Überdeckung, die eine endliche Teilüberdeckung besitzt (wieso gilt dies im Beispiel?). In der Tat ist $[0, 1]$ überdeckungskompakt (siehe Satz 9.66 weiter unten).*

Überdeckungen spielen bei vielen Beweisen eine wichtige Rolle, wie wir später sehen werden. Überdeckungskompakte Räume sind in diesem Zusammenhang einfach zu handhaben, weil man für eine gegebene Überdeckung (nach Annahme) meist nur endlich viele offene Mengen zu „kontrollieren“ braucht.

**Übung 9.63.** *Sei $X$ ein metrischer Raum. Zeigen Sie, dass $X$ genau dann überdeckungskompakt ist, wenn jede offene Überdeckung von $X$ bestehend aus offenen Bällen eine endliche Teilüberdeckung besitzt.*

**Definition 9.64** (Lebesgue-Zahl)**.** Sei $X$ ein metrischer Raum und $\mathcal{O}$ eine offene Überdeckung von $X$. Wir sagen, dass eine Zahl $r > 0$ eine **Lebesgue-Zahl** für $\mathcal{O}$ ist, falls es für jedes $x \in X$ eine offene Menge $O \in \mathcal{O}$ gibt mit $B_r(x) \subseteq O$.

Wir bemerken, dass in obiger Definition die gleichmässige Wahl von der Zahl $r$ unabhängig von $x \in X$ entscheidend für die Definition ist. Wir benötigen noch eine weitere Definition, die den Begriff der Beschränkheit auf eine wesentliche Art und Weise verstärkt.

**Definition 9.65** (Total beschränkt)**.** Sei $X$ ein metrischer Raum. Wir sagen, dass $X$ **total beschränkt** ist, falls es für jedes $\varepsilon > 0$ endlich viele $x_1, \ldots, x_n \in X$ gibt mit $X = \bigcup_{j=1}^{n} B_\varepsilon(x_j)$.

Beispielsweise ist $[0, 1]$ total beschränkt und $\mathbb{R}$ nicht total beschränkt (wieso?).

### 9.5.2 Äquivalenzen

Der folgende Satz stellt obige Begriffe in direkte Beziehung zu einander.

**Satz 9.66** (Kompaktheit)**.** *Sei $X$ ein metrischer Raum. Dann sind folgende neun Eigenschaften äquivalent. Wenn diese erfüllt sind, so nennen wir $X$ einen* ***kompakten metrischen Raum****.*

*(1) Jede unendliche Teilmenge von $X$ besitzt einen Häufungspunkt.*

*(2) $X$ ist folgenkompakt (das heisst, jede Folge in $X$ hat eine in $X$ konvergente Teilfolge).*

*(3) $X$ ist abzählbar überdeckungskompakt (das heisst, für jede Folge $O_1, O_2, \ldots$ offener Teilmengen in $X$ mit $X = \bigcup_{n\in\mathbb{N}} O_n$ gibt es ein $N \in \mathbb{N}$ mit $X = \bigcup_{n=1}^{N} O_n$).*

*(4) Jede stetige, komplexwertige Funktion auf $X$ ist beschränkt.* VO: 02.03.18

*(5) Jede stetige, reellwertige Funktion auf $X$ nimmt ein Maximum und ein Minimum an.*

*(6) Jede offene Überdeckung von $X$ besitzt eine Lebesgue-Zahl und $X$ ist total beschränkt.*

*(7) $X$ ist überdeckungskompakt.*

*(8) $X$ erfüllt das Schachtelungsprinzip.*

*(9) $X$ ist total beschränkt und vollständig.*

**Die Bedingungen (6)-(7) wurden in der Vorlesung nicht besprochen, sind daher auch nicht Prüfungsstoff. Proposition 9.67 wurde hingegen bewiesen und war auch mehrmals in den folgenden Kapiteln von Relevanz.*

*Beweis der Äquivalenz von (1)-(5).* Wir beweisen hier die Implikationen (1) $\implies$ (2) $\implies$ (3) $\implies$ (4) $\implies$ (5) $\implies$ (1).

Für die Implikation (1) $\implies$ (2) betrachten wir eine Folge $(x_n)_n$ in $X$ und die Bildmenge $D = \{x_n \mid n \in \mathbb{N}\}$. Falls $D$ endlich ist, so hat $(x_n)_n$ eine konstante und insbesondere konvergente Teilfolge. Wir nehmen nun also an, dass $D$ unendlich ist, womit $D$ nach (1) einen Häufungspunkt $x_0 \in X$ besitzt. Nach dem Analogon von Proposition 5.23 für metrische Räume – Übung 9.27 – reicht es zu zeigen, dass für $\varepsilon > 0$ und $N \in \mathbb{N}$ ein $n \geq N$ mit $x_n \in B_\varepsilon(x_0)$ existiert. Sei $\varepsilon' > 0$ so gewählt, dass

$$B_{\varepsilon'}(x_0) \setminus \{x_0\} \subseteq B_\varepsilon(x_0) \setminus \{x_1, \ldots, x_N\}\,.$$

Da $x_0$ ein Häufungspunkt ist, gibt es ein $n \in \mathbb{N}$ mit $x_n \in B_{\varepsilon'}(x_0) \setminus \{x_0\} \subseteq B_\varepsilon(x_0)$, womit insbesondere $n > N$ sein muss.

Wir zeigen (2) $\implies$ (3). Sei also $\mathcal{O} = \{O_1, O_2, \ldots\}$ eine abzählbare offene Überdeckung von $X$. Angenommen $\mathcal{O}$ besitzt keine endliche Teilüberdeckung. Dann gibt es für jedes $n \in \mathbb{N}$ einen Punkt $x_n \in X \setminus \bigcup_{j=1}^n O_n$. Nach Vorraussetzung hat die Folge $(x_n)_n$ eine konvergente Teilfolge $(x_{n_k})_k$; sei $x_0 \in X$ deren Grenzwert. Nach Konstruktion ist $x_n \in X \setminus O_\ell$ für alle $n \geq \ell$ und aus Lemma 9.19 folgt $x_0 \in X \setminus O_\ell$ für alle $\ell \in \mathbb{N}$. Dies widerspricht aber der Annahme, dass $\mathcal{O}$ eine offene Überdeckung von $X$ ist.

Für die Implikation (3) $\implies$ (4) sei $f : X \to \mathbb{C}$ eine stetige Funktion. Es gilt $\mathbb{C} = \bigcup_{n\in\mathbb{N}} B_n(0)$ und somit

$$X = f^{-1}(\mathbb{C}) = \bigcup_{n\in\mathbb{N}} f^{-1}(B_n(0)).$$

Nach Lemma 9.17 ist $\mathcal{O} = \left\{f^{-1}(B_n(0)) \mid n \in \mathbb{N}\right\}$ eine abzählbare offene Überdeckung, womit nach (3) ein $n \in \mathbb{N}$ mit $f^{-1}(B_n(0)) = \bigcup_{k=1}^n f^{-1}(B_k(0)) = X$ existiert. Es gilt also $|f(x)| < n$ für alle $x \in X$ wie behauptet.

Der Beweis der Implikation (4) $\implies$ (5) ist analog zum Beweis von Korollar 3.72; wir geben ihn hier der Vollständigkeit halber an. Sei $f : X \to \mathbb{R}$ stetig. Wir setzen $M = \sup f(X)$ und nehmen per Widerspruch an, dass $f$ sein Maximum nicht annimmt. Damit ist die Funktion

$$g : X \to \mathbb{R},\ x \mapsto \frac{1}{M - f(x)}$$

wohldefiniert und stetig. Nach Annahme in (4) gibt es nun ein $S > 0$ mit $g(x) \leq S$ oder äquivalenterweise $f(x) \leq M - \frac{1}{S}$ für alle $x \in X$. Dies widerspricht aber der Definition von $M$ als Supremum.

Wir zeigen nun (5) $\implies$ (1). Sei $D \subseteq X$ eine Teilmenge. Wir nehmen an, dass $D$ keine Häufungspunkte besitzt, und möchten nachweisen, dass $D$ endlich ist. Dazu möchten wir zuerst zeigen, dass es einen Radius $r > 0$ gibt, so dass für $x \in D$ der Ball $B_r(x)$ keinen weiteren Punkt von $D$ enthält. Hierfür betrachten wir die Funktion $\eta : X \to \mathbb{R}_{\geq 0}$ gegeben durch

$$\eta(x) = \sup\{\delta \geq 0 \mid |B_\delta(x) \cap D| \leq 1\}$$

für $x \in X$. Da $D$ keine Häufungspunkte besitzt, können wir für jedes $x \in X$ ein $\delta > 0$ finden mit $|B_\delta(x) \cap D| \leq 1$. Somit ist $\eta(x) > 0$ für alle $x \in X$. Wir bemerken noch, dass daher

$$\{\delta \geq 0 \mid |B_\delta(x) \cap D| \leq 1\} = [0, \eta(x)]$$

für alle $x \in X$ gilt. (Wieso? )

Wir behaupten nun, dass $\eta : X \to \mathbb{R}_{>0}$ Lipschitz-stetig mit Lipschitz-Konstante 1 (und insbesondere stetig) ist. Daraus folgt, dass $r = \min \eta(X)$ grösser Null ist und die oben gewünschte Eigenschaft erfüllt. Seien $x_1, x_2 \in X$. Falls $\varepsilon = \eta(x_1) - \mathrm{d}(x_1, x_2) > 0$ ist, gilt $B_\varepsilon(x_2) \subseteq B_{\eta(x_1)}(x_1)$ auf Grund der Dreiecksungleichung, womit $\eta(x_2) \geq \varepsilon = \eta(x_1) - \mathrm{d}(x_1, x_2)$.

Falls hingegen $\varepsilon = \eta(x_1) - \mathrm{d}(x_1, x_2) \leq 0$ ist, gilt dies ebenfalls. Es gilt also in jedem Fall $\eta(x_1) - \eta(x_2) \leq \mathrm{d}(x_1, x_2)$. Auf Grund der Symmetrie zwischen $x_1$ und $x_2$ schliessen wir also auf $|\eta(x_1) - \eta(x_2)| \leq \mathrm{d}(x_1, x_2)$ wie behauptet.

Wir nehmen nun zusätzlich an, dass $D$ unendlich ist und werden einen Widerspruch ableiten. Sei also $(x_n)_n$ eine Folge in $D$ mit $x_n \neq x_m$ für alle $n \neq m$. Zu $n \in \mathbb{N}$ und $x \in X$ setzen wir

$$f(x) = \begin{cases} n\big(\frac{1}{4}r - \mathrm{d}(x, x_n)\big) & \text{falls } x \in B_{r/4}(x_n) \text{ für ein } n \in \mathbb{N} \text{ und} \\ 0 & \text{sonst.} \end{cases}$$

Nach Wahl von $r$ gibt es für ein gegebenes $x \in X$ höchstens ein $n \in \mathbb{N}$ mit $x_n \in B_{r/2}(x)$, womit insbesondere die Funktion $f : X \to \mathbb{R}_{\geq 0}$ wohldefiniert ist. Des Weiteren ist $f$ stetig bei jedem $x \in X$, denn falls es kein $n \in \mathbb{N}$ mit $x_n \in B_{r/2}(x)$ gibt, so ist $f(x) = 0 = f(y)$ für alle $y \in B_{r/4}(x)$, und falls es doch ein eindeutig bestimmtes $n \in N$ mit $x_n \in B_{r/2}(x)$ gibt, so ist $f(y) = \max(0, n(\frac{1}{4}r - \mathrm{d}(y, x_n)))$ für alle $y \in B_{r/4}(x)$ und damit ebenso bei $x$ stetig. Es gilt aber $f(x_n) = \frac{1}{4}nr$ für alle $n \in \mathbb{N}$, womit $f$ unbeschränkt ist, was (5) widerspricht. Dieser Widerspruch zeigt, dass eine unendliche Teilmenge einen Häufungspunkt besitzen muss. □

**Proposition 9.67** (Lebesgue-Zahl)**.** *Sei $X$ ein folgenkompakter metrischer Raum. Dann hat jede offene Überdeckung eine Lebesgue-Zahl.*

*Beweis.* Sei $\mathcal{O}$ eine offene Überdeckung von $X$, für welche wir eine Lebesgue-Zahl finden wollen. Dazu betrachten wir die Hilfsfunktion $\eta : X \to \mathbb{R}_{>0}$ gegeben durch

$$\eta(x) = \sup\{\delta > 0 \mid \exists O \in \mathcal{O} : B_\delta(x) \subseteq O\}$$

Wir behaupten, dass diese Lipschitz-stetig ist mit Lipschitz-Konstante 1. Seien also $x_1, x_2 \in X$ und sei $\delta < \eta(x_1)$. Falls $\varepsilon = \delta - \mathrm{d}(x_1, x_2) > 0$ ist, gibt es nach Wahl von $\delta$ ein $O \in \mathcal{O}$ mit $B_\varepsilon(x_2) \subseteq B_\delta(x_1) \subseteq O$. Somit ist $\eta(x_2) \geq \varepsilon = \delta - \mathrm{d}(x_1, x_2)$, was auch für $\varepsilon \leq 0$ erfüllt ist. Nach Übergang zum Supremum ist $\eta(x_2) \geq \eta(x_1) - \mathrm{d}(x_1, x_2)$ oder in anderen Worten $\mathrm{d}(x_1, x_2) \geq \eta(x_1) - \eta(x_2)$. Aus Symmetriegründen folgt daraus $\mathrm{d}(x_1, x_2) \geq |\eta(x_1) - \eta(x_2)|$ wie gewünscht.

Nach Annahme an $X$ (und der schon bewiesenen Äquivalenz der Eigenschaften (1)-(5) in Satz 9.66) existiert somit ein $x_0 \in X$ mit $\eta(x_0) = \min \eta(X)$. Wir zeigen nun, dass $r_{\mathrm{Leb}} = \frac{1}{2}\eta(x_0) > 0$ eine Lebesgue-Zahl für $\mathcal{O}$ darstellt. In der Tat gilt für jedes $x \in X$ die Ungleichung $r_{\mathrm{Leb}} < \eta(x_0) \leq \eta(x)$, womit ein $O \in \mathcal{O}$ mit $B_{r_{\mathrm{Leb}}}(x) \subseteq O$ existiert. □

VO: Die verbleibenden Äquivalenzen sind hier nur für Interessierte inkludiert.

*Beweis der restlichen Äquivalenzen in Satz 9.66*.* Wir beginnen mit dem Beweis von (5) $\Longrightarrow$ (6). Nach obiger Proposition hat jede offene Überdeckung eine Lebesgue-Zahl. Es bleibt noch zu zeigen, dass $X$ total beschränkt ist. Dazu dürfen wir auf Grund der Äquivalenz von (5) und (2) verwenden, dass $X$ folgenkompakt ist. Sei $\varepsilon > 0$ und angenommen $X$ lässt sich nicht durch endlich viele Bälle vom Radius $\varepsilon > 0$ überdecken. Sei $x_1 \in X$ beliebig. Wir wählen rekursiv

$$x_n \in X \setminus (B_\varepsilon(x_1) \cup \ldots \cup B_\varepsilon(x_{n-1}))$$

für alle $n \geq 2$ (was nach der indirekten Annahme jeweils möglich ist). Somit erhalten wir eine Folge $(x_n)_n$ mit $\mathrm{d}(x_m, x_n) \geq \varepsilon$ für alle $m, n \in \mathbb{N}$ mit $m \neq n$. Sei $(x_{n_k})_k$ eine konvergente Teilfolge von $(x_n)_n$. Da aber jede konvergente Folge auch eine Cauchy-Folge ist, erhalten wir mit $\mathrm{d}(x_{n_k}, x_{n_\ell}) \geq \varepsilon$ für $k \neq \ell$ einen Widerspruch. Also lässt sich $X$ doch durch endlich viele Bälle vom Radius $\varepsilon > 0$ überdecken. Da $\varepsilon > 0$ beliebig war, ist $X$ total beschränkt.

Für die Implikation (6) $\implies$ (7) sei $\mathcal{O}$ eine beliebige offene Überdeckung von $X$ und $r > 0$ eine Lebesgue-Zahl für $\mathcal{O}$ nach (6). Des Weiteren existieren nach totaler Beschränktheit $x_1, \ldots, x_n \in X$ mit $X = \bigcup_{k=1}^n B_r(x_k)$. Für jeden solchen Ball $B_r(x_k)$ existiert nach Definition der Lebesgue-Zahl ein $O_k \in \mathcal{O}$ mit $B_r(x_k) \subseteq O_k$. Damit gilt also $X = \bigcup_{k=1}^n O_k$, womit $\{O_1, \ldots, O_n\}$ eine endliche Teilüberdeckung von $\mathcal{O}$ darstellt.

Wir zeigen nun (7) $\implies$ (8). Sei also $\mathcal{A}$ eine Kollektion abgeschlossener Teilmengen mit $\bigcap_{A \in \mathcal{A}} A = \emptyset$. Wir wollen zeigen, dass endlich viele $A_1, \ldots, A_N \in \mathcal{A}$ existieren mit $\bigcap_{n=1}^N A_n = \emptyset$. Nach Definition von $\mathcal{A}$ ist $\mathcal{O} = \{X \setminus A \mid A \in \mathcal{A}\}$ eine offene Überdeckung von $X$, womit nach (7) $O_1 = X \setminus A_1, \ldots, O_N = X \setminus A_N$ in $\mathcal{O}$ existieren mit $X = \bigcup_{n=1}^N O_n$. In anderen Worten ist $\bigcap_{n=1}^N A_n = \emptyset$ wie gewünscht.

Für die Implikationen (8) $\implies$ (9) beweisen wir zuerst, dass $X$ total beschränkt ist. Sei $\varepsilon > 0$ beliebig und $\mathcal{A} = \{X \setminus B_\varepsilon(x) \mid x \in X\}$. Dann ist $\bigcap_{A \in \mathcal{A}} A = \emptyset$, womit nach dem Schachtelungsprinzip $x_1, \ldots, x_n \in X$ mit $\bigcap_{k=1}^n (X \setminus B_\varepsilon(x_k)) = \emptyset$ oder in anderen Worten mit $\bigcup_{k=1}^n B_\varepsilon(x_k) = X$ existieren müssen.

Um die Vollständigkeit zu verifizieren, betrachten wir eine Cauchy-Folge $(x_n)_n$ und nehmen an, dass $(x_n)_n$ nicht konvergiert. Da eine Cauchy-Folge genau dann konvergiert, wenn sie eine konvergente Teilfolge besitzt, ist für jedes $n \in \mathbb{N}$ die Teilmenge $D_n = \{x_k \mid k \geq n\}$ abgeschlossen. Weiter gilt für alle $n \in \mathbb{N}$

$$\bigcap_{k=1}^n D_k = D_n \neq \emptyset,$$

womit nach dem Schachtelungsprinzip $\bigcap_{k \in \mathbb{N}} D_k \neq \emptyset$, was einen Widerspruch darstellt.

Für die verbleibende Implikation (9) $\implies$ (1) sei $D \subseteq X$ eine unendliche Teilmenge. Wir möchten eine Cauchy-Folge $(x_n)_n$ in $D$ mit paarweise verschiedenen Folgenglieder konstruieren, welche dann nach (9) konvergieren muss. Der Grenzwert von $(x_n)_n$ ist ein Häufungspunkt von $D$. (Wieso?). Da $X$ total beschränkt ist, gibt es eine endliche Teilmenge $F_1 \subseteq X$ mit

$$X = \bigcup_{y \in F_1} B_1(y).$$

Sei $y_1 \in X$ mit $|B_1(y_1) \cap D| = \infty$ und sei $D_1 = B_1(y_1) \cap D$. Im nächsten Schritt schreiben wir $X = \bigcup_{y \in F_2} B_{\frac{1}{2}}(y)$ für eine endliche Teilmenge $F_2 \subseteq X$, wählen ein $y_2 \in F_2$ mit $|B_{\frac{1}{2}}(y_2) \cap D_1| = \infty$ und setzen $D_2 = B_{\frac{1}{2}}(y_2) \cap D_1$. Fährt man so fort, erhält man eine absteigende Folge

$$D_1 \supseteq D_2 \supseteq D_3 \supseteq \ldots$$

von Teilmengen von $D$ mit $|D_k| = \infty$ sowie $D_k \subseteq B_{\frac{1}{k}}(y_k)$ für alle $k \in \mathbb{N}$. Für jedes $n \in \mathbb{N}$ sei

nun $x_n \in D_n$ ein beliebiger Punkt. Die resultierende Folge $(x_n)_n$ ist eine Cauchy-Folge, da für $n \geq m$ gilt $x_n, x_m \in B_{\frac{1}{n}}(y_n)$, womit $\mathrm{d}(x_n, x_m) < \frac{2}{n}$. □

Der Beweis der Überdeckungskompaktheit eines kompakten Intervalles in $\mathbb{R}$ ist nicht wesentlich einfacher als die allgemeine Äquivalenz zur Folgenkompaktheit in Satz 9.66. Trotzdem mag es hilfreich sein, diesen Fall getrennt zu untersuchen, da dadurch die Ähnlichkeiten zwischen totaler Beschränktheit und dem Intervallschachtelungsprinzip aufgezeigt werden können.

**Übung 9.68.** *Sei $[a, b] \subseteq \mathbb{R}$ ein nicht-leeres, kompaktes Intervall. Verwenden Sie das Intervallschachtelungsprinzip, um die Überdeckungskompaktheit von $[a, b]$ zu beweisen.*

**Übung 9.69** (Alternative Beweise)**.** *Beweisen Sie Lemma 9.60 unter Verwendung von Überdeckungskompaktheit.*

### 9.5.3 Endlich-dimensionale Vektorräume

Wie wir bei Lemma 9.60 schon gesehen haben, ist eine abgeschlossene und beschränkte Teilmenge eines metrischen Raumes nicht zwingend kompakt. Betrachtet man aber $\mathbb{R}$ oder allgemeiner $\mathbb{R}^d$ mit der Euklidschen Metrik, dann stimmt auch die Umkehrung. Wenn nichts anderes angeben ist, so statten wir $\mathbb{R}^d$ für $d \geq 1$ mit der Euklidschen Metrik aus.

**Satz 9.70** (Heine-Borel)**.** *Eine Teilmenge $K \subseteq \mathbb{R}^d$ für $d \geq 1$ ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.*

*Beweis.* Ist $K$ kompakt, dann ist $K$ nach Lemma 9.60 abgeschlossen und beschränkt.

Für die Umkehrung zeigen wir, dass $K$ folgenkompakt ist. Sei also $(x_n)_n$ eine Folge in $K$. Da $K$ beschränkt ist, ist die Folge beschränkt und besitzt nach Theorem 5.94 eine konvergente Teilfolge $(x_{n_k})_k$. Der Grenzwert $x \in \mathbb{R}^d$ dieser Teilfolge muss in $K$ liegen, da $K$ abgeschlossen ist (Lemma 9.19). □

Wie nützlich die Aussage von Satz 9.70 gemeinsam mit Satz 9.66(5) sein kann, illustriert beispielsweise die folgende Übung. In dieser möchten wir zeigen, dass je zwei Normen auf $\mathbb{R}^d$ äquivalent sind, wie schon in Abschnitt 5.7.2 versprochen wurde.

**Wichtige Übung 9.71** (Alle Normen auf $\mathbb{R}^d$ sind äquivalent)**.** *Sei $d \geq 1$ und seien $\|\cdot\|, \|\cdot\|'$ zwei Normen auf $\mathbb{R}^d$. Wir möchten in folgenden Schritten zeigen, dass $\|\cdot\|$ und $\|\cdot\|'$ äquivalent sind.*

*(i) Erklären Sie, wieso Sie annehmen können, dass $\|\cdot\|' = \|\cdot\|_1$ die Einsnorm ist.*

*(ii) Finden Sie eine Konstante $C > 0$ mit $\|x\| \leq C\|x\|_1$ für alle $x \in \mathbb{R}^d$. Schliessen Sie daraus auch, dass $\|\cdot\| : \mathbb{R}^d \to \mathbb{R}$ eine bezüglich der euklidischen Metrik $\mathrm{d}_2$ stetige Abbildung ist.*

*(iii) Sei $A = \{x \in \mathbb{R}^d \mid \|x\|_1 = 1\}$. Zeigen Sie, dass die Einschränkung von $\|\cdot\|$ auf $A$ ein Minimum annimmt.*

*(iv) Verwenden Sie (iii) um eine Konstante $C' > 0$ mit $C'\|x\|_1 \leq \|x\|$ für alle $x \in \mathbb{R}^d$ zu finden und schliessen Sie damit auf die zu beweisende Aussage.*

### 9.5.4 Bilder von kompakten Mengen

Eine weitreichende Verallgemeinerung der Existenz eines Maximums und eines Minimums stetiger, reellwertiger Funktionen auf kompakten Räumen (in Satz 9.66(5)) ist das folgende Resultat.

**Satz 9.72** (Kompaktes Bild)**.** *Seien $X, Y$ metrische Räume und sei $f : X \to Y$ eine stetige Abbildung. Falls $X$ kompakt ist, so ist auch $f(X)$ kompakt.*

*Beweis.* Gegeben eine Folge $(y_n)_n$ in $f(X)$ schreiben wir $f(x_n) = y_n$ für $n \in \mathbb{N}$ und $x_n \in X$. Dann existiert eine konvergente Teilfolge $(x_{n_k})_k$ von $(x_n)_n$ mit Grenzwert $x_0$, da $X$ kompakt ist. Wegen Folgenstetigkeit von $f$ konvergiert also die Teilfolge $(f(x_{n_k}))_k = (y_{n_k})_k$ gegen $f(x_0)$. Die Folge $(y_n)_n$ war aber beliebig. Somit ist $f(X)$ kompakt. □

**Übung 9.73.** *Beweisen Sie Satz 9.72 jeweils mit den Charakterisierungen (5) und (7) von Kompaktheit aus Satz 9.66.*

### 9.5.5 Operatornorm von Matrizen

Als eine weitere Anwendung von Satz 9.66 möchten wir in diesem Abschnitt eine natürliche Norm auf dem Vektorraum $\mathrm{Mat}_{m,n}(\mathbb{R})$ der reellen $m \times n$-Matrizen definieren. Wir bemerken zuerst, dass $\mathrm{Mat}_{m,n}(\mathbb{R})$ als reeller Vektorraum zu $\mathbb{R}^{mn}$ isomorph ist. Insbesondere kann man mit einer beliebigen Norm auf $\mathbb{R}^{mn}$ eine Norm auf $\mathrm{Mat}_{m,n}(\mathbb{R})$ induzieren (und umgekehrt). Alle solche Normen sind nach Übung 9.71 äquivalent, womit wir eine natürliche induzierte Topologie auf $\mathrm{Mat}_{m,n}(\mathbb{R})$ erhalten. Selbstverständlich kann man auch direkt auf $\mathrm{Mat}_{m,n}(\mathbb{R})$ eine Topologie definieren.

**Übung 9.74** (Hilbert-Schmidt-Norm)**.** *Zeigen Sie, dass $\langle A, B\rangle = \mathrm{Tr}(AB^T)$ ein Skalarprodukt auf $\mathrm{Mat}_{n,n}(\mathbb{R})$ definiert. Die induzierte Norm wird die Hilbert-Schmidt-Norm auf $\mathrm{Mat}_{n,n}(\mathbb{R})$. Identifizieren Sie die dazugehörige Norm auf $\mathbb{R}^{n^2}$.*

In unserem Kontext ist vor allem die folgende Norm auf $\mathrm{Mat}_{m,n}(\mathbb{R})$ nützlich.

**Definition 9.75** (Operatornorm)**.** Die **Operatornorm** einer Matrix $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$ ist durch

$$\|A\|_{\mathrm{op}} = \sup_{x \in \mathbb{R}^n : \|x\|_2 \leq 1} \|Ax\|_2$$

definiert.

Wir bemerken an dieser Stelle, dass sowohl auf $\mathbb{R}^n$ als auch $\mathbb{R}^m$ in obiger Definition eine andere Norm hätte verwendet werden können, womit man wiederum eine andere Operatornorm erhalten hätte. Allerdings werden wir nur die obige Operatornorm verwenden.

VO: 05.03.18

**Lemma 9.76** (Eigenschaften der Operatornorm)**.** *Seien $m, n \in \mathbb{N}$. Dann definiert $\|\cdot\|_{\mathrm{op}}$ in der Tat eine (wohldefinierte) Norm auf $\mathrm{Mat}_{m,n}(\mathbb{R})$ und erfüllt*

$$\|Ax\|_2 \leq \|A\|_{\mathrm{op}} \|x\|_2 \tag{9.2}$$

*für alle* $x \in \mathbb{R}^n$ *und alle* $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$. *Des Weiteren gilt für* $k \in \mathbb{N}$, $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$ *und* $B \in \mathrm{Mat}_{n,k}(\mathbb{R})$

$$\|AB\|_{\mathrm{op}} \leq \|A\|_{\mathrm{op}}\|B\|_{\mathrm{op}}. \tag{9.3}$$

*Beweis.* Nach dem Satz von Heine Borel (Satz 9.70) ist $\overline{B_1(0)} = \{x \in \mathbb{R}^n \mid \|x\|_2 \leq 1\}$ kompakt. Da die Abbildung $x \in \overline{B_1(0)} \mapsto \|Ax\|_2$ (als Verknüpfung von stetigen Funktionen) stetig ist, ist sie auf Grund von Satz 9.66(4) beschränkt, womit $\|A\|_{\mathrm{op}} < \infty$.

Für die Definitheit sei $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$ mit $\|A\|_{\mathrm{op}} = 0$. Dann ist $\|Ae_i\|_2 \leq \|A\|_{\mathrm{op}} = 0$ und somit $Ae_i = 0$ für alle $i \in \{1, \ldots, n\}$. Also ist $A = 0$. Für $\lambda \in \mathbb{R}$ und $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$ ist

$$\begin{aligned}\|\lambda A\|_{\mathrm{op}} &= \sup_{x \in \mathbb{R}^n : \|x\|_2 \leq 1} \|\lambda A x\|_2 = \sup_{x \in \mathbb{R}^n : \|x\|_2 \leq 1} |\lambda| \|Ax\|_2 \\ &= |\lambda| \sup_{x \in \mathbb{R}^n : \|x\|_2 \leq 1} \|Ax\|_2 = |\lambda| \|A\|_{\mathrm{op}}.\end{aligned}$$

Für $A, B \in \mathrm{Mat}_{m,n}(\mathbb{R})$ und $x \in \overline{B_1(0)}$ gilt

$$\|(A+B)x\|_2 = \|Ax + Bx\|_2 \leq \|Ax\|_2 + \|Bx\|_2 \leq \|A\|_{\mathrm{op}} + \|B\|_{\mathrm{op}}$$

und somit nach Übergang zum Supremum $\|A + B\|_{\mathrm{op}} \leq \|A\|_{\mathrm{op}} + \|B\|_{\mathrm{op}}$.

Für die Ungleichung (9.2) stellt man zuerst fest, dass im Falle $x = 0$ nichts zu beweisen ist. Ist $x \neq 0$, so gilt

$$\|Ax\|_2 = \left\|A \tfrac{x}{\|x\|_2}\right\| \|x\|_2 \leq \|A\|_{\mathrm{op}} \|x\|_2.$$

Für die Ungleichung (9.3) berechnet man zu $x \in \overline{B_1(0)}$ mit (9.2)

$$\|ABx\|_2 \leq \|A\|_{\mathrm{op}} \|Bx\|_2 \leq \|A\|_{\mathrm{op}} \|B\|_{\mathrm{op}},$$

womit die Behauptung folgt. □

Analog lässt sich auch eine Topologie sowie eine Operatornorm auf $\mathrm{Mat}_{m,n}(\mathbb{C})$ durch

$$\|A\|_{\mathrm{op}} = \sup_{x \in \mathbb{C}^n : \|x\|_2 \leq 1} \|Ax\|_2$$

für $A \in \mathrm{Mat}_{m,n}(\mathbb{C})$ definieren.

### 9.5.6 Gleichmässige Stetigkeit und Oszillation

Stetige Abbildungen auf kompakten Mengen haben die (wichtige) Eigenschaft der gleichmässigen Stetigkeit.

**Proposition 9.77.** *Seien* $X, Y$ *zwei metrische Räume und* $f : X \to Y$ *eine stetige Funktion. Falls* $X$ *kompakt ist, so ist* $f$ *gleichmässig stetig.*

Dies ist eine Verallgemeinerung des Satzes von Heine (Satz 3.78), welcher den Spezialfall $X = [a, b] \subseteq \mathbb{R}$ für $a < b$ behandelt. Wir ermutigen Leserinnen und Leser allerdings dazu, den Beweis in diesem, allgemeinen Fall durchzuführen. Da der Beweis weitgehend analog ist, möchten wir an dieser Stelle darauf verzichten und stattdessen einen Beweis in einem noch allgemeineren und für uns später wichtigen Kontext liefern.

**Definition 9.78.** Sei $f : X \to \mathbb{R}$ eine beschränkte, reellwertige Funktion auf einem metrischen Raum $X$. Für $x \in X$ ist die **Oszillation** oder **Schwankung** von $f$ bei $x$ durch

$$\omega(f, x) = \lim_{\delta \searrow 0} \omega(f, x, \delta)$$

definiert, wobei für $\delta > 0$

$$\omega(f, x, \delta) = \sup f(B_\delta(x)) - \inf f(B_\delta(x)).$$

Wir bemerken dazu, dass für $f$ wie oben und für $0 < \delta_1 < \delta_2$ die Inklusion $B_{\delta_1}(x) \subseteq B_{\delta_2}(x)$ gilt und dies

$$\sup f(B_{\delta_1}(x)) \leq \sup f(B_{\delta_2}(x)), \quad \inf f(B_{\delta_1}(x)) \geq \inf f(B_{\delta_2}(x)),$$

und damit auch $\omega(f, x, \delta_1) \leq \omega(f, x, \delta_2)$ impliziert. Insbesondere zeigt dies, dass der Grenzwert für $\delta \searrow 0$ in Definition 9.78 tatsächlich existiert. Des Weiteren bemerken wir, dass $f$ genau dann bei $x \in X$ stetig ist, wenn $\omega(f, x) = 0$ ist. (Wieso?).

**Proposition 9.79** (Gleichmässig kleine Oszillation)**.** *Sei $X$ eine kompakter metrischer Raum und sei $f : X \to \mathbb{R}$ eine beschränkte Funktion. Angenommen es gibt $\eta \geq 0$, so dass $\omega(f, x) \leq \eta$ für alle $x \in X$. Dann existiert für jedes $\varepsilon > 0$ ein $\delta > 0$, so dass für alle $x \in X$*

$$\omega(f, x, \delta) < \eta + \varepsilon$$

*gilt.*

Wir bemerken, dass der Fall $\eta = 0$ gerade der Aussage, dass eine stetige Funktion auf einer kompakten Menge gleichmässig stetig ist, entspricht (Proposition 9.77).

*Beweis.* Angenommen die Aussage der Proposition trifft nicht zu. Sei also $\varepsilon > 0$, so dass für alle $\delta > 0$ ein $x \in X$ mit $\omega(f, x, \delta) \geq \eta + \varepsilon$ existiert. Zu $m \in \mathbb{N}$ und $\delta_m = \frac{1}{m}$ können wir also $x_m \in X$ mit $\omega(f, x_m, \frac{1}{m}) \geq \eta + \varepsilon$ wählen. Nach Definition von $\omega(f, x_m, \frac{1}{m})$ zeigt dies, dass es auch $x_m^+, x_m^- \in B_{\frac{1}{m}}(x_m)$ mit

$$f(x_m^+) - f(x_m^-) > \eta + \frac{\varepsilon}{2} \tag{9.4}$$

gibt. Da $X$ als kompakt vorausgesetzt wurde, existiert eine konvergente Teilfolge $(x_{m_\ell})$ in $X$. Sei

$$x = \lim_{\ell \to \infty} x_{m_\ell} = \lim_{\ell \to \infty} x_{m_\ell}^+ = \lim_{\ell \to \infty} x_{m_\ell}^- \in X$$

deren Grenzwert.

Zu jedem $\delta > 0$ gibt es insbesondere ein $\ell$ mit $x^+_{m_\ell}, x^-_{m_\ell} \in B_\delta(x)$. Die Ungleichung (9.4) impliziert $\omega(f, x, \delta) > \eta + \frac{\varepsilon}{2}$. Da $\delta > 0$ beliebig war, erhalten wir $\omega(f, x) \geq \eta + \frac{\varepsilon}{2}$, was unserer Voraussetzung widerspricht. $\square$

**Übung 9.80.** *Beweisen Sie Proposition 9.79 unter Verwendung einer offenen Überdeckung und der Existenz einer Lebesgue-Zahl in Proposition 9.67.*

## 9.6 Fundamentalsatz der Algebra

**Theorem 9.81** (Fundamentalsatz der Algebra)**.** *Jedes komplexe Polynom $f \in \mathbb{C}[z]$ mit positivem Grad $n = \deg(f) \geq 1$ hat mindestens eine Nullstelle in $\mathbb{C}$. Des Weiteren lässt sich $f$ als Produkt eines Skalars $a_n \in \mathbb{C}^\times$ und $n$ Linearfaktoren der Form $(z - \alpha_j)$ schreiben, das heisst*

$$f = a_n \prod_{j=1}^{n} (z - \alpha_j)$$

*mit $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$.*

VO: 07.03.18

Es genügt dabei, den ersten Teil des Theorems zu beweisen. Der zweite Teil folgt dann aus dem ersten, Division mit Rest (siehe Abschnitt 3.2) und Induktion. Der Fundamentalsatz der Algebra hat viele verschiedene Beweise; einige davon beruhen auf Funktionentheorie (siehe zweites Studienjahr) oder algebraischer Topologie (siehe drittes oder viertes Jahr des Mathematikstudiums). Wir werden hier das asymptotische Verhalten von Polynomen (Proposition 3.15), den Satz von Heine-Borel (Satz 9.70) und die Polarzerlegung komplexer Zahlen (Lemma 6.81) verwenden.

*Beweis.* Sei $f \in \mathbb{C}[z]$ ein komplexes Polynom mit Grad $n = \deg(f) \geq 1$. Falls $f(0) = 0$ ist, so haben wir bereits eine Nullstelle gefunden. Also angenommen $M = 2|f(0)| > 0$. Nach Proposition 3.15 gibt es daher ein $R \geq 1$, so dass für alle $z \in \mathbb{C}$ mit $|z| \geq R$ auch $|f(z)| \geq M$ gilt. Nach dem Satz von Heine-Borel (Satz 9.70) ist

$$K = \{z \in \mathbb{C} \mid |z| \leq R\}$$

eine kompakte Teilmenge. Wir wenden nun die Existenz des Minimalwertes in Satz 9.66(5) auf die stetige Funktion $z \in K \mapsto |f(z)| \in \mathbb{R}$ an und finden ein $z_0 \in K$ mit $|f(z_0)| = \min_{z \in K} |f(z)|$.

Da $|f(0)| = \frac{M}{2}$, gilt die Ungleichung $|f(z_0)| < M$. Da auch $|f(z)| \geq M$ für alle $z \in \mathbb{C} \setminus K$ gilt, erhalten wir damit $|f(z)| \geq |f(z_0)|$ für alle $z \in \mathbb{C}$. Die Behauptung ist nun, dass $z_0$ eine Nullstelle von $f$ ist. Wir nehmen stattdessen an, dass $|f(z_0)| > 0$ ist.

Wir schreiben $f$ als Potenzreihe um $z_0$ (für das Polynom folgt die Existenz dieser Darstellung aber auch einfach mittels Division mit Rest und Induktion nach $n = \deg(f)$)

$$f(z) = \sum_{k=0}^{n} b_k (z - z_0)^k$$

mit $b_0 = f(z_0)$, $b_1 = f'(z_0)$, $\ldots$ , $b_n = \frac{f^{(n)}(z_0)}{n!} \in \mathbb{C}$. Per Annahme an $z_0$ gilt $b_0 \neq 0$, was wir nun auf einen Widerspruch führen wollen, indem wir einen Punkt $z$ in der Nähe von $z_0$ mit $|f(z)| < |b_0|$ finden. Sei $\ell \geq 1$ der kleinste Index grösser gleich 1 mit $b_\ell \neq 0$. Dann gilt

$$f(z) = b_0 + b_\ell (z - z_0)^\ell + \sum_{k=\ell+1}^{n} b_k (z - z_0)^k = b_0 + b_\ell (z - z_0)^\ell + O(|z - z_0|^{\ell+1})$$

für $z \to z_0$. Wir setzen $z = z_0 + r\mathrm{e}^{\mathrm{i}\varphi}$ für ein festes $\varphi \in \mathbb{R}$ (welches wir später wählen wollen) und ein variierendes $r > 0$, womit

$$f(z_0 + r\mathrm{e}^{\mathrm{i}\varphi}) = b_0 + b_\ell r^\ell \mathrm{e}^{\mathrm{i}\ell\varphi} + O(r^{\ell+1}) = b_0 \left(1 + \tfrac{b_\ell}{b_0} r^\ell \mathrm{e}^{\mathrm{i}\ell\varphi}\right) + O(r^{\ell+1})$$

für $r \searrow 0$. Wir schreiben $\frac{b_\ell}{b_0} = s\mathrm{e}^{\mathrm{i}\psi}$ für ein $s > 0$ und $\psi \in [0, 2\pi)$ und erhalten

$$f(z_0 + r\mathrm{e}^{\mathrm{i}\varphi}) = b_0 \left(1 + s\mathrm{e}^{\mathrm{i}\psi} r^\ell \mathrm{e}^{\mathrm{i}\ell\varphi}\right) + O(r^{\ell+1}) = b_0 \left(1 + s r^\ell \mathrm{e}^{\mathrm{i}(\ell\varphi+\psi)}\right) + O(r^{\ell+1})$$

für $r \searrow 0$. Wir möchten nun den obigen Term bei $r^\ell$ so arrangieren, dass er negativ und reell ist. Dafür setzt man $\varphi = \frac{-\psi+\pi}{\ell}$, so ist der obige Winkel $\ell\varphi + \psi$ gleich $\pi$ und $\mathrm{e}^{\mathrm{i}(\ell\varphi+\psi)} = -1$. Zusammenfassend gilt

$$|f(z_0 + r\mathrm{e}^{\mathrm{i}\varphi})| = \left|b_0 \left(1 - s r^\ell\right) + O(r^{\ell+1})\right| \leq |b_0| \left(1 - s r^\ell\right) + O(r^{\ell+1})$$

für $r \to 0$. Für genügend kleine $r > 0$ ist diese obere Schranke aber kleiner als $|b_0|$, was einen Widerspruch zu $|f(z_0)| = |b_0| = \min_{z \in K} |f(z)|$ darstellt. Dies beweist, dass $f(z_0) = 0$ gelten muss. □

Auch für reelle Polynome kann ein Fundamentalsatz formuliert werden. Wir bemerken hierfür, dass für jedes $f \in \mathbb{R}[x]$ mit einer Nullstelle $\alpha \in \mathbb{C}$ auch $\overline{\alpha}$ eine Nullstelle ist (wieso?). In der Zerlegung von $f$ in Linearfaktoren in Theorem 9.81 sieht man also zu jeder echt komplexen Nullstelle $\alpha \in \mathbb{C} \setminus \mathbb{R}$ den Faktor

$$(x - \alpha)(x - \overline{\alpha}) = x^2 - (\alpha + \overline{\alpha})x + |\alpha|^2 = x^2 - 2\,\mathrm{Re}(\alpha)x + |\alpha|^2$$

stehen. Wendet man diese Einsicht auf jede echt komplexe Nullstelle von $f$ an, so erhält man, dass $f$ als Produkt von reellen Polynomen vom Grad 1 und vom Grad 2 geschrieben werden kann. Wir haben diese Einsichten implizit bereits bei der Besprechung der Integration von rationalen Funktionen verwendet.

**Übung 9.82** (Endlich-dimensionale Körpererweiterungen von $\mathbb{C}$)**.** *Sei $\mathbb{K} \supseteq \mathbb{C}$ ein Körper, für den die Körperoperationen auf $\mathbb{C}$ mit den üblichen Körperoperationen von $\mathbb{C}$ selbst übereinstimmen (eine Körpererweiterung von $\mathbb{C}$). Dann lässt sich $\mathbb{K}$ auf natürliche Weise als Vektorraum über $\mathbb{C}$ auffassen. Zeigen Sie, dass der Körper $\mathbb{K}$ automatisch gleich $\mathbb{C}$ selbst ist, falls die Dimension von $\mathbb{K}$ über $\mathbb{C}$ endlich ist.*

**Übung 9.83** (Minimumsprinzip)**.** *Sei $f$ eine komplexwertige Funktion auf einer offenen Teilmenge von $\mathbb{C}$, die sich lokal durch Potenzreihen darstellen lässt (eine analytische Funktion). Genauer soll für jedes $x_0 \in U$ ein $R > 0$ mit $B_R(x_0) \subseteq U$ existieren, so dass $f$ auf $B_R(x_0)$ gleich einer Potenzreihe um $x_0$ mit Konvergenzradius grösser gleich $R$ ist. Zeigen Sie, dass $f$ kein Minimum annimmt.*

## 9.7 Der Raum der stetigen Funktionen

Wie wir bereits erwähnt haben, werden wir den Banachschen Fixpunktsatz unter anderem in der Diskussion von Differentialgleichungen verwenden. Wir wollen hier noch den dafür geeigneten vollständigen metrischen Raum besprechen.

**Lemma 9.84** (Vektorraum der stetigen Funktionen)**.** *Sei $X$ ein metrischer Raum. Dann ist für jedes $d \geq 1$ die Menge*

$$C(X, \mathbb{R}^d) = \Big\{ f : X \to \mathbb{R}^d \mid f \text{ ist stetig} \Big\}$$

*mit den (punktweisen) Verknüpfungen*

$$\begin{aligned} f + g : x \in X &\mapsto f(x) + g(x), \\ \lambda f : x \in X &\mapsto \lambda f(x) \end{aligned}$$

*für $f, g$ in $C(X, \mathbb{R}^d)$ und $\lambda \in \mathbb{R}$ ein Vektorraum.*

**Übung 9.85.** *Beweisen Sie Lemma 9.84.*

Ist der metrische Raum $X$ in obigem Lemma kompakt, so lässt sich der Vektorraum $C(X, \mathbb{R}^d)$ mit einer Norm ausstatten.

**Proposition 9.86** (Vollständigkeit des Raumes der stetigen Funktionen)**.** *Sei $K$ ein kompakter metrischer Raum. Dann definiert*

$$\|f\|_\infty = \sup_{x \in K} \|f(x)\|_2 = \max_{x \in K} \|f(x)\|_2$$

*eine Norm auf $C(K, \mathbb{R}^d)$, die sogenannte **Supremumsnorm**, und $C(K, \mathbb{R}^d)$ ist bezüglich dieser Norm vollständig. Des Weiteren konvergiert $f_n \in C(K)$ bezüglich der Norm $\| \cdot \|_\infty$ gegen $f \in C(K)$ genau dann, wenn $f_n$ gleichmässig gegen $f$ konvergiert, das heisst wenn es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt so dass für alle $n \in \mathbb{N}$ mit $n \geq N$ und alle $x \in X$ die Abschätzung $\|f_n(x) - f(x)\|_2 < \varepsilon$ gilt.*

In Beispiel 9.11 wurde für den Fall, wo $K \subseteq \mathbb{R}$ ein kompaktes Intervall ist und $d = 1$ ist, schon gezeigt, dass $\| \cdot \|_\infty$ eine Norm definiert. Der allgemeine Fall ist analog. Wir müssen also noch zeigen, dass $C(K, \mathbb{R}^d)$ mit der Supremumsnorm vollständig ist. Auf Grund des Zusammenhangs mit gleichmässiger Konvergenz ist dies aber im Wesentlichen eine Wiederholung von Satz 6.48.

*Beweis.* Sei $(f_n)_n$ eine Cauchy-Folge in $C(K, \mathbb{R}^d)$. Für $x \in K$ ist dann wegen

$$\|f_n(x) - f_m(x)\|_2 \leq \|f_m - f_n\|_\infty$$

für alle $m, n \in \mathbb{N}$ die Folge $(f_n(x))_n$ eine Cauchy-Folge in $\mathbb{R}^d$. Da $\mathbb{R}$ vollständig ist (Satz 5.48) existiert also der Grenzwert dieser Folge, welchen wir als $f(x) = \lim_{n \to \infty} f_n(x)$ bezeichnen.

Da $x \in K$ beliebig war (und Grenzwerte eindeutig bestimmt sind), definiert dies also eine Funktion $f : K \to \mathbb{R}^d$.

Wir behaupten nun, dass $f_n$ gleichmässig gegen $f$ strebt. Sei also $\varepsilon > 0$. Dann existiert ein $N \in \mathbb{N}$, so dass $\|f_m(x) - f_n(x)\|_2 \leq \|f_m - f_n\|_\infty < \varepsilon$ gilt für alle $n, m \geq N$ und $x \in K$. Mit $m \to \infty$ folgt daraus $\|f_n(x) - f(x)\|_2 \leq \varepsilon$ für alle $x \in X$ und $n \geq N$, was zu zeigen war. Auch zeigt dies, dass $\|f_n - f\|_\infty \leq \varepsilon$ für alle $n \geq N$ erfüllt ist.

Zuletzt zeigen wir noch, dass $f \in C(K, \mathbb{R}^d)$ ist. Sei also $\varepsilon > 0$ und $N \in \mathbb{N}$ mit der Eigenschaft $\|f_N - f\|_\infty \leq \varepsilon$. Für jedes $x_0 \in K$ gibt es auf Grund der Stetigkeit von $f_N$ bei $x_0$ ein $\delta > 0$, so dass $\|f_N(x) - f_N(x_0)\|_2 \leq \varepsilon$ für alle $x \in K$ mit $\mathrm{d}(x, x_0) < \delta$ gilt. In Kombination erhält man damit für alle $x \in K$ mit $\mathrm{d}(x, x_0) < \delta$

$$\|f(x) - f(x_0)\|_2 \leq \|f(x) - f_N(x)\|_2 + \|f_N(x) - f_N(x_0)\|_2 + \|f_N(x_0) - f(x_0)\|_2 \leq 3\varepsilon$$

Da $\varepsilon > 0$ beliebig war, beweist dies die Stetigkeit von $f$ bei $x_0 \in K$.

Schlussendlich zeigen wir noch, dass $f_n$ für $n \to \infty$ genau dann bezüglich Norm $\|\cdot\|_\infty$ gegen $f$ strebt, wenn $f_n$ für $n \to \infty$ gleichmässig gegen $f$ strebt. In der Tat ist die Bedingung $\|f_n - f\|_\infty \leq \varepsilon$ zu $\forall x \in X : |f_n(x) - f(x)| \leq \varepsilon$ äquivalent. Setzt man dies in die Definitionen der beiden Konvergenzen ein, so erhält man die behauptete Äquivalenz. $\square$

## 9.8 Konstruktion der reellen Zahlen*

Wir haben bis jetzt all unsere Diskussionen auf den Axiomen der reellen Zahlen (siehe Abschnitt 2.1) aufgebaut. Unter anderem wurden aus diesen auch die natürlichen, die ganzen und die rationalen Zahlen konstruiert. Die etwas abstrakteren Begriffe dieses Kapitels lassen sich jedoch auch dazu verwenden, eine Konstruktion der reellen Zahlen aus den natürlichen Zahlen anzugeben. Wir wollen dies hier andeuten, ohne aber auf alle Details einzugehen.

**Satz 9.87** (Existenz der reellen Zahlen)**.** *Man kann ein Modell* $(\mathbb{R}, +, \cdot, \leq)$ *der Axiome der reellen Zahlen (in Abschnitt 2.1) unter Verwendung der natürlichen Zahlen (wie in Abschnitt 1.5) definieren.*

*Beweis.* Wie in Abschnitt 1.5.1 besprochen wurde, kann man aus $\mathbb{N}$ zuerst $\mathbb{Z}$ und dann auch den Körper $\mathbb{Q}$ mittels rein algebraischer Überlegungen und geeigneten Äquivalenzrelationen konstruieren. Mit der Ordnung auf den natürlichen Zahlen (definiert für $m, n \in \mathbb{N}$ durch $m \leq n \Leftrightarrow \exists k \in \mathbb{N}_0 : m + k = n$) lässt sich auch eine Ordnung auf $\mathbb{Q}$ definieren, die alle gewünschten Eigenschaften erfüllt. Wir nehmen nun im Folgenden also den Körper $\mathbb{Q}$ und die Relation $\leq_{\mathbb{Q}}$ als gegeben an und werden die bekannten Eigenschaften (die Axiome (1)-(15) aus Abschnitt 2.1 für $(\mathbb{Q}, +, \cdot)$ und ihre Konsequenzen) ohne Verweis verwenden.

Definition der reellen Zahlen: Um $\mathbb{R}$ aus $\mathbb{Q}$ zu konstruieren, definieren wir zuerst

$$V = \Big\{(x_n)_n \in \mathbb{Q}^{\mathbb{N}} \mid (x_n)_n \text{ ist eine Cauchy-Folge in } \mathbb{Q}\Big\},$$

wobei $(x_n)_n \in \mathbb{Q}^{\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{Q}$ ist, falls es für jedes $\varepsilon \in \mathbb{Q}$ mit $\varepsilon >_{\mathbb{Q}} 0$ ein $N \in \mathbb{N}$ gibt mit $|x_m - x_n| <_{\mathbb{Q}} \varepsilon$ für alle $m, n \geq N$. Die Menge $V$ ist bezüglich der komponentenweisen Operationen ein Vektorraum über $\mathbb{Q}$ (wieso?) und die Menge

$$V_0 = \Big\{(z_n)_n \in \mathbb{Q}^{\mathbb{N}} \mid \lim_{n\to\infty} z_n = 0\Big\}$$

der Nullfolgen in $\mathbb{Q}$ ist ein Teilraum. Wir definieren die Menge der reellen Zahlen

$$\mathbb{R} = V/V_0 = \Big\{[(x_n)_n] \mid (x_n)_n \in V\Big\}$$

im Sinne der Linearen Algebra, wobei

$$[(x_n)_n] = (x_n)_n + V_0 = \Big\{(x_n + z_n)_n \mid (z_n)_n \in \mathbb{Q}^{\mathbb{N}} \text{ mit } \lim_{n\to\infty} z_n = 0\Big\}$$

die Nebenklasse von $(x_n)_n \in V$ bezeichnet.

Operationen auf den reellen Zahlen: Da $V_0$ ein Teilraum von $V$ ist, ist $\mathbb{R} = V/V_0$ ein Vektorraum über $\mathbb{Q}$. Insbesondere erfüllt $\mathbb{R}$ bezüglich der induzierten Addition die Axiome (1)-(4) der Addition.

Um die Multiplikation auf $\mathbb{R}$ zu definieren, setzen wir analog zur Addition zuerst

$$(x_n)_n \cdot (y_n)_n = (x_n y_n)_n$$

für alle $(x_n)_n, (y_n)_n \in V$. Hierbei ist $(x_n y_n)_n \in V$ in der Tat wieder eine Cauchy-Folge (überprüfen Sie dies). Des Weiteren gilt für eine Cauchy-Folge $(x_n)_n \in V$ und eine Nullfolge $(z_n)_n \in V_0$, dass $|x_n|$ beschränkt ist (unabhängig von $n$) und somit $\lim_{n\to\infty} x_n z_n = 0$ oder äquivalent $(x_n)_n \cdot (z_n)_n \in V_0$ gilt. Wir definieren nun eine Multiplikation $\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ durch

$$[(x_n)_n] \cdot [(y_n)_n] = [(x_n y_n)_n]$$

für $[(x_n)_n], [(y_n)_n] \in \mathbb{R}$. Diese ist in der Tat wohldefiniert, denn es gilt für $(z_n)_n, (z'_n)_n \in V_0$

$$(x_n + z_n)_n \cdot (y_n + z'_n)_n = (x_n y_n + \underbrace{z_n y_n + x_n z'_n + z_n z'_n}_{\text{eine Nullfolge}})_n \in (x_n y_n)_n + V_0.$$

Da $\mathbb{Q}$ ein Körper ist, kann man nun die weiteren Körperaxiome von $(\mathbb{R}, +, \cdot)$ überprüfen. Unter anderem zeigt man, dass $[(1)_n]$ das multiplikative Einselement ist. Wir wollen hier bloss die Existenz einer multiplikativen Inversen (Axiom (6)) genauer besprechen.

Existenz einer multiplikativen Inversen: Sei $[(x_n)_n] \in \mathbb{R} \setminus \{0\}$. Dann ist $(x_n)_n \in V$ eine Cauchy-Folge in $\mathbb{Q}$, die nicht gegen Null strebt. Insbesondere gibt es ein rationales $\delta >_{\mathbb{Q}} 0$, so dass $|x_n| \geq_{\mathbb{Q}} 2\delta$ für unendlich viele $n \in \mathbb{N}$. Da $(x_n)_n$ eine Cauchy-Folge ist, gibt es auch ein $N \in \mathbb{N}$, so dass $|x_m - x_n| <_{\mathbb{Q}} \delta$ für alle $m, n \geq N$. Wir wählen ein $n \geq N$ mit $|x_n| \geq_{\mathbb{Q}} 2\delta$ und erhalten

$$|x_m| \geq_{\mathbb{Q}} |x_n| - |x_m - x_n| \geq_{\mathbb{Q}} \delta$$

für alle $m \geq N$. Damit definieren wir nun eine Folge $(y_m)_m$ durch

$$y_m = \begin{cases} 1 & \text{falls } m < N \\ x_m^{-1} & \text{falls } m \geq N \end{cases}.$$

zu $m \in \mathbb{N}$. Dies ist eine Cauchy-Folge. In der Tat existiert für jedes rationale $\varepsilon > 0$ ein $N_\varepsilon \in \mathbb{N}$, so dass für $m, n \geq N_\varepsilon$ die Abschätzung $|x_m - x_n| <_{\mathbb{Q}} \delta^2 \varepsilon$ und damit

$$|y_m - y_n| = |x_m^{-1} - x_n^{-1}| = |x_m^{-1}||x_n^{-1}||x_m - x_n| \leq_{\mathbb{Q}} \delta^{-2}|x_m - x_n| <_{\mathbb{Q}} \varepsilon$$

für alle $m, n \geq \max(N, N_\varepsilon)$ gilt. Nun folgt

$$[(x_n)_n] \cdot [(y_n)_n] = [(x_n y_n)_n] = [(1)_n],$$

da $(x_n y_n - 1)_n \in V_0$ eine Nullfolge ist (sie ist schliesslich konstant Null). Da $[(1)_n]$ das multiplikative Einselement in $\mathbb{R}$ ist, haben wir die Existenz der multiplikativen Inversen gezeigt. Also ist $(\mathbb{R}, +, \cdot)$ ein Körper.

Die Ordnung auf den reellen Zahlen: Wir verwenden nun die Ordnung $\leq_{\mathbb{Q}}$ auf $\mathbb{Q}$, um aus $\mathbb{R}$ einen geordneten Körper zu machen. Für $[(x_n)_n], [(y_n)_n] \in \mathbb{R}$ soll per Definition

$[(x_n)_n] \leq_{\mathbb{R}} [(y_n)_n]$ genau dann gelten, wenn es eine Nullfolge $(z_n)_n \in V_0$ gibt mit

$$x_n \leq_{\mathbb{Q}} y_n + z_n$$

für alle $n \in \mathbb{N}$. Dies ist in der Tat wohldefiniert und erfüllt die Axiome der Anordnung (die Axiome (10)-(13)) und der Kompatibilität mit der Körperstruktur (Axiome (14)-(15)). Wir beweisen als Beispiel die Linearität der Ordnung $\leq_{\mathbb{R}}$ (Axiom (13)). Seien also $[(x_n)_n], [(y_n)_n]$ in $\mathbb{R}$.

- Falls es ein $N \in \mathbb{N}$ gibt, so dass $x_n \leq_{\mathbb{Q}} y_n$ für alle $n \geq N$ gilt, dann definieren wir eine Nullfolge $(z_n)_n \in V_0$ durch

  $$z_n = \begin{cases} x_n - y_n & \text{falls } n < N \\ 0 & \text{falls } n \geq N \end{cases}$$

  für $n \in \mathbb{N}$. Wir erhalten also $x_n \leq_{\mathbb{Q}} y_n + z_n$ für alle $n \in \mathbb{N}$ und damit $[(x_n)_n] \leq_{\mathbb{R}} [(y_n)_n]$.

- Falls $x_n \geq_{\mathbb{Q}} y_n$ für fast alle $n \in \mathbb{N}$ gilt, so folgt analog $[(x_n)_n] \geq_{\mathbb{R}} [(y_n)_n]$.

- Es verbleibt den Fall zu behandeln, wo $x_n \leq_{\mathbb{Q}} y_n$ für unendlich viele $n \in \mathbb{N}$ und ebenso $x_n \geq_{\mathbb{Q}} y_n$ für unendlich viele $n \in \mathbb{N}$ gilt. In diesem Fall behaupten wir, dass $z_n = x_n - y_n$ für $n \in \mathbb{N}$ eine Nullfolge $(z_n)_n \in V_0$ definiert und damit $[(x_n)_n] = [(y_n)_n]$ und insbesondere $[(x_n)_n] \leq_{\mathbb{R}} [(y_n)_n]$ gilt. Sei also $\varepsilon >_{\mathbb{Q}} 0$ rational. Dann existiert ein $N \in \mathbb{N}$, so dass $|x_m - x_n| <_{\mathbb{Q}} \varepsilon$ und $|y_m - y_n| <_{\mathbb{Q}} \varepsilon$ für alle $n \geq N$ gilt. Sei $m_1 \geq N$, so dass $x_{m_1} \leq_{\mathbb{Q}} y_{m_1}$ und $m_2 \geq N$, so dass $x_{m_2} \geq_{\mathbb{Q}} y_{m_2}$. Dann folgt für $n \geq N$ beliebig

  $$-2\varepsilon \leq_{\mathbb{Q}} x_{m_2} - y_{m_2} - 2\varepsilon \leq_{\mathbb{Q}} x_n - y_n \leq_{\mathbb{Q}} x_{m_1} - y_{m_1} + 2\varepsilon \leq_{\mathbb{Q}} 2\varepsilon.$$

  Da $\varepsilon >_{\mathbb{Q}} 0$ beliebig war, folgt die Behauptung und damit, dass $\leq_{\mathbb{R}}$ eine lineare Ordnung (Axiom (13)) ist.

Zusammenfassend können wir also sagen, dass wir die Axiome (1)-(15) mehr oder weniger direkt von $\mathbb{Q}$ auf die hier aus $\mathbb{Q}$ definierten reellen Zahlen $\mathbb{R}$ übertragen können.

Wir betrachten $\mathbb{Q}$ als eine Teilmenge von $\mathbb{R}$, indem wir $q \in \mathbb{Q}$ mit $[(q)_n] \in \mathbb{R}$ identifizieren (die Abbildung $q \in \mathbb{Q} \mapsto [(q)_n] \in \mathbb{R}$ ist injektiv). Da $q \leq_{\mathbb{Q}} q'$ für $q, q' \in \mathbb{Q}$ genau dann gilt, wenn $[(q)_n] \leq_{\mathbb{R}} [(q')_n]$ gilt, ist diese Identifikation mit der von uns definierten Ordnung auf $\mathbb{R}$ kompatibel und wir können zur Vereinfachung der Notation wieder schlicht $\leq$ für $\leq_{\mathbb{Q}}$ und $\leq_{\mathbb{R}}$ schreiben. Weiter schreiben wir im Folgenden $\overline{x} \in \mathbb{R}$ anstelle von $[(x_n)_n] \in \mathbb{R}$.

Wir wollen zuletzt das entscheidende Axiom der Vollständigkeit (Axiom (16)) sorgfältiger beweisen, da dieses ja nicht von $\mathbb{Q}$ nicht erfüllt wird.

Das Archimedische Prinzip: Wir behaupten zuerst, dass zu $\overline{x} \in \mathbb{R}$ mit $\overline{x} > 0$ ein $n \in \mathbb{N}$ mit $0 < \frac{1}{n} < x$ existiert. In der Verifikation der Existenz einer multiplikativen Inversen wurde bereits gezeigt, dass es eine rationale Zahl $\delta > 0$ mit $|x_n| > \delta$ für fast alle $n \in \mathbb{N}$ gibt. Jedes $n \in \mathbb{N}$ mit $\frac{1}{n} < \delta$ erfüllt nun $0 < \frac{1}{n} < \overline{x}$. Um ein solches $n$ zu finden, wählt man für $\delta = \frac{p}{q}$ schlicht $n = q$.

Vollständigkeit von $\mathbb{R}$: Seien $X, Y \subseteq \mathbb{R}$ zwei nicht-leere Teilmengen, so dass $\overline{x} \leq \overline{y}$ für alle $\overline{x} \in X$ und $\overline{y} \in Y$. Wir wollen nun eine reelle Zahl zwischen $X$ und $Y$ finden indem wir eine Cauchy-Folge in $\mathbb{Q}$ konstruieren.

Sei $\overline{x} = [(x_n)_n] \in X$ beliebig. Da $(x_n)_n$ eine Cauchy-Folge ist, gibt es ein $N \in \mathbb{N}$, so dass $|x_m - x_n| < 1$ für alle $m, n \geq N$. Wir setzen $a_1 = x_N - 1 \in \mathbb{Q}$. Dann gilt $a_1 \leq x_m$ für alle $m \geq N$ und damit wie im Beweis der Linearität der Ordnung $a_1 \leq \overline{x}$. Insbesondere gilt auf Grund unserer Annahme an $X$ und $Y$ auch $a_1 \leq \overline{y}$ für alle $\overline{y} \in Y$, was wir kurz als

$$a_1 \leq Y$$

schreiben werden. Genauso kann man unter Verwendung eines $\overline{y} \in Y$ ein $b_1 \in \mathbb{Q}$ finden mit $\overline{x} \leq b_1$ für alle $\overline{x} \in X$, was wir kurz als

$$X \leq b_1$$

schreiben werden.

Wir nehmen nun an, dass wir bereits $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{Q}$ konstruiert haben mit

$$a_1 \leq a_2 \leq \ldots \leq a_m \leq b_m \leq b_{m-1} \leq \ldots \leq b_1, \quad b_k - a_k = \frac{b_1 - a_1}{2^{k-1}}$$

für alle $k = 1, \ldots, m$ und der Eigenschaft $a_m \leq Y$, $X \leq b_m$. Wir möchten nun $a_{m+1}, b_{m+1} \in \mathbb{Q}$ definieren. Dabei unterscheiden wir drei Möglichkeiten, wie $c_m = \frac{a_m + b_m}{2}$ gegenüber $X$ und $Y$ stehen kann.

- Es gibt ein $\overline{x} \in X$ mit $c_m < \overline{x}$. In diesem Fall definieren wir $a_{m+1} = c_m$, $b_{m+1} = b_m$ sowie $c_{m+1} = \frac{a_{m+1} + b_{m+1}}{2}$, und erhalten $a_{m+1} < \overline{x} \leq Y$ und $X \leq b_{m+1}$.
- Es gibt ein $\overline{y} \in Y$ mit $\overline{y} < c_m$. In diesem Fall definieren wir $a_{m+1} = a_m$, $b_{m+1} = c_m$ sowie $c_{m+1} = \frac{a_{m+1} + b_{m+1}}{2}$, und erhalten $a_{m+1} \leq Y$ und $X \leq \overline{y} < b_{m+1}$.
- Es gilt $\overline{x} \leq c_m \leq \overline{y}$ für alle $\overline{x} \in X$ und $\overline{y} \in Y$. In diesem Fall erfüllt $c_m \in \mathbb{Q}$ bereits die Existenzaussage des Vollständigkeitsaxioms.

Falls obige Rekursion nicht bereits nach endlich vielen Schritten zum Ziel führt (das heisst, immer entweder der erste oder der zweite Fall eintritt), so konstruiert diese stattdessen zwei Folgen $(a_n)_n, (b_n)_n$ in $\mathbb{Q}$ mit $a_m \leq b_n$, $a_m \leq \overline{y}$, $\overline{x} \leq b_n$ sowie $b_n - a_n = \frac{b_1 - a_1}{2^{n-1}}$ für alle $m, n \in \mathbb{N}$ und $\overline{x} \in X$, $\overline{y} \in Y$. Es folgt insbesondere, dass $z_n = b_n - a_n \in \mathbb{Q}$ für $n \in \mathbb{N}$ eine Nullfolge $(z_n)_n$ definiert (wieso?). Für $\varepsilon \in \mathbb{Q}$ mit $\varepsilon > 0$ gibt es also ein $N$, so dass $0 < b_n - a_n < \varepsilon$ für $n \geq N$. Da $a_m \leq b_n$ für alle $m, n \geq N$ gilt, folgt daraus $a_m - a_n < \varepsilon$ und aus Symmetrie $|a_m - a_n| < \varepsilon$ für $m, n \geq N$. Dies zeigt, dass $(a_n)_n$ eine Cauchy-Folge in $\mathbb{Q}$ ist. Sei

$$\overline{c} = [(a_n)_n]) = [(a_n + z_n)_n] = [(b_n)_n] \in \mathbb{R}$$

Wir behaupten nun, dass $\overline{c}$ zwischen $X$ und $Y$ liegt.

Für dies bemerken wir, dass $a_m \leq \overline{c}$ (da $a_m \leq b_n$ für alle $n \in \mathbb{N}$) und $\overline{c} \leq b_m$ (da ebenso $a_n \leq b_m$ für alle $n \in \mathbb{N}$) für alle $m \in \mathbb{N}$. Weiters ist $(b_m - a_m)_m$ eine Nullfolge in $\mathbb{Q}$

und auf Grund des Archimedischen Prinzips auch in $\mathbb{R}$. Daraus folgt

$$\overline{c} = \lim_{m\to\infty} a_m = \lim_{m\to\infty} b_m$$

und mit $a_m \leq Y$ und $X \leq b_m$ für alle $m \in \mathbb{N}$ auch $\overline{c} \leq Y$ und $X \leq \overline{c}$. Dies beweist die Existenz im Vollständigkeitsaxiom und beendet den Beweis. $\square$

## 9.9 Weitere Lernmaterialien

### 9.9.1 Verwendung des Kapitels

Wie Sie sicher schon bemerkt haben, ist dieses Kapitel vielmehr ein theoretisches Hilfsmittel als ein praktisches Werkzeug für Berechnungen. Die zentralen Begriffe dieses Kapitels sind Konvergenz, Stetigkeit, Kompaktheit und Vollständigkeit, wobei wir diese Begriffe eben für metrische Räume besprochen haben und diese Begriffe damit universeller einsetzbar geworden sind. Insbesondere werden diese Begriffe für die mehrdimensionale Analysis von grundlegender Bedeutung sein. Wir bemerken noch, dass viele der Sätze des ersten Semesters bei Verallgemeinerung ins Mehrdimensionale den Begriff der zusammenhängenden offenen Teilmenge benötigen werden. Des Weiteren haben wir auch die Vollständigkeit des Funktionenraums $C(K, \mathbb{R}^d)$ bewiesen und werden dies später gemeinsam mit dem auch sehr allgemein gehaltenen Banachschen Fixpunktsatz verwenden, um implizite Gleichungen und Differentialgleichungen abstrakt zu lösen.

### 9.9.2 Weitere Normen und Metriken

Wir möchten Ihnen in diesem Abschnitt die Gelegenheit geben, andere Normen oder Metriken kennenzulernen, die in gewissen Situationen von Interesse sein können.

**Übung** ($p$-Normen)**.** *In dieser Übung möchten wir für jedes $p \geq 1$ eine Norm, die sogenannte $p$-Norm definieren, die für $p = 1$ gerade die Einsnorm und für $p = 2$ gerade die Euklidsche Norm ist. Wir setzen für $x = (x_1, \ldots, x_d)^t \in \mathbb{R}^d$*

$$\|x\|_p = \left( \sum_{k=1}^{d} |x_k|^p \right)^{\frac{1}{p}}.$$

*Verfizieren Sie zuerst, dass $\|\cdot\|_p$ definit und homogen ist. Zeigen Sie dann die Dreiecksungleichung in folgenden Schritten. Sei $q = \frac{p}{p-1}$.*

*(i) (Young-Ungleichung) Zeigen Sie für $a, b \geq 0$*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

*(ii) (Hölder-Ungleichung) Zeigen Sie für $x = (x_1, \ldots, x_d)^t$ und $y = (y_1, \ldots, y_d)^t$*

$$\sum_{k=1}^{d} |x_k y_k| \leq \|x\|_p \|y\|_q.$$

*(iii) (Dreiecksungleichung/Minkowski-Ungleichung) Zeigen Sie für $x, y \in \mathbb{R}^d$*

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p.$$

**Übung** (Hexagon-Metrik). *Finden Sie eine Norm auf $\mathbb{R}^2$, so dass der Einheitsball $B_1(0)$ bezüglich der induzierten Metrik das reguläre Hexagon mit Eckpunkt $(1,0)$ ist. Gibt es eine Norm auf $\mathbb{R}^2$, so dass der Einheitsball durch ein reguläres Pentagon gegeben ist?*

**Übung** (Ultrametriken). *Eine Ultrametrik auf einer Menge $X$ ist eine Abbildung $\mathrm{d} : X \times X \to \mathbb{R}_{\geq 0}$, die die gleichen Eigenschaften wie eine Metrik hat, abgesehen davon, dass sie anstelle der Dreiecksungleichung die Ungleichung*

$$\mathrm{d}(x_1, x_3) \leq \max\{\mathrm{d}(x_1, x_2), \mathrm{d}(x_2, x_3)\}$$

*für alle $x_1, x_2, x_3 \in X$ erfüllt.*

*(i) Zeigen Sie, dass jede Ultrametrik eine Metrik ist.*

*(ii) (p-adische Metrik auf $\mathbb{Z}$) Sei $p \in \mathbb{Z}$ eine Primzahl. Wir definieren*

$$\nu_p(x) = \max\left\{k \mid p^k \text{ teilt } x\right\}$$

*für $x \in \mathbb{Z} \setminus \{0\}$ und $\mathrm{d}_p : \mathbb{Z} \times \mathbb{Z} \to \mathbb{R}_{\geq 0}$ durch*

$$\mathrm{d}_p(x, y) = \begin{cases} p^{-\nu_p(x-y)} & \text{falls } x \neq y \\ 0 & \text{falls } x = y \end{cases}.$$

*für $x, y \in \mathbb{Z}$. Zeigen Sie, dass $\mathrm{d}_p$ eine Ultrametrik auf $\mathbb{Z}$ definiert und beschreiben Sie die Bälle in dieser Metrik.*

*(iii) Sei $X$ eine Menge und $\mathrm{d}$ eine Ultrametrik. Zeigen Sie, dass für alle $r > 0$, $x_0 \in X$ und $x \in B_r(x_0)$ der Ball von Radius $r$ um $x$ gleich $B_r(x_0)$ ist. In anderen Worten ist jeder Punkt in einem Ball Zentrum dieses Balles.*

**Übung** (Rangmetrik). *Sei $\mathbb{K}$ ein Körper und $X$ die Menge der $m \times n$-Matrizen über $\mathbb{K}$. Wir definieren $\mathrm{d}(A, B) = \mathrm{rang}(A - B)$ für $A, B \in X$. Zeigen Sie, dass $\mathrm{d}$ eine Ultrametrik auf $X$ ist.*

**Übung** (Verschiedene Produktmetriken und die Produkttopologie). *In dieser Übung möchten wir kartesische Produkte von metrischen Räumen mit Metriken ausstatten, so dass „horizontale und vertikale“ Abstände gewissermassen natürlich sind. Seien $(X, \mathrm{d}_X)$ und $(Y, \mathrm{d}_Y)$ metrische Räume. Für $(x_1, y_1), (x_2, y_2) \in X \times Y$ definieren wir*

$$\begin{aligned} \mathrm{d}_\infty((x_1, y_1), (x_2, y_2)) &= \max\{\mathrm{d}_X(x_1, x_2), \mathrm{d}_Y(y_1, y_2)\} \\ \mathrm{d}_1((x_1, y_1), (x_2, y_2)) &= \mathrm{d}_X(x_1, x_2) + \mathrm{d}_Y(y_1, y_2) \\ \mathrm{d}_2((x_1, y_1), (x_2, y_2)) &= \sqrt{\mathrm{d}_X(x_1, x_2)^2 + \mathrm{d}_Y(y_1, y_2)^2}. \end{aligned}$$

*Verfizieren Sie, dass $\mathrm{d}_\infty, \mathrm{d}_1$ und $\mathrm{d}_2$ Metriken auf $X \times Y$ definieren und dass diese Metriken die gleiche Topologie induzieren. Diese Topologie nennt sich die Produkttopologie auf $X \times Y$. Zeigen Sie, dass eine Teilmenge $U \subseteq X \times Y$ genau dann offen ist, wenn für jedes $x \in U$ offene Teilmenge $U_1 \subseteq X$ und $U_2 \subseteq Y$ mit $x \in U_1 \times U_2 \subseteq U$ existieren.*

**Übung** (Metriken auf Folgenräumen). *Für einen beschränkten metrischen Raum* $(X, \mathrm{d})$ *und Punkte* $(x_n)_n, (y_n)_n \in X^{\mathbb{N}}$ *setzen wir*

$$\mathrm{d}'((x_n)_n), (y_n)_n) = \sum_{n=1}^{\infty} 2^{-n} \mathrm{d}(x_n, y_n).$$

*Zeigen Sie, dass* $\mathrm{d}'$ *eine (wohldefinierte) Metrik auf* $X^{\mathbb{N}}$ *darstellt. Fast in Analogie zu obiger Übung kann man auch hier die offenen Mengen charakterisieren. Zeigen Sie, dass eine Teilmenge* $U \subseteq X^{\mathbb{N}}$ *genau dann offen ist, wenn für alle* $x \in U$ *offene Mengen* $U_1, U_2, \ldots$ *in* $X$ *existieren mit* $U_n = X$ *für fast alle* $n \in \mathbb{N}$ *und*

$$x \in \prod_{n \in \mathbb{N}} U_n = \left\{ (y_n)_n \in X^{\mathbb{N}} \mid y_n \in U_n \text{ für alle } n \in \mathbb{N} \right\} \subseteq U.$$

**Übung.** *Sei* $f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ *eine monoton wachsende, konkave Funktion mit* $f(x) = 0 \iff x = 0$ *für alle* $x \in \mathbb{R}_{\geq 0}$ *und sei* $(X, \mathrm{d})$ *ein metrischer Raum. Zeigen Sie wie folgt, dass* $f \circ \mathrm{d}$ *eine Metrik auf* $X$ *definiert.*

*(i) Beweisen Sie die Ungleichung* $f(tx) \geq tf(x)$ *für alle* $x \in \mathbb{R}_{\geq 0}$ *und* $t \in [0, 1]$.

*(ii) Zeigen Sie, dass für alle* $x, y \in \mathbb{R}_{\geq 0}$ *die Ungleichung* $f(x) + f(y) \geq f(x + y)$ *gilt (*$f$ *ist subadditiv).*

*(iii) Schliessen Sie auf die zu beweisende Aussage.*

*Geben Sie des Weiteren ein Beispiel einer solchen Abbildung* $f$ *und der Metrik* $f \circ \mathrm{d}$ *für* $X = \mathbb{R}$ *und die Standardmetrik* $\mathrm{d}$ *auf* $\mathbb{R}$ *an.*

**Übung** (Hausdorff-Metrik). *Sei* $(X, \mathrm{d})$ *ein metrischer Raum. In dieser Übung möchten wir einen Abstand von Teilmengen von* $X$ *zueinander einführen und verwenden die in Übung 9.41 definierten Distanzfunktionen. Für nicht-leere, abgeschlossene Teilmengen* $A, B \subseteq X$ *sei*

$$\mathrm{d}(A, B) = \max \left\{ \sup_{a \in A} \mathrm{d}(a, B), \sup_{b \in B} \mathrm{d}(b, A) \right\}$$

*Für die Wohldefiniertheit wollen wir hierbei annehmen, dass* $\mathrm{d}(x, y) \leq M$ *für ein* $M > 0$ *und alle* $x, y \in X$ *(siehe Übung 9.6). Zeigen Sie, dass obiges eine Metrik auf der Menge der nicht-leeren abgeschlossenen Teilmengen von* $X$ *definiert.*

### 9.9.3 Weitere Übungen

**Übung** (Teilmengen gegeben durch Gleichungen).

*(i) Sei* $X$ *ein metrischer Raum und* $\mathcal{I}$ *eine Menge. Gegeben sei für jedes* $\alpha \in \mathcal{I}$ *ein metrischer Raum* $Y_\alpha$*, eine stetige Abbildung* $f_\alpha : X \to Y_\alpha$ *und ein Punkt* $y_\alpha \in Y_\alpha$*. Zeigen Sie, dass die durch Gleichungen gegebene Teilmenge*

$$\{x \in X \mid \forall \alpha \in \mathcal{I} : f_\alpha(x) = y_\alpha\} \subseteq X$$

*abgeschlossen ist. Zeigen Sie auch, dass die Teilmenge* $\{x \in X \mid \forall \alpha \in \mathcal{I} : f_\alpha(x) \neq y_\alpha\}$ *offen ist, falls* $\mathcal{I}$ *eine endliche Menge ist.*

*Wenden Sie dies nun auf einige anschauliche Beispiele an.*

*(ii) Zeigen Sie, dass die Sphäre* $\mathbb{S}^{d-1} = \left\{x = (x_1, \ldots, x_d) \in \mathbb{R}^d \mid x_1^2 + \ldots + x_d^2 = 1\right\}$ *für* $d \geq 1$ *abgeschlossen ist.*

*(iii) Zeigen Sie, dass alle (linearen) Unterräume von* $\mathbb{C}^d$ *für* $d \geq 1$ *abgeschlossen sind.*

*(iv) Sei* $I \subseteq \mathbb{R}$ *ein Intervall und* $\gamma : I \to \mathbb{R}^d$ *eine stetige Funktion für* $d \geq 1$ *(ein stetiger Weg in* $\mathbb{R}^d$*). Zeigen Sie, dass der Graph von* $\gamma$

$$\left\{x = (t, v) \in \mathbb{R}^{d+1} \mid \gamma(t) = v\right\} \subseteq \mathbb{R}^{d+1}$$

*(auch die Spur des Weges* $\gamma$ *genannt) abgeschlossen ist.*

**Übung** (Urysohn's Lemma)**.** *Sei* $X$ *ein metrischer Raum. Wir möchten zeigen, dass sich disjunkte, abgeschlossene Teilmengen von* $X$ *mit einer stetigen Funktion „trennen" lassen. Seien also* $A, B \subseteq X$ *abgeschlossene, nicht-leere, disjunkte Teilmengen. Zeigen Sie, dass eine stetige Funktion* $f : X \to [0, 1]$ *mit* $f(a) = 0$ *für alle* $a \in A$ *und* $f(b) = 1$ *für alle* $b \in B$ *existiert. Betrachten Sie dazu die Funktion*

$$x \in X \mapsto \frac{\mathrm{d}(x, A)}{\mathrm{d}(x, A) + \mathrm{d}(x, B)}$$

*und verifizieren Sie auch, dass diese wohldefiniert ist.*

**Übung** (Fortsetzungssatz von Tietze)**.** *Sei* $X$ *ein metrischer Raum, sei* $A \subseteq X$ *eine nicht-leere abgeschlossene Teilmenge und sei* $f : A \to \mathbb{R}_{\geq 0}$ *eine stetige, beschränkte Funktion. Für alle* $x \in X$ *setzen wir*

$$F(x) = \begin{cases} f(x) & \text{falls } x \in A \\ \inf\left\{f(a) + \frac{\mathrm{d}(x,a)}{\mathrm{d}(x,A)} - 1 \mid a \in A\right\} & \text{falls } x \notin A \end{cases}.$$

*Zeigen Sie, dass* $F : X \to \mathbb{R}_{\geq 0}$ *eine stetige Funktion mit* $F|_A = f$ *ist.*

**Übung** (Stetigkeit der Matrixmultiplikation)**.** *Seien* $d, d' \geq 1$.

*(a) Zeigen Sie, dass* $\mathrm{Mat}_{d',d}(\mathbb{C}) \times \mathbb{C}^d,\ (A, x) \mapsto Ax \in \mathbb{C}^{d'}$ *stetig ist.*

*(b) Zeigen Sie, dass die Gruppe der invertierbaren* $d \times d$*-Matrizen* $\mathrm{GL}_d(\mathbb{C})$ *eine offene Teilmenge von* $\mathrm{Mat}_{d,d}(\mathbb{C})$ *bildet.*

*(c) Beweisen Sie, dass die Abbildungen*

$$(A, B) \in \mathrm{GL}_d(\mathbb{C}) \mapsto AB \in \mathrm{GL}_d(\mathbb{C}), \quad A \in \mathrm{GL}_d(\mathbb{C}) \mapsto A^{-1} \in \mathrm{GL}_d(\mathbb{C})$$

*stetig sind.*

**Übung** (Topologie von Matrixgruppen). *Sei $d \geq 1$.*

*(i) Zeigen Sie, dass $\mathrm{GL}_d(\mathbb{C})$ wegzusammenhängend und somit auch zusammenhängend ist. Verifizieren Sie auch, dass $\mathrm{GL}_d(\mathbb{R})$ nicht zusammenhängend ist.*

*(ii) Zeigen Sie, dass $\mathrm{SL}_d(\mathbb{R}) = \{A \in \mathrm{GL}_d(\mathbb{R}) \mid \det(A) = 1\}$ wegzusammenhängend ist.*

*(iii) Zeigen Sie, dass $\mathrm{SO}(d) = \left\{A \in \mathrm{SL}_d(\mathbb{R}) \mid A^T A = \mathrm{id}\right\}$ kompakt und zusammenhängend ist.*

**Übung** (Topologische Eigenschaften von Produkträumen). *Seien $X$ und $Y$ nicht-leere metrische Räume. In einer Übung in Abschnitt 9.9.2 wurden Metriken $\mathrm{d}_\infty, \mathrm{d}_1, \mathrm{d}_2$ auf dem Produkt $X \times Y$ definiert, welche allesamt die gleiche Topologie (die Produkttopologie) induzieren. Im Folgenden möchten wir topologische Eigenschaften des Produktraumes $X \times Y$ charakterisieren.*

*(i) Zeigen Sie, dass eine Folge in $X \times Y$ genau dann konvergiert, wenn sie komponentenweise konvergiert.*

*(ii) Beweisen Sie, dass $X \times Y$ genau dann kompakt ist, wenn $X$ und $Y$ kompakt sind.*

*(iii) Zeigen Sie, dass $X \times Y$ genau dann zusammenhängend ist, wenn $X$ und $Y$ zusammenhängend sind.*

*(iv) Zeigen Sie, dass $X \times Y$ genau dann wegzusammenhängend ist, wenn $X$ und $Y$ wegzusammenhängend sind.*

**Übung** (Der abzählbare Diagonalfolgentrick). *Sei $[0,1]^{\mathbb{N}}$ die Menge der $[0,1]$-wertigen Folgen und sei*

$$\mathrm{d}(x,y) = \sum_{n=1}^{\infty} |x_n - y_n| 2^{-n}$$

*für $x = (x_n)_n, y = (y_n)_n \in [0,1]^{\mathbb{N}}$. In einer Übung in Abschnitt 9.9.2 wurde gezeigt, dass $\mathrm{d}$ in der Tat eine Metrik auf $[0,1]^{\mathbb{N}}$ definiert. Auch wurde die Topologie auf $[0,1]^{\mathbb{N}}$ beschrieben. In dieser Übung möchten wir beweisen, dass $[0,1]^{\mathbb{N}}$ kompakt ist. Sei also $(x^{(n)})_n$ eine Folge in $[0,1]^{\mathbb{N}}$.*

*(i) Zeigen Sie, dass eine Folge in $[0,1]^{\mathbb{N}}$ genau dann konvergiert, wenn sie komponentenweise konvergiert.*

*Wir konstruieren nun eine konvergente Teilfolge von $(x^{(n)})_n$ wie folgt.*

*(ii) Finden Sie eine unendliche Teilmenge $\mathcal{J}_1 \subseteq \mathbb{N}$, so dass $(\pi_1(x^{(n)}))_{n \in \mathcal{J}_1}$ konvergiert. Genau genommen wählt man hier für die korrekte Formulierung eine streng monotone Bijektion $\varphi_1 : \mathbb{N} \to \mathcal{J}_1$, womit die Teilfolge $(\pi_1(x^{(\varphi_1(n))}))_n$ konvergieren soll.*

*(iii) Wenden Sie obige Idee rekursiv an, um unendliche Teilmengen $\mathcal{J}_1 \supseteq \mathcal{J}_2 \supseteq \ldots$ zu konstruieren, so dass $(\pi_k(x^{(n)}))_{n \in \mathcal{J}_k}$ für jedes $k \in \mathbb{N}$ konvergiert.*

*Nun würde man gerne den Schnitt $\bigcap_{k \in \mathbb{N}} \mathcal{J}_k$ zum Auffinden einer konvergenten Teilfolge betrachten. Da hier aber über unendliche viele Teilmengen geschnitten wird, könnte dieser leer sein. Stattdessen geht man wie folgt vor.*

(v) *Sei* $(n_k)_k$ *eine aufsteigende Folge in* $\mathbb{N}$*, so dass* $n_k \in \mathcal{J}_k$ *für jedes* $k \in \mathbb{N}$*. Zeigen Sie, dass* $(x^{(n_k)})_k$ *konvergiert.*

**Übung** (Umgebungsfilter)**.** *Sei* $(X, \mathrm{d})$ *ein metrischer Raum. Der Umgebungsfilter eines Punktes* $x \in X$ *ist die Menge*

$$\mathcal{U}_x = \{U \subseteq X \mid \exists \varepsilon > 0 : B_\varepsilon(x) \subseteq U\}.$$

*Zeigen Sie, dass* $\mathcal{U}_x$ *für jeden Punkt* $x \in X$ *einen Filter auf* $X$ *definiert.*

**Übung** (Regeln für den Abschluss)**.** *Sei* $(X, \mathrm{d})$ *ein metrischer Raum und* $Y \subseteq X$ *eine Teilmenge. Verifizieren Sie die Regeln*

$$\overline{Y_1 \cup Y_2} = \overline{Y_1} \cup \overline{Y_2}, \quad \overline{Y_1 \cap Y_2} \subseteq \overline{Y_1} \cap \overline{Y_2}$$

*für Teilmengen* $Y, Y_1, Y_2 \subseteq X$ *und erklären Sie anhand eines Beispiels, dass die Gleichheit* $\overline{Y_1 \cap Y_2} = \overline{Y_1} \cap \overline{Y_2}$ *nicht immer gelten muss.*

**Übung** (Zusammenhangskomponenten eines metrischen Raumes)**.** *Sei* $(X, \mathrm{d})$ *ein metrischer Raum.*

(i) *Zeigen Sie, dass jeder Punkt* $x \in X$ *in einem maximalen zusammenhängenden Teilraum* $Z(x) \subseteq X$ *enthalten ist.*

(ii) *Zeigen Sie, dass für* $x, y \in X$ *entweder* $Z(x) = Z(y)$ *oder* $Z(x) \cap Z(y) = \emptyset$ *gilt.*

(iii) *Schliessen Sie, dass die Mengen* $Z(x)$ *– die sogenannten Zusammenhangskomponenten von* $X$*– für* $x \in X$ *eine Partition von* $X$ *bilden.*

**Übung** (Abschluss einer zusammenhängenden Teilmenge)**.** *Sei* $(X, \mathrm{d})$ *ein metrischer Raum und sei* $Y \subseteq X$ *ein zusammenhängender Teilraum. Zeigen Sie, dass dann auch jeder Teilraum* $B \subseteq X$ *mit* $Y \subseteq B \subseteq \overline{Y}$ *zusammenhängend ist. Verwenden Sie dies, um zu zeigen, dass die Zusammenhangskomponenten von* $X$ *aus der vorherigen Übung abgeschlossen sind.*

**Übung** (Direkter Beweis gewisser Implikation in Satz 9.66)**.** *Sei* $X$ *ein metrischer Raum. Beweisen Sie folgende Aussagen ohne auf Satz 9.66 zurückzugreifen.*

(i) *Ist* $X$ *folgenkompakt, so besitzt* $X$ *eine Lebesgue-Zahl und* $X$ *ist total beschränkt.*

(ii) *Erfüllt* $X$ *das Schachtelungsprinzip, so ist* $X$ *überdeckungskompakt.*

(iii) *Ist* $X$ *überdeckungskompakt, so nimmt jede stetige, reellwertige Funktion ein Maximum und ein Minimum an.*

**Übung.** *In dieser Übung möchten wir zeigen, dass für gewisse stetige, bijektive Funktion auch deren Inverse stetig ist.*

(i) *Wir beginnen mit einem Gegenbeispiel zur allgemeinen Aussage. Finden Sie metrische Räume* $X, Y$ *und eine bijektive, stetige Abbildung* $f : X \to Y$*, so dass* $f^{-1}$ *nicht stetig ist.*

*(ii) Seien $X, Y$ metrische Räume und $f : X \to Y$ bijektiv und stetig. Zeigen Sie, dass $f^{-1}$ stetig ist, falls $X$ kompakt ist.*

**Übung** (Diskrete Teilmengen von kompakten Räumen). *Sei $X$ ein metrischer Raum. Wir nennen eine Teilmenge $A \subseteq X$ diskret, falls es um jeden Punkt $x_0$ in $A$ einen Ball $B_r(x_0) \subseteq X$ gibt mit $B_r(x_0) \cap A = \{x_0\}$. Beispielweise ist $\mathbb{Z} \subseteq \mathbb{R}$ diskret. Zeigen Sie, dass eine diskrete, abgeschlossene Teilmenge eines kompakten metrischen Raumes endlich ist.*

**Übung** (Separable Räume). *Wir nennen einen metrischen Raum $X$ separabel, falls er eine abzählbare dichte Teilmenge enthält. Beispielsweise ist $\mathbb{R}$ separabel, da $\mathbb{Q} \subseteq \mathbb{R}$ abzählbar und dicht ist. Zeigen Sie, dass jeder kompakte metrische Raum separabel ist.*

**Übung** (Challenge – Satz von Arzelà-Ascoli). *Der Satz von Heine-Borel beschreibt die kompakten Teilmengen von $\mathbb{R}^d$ für $d \geq 1$. In dieser Übung möchten wir die kompakten Teilmengen in einer weiteren Situation beschreiben. Sei $K$ ein kompakter metrischer Raum und sei $\mathcal{F} \subseteq C(K)$ eine Familie stetiger Funktionen, wobei wir $C(K)$ mit der Norm $\|\cdot\|_\infty$ ausstatten.*

*Zeigen Sie, dass $\mathcal{F}$ genau dann kompakt ist, wenn die folgenden drei Bedingungen erfüllt sind:*

- *$\mathcal{F}$ ist abgeschlossen.*
- *$\mathcal{F}$ ist beschränkt.*
- *$\mathcal{F}$ ist gleichstetig. Das heisst, es gibt für alle $\varepsilon > 0$ ein $\delta > 0$, so dass für alle $f \in \mathcal{F}$ und für alle $x, y \in K$ gilt $\mathrm{d}(x, y) < \delta \implies |f(x) - f(y)| < \varepsilon$.*

**Übung** (Ein unendlich-dimensionaler vollständiger normierter Vektorraum).

*(a) Wir betrachten den Vektorraum*

$$\ell^1 = \left\{ (x_k)_k \in \mathbb{C}^{\mathbb{N}} \mid \sum_{k=1}^{\infty} x_k \text{ konvergiert absolut} \right\}$$

*mit der Norm $\|(x_k)_k\|_1 = \sum_{k=1}^{\infty} |x_k|$. Zeigen Sie, dass $\ell^1$ vollständig ist. (Dies ist ein erstes Beispiel eines sogenannten Banachraums.)*

*(b) Zeigen Sie, dass der Unterraum $U \subseteq \ell^1$ der Folgen, die schliesslich Null sind, dicht in $\ell^1$ liegt und nicht vollständig ist.*

**Übung** (Ein unendlich-dimensionaler, nicht-vollständiger, normierter Vektorraum). *Sei $[a, b]$ ein kompaktes Intervall mit Endpunkten $a < b$ und sei $C([a, b])$ ausgestattet mit der Norm $\|\cdot\|_1$ aus Übung 9.12. Zeigen Sie, dass $C([a, b])$ mit dieser Norm nicht vollständig ist.*

**Übung.** *Seien $(V, \|\cdot\|_V), (W, \|\cdot\|_W)$ normierte Vektorräume und sei $F : V \to W$ eine lineare Abbildung. Genau wie in Definition 9.75 setzen wir nun*

$$\|F\|_{\mathrm{op}} = \sup_{v \in V : \|v\|_V = 1} \|F(v)\|_W.$$

*Zeigen Sie, dass $F$ genau dann stetig ist, wenn $\|F\|_{\mathrm{op}} < \infty$.*

**Übung.** *Sei $\alpha > 1$. Verwenden Sie den Banachschen Fixpunktsatz und dessen Beweis, um zu zeigen, dass die durch*

$$x_{n+1} = \frac{1}{x_n + \alpha} \text{ für } n \geq 2,\ x_1 = 0$$

*rekursiv definierte Folge $(x_n)_n$ konvergent ist.*

**Übung.** *Zeigen Sie, dass es für jedes $g \in C([0,1])$ ein $f \in C([0,1])$ mit*

$$f(x) - \int_0^x e^{-y} f(y)\,\mathrm{d}x = g(x)$$

*gibt.*

### 9.9.4 Lernkarten

Sie können wiederum die Lernkarten oder den Graphen für Ihre Wiederholung der Themen des Kapitels verwenden.

# Kapitel 10

# Mehrdimensionale Differentialrechnung

Wir erweitern in diesem Kapitel den Begriff der Ableitung, um auch Funktionen auf (offenen Teilmengen von) $\mathbb{R}^n$ nach $\mathbb{R}^m$ zu erlauben. Diese sind für verschiedene $n \in \mathbb{N}$ und auch verschiedene Dimensionen $m \in \mathbb{N}$ des Zielraums von Nützen, weshalb wir in dieser Hinsicht keine Einschränkungen treffen wollen. Ist beispielsweise eine Funktion $f : U \to \mathbb{R}^m$ für $U \subseteq \mathbb{R}^n$ und $n, m \geq 1$ gegeben, so könnte man sich für Folgendes interessieren.

- Ist $m = 1$ und $n$ beliebig, so kann man sich zum Beispiel fragen, ob $f$ ein Minimum oder ein Maximum annimmt. Weiter möchte man für $n = 2$ manchmal einen Graphen von $f$ zeichnen, der dann eine Fläche im $\mathbb{R}^3$ darstellt. Beispielsweise ist der Graph der Funktion $(x, y)^t \in B_1(0) \subseteq \mathbb{R}^2 \mapsto \sqrt{1 - x^2 - y^2} \in \mathbb{R}$ gerade die obere Hemisphäre der Sphäre $\mathbb{S}^2 = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$.

- Ist $n = 1$ und $m$ beliebig, so behandelt man Wege im $\mathbb{R}^m$, zu welchen verschiedene Begriffe von Interesse sind (Geschwindigkeit, Krümmung, Wegintegrale und so weiter).

- Sind $m = n$ beliebig, so lässt sich $f$ als Vektorfeld auffassen. Ein solches Vektorfeld lässt sich beispielsweise als Krafteinwirkung auffassen, womit man sich vielleicht für die Arbeit des Kraftfelds entlang eines Weges interessieren könnte.

- Ist immer noch $m = n$ beliebig, so ist $f$ möglicherweise ein Koordinatenwechsel. Wichtige konkrete Beispiele dafür wären Polarkoordinaten im $\mathbb{R}^2$, Zylinderkoordinaten im $\mathbb{R}^3$ oder sphärische Koordinaten im $\mathbb{R}^3$. (Wir werden allgemeine derartige glatten Koordinatensysteme im nächsten Kapitel betrachten.)

- Sind $m$ und $n$ beliebig, so kann $f$ ein Gleichungssystem der Form $f(x) = 0$ definieren, zu denen man die Lösungsmenge untersuchen will. Beispielsweise stellen die Lösungen von $f(x) = 0$ für $f : \mathbb{R}^3 \to \mathbb{R}^2, (x, y, z)^t \mapsto (x^2 + y^2 + z^2 - 1, (x - \frac{3}{2})^2 + y^2 + z^2 - 1)$ einen Kreis im $\mathbb{R}^3$ dar. (Lösungsmengen derartiger glatter Gleichungssysteme ergeben oft Beispiele für den Begriff der Teilmannigfaltigkeit, den wir ebenso im nächsten Kapitel besprechen werden.)

## 10.1 Die Ableitung

Wir betrachten im Folgenden Funktionen $f : U \to \mathbb{R}^m$, wobei $m, n \geq 1$ und der Definitionsbereich $U$ eine Teilmenge von $\mathbb{R}^n$ ist. Wir möchten nun kurz die Eigenschaften des Definitionsbereiches $U$ ansprechen, die dieser haben soll oder kann.

### 10.1.1 Der Definitionsbereich

Bis auf weiteres wird $U \subseteq \mathbb{R}^n$ eine offene Teilmenge sein, wobei wir, wenn nicht spezifisch anders angegeben, sowohl auf $\mathbb{R}^n$ als auch auf $\mathbb{R}^m$ die Euklidsche Norm $\|\cdot\| = \|\cdot\|_2$ verwenden werden (und verwendet haben). Offenheit von $U$ wird es uns erlauben, Punkte in $U$ von allen Richtungen annähern zu können. Auch wenn es nicht immer explizit erwähnt ist, wird $U$ im Folgenden stets nicht-leer sein.

Zusätzlich zur Annahme, dass $U$ offen ist, werden wir mitunter auch folgende Eigenschaften benötigen.

- $U$ ist **zusammenhängend**, falls sich $U$ nicht als disjunkte Vereinigung zweier offener, nicht-leerer Teilmengen von $U$ schreiben lässt (siehe Abschnitt 9.2.3).
- $U$ ist **wegzusammenhängend**, falls es zu je zwei Punkten $x_0, x_1 \in U$ einen Weg in $U$ gibt, der $x_0$ und $x_1$ verbindet (siehe Definition 9.46).
- $U$ ist **sternförmig**, falls es ein Zentrum $z \in U$ gibt, so dass für alle $x \in U$ und $t \in [0, 1]$ auch $(1 - t)z + tx \in U$ ist.
- $U$ ist **konvex**, falls für alle $x_0, x_1 \in U$ und $t \in [0, 1]$ auch $(1 - t)x_0 + tx_1 \in U$ ist.

Wir haben bereits in Proposition 9.49 gesehen, dass Zusammenhang und Wegzusammenhang für offene Mengen in $\mathbb{R}^n$ äquivalent sind. Wir bemerken, dass sternförmige Teilmengen von $\mathbb{R}^n$ automatisch wegzusammenhängend sind, da sich je zwei Punkte $x_0, x_1$ durch Aneinanderhängen der Geradensegmente zwischen $x_0$ und einem Zentrum $z$ (wie oben) und zwischen $z$ und $x_1$ verbinden lassen. Weiter ist jede konvexe Teilmenge sternförmig (jeder Punkt in $U$ lässt sich als Zentrum wählen). Umgekehrt braucht eine wegzusammenhängende Menge (auch wenn sie offen ist) nicht sternförmig zu sein und eine sternförmige Menge nicht konvex zu sein (finden Sie hier elementare Beispiele).

In Analogie zu Intervallen sind wir unter anderem an wegzusammenhängenden offenen Mengen interessiert, da wir zeigen werden, dass differenzierbare Funktionen mit Ableitung Null auf einem derartigen Gebiet konstant sind.

**Definition 10.1.** Ein **Gebiet** in $\mathbb{R}^n$ ist eine nicht-leere, offene, zusammenhängende Teilmenge von $\mathbb{R}^n$.

### 10.1.2 Lineare Abbildungen

Wir schreiben $e_1, \ldots, e_m$ für die Standardbasis des $\mathbb{R}$-Vektorraums $\mathbb{R}^m$, das heisst, für jedes $j \in \{1, \ldots, m\}$ ist

$$e_j = (0, 0, \ldots, 0, \underset{\substack{\uparrow \\ j}}{1}, 0, \ldots, 0)^t$$

der Vektor, der in der $j$-ten Zeile eine Eins und sonst überall nur Nulleinträge besitzt. Wie Sie bereits aus der Linearen Algebra I wissen, sind lineare Abbildungen $A : \mathbb{R}^n \to \mathbb{R}^m$ durch die Eigenschaft

$$A(ax + by) = aA(x) + bA(y)$$

für alle $a, b \in \mathbb{R}$ und $x, y \in \mathbb{R}^n$ definiert. Des Weiteren können diese eindeutig durch eine Matrix in $\mathrm{Mat}_{m,n}(\mathbb{R})$ beschrieben werden, wobei die $j$-te Spalte der Matrix gerade $A(e_j)$ ist.

Des Öfteren werden wir nicht so sehr auf die Unterscheidung zwischen der linearen Abbildung und ihrer Matrixdarstellung bestehen und die Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$ mit ihrer Darstellungsmatrix $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$ bezüglich der Standardbasis identifizieren.

Der Spezialfall $m = 1$ ist noch erwähnenswert. In diesem Fall lässt sich eine lineare Abbildung $A : \mathbb{R}^n \to \mathbb{R}$ auch als ein inneres Produkt mit einem fest gewählten Vektor $v \in \mathbb{R}^n$ interpretieren, so dass $A(x) = \langle x, v \rangle$ für alle $x \in \mathbb{R}^n$ gilt. Dies erlaubt eine geometrische Interpretation der Abbildung.

**Übung 10.2.** *Die Menge der linearen Abbildungen $\mathbb{R}^n \to \mathbb{R}$ nennt sich auch der Dualraum $(\mathbb{R}^n)^*$ von $\mathbb{R}^n$. Zeigen Sie, dass die Abbildung $v \in \mathbb{R}^n \mapsto \langle \cdot, v \rangle \in (\mathbb{R}^n)^*$ linear und bijektiv ist wie oben behauptet, wobei $\langle \cdot, v \rangle$ zu $v \in \mathbb{R}^n$ die lineare Abbildung $x \in \mathbb{R}^n \mapsto \langle x, v \rangle \in \mathbb{R}$ bezeichnet. Für welche $x$ in $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n \mid \|x\|_2 = 1\}$ ist $\langle x, v \rangle$ maximal (oder minimal)?*

### 10.1.3 Definitionen

Das Hauptziel bei der Definition der Ableitung für reellwertige Funktionen auf Teilmengen von $\mathbb{R}$ (Definition 7.1) war Funktionen (lokal) durch Geraden – auch affin lineare Funktionen von $\mathbb{R}$ nach $\mathbb{R}$ genannt – approximieren zu können. Genauso möchten wir nun für Funktionen $U \to \mathbb{R}^m$ mit $U \subseteq \mathbb{R}^n$ vorgehen. Dabei ist eine **affin lineare Funktion** $F : \mathbb{R}^n \to \mathbb{R}^m$ durch $F(x) = y_0 + L(x)$ für alle $x \in \mathbb{R}^n$ gegeben, wobei ein Punkt $y_0 \in \mathbb{R}^m$ und eine lineare Abbildung $L : \mathbb{R}^n \to \mathbb{R}^m$ fest gewählt sind.

**Definition 10.3** (Totale Ableitung)**.** Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$ eine Funktion. Dann heisst $f$ bei $x_0 \in U$ **differenzierbar** (oder **ableitbar**), falls es eine lineare Abbildung $L : \mathbb{R}^n \to \mathbb{R}^m$ gibt, so dass

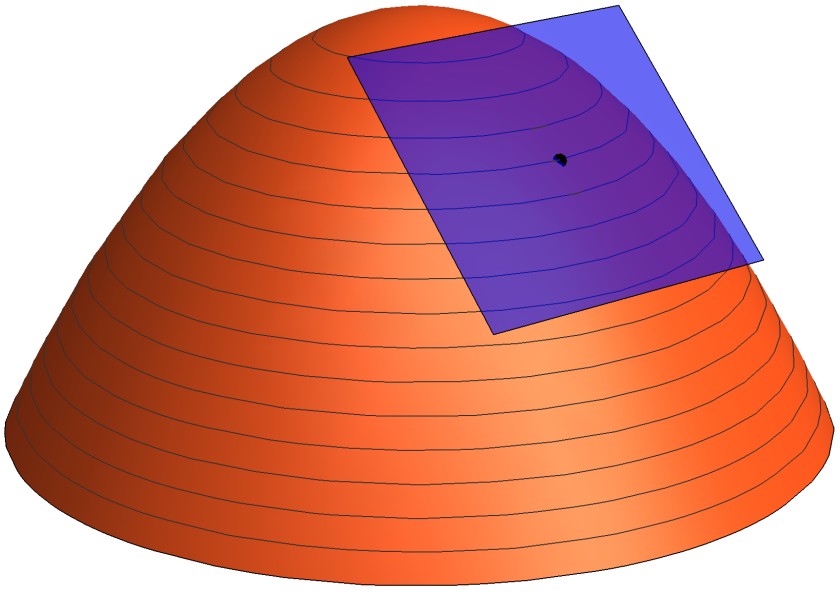
$$f(x_0 + h) = f(x_0) + L(h) + \alpha_f(x_0, h)$$

und $\alpha_f(x_0, h) = o(\|h\|)$ für $h \to 0$ oder äquivalenterweise

$$\lim_{h \to 0} \frac{\|f(x_0 + h) - f(x_0) - L(h)\|}{\|h\|} = 0$$

gilt. Die lineare Abbildung $L$ wird die **totale Ableitung**, das **Differential** oder die **Tangentialabbildung** genannt und als $\mathrm{D}_{x_0} f$, $\mathrm{d}f(x_0)$, $\mathrm{D}f(x_0)$ oder auch $f'(x_0)$ geschrieben. Weiter heisst $f$ differenzierbar, falls $f$ bei jedem Punkt in $U$ differenzierbar ist.

In diesem Zusammenhang wird, wie schon zuvor, $h = \triangle x$ das **Inkrement des Arguments** und $\triangle f(x_0, h) = f(x_0 + h) - f(x_0)$ das **Inkrement der Funktion** genannt. Wir bemerken, dass auf Grund der Offenheit von $U$ für $x_0 \in U$ und jedes hinreichend kleine $h \in \mathbb{R}^d$ ebenso $x_0 + h \in U$ gilt und damit in der Tat $f(x_0 + h)$ definiert ist. Es empfiehlt sich, die totale Ableitung als den linearen Teil der **besten affinen Approximation** $x_0 + h \mapsto f(x_0) + \mathrm{D}_{x_0} f h$ der Funktion zu sehen. Insbesondere wird dadurch die Analogie zur Differenzierbarkeit einer Funktion auf $\mathbb{R}$ in Definition 7.1 sichtbar (siehe insbesondere (7.2)). Wir bemerken noch, dass Differenzierbarkeit in $x_0$ Stetigkeit in $x_0$ impliziert (wieso?).



Figur 10.1: Für eine Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ entspricht die beste affine Approximation der Tangentialebene des Graphen, der wiederum eine Fläche im $\mathbb{R}^3$ darstellt.

**Applet 10.4** (Tangentialebene)**.** *Wir stellen wie bereits in obigem Bild die Tangentialebenen für die Graphen von zwei Funktionen $f : \mathbb{R}^2 \to \mathbb{R}$ dar. Des Weiteren werden die partiellen Ableitungen und Richtungsableitungen in Definition 10.5 visualisiert. Gibt es zu jedem Punkt eine Richtungsableitung die verschwindet?*

Für $v \in \mathbb{R}^m$ hat die konstante Abbildung $f(x) = v \in \mathbb{R}^m$ für alle $x \in \mathbb{R}^n$ bei jedem Punkt die totale Ableitung $0 \in \mathrm{Mat}_{m,n}(\mathbb{R})$. Eine affine Abbildung $f(x) = v + A(x) \in \mathbb{R}^m$ für alle $x \in \mathbb{R}^n$ und eine vorgegebene Matrix $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$ hat hingegen die Ableitung $D_x f = A$ für alle $x \in \mathbb{R}^n$ (wieso?).

Wie bereits betont, darf man bei der Bewegung $h \to 0$ im $\mathbb{R}^n$ keinerlei Einschränkungen vornehmen. Schränken wir die Bewegung dennoch auf Geraden ein (so dass das Inkrement des Arguments die Form $h = sv$ für $s \in \mathbb{R}$ und einen festen Vektor $v \in \mathbb{R}^n$ hat), so ergibt sich folgender neuer Begriff.

**Definition 10.5** (Ableitung entlang eines Vektors)**.** Sei $U \subseteq \mathbb{R}^n$ eine offene Teilmenge und $f : U \to \mathbb{R}^m$ eine Funktion. Die **Ableitung** von $f$ **entlang eines Vektors** $v \in \mathbb{R}^n$ ist an einer Stelle $x_0 \in U$ durch

$$\partial_v f(x_0) = \lim_{s \to 0} \frac{f(x_0 + sv) - f(x_0)}{s}$$

definiert, falls der Grenzwert existiert. Falls $\|v\| = 1$ gilt, so spricht man auch von der **Richtungsableitung** in der Richtung $v$ bei $x_0$.

VO: 09.03.18

Im Spezialfall, wo $v = e_j$ für ein $j \in \{1, \ldots, n\}$ ist, wird der obige Grenzwert

$$\partial_j f(x_0) = \partial_{e_j} f(x_0) = \lim_{s \to 0} \frac{f(x_0 + se_j) - f(x_0)}{s}$$

auch die **partielle Ableitung** in der $j$-ten Koordinate (oder der Variable $x_j$) bei $x_0$ genannt, falls er existiert. Wir schreiben mitunter auch $\frac{\partial f}{\partial x_j}(x_0)$ oder $\partial_{x_j} f(x_0)$. Existiert die partielle Ableitung in der $j$-ten Koordinate an jedem Punkt in $U$, so erhält man also eine Funktion $\partial_j f : U \to \mathbb{R}^m$.

Die partielle Ableitung (und die Richtungsableitung entlang eines beliebigen Vektors) ist also eine Ableitung nach einer der unabhängigen Variablen, wobei alle anderen Variablen quasi als Konstanten erachtet werden. Zum Beispiel existieren für die Funktion $f : \mathbb{R}^3 \to \mathbb{R}$ mit $f(x, y, z) = x(y^2 + \sin(z))$ für $x, y, z \in \mathbb{R}$ die partiellen Ableitungen bezüglich allen Koordinatenrichtungen und sind gegeben durch

$$\begin{aligned}
\partial_x f(x, y, z) &= y^2 + \sin(z) \\
\partial_y f(x, y, z) &= 2xy \\
\partial_z f(x, y, z) &= x \cos(z)
\end{aligned}$$

für alle $(x, y, z)^t \in \mathbb{R}^3$, da wir einfach alle uns bekannten Regeln aus Abschnitt 7.1.2 anwenden können.

Existiert die totale Ableitung, so lässt sich diese mittels folgender Proposition mit partiellen Ableitungen und Ableitung entlang beliebigen Vektoren in Verbindung bringen.

**Proposition 10.6** (Matrixdarstellung des totalen Differentials)**.** *Sei $U \subseteq \mathbb{R}^n$ offen und sei $f : U \to \mathbb{R}^m$ bei $x_0 \in U$ differenzierbar. Dann existiert für jedes $v \in \mathbb{R}^n$ die Ableitung von $f$*

*entlang $v$ und es gilt*

$$\partial_v f(x_0) = \mathrm{D}_{x_0} f(v).$$

*Insbesondere ist die totale Ableitung $\mathrm{D}_{x_0} f$ eindeutig durch die partiellen Ableitungen bestimmt und es gilt*

$$\mathrm{D}_{x_0} f = (\partial_1 f(x_0), \ldots, \partial_n f(x_0)) \in \mathrm{Mat}_{m,n}(\mathbb{R}),$$

*wobei letzteres auch die **Jacobi-Matrix** von $f$ bei $x_0$ genannt wird.*

*Beweis.* Nach Annahme existiert die totale Ableitung $\mathrm{D}_{x_0} f$ und es gilt

$$f(x_0 + h) = f(x_0) + \mathrm{D}_{x_0} f(h) + o(\|h\|)$$

für $h \to 0$. Wir setzen $h = sv$ für $s \to 0$ und $v \in \mathbb{R}^n$, womit gilt

$$\partial_v f(x_0) = \lim_{s \to 0} \frac{f(x_0 + sv) - f(x_0)}{s} = \lim_{s \to 0} (\mathrm{D}_{x_0} f(v) + o(1)) = \mathrm{D}_{x_0} f(v).$$

Also existiert die Ableitungen von $f$ entlang dem (beliebigen) Vektor $v$ bei $x_0$. Insbesonderen existieren alle partiellen Ableitung von $f$ bei $x_0$ und die partielle Ableitung in der $j$-ten Richtung für $j \in \{1, \ldots, n\}$ stellt die $j$-te Spalte der Matrix $\mathrm{D}_{x_0} f$ dar, wie behauptet. □

Interessant wäre auch die Umkehrung von Proposition 10.6; unter anderem da a priori nicht klar ist, wie man die Existenz einer totalen Ableitung in konkreten Situationen nachweisen kann. Wie wir sehen werden, existiert unter gewissen, nicht allzu starken Annahmen, eine solche Umkehrung.

Wie schon im Eindimensionalen gelten auch hier Summen- und Produktregel, wie folgende Übung zeigt.

**Wichtige Übung 10.7** (Summen- und Produktregel)**.** *Sei $U \subseteq \mathbb{R}^n$ offen und seien $f_1, f_2 : U \to \mathbb{R}^m$ Funktionen. Angenommen $f_1$ und $f_2$ sind differenzierbar bei $x_0 \in U$.*

*(i) Zeigen Sie, dass $f_1 + f_2$ bei $x_0$ differenzierbar ist und*

$$\mathrm{D}_{x_0}(f_1 + f_2) = \mathrm{D}_{x_0} f_1 + \mathrm{D}_{x_0} f_2$$

*erfüllt.*

*(ii) Sei jetzt $m = 1$. Zeigen Sie, dass $f_1 \cdot f_2$ bei $x_0$ differenzierbar ist und*

$$\mathrm{D}_{x_0}(f_1 f_2) = f_2(x_0)\, \mathrm{D}_{x_0} f_1 + f_1(x_0)\, \mathrm{D}_{x_0} f_2$$

*erfüllt.*

*(iii) Nun nehmen wir stattdessen an, dass bloss die Ableitungen $\partial_v f_1$, $\partial_v f_2$ entlang $v \in \mathbb{R}^n$ existieren. Formulieren und beweisen Sie für diese Ableitung analoge Aussagen wie in (i) und (ii).*

Landau-Notation: Wir merken an dieser Stelle kurz an, dass keine formale Definition der oben verwendeten Landau-Symbole (siehe zum Beispiel Definition 10.3) gegeben wurde, da dies nur eine kleine Anpassung der Diskussionen in den Abschnitten 5.6 und 8.6.1 darstellt. Beispielsweise ist für eine Funktion $f : U \to \mathbb{R}^m$ auf einer offenen Menge $U \subseteq \mathbb{R}^n$ und $x_0 \in U$

$$f(x) = o(\|x - x_0\|)$$

für $x \to x_0$, falls $\frac{\|f(x)\|}{\|x-x_0\|}$ für $x \to x_0$ gegen Null geht.

### 10.1.4 Reduktion der Dimension

Es wäre von Vorteil, wenn wir die Differenzierbarkeit einer Funktion von $\mathbb{R}^n$ nach $\mathbb{R}^m$ auf die Differenzierbarkeit reellwertiger Funktionen von $\mathbb{R}$ nach $\mathbb{R}$ zurückführen könnten. Denn dann würden uns sämtliche Hilfsmittel aus den Kapiteln 7 und 8 zur Verfügung stehen. Hierbei ist die Reduktion der Dimension des Zielraumes auf $m = 1$ eine Charakterisierung.

**Lemma 10.8** (Differenzierbarkeit via Komponenten)**.** *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$ eine Funktion. Dann ist $f$ genau dann bei $x_0 \in U$ differenzierbar, wenn die Komponenten $f_k = \pi_k \circ f$ für jedes $k \in \{1, \dots, m\}$ bei $x_0$ differenzierbar sind. In diesem Fall gilt*

$$\mathrm{D}_{x_0} f = \begin{pmatrix} \mathrm{D}_{x_0} f_1 \\ \vdots \\ \mathrm{D}_{x_0} f_m \end{pmatrix}.$$

*Beweis.* Wir beweisen nur eine der beiden Implikationen und überlassen die zweite den Leserinnen und Lesern (Übung 10.9). Angenommen $f_k$ ist für jedes $k \in \{1, \dots, m\}$ bei $x_0$ differenzierbar. Dann gilt für $k \in \{1, \dots, m\}$

$$f_k(x_0 + h) = f_k(x_0) + \mathrm{D}_{x_0} f_k(h) + \alpha_k(x_0, h)$$

für gewisse Funktionen $\alpha_k$ mit $\alpha_k(x_0, h) = o(\|h\|)$ für $h \to 0$. Daraus folgt aber

$$f(x_0 + h) = \begin{pmatrix} f_1(x_0 + h) \\ \vdots \\ f_m(x_0 + h) \end{pmatrix} = \begin{pmatrix} f_1(x_0) \\ \vdots \\ f_m(x_0) \end{pmatrix} + \begin{pmatrix} \mathrm{D}_{x_0} f_1 \\ \vdots \\ \mathrm{D}_{x_0} f_m \end{pmatrix} h + \alpha(x_0, h),$$

wobei

$$\alpha(x_0, h) = \begin{pmatrix} \alpha_1(x_0, h) \\ \vdots \\ \alpha_m(x_0, h) \end{pmatrix} = o(\|h\|)$$

für $h \to 0$ gilt. Also ist $f$ differenzierbar und es gilt die behauptete Formel für $\mathrm{D}_{x_0} f$. □

**Übung 10.9.** *Beweisen Sie die zweite Richtung in Lemma 10.8.*

Die Reduktion der Existenz der totalen Ableitungen auf die Existenz der partiellen Ableitung ist etwas schwieriger und im Allgemeinen schlicht nicht möglich. Betrachten wir beispielsweise die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch

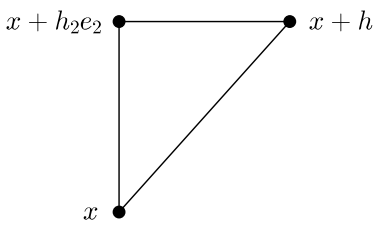
$$f(x,y) = \begin{cases} \frac{xy}{\sqrt{x^2+y^2}} & \text{falls } (x,y)^t \in \mathbb{R}^2 \setminus \{0\} \\ 0 & \text{falls } (x,y)^t = 0 \end{cases} \tag{10.1}$$

für $(x,y)^t \in \mathbb{R}^2$, so existieren beide partiellen Ableitungen $\partial_x f$, $\partial_y f$ auf ganz $\mathbb{R}^2$, aber die totale Ableitung $\mathrm{D}_0 f$ existiert trotzdem nicht. Denn für $x = y$ gilt $f(x,y) = \frac{|x|}{\sqrt{2}}$ (wieso impliziert dies, dass die Ableitung nicht existiert?).

Nimmt man jedoch etwas schönere Eigenschaften (als die blosse Existenz) der partiellen Ableitungen an, so erhält man folgende Aussage.

**Satz 10.10** (Existenz der totalen Ableitung)**.** *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$ eine Funktion. Falls für jedes $j \in \{1,\dots,n\}$ die partielle Ableitung $\partial_j f$ auf ganz $U$ existiert und eine stetige Funktion definiert, so ist $f$ auf ganz $U$ differenzierbar.*

Es empfiehlt sich an dieser Stelle zu überprüfen, dass das Beispiel in (10.1) die Stetigkeitsvoraussetzung des Satzes nicht erfüllt.




Figur 10.2: Illustration des Beweis von Satz 10.10 für $n = 2$. Der Term $f(x+h)-f(x)$ lässt sich mittels des Hilfspunktes $x+h_2e_2$ als Summe der beiden Ausdrücke $f(x+h)-f(x+h_2e_2)$ und $f(x+h_2e_2)-f(x)$ schreiben. Diese sind aber jeweils Differenzen von Auswertungen von $f$ bei Punkten, die sich jeweils nur in einer Koordinate unterscheiden. Dies erlaubt die Anwendung des Mittelwertsatzes der eindimensionalen Differentialrechnung.

*Beweis.* Auf Grund von Lemma 10.8 können wir $m = 1$ annehmen. Für $x \in U$ und hinreichend kleine $h = (h_1,\dots,h_n)^t \in \mathbb{R}^n$ gilt dann

$$\begin{aligned} f(x+h) - f(x) &= f(x_1+h_1, x_2+h_2, \dots, x_n+h_n) - f(x_1, x_2+h_2, \dots, x_n+h_n) \\ &\quad + f(x_1, x_2+h_2, x_3+h_3 \dots, x_n+h_n) - f(x_1, x_2, x_3+h_3 \dots, x_n+h_n) \\ &\quad + \dots + f(x_1, \dots, x_{n-1}, x_n+h_n) - f(x) \\ &= \partial_1 f(x_1+\xi_1(h), x_2+h_2, \dots, x_n+h_n)h_1 \\ &\quad + \partial_2 f(x_1, x_2+\xi_2(h), x_3+h_3, \dots, x_n+h_n)h_2 \\ &\quad + \dots + \partial_n f(x_1, \dots, x_{n-1}, x_n+\xi_n(h))h_n, \end{aligned}$$

wobei für jedes $j \in \{1,\dots,n\}$ nach dem Mittelwertsatz (Theorem 7.29) ein Zwischenpunkt $\xi_j(h)$ zwischen 0 und $h_j$ gewählt wurde. Wegen Stetigkeit der partiellen Ableitungen können wir nun in obigen Ausdrücken stattdessen die partiellen Ableitungen bei $x$ betrachten.

Tatsächlich gilt für

$$\begin{aligned}\alpha(x,h) = &\big(\partial_1 f(x_1+\xi_1(h), x_2+h_2, \ldots, x_n+h_n) - \partial_1 f(x)\big)h_1 \\ &+ \big(\partial_2 f(x_1, x_2+\xi_2(h), x_3+h_3, \ldots, x_n+h_n) - \partial_2 f(x)\big)h_2 \\ &+ \ldots + \big(\partial_n f(x_1, \ldots, x_{n-1}, x_n+\xi_n(h)) - \partial_n f(x)\big)h_n,\end{aligned}$$

nach den Annahmen des Satzes und wegen $\frac{|h_k|}{\|h\|} \leq 1$ für alle $h \in \mathbb{R}^n$ und $k \in \{1, \ldots, n\}$ die Asymptotik

$$\lim_{h \to 0} \frac{\alpha(x,h)}{\|h\|} = 0.$$

Daher ist schlussendlich

$$f(x+h) - f(x) = \partial_1 f(x)h_1 + \ldots + \partial_n f(x)h_n + \alpha(x,h) = Lh + \alpha(x,h),$$

wobei $L = (\partial_1 f(x), \ldots, \partial_n f(x)) \in \mathrm{Mat}_{1,n}(\mathbb{R})$. Also ist $f$ bei $x$ differenzierbar und da $x \in U$ beliebig war, ist $f$ also differenzierbar. □

**Definition 10.11.** Wir nennen eine Funktion $f : U \to \mathbb{R}^m$ auf einer offenen Teilmenge $U \subseteq \mathbb{R}^n$ **stetig differenzierbar**, wenn $f$ differenzierbar ist und die Ableitung

$$x \in U \mapsto \mathrm{D}_x f \in \mathrm{Mat}_{m,n}(\mathbb{R})$$

stetig ist.

Nach Satz 10.10 und Proposition 10.6 ist $f : U \to \mathbb{R}^m$ genau dann stetig differenzierbar, wenn alle partiellen Ableitungen von $f$ existieren und stetig sind. Kombinieren wir die Aussagen dieses Abschnitts so können wir die Differenzierbarkeit vieler Abbildungen $f$ von $U \subseteq \mathbb{R}^n$ nach $\mathbb{R}^m$ beweisen. Die Ableitung $D_x f$ ist in diesem Fall immer eine lineare Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^m$, die wir, wie bereits erwähnt, mit der Jacobi-Matrix in $\mathrm{Mat}_{mn}(\mathbb{R})$ (bestehend aus allen partiellen Ableitungen) identifizieren.

**Beispiel 10.12.** *Sei* $f : \mathbb{R}^2 \to \mathbb{R}^2$ *definiert durch* $f : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x^2 - \cos(xy) \\ y^4 - \exp(x) \end{pmatrix}$. *Die totale Ableitung oder Jacobi-Matrix von* $f$ *bei* $\begin{pmatrix} x \\ y \end{pmatrix}$ *ist dann gegeben durch*

$$\begin{pmatrix} 2x + \sin(xy)y & \sin(xy)x \\ -\exp(x) & 4y^3 \end{pmatrix}.$$

## 10.2 Die Kettenregel und der Mittelwertsatz

### 10.2.1 Verknüpfungen differenzierbarer Funktionen

**Satz 10.13** (Kettenregel der mehrdimensionalen Differentialrechnung)**.** *Seien $k, m, n \geq 1$, $U \subseteq \mathbb{R}^n$ offen, $V \subseteq \mathbb{R}^m$ offen. Weiter sei $f : U \to V$ bei $x_0$ differenzierbar und $g : V \to \mathbb{R}^k$ bei $f(x_0)$ differenzierbar. Dann ist $g \circ f$ bei $x_0$ differenzierbar und die totale Ableitung $\mathrm{D}_{x_0}(g \circ f)$ bei $x_0$ ist durch die Verknüpfungen der linearen Abbildungen*

$$\mathrm{D}_{x_0}(g \circ f) = \mathrm{D}_{f(x_0)}g \circ \mathrm{D}_{x_0}f \tag{10.2}$$

*gegeben.*

Wir erinnern daran, dass $\mathrm{D}_{f(x_0)}g$ mit einer $k \times m$-Matrix und $\mathrm{D}_{x_0}f$ mit einer $m \times n$-Matrix identifiziert werden kann, womit die Verknüpfung (10.2) von den Dimensionen her Sinn macht.

*Beweis.* Wir verwenden die Definition der Differenzierbarkeit von $f$ bei $x_0$, womit gilt

$$f(x_0 + h) = f(x_0) + \mathrm{D}_{x_0}f(h) + \alpha_f(x_0, h)$$

und $\alpha_f(x_0, h) = o(\|h\|)$ für $h \to 0$. Nach Differenzierbarkeit von $g$ bei $y_0 = f(x_0)$ gilt ebenso

$$g(y_0 + \tilde{h}) = g(y_0) + \mathrm{D}_{y_0}g(\tilde{h}) + \alpha_g(y_0, \tilde{h})$$

mit $\alpha_g(y_0, \tilde{h}) = o(\|\tilde{h}\|)$ für $\tilde{h} \to 0$. Gemeinsam erhalten wir für $h \in \mathbb{R}^n$ klein genug und

$$\tilde{h} = f(x_0 + h) - f(x_0) = \mathrm{D}_{x_0}f(h) + \alpha_f(x_0, h)$$

die Gleichung

$$\begin{aligned} g(f(x_0 + h)) &= g(f(x_0)) + \mathrm{D}_{y_0}g\big(\mathrm{D}_{x_0}f(h) + \alpha_f(x_0, h)\big) + \alpha_g(y_0, \tilde{h}) \\ &= g(f(x_0)) + \mathrm{D}_{y_0}g \circ \mathrm{D}_{x_0}f(h) + \alpha_{g\circ f}(x_0, h), \end{aligned}$$

wobei wir

$$\alpha_{g\circ f}(x_0, h) = \mathrm{D}_{y_0}g(\alpha_f(x_0, h)) + \alpha_g(y_0, f(x_0 + h) - f(x_0))$$

gesetzt haben. Wir möchten nun zeigen, dass $\alpha_{g\circ f}(x_0, h) = o(\|h\|)$ für $h \to 0$. Da $\alpha_f(x_0, h) = o(\|h\|)$ für $h \to 0$ gilt, ist auch $\|\mathrm{D}_{y_0}g(\alpha_f(x_0, h))\| \leq \|\mathrm{D}_{y_0}g\|_{\mathrm{op}}\|\alpha_f(x_0, h)\| = o(\|h\|)$ für $h \to 0$.

Es bleibt zu zeigen, dass $\alpha_g(y_0, f(x_0 + h) - f(x_0)) = o(\|h\|)$ für $h \to 0$ ist. Nach Differenzierbarkeit von $g$ bei $y_0$ gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$, so dass für $\tilde{h} \in \mathbb{R}^m$ mit $\|\tilde{h}\| < \delta$ die Abschätzung

$$\|\alpha_g(y_0, \tilde{h})\| \leq \varepsilon\|\tilde{h}\| \tag{10.3}$$

gilt. Nach vorrausgesetzer Differenzierbarkeit von $f$ bei $x_0$ gilt für $\tilde{h} = f(x_0 + h) - f(x)$ die Abschätzung

$$\begin{aligned}\|\tilde{h}\| &= \|\,\mathrm{D}_{x_0} f(h) + \alpha_f(x_0, h)\| \\ &\leq \|\,\mathrm{D}_{x_0} f(h)\| + \|\alpha_f(x_0, h)\| \\ &\leq \|\,\mathrm{D}_{x_0} f\|_{\mathrm{op}}\|h\| + o(\|h\|) = O(\|h\|)\end{aligned}$$

für $h \to 0$. Also gibt es eine offene Umgebung $O$ von $0 \in \mathbb{R}^n$ und eine Konstante $C > 0$ (zum Beispiel $C = \|\,\mathrm{D}_{x_0} f\|_{\mathrm{op}} + 1$) mit $\|\tilde{h}\| \leq C\|h\|$ für alle $h \in O$. Für $h \in O$ mit $\|h\| < \frac{\delta}{C}$ folgt also $\|\tilde{h}\| < \delta$ und mit (10.3) auch

$$\|\alpha_g(y_0, \underbrace{f(x_0 + h) - f(x_0)}_{=\tilde{h}})\| \leq \varepsilon\|\tilde{h}\| \leq C\varepsilon\|h\|.$$

Da die Konstante $C$ von $\varepsilon$ unabhängig ist, folgt die Differenzierbarkeit von $g \circ f$ bei $x_0$ und die Kettenregel in Gleichung (10.2). □

VO: 12.03.18

Wir betrachten nun den Spezialfall $n = 1$ für die Kettenregel. Sei also $\gamma : I \to V \subseteq \mathbb{R}^m$ ein differenzierbarer Weg von einem offenen Intervall $I$ in eine offene Teilmenge $V \subseteq \mathbb{R}^m$. Sei weiter $f : V \to \mathbb{R}^k$ differenzierbar. Dann ergibt die Kettenregel (Satz 10.13), dass $f \circ \gamma$ differenzierbar ist und die Formel

$$(f \circ \gamma)'(t) = \mathrm{D}_{\gamma(t)} f \cdot \gamma'(t)$$

für alle $t \in I$ gilt. Sollte noch zusätzlich $k = 1$ sein, so ist $f \circ \gamma : I \to \mathbb{R}$ und $(\,\mathrm{D}_{\gamma(t)} f)\gamma'(t)$ ist für $t \in I$ das Matrixprodukt der $1 \times m$-Matrix $\mathrm{D}_{\gamma(t)} f$ mit der $m \times 1$-Matrix $\gamma'(t)$ (ein Vektor in $\mathbb{R}^m$). Wir interpretieren in diesem Fall $\mathrm{D}_x f$ für $x \in V$ auch als den Spaltenvektor

$$\mathrm{grad} f(x) = \nabla f(x) = (\,\mathrm{D}_x f)^t \in \mathbb{R}^m$$

und nennen dies den **Gradienten der Funktion** $f$ bei der Stelle $x$. In dieser Schreibweise erhalten wir die Formel

$$(f \circ \gamma)'(t) = \mathrm{D}_{\gamma(t)} f \cdot \gamma'(t) = \big\langle \nabla f(\gamma(t)), \gamma'(t) \big\rangle \tag{10.4}$$

für alle $t \in I$.

Der Begriff der Richtungsableitung und der Fall der Gleichheit in der Cauchy-Schwarz-Ungleichung erlauben es uns auch, eine geometrische Interpretation des Gradienten einer Funktion anzugeben. Ist $f : U \to \mathbb{R}$ eine differenzierbare Funktion auf einer offenen Teilmenge $U \subseteq \mathbb{R}^n$ und ist $v \in \mathbb{R}^n$ ein Vektor der Länge 1, so gilt nach Proposition 10.6 und vorherigem bei $x \in U$

$$\partial_v f(x) = \mathrm{D}_x f(v) = \langle \nabla f(x), v \rangle\,.$$

Angenommen $\nabla f(x) \neq 0$. Nach der Ungleichung von Cauchy-Schwarz (Proposition 5.86) ist obiger Ausdruck genau dann maximal (das heisst, gleich $\|\nabla f(x)\|$), wenn $v$ in dieselbe Richtung wie $\nabla f(x)$ zeigt (mit positivem skalarem Vielfachen – also $v = \frac{1}{\|\nabla f(x)\|}\nabla f(x)$). In Worten ausgedrückt heisst dies, dass der Gradient von $f$ an jedem Punkt in die Richtung der grössten Richtungsableitung zeigt, das heisst, die Richtung des grössten Anstiegs um $x$ kennzeichnet. Des Weiteren gibt $\|\nabla f(x)\|$ die Steigung in dieser Richtung an.

### 10.2.2 Geometrische Interpretation der mehrdimensionalen Kettenregel

Für $x \in \mathbb{R}^n$ definieren wir den **Tangentenraum** $\mathrm{T}_x\mathbb{R}^n$ von $\mathbb{R}^n$ bei $x$ durch

$$\mathrm{T}_x\mathbb{R}^n = \{x\} \times \mathbb{R}^n,$$

wobei wir Elemente von $\mathrm{T}_x\mathbb{R}^n$ als Vektoren mit Fusspunkt $x$ visualisieren. Via

$$(x, v) + (x, w) = (x, v + w), \quad \alpha(x, v) = (x, \alpha v)$$

für $(x, v), (x, w) \in \mathrm{T}_x\mathbb{R}^n$ und $\alpha \in \mathbb{R}$ statten wir den Tangentenraum $\mathrm{T}_x\mathbb{R}^n$ bei $x$ mit einer natürlichen Vektorraumstruktur aus, bezüglich der Vektoren wie üblich addiert und skaliert werden, aber der Ort hierbei unverändert bleibt.

Die disjunkte Vereinigung aller Tangentenräume ist das sogenannte **Tangentenbündel**

$$\mathrm{T}\mathbb{R}^n = \bigsqcup_{x \in \mathbb{R}^n} \mathrm{T}_x\mathbb{R}^n = \bigsqcup_{x \in \mathbb{R}^n} \{x\} \times \mathbb{R}^n = \mathbb{R}^n \times \mathbb{R}^n$$

von $\mathbb{R}^n$, welches man sich auch als Phasenraum vorstellen kann. Da wir für das Tangentenbündel verschiedene Fusspunkte betrachten, gibt es keine natürliche Weise, auf diesem eine Vektorraumstruktur zu definieren (wir möchten nur „Vektoren“ und nicht Fusspunkte addieren).

Weiter definieren wir für eine offene Teilmenge $U \subseteq \mathbb{R}^n$ genau gleich den Tangentenraum von $U$ bei $x \in U$ als

$$\mathrm{T}_x U = \{x\} \times \mathbb{R}^n,$$

welcher sich wie zuvor als Vektorraum aller möglichen Ableitungen $\gamma'(0)$ für differenzierbare Wege $\gamma$ mit Werten in $U$, die $\gamma(0) = x$ erfüllen, auffassen lässt. Des Weiteren ist das Tangentenbündel von $U$ durch

$$\mathrm{T}U = \bigsqcup_{x \in U} \mathrm{T}_x U = \bigsqcup_{x \in U} \{x\} \times \mathbb{R}^n = U \times \mathbb{R}^n$$

definiert.

In diesem Kontext können wir zum Beispiel die Ableitung zur Zeit $t$ eines differenzierbaren Weges $\gamma : I \to U$ als

$$\mathrm{D}_t\gamma = (\gamma(t), \gamma'(t)) \in \mathrm{T}_{\gamma(t)}U \subseteq \mathrm{T}U$$

interpretieren, wobei $\gamma(t)$ als der Ort zum Zeitpunkt $t$ und $\gamma'(t)$ als die gerichtete Geschwindigkeit zum Zeitpunkt $t$ aufgefasst wird. Der Vorteil dieses Gesichtspunktes ist gewissermassen, dass man gleichzeitig beide interessanten „Daten“ Ort und Geschwindigkeit zur Verfügung hat.

Wir definieren nun auch für allgemeine differenzierbare Abbildungen $f : U \to V$ und $g : V \to \mathbb{R}^k$ auf offenen Teilmengen $U \subseteq \mathbb{R}^n$ und $V \subseteq \mathbb{R}^m$ die Ableitungen als die Abbildungen

$$\begin{aligned} \mathrm{D}f : U \times \mathbb{R}^n = \mathrm{T}U &\to V \times \mathbb{R}^m = \mathrm{T}V \\ (x, v) &\mapsto (f(x),\, \mathrm{D}_x f(v)) \end{aligned}$$

und

$$\begin{aligned} \mathrm{D}g : V \times \mathbb{R}^m = \mathrm{T}V &\to \mathbb{R}^k \times \mathbb{R}^k = \mathrm{T}\mathbb{R}^k \\ (y, w) &\mapsto (g(y),\, \mathrm{D}_y g(w)), \end{aligned}$$

die wiederum sowohl eine Ort- als auch eine Geschwindigkeitskomponente betrachten. Dann nimmt die Kettenregel auf ganz $U$ die einfachere Form

$$\mathrm{D}(g \circ f) = \mathrm{D}g \circ \mathrm{D}f$$

an. In der Tat gilt für $(x, v) \in \mathrm{T}U$ nach der Kettenregel

$$\mathrm{D}(g \circ f)(x, v) = \big(g \circ f(x),\, \mathrm{D}_x(g \circ f)(v)\big) = \big(g(f(x)),\, \mathrm{D}_{f(x)}g\, \mathrm{D}_x f(v)\big) = \mathrm{D}g(\,\mathrm{D}f(x, v)).$$

Diese Unterscheidung zwischen Ort und Tangentenvektor macht aus vielerlei Sicht Sinn. Betrachten wir zum Beispiel die offene Menge $U$ als unser Universum, so lassen wir natürlich nur Wege in $U$ zu, womit alle möglichen Ableitungen eines differenzierbaren Weges Tangentenvektoren zu Punkten in $U$ sind. Des Weiteren hat das Verdoppeln eines Tangenvektors in $T_x\mathbb{R}^n$ bei $x$ die klare physikalische Interpretation einer Verdoppelung der Geschwindigkeit eines Weges durch $x$, doch hat das Verdoppeln der Ortskoordinaten keine natürliche Interpretation, da es keinen physikalisch sinnvollen Ursprung des Koordinatensystems gibt. Wir werden diese Sichtweise nicht sehr oft für offene Teilmengen des Euklidschen Raum $\mathbb{R}^n$ aber später für Teilmannigfaltigkeiten von $\mathbb{R}^n$ verwenden.

### 10.2.3 Der Mittelwertsatz

Wir formulieren nun den Mittelwertsatz für reellwertige differenzierbare Funktionen auf einer offenen Menge in $\mathbb{R}^n$, der eine Verallgemeinerung des Mittelwertsatzes für Funktionen

auf $\mathbb{R}$ darstellt (Theorem 7.29). Dazu betrachten wir eine gegebene Funktion $f$ entlang eines Geradenstücks in der offenen Menge.

**Satz 10.14** (Mittelwertsatz für reellwertige Funktionen auf $\mathbb{R}^n$)**.** *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ differenzierbar. Sei $x_0 \in U$ und $h \in \mathbb{R}^n$. Falls $x_0 + th \in U$ für alle $t \in [0,1]$, dann gilt*

$$f(x_0 + h) - f(x_0) = \mathrm{D}_\xi f(h) = \partial_h f(\xi)$$

*für ein $\xi = x_0 + t_\xi h$ mit $t_\xi \in (0,1)$.*

In Worten ausgedrückt existiert also entlang des geraden Weges zwischen $x_0$ und $x_0 + h$ ein Punkt, wo die Ableitung entlang des durch den geraden Weg gegebenen Vektors gerade die Differenz der Funktionswerte an den Randpunkten des Weges ist.

*Beweis.* Wir bemerken, dass die Ableitung des geraden Weges $t \in \mathbb{R} \mapsto x_0 + th$ für vorgegebene $x_0, h \in \mathbb{R}^n$ bei jedem $t$ gleich $h$ ist. Daher erfüllt die Funktion

$$g : t \in [0,1] \mapsto f(x_0 + th) \in \mathbb{R}$$

auf Grund der Kettenregel in Satz 10.13 alle Voraussetzungen des eindimensionalen Mittelwertsatzes (Theorem 7.29). Also existiert $t_\xi \in (0,1)$ mit $g(1) - g(0) = g'(t_\xi) = \mathrm{D}_{x_0 + t_\xi h} f(h)$ nach der Kettenregel und somit

$$f(x_0 + h) - f(x) = g(1) - g(0) = g'(t_\xi) = \mathrm{D}_\xi f(h)$$

für $\xi = x_0 + t_\xi h$. □

**Korollar 10.15.** *Sei $U \subseteq \mathbb{R}^n$ ein Gebiet und sei $f : U \to \mathbb{R}^m$ differenzierbar mit $\mathrm{D}_x f = 0$ für alle $x \in U$. Dann ist $f$ konstant.*

*Beweis.* Es genügt den Fall $m = 1$ zu betrachten (wieso?). Wir nehmen an, dass $U$ nichtleer ist und wählen ein $x_0 \in U$. Wir betrachten

$$U' = \{x \in U \mid f(x) = f(x_0)\}.$$

Da $f$ stetig ist, ist $U'$ eine abgeschlossene Teilmenge von $U$ (siehe Proposition 9.37). Des Weiteren folgt aus der Annahme und Satz 10.14, dass $U'$ offen ist: In der Tat existiert zu $x \in U'$ ein $\varepsilon > 0$ mit $B_\varepsilon(x) \subseteq U$ und da sich jeder Punkt $y \in B_\varepsilon(x)$ mit einem geraden Weg zu $x$ verbinden lässt, gilt nach Satz 10.14 auch $f(y) = f(x) = f(x_0)$. Also ist $y \in U'$ und da $y \in B_\varepsilon(x)$ beliebig war, ist $B_\varepsilon(x) \subseteq U'$.

Da aber $U$ zusammenhängend ist und $U'$ nicht-leer ist, folgt $U' = U$ und damit auch das Korollar. □

**Definition 10.16** (Lokale Lipschitz-Stetigkeit)**.** Eine Funktion $f : X \to Y$ zwischen zwei metrischen Räumen $X, Y$ heisst **lokal Lipschitz-stetig**, falls für jedes $x_0 \in X$ ein $\varepsilon > 0$ existiert, so dass $f|_{B_\varepsilon(x_0)}$ Lipschitz-stetig ist.

**Korollar 10.17.** *Sei $U \subseteq \mathbb{R}^n$ offen und sei $f : U \to \mathbb{R}^m$ eine stetig differenzierbare Funktion. Dann ist $f$ lokal Lipschitz-stetig. Falls $U$ zusätzlich konvex und die Ableitung beschränkt ist, dann ist $f$ sogar Lipschitz-stetig.*

*Beweis.* Es genügt den Fall $m = 1$ zu betrachten (wieso?). Wir nehmen zuerst an, dass $U$ konvex ist und die Ableitung beschränkt ist. Letzteres bedeutet, dass es ein $M \geq 0$ gibt, so dass $\|\mathrm{D}_\xi f\|_{\mathrm{op}} \leq M$ für alle $\xi \in U$. Aus dem Mittelwertsatz (Satz 10.14) folgt damit für $x, y \in U$

$$\|f(x) - f(y)\| = \|\mathrm{D}_\xi f(x - y)\| \leq M\|x - y\|$$

für ein $\xi \in U$, da $U$ konvex ist und somit das Geradenstück zwischen $x$ und $y$ enthält. Dies beweist die zweite Aussage im Korollar.

Die erste Aussage folgt aus der zweiten angewendet auf den Ball $U_0 = B_\varepsilon(x_0)$ und $f_0 = f|_{U_0}$, wobei $\varepsilon > 0$ so gewählt ist, dass $\overline{B_\varepsilon(x_0)} \subseteq U$. In der Tat ist dann $U_0$ konvex (wieso?) und die Abbildung $\xi \in \overline{B_\varepsilon(x_0)} \mapsto D_\xi f$ ist als stetige Funktion auf der kompakten Menge $\overline{B_\varepsilon(x_0)}$ (siehe Satz 9.66) beschränkt, was die Beschränktheit von der Ableitung auf $B_\varepsilon(x_0)$ impliziert. □

**Übung 10.18** (Eine Distanzfunktion auf $U$)**.** *Sei $U \subseteq \mathbb{R}^n$ offen und zusammenhängend. Wir sagen, dass ein Weg (d.h. eine stetige Abbildung) $\gamma : [0, 1] \to U$ stückweise differenzierbar ist, falls es eine Zerlegung $\mathfrak{Z} = \{t_0 = 0 < t_1 < \ldots < t_K = 1\}$ von $[0, 1]$ gibt, sodass für alle $k \in \{1, \ldots, K\}$ die Einschränkung $\gamma_{|[t_{k-1},t_k]}$ stetig differenzierbar ist.*

*(i) Zeigen Sie, dass es zu je zwei Punkten $x, y \in U$ einen stückweise differenzierbaren Weg von $x$ nach $y$ gibt.*

*Definieren Sie die Länge eines stückweise differenzierbaren Weges wie oben als*

$$L(\gamma) = \sum_{k=1}^{K} \int_{t_{k-1}}^{t_k} \|\gamma'(s)\| \, \mathrm{d}s.$$

*Wir behaupten, dass die Wegmetrik $\mathrm{d}_{\mathrm{Weg}}(x, y)$ für $x, y \in U$, welche durch*

$$\mathrm{d}_{\mathrm{Weg}}(x, y) = \inf\{L(\gamma) \mid \gamma \text{ ist ein stückweise differenzierbarer Weg von } x \text{ nach } y\}.$$

*definiert ist, tatsächlich eine Metrik ist und dass diese die übliche Topologie definiert.*

*(i) Sei $f : U \to \mathbb{R}^m$ stetig differenzierbar mit beschränkten Ableitungen. Zeigen Sie, dass $f$ Lipschitz-stetig ist, wenn man $U$ mit der Wegmetrik $\mathrm{d}_{\mathrm{Weg}}(x, y)$ ausstattet.*

*(ii) Finden Sie ein Beispiel einer zusammenhängenden, nicht konvexen Menge und einer differenzierbaren Funktion mit beschränkten Ableitungen, die bezüglich $\|\cdot\|$ nicht Lipschitz-stetig ist.*

VO: VO 14. März

## 10.3 Höhere Ableitungen und Taylor-Approximation

### 10.3.1 Definition und Eigenschaften der höheren partiellen Ableitungen

Auf Grund von Satz 10.10 über die Existenz der totalen Ableitung werden wir im Folgenden eigentlich immer die Stetigkeit der partiellen Ableitungen (das heisst, stetige Differenzierbarkeit) verlangen. Wie wir hier sehen werden, ist diese Annahme auch für „höhere Ableitungen“ von Bedeutung.

**Definition 10.19** (Höhere stetige Differenzierbarkeit)**.** Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$ eine Funktion. Wir sagen, dass $f$ **zweimal stetig differenzierbar** ist, falls $f$ stetig differenzierbar ist und für alle $j, k \in \{1, \dots, n\}$ die partielle Ableitung $\partial_k f : U \to \mathbb{R}^m$ wiederum eine stetige partielle Ableitung $\partial_j \partial_k f$ besitzt. Im Allgemeinen heisst $f$ **$d$-mal stetig differenzierbar** für ein $d \geq 2$, falls $f$ stetig differenzierbar ist und für jedes $j \in \{1, \dots, n\}$ die partielle Ableitung $(d-1)$-mal stetig differenzierbar ist. Weiter sei

$$C^d(U) = \{f : U \to \mathbb{R} \mid f \text{ ist } d\text{-mal stetig differenzierbar}\}$$

die Menge der $d$-mal stetig differenzierbaren reellwertigen Funktionen auf $U$. Wir sagen, dass eine iterierte partielle Ableitung einer $d$-mal stetig differenzierbaren Funktion $f : U \to \mathbb{R}$ **Ordnung** $\ell$ für $\ell \in \{1, \dots, d\}$ hat, falls genau $\ell$ partielle Ableitungen auf $f$ angewandt wurden. Des Weiteren nennt man die Funktion $f$ **glatt**, falls sie beliebig oft (also für alle $d \in \mathbb{N}$ $d$-mal) stetig differenzierbar ist.

Wir möchten an dieser Stelle anmerken, dass wir ab jetzt oft nur $\mathbb{R}$-wertige statt $\mathbb{R}^m$-wertige Funktionen betrachten werden. Viele der folgenden Aussagen übertragen sich allerdings wegen der Reduktionseigenschaft in Lemma 10.8 auch auf den Fall $m > 1$.

**Satz 10.20** (Satz von Schwarz)**.** *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine zweimal stetig differenzierbare Funktion. Dann gilt für alle $j, k \in \{1, \dots, n\}$*

$$\partial_j \partial_k f = \partial_k \partial_j f$$

*auf ganz $U$.*

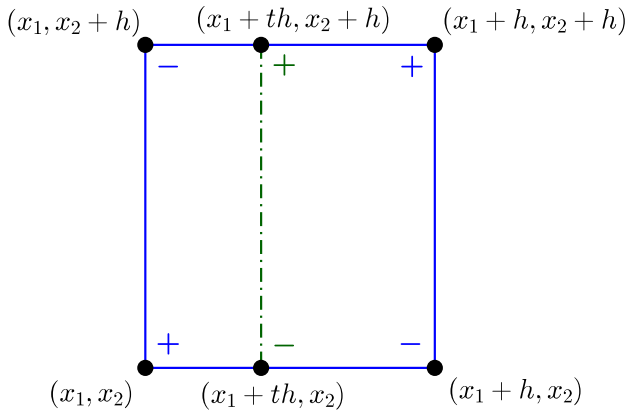
*Beweis.* Es genügt den Fall $n = 2$ und $j = 1$, $k = 2$ zu betrachten, der allgemeine Fall ist nur in der Notation schwieriger und folgt auch aus dem betrachteten Spezialfall. Für $x \in U$ und ein genügend kleines $h > 0$ (so dass $(x_1 + t_1 h, x_2 + t_2 h) \in U$ für alle $t_1, t_2 \in [0, 1]$) definieren wir eine Funktion $F$ durch

$$F(h) = f(x_1 + h, x_2 + h) - f(x_1 + h, x_2) - f(x_1, x_2 + h) + f(x_1, x_2).$$

Weiter betrachten wir für ein genügend kleines aber festes $h \in (0, 1)$ die nach der Kettenregel differenzierbare Funktion $t \in [0, 1] \mapsto \varphi(t) = f(x_1 + th, x_2 + h) - f(x_1 + th, x_2)$ und erhalten

$$F(h) = \varphi(1) - \varphi(0) = \varphi'(\xi_1) = \big(\partial_1 f(x_1 + \xi_1 h, x_2 + h) - \partial_1 f(x_1 + \xi_1 h, x_2)\big) h$$

für ein $\xi_1 \in (0,1)$ nach dem eindimensionalen Mittelwertsatz (Theorem 7.29) angewendet auf die Hilfsfunktion $\varphi$.




Figur 10.3: Die Funktion $h \mapsto F(h)$ ist eine vorzeichenbehaftete Summe von Funktionswerten von $f$ bei den Ecken eines Quadrats (hier gekennzeichnet durch eine durchgezogene Linie). Die Funktion $t \mapsto \varphi(t)$ entspricht der Differenz der Funktionswerte auf einer vertikalen Strecke durch das Quadrat.

Eine nochmalige Anwendungen des eindimensionalen Mittelwertsatzes auf die Funktion $\psi : t \in [0,1] \mapsto \partial_1 f(x_1 + \xi_1 h, x_2 + th)$ ergibt (gemeinsam mit der Kettenregel)

$$\begin{aligned} F(h) &= \big(\partial_1 f(x_1 + \xi_1 h, x_2 + h) - \partial_1 f(x_1 + \xi_1 h, x_2)\big)h \\ &= \partial_2 \partial_1 f(x_1 + \xi_1 h, x_2 + \xi_2 h)h^2 \end{aligned}$$

für einen Zwischenpunkt $\xi_2 \in (0,1)$.

Da aber in der Funktion $h \mapsto F(h)$ die beiden Komponenten symmetrisch verwendet wurden, können wir das Argument auch mit vertauschten Rollen der ersten und zweiten Komponenten durchführen. Dies ergibt analog gewisse $\xi_1', \xi_2' \in (0,1)$ mit

$$F(h) = \partial_1 \partial_2 f(x_1 + \xi_1' h, x_2 + \xi_2' h)h^2.$$

Wir dividieren nun durch $h^2 > 0$ und erhalten

$$\partial_2 \partial_1 f(x_1 + \xi_1 h, x_2 + \xi_2 h) = \partial_1 \partial_2 f(x_1 + \xi_1' h, x_2 + \xi_2' h).$$

Des Weiteren gilt wegen $\xi_1, \xi_2, \xi_1', \xi_2' \in (0,1)$, dass $(\xi_1 h, \xi_2 h)$ und $(\xi_1' h, \xi_2' h)$ beide gegen $(0,0)$ streben wenn $h \searrow 0$. Also folgt auf Grund der Stetigkeit beider partiellen Ableitungen $\partial_2 \partial_1 f(x) = \partial_1 \partial_2 f(x)$ wie gewünscht. $\square$

Wir bemerken, dass die Annahme der Stetigkeit im Satz von Schwarz notwendig ist – siehe die entsprechende Übung in Abschnitt 10.8.2.

Die **Hesse-Matrix** $H(x) = (H_{ij}(x))_{ij} \in \mathrm{Mat}_{n,n}(\mathbb{R})$ bei $x \in U$ einer zweimal stetig differenzierbaren Funktion $f : U \to \mathbb{R}$ ist gegeben durch

$$H_{ij}(x) = \partial_i \partial_j f(x)$$

für $i, j \in \{1, \ldots, n\}$. Der Satz von Schwarz (Satz 10.20) besagt nun genau $H_{ij}(x) = H_{ji}(x)$ für alle $i, j \in \{1, \ldots, n\}$, also dass $H(x)$ eine symmetrische Matrix ist.

Eine direkte Konsequenz und Verallgemeinerung des Satzes von Schwarz (Satz 10.20) ist das folgende Korollar.

**Korollar 10.21** (Satz von Schwarz)**.** *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$ $d$-mal stetig differenzierbar. Dann spielt die Reihenfolge der partiellen Ableitungen (bis zur Ordnung $d$) keine Rolle.*

Die höheren partiellen Ableitungen einer stetig differenzierbaren Funktion $f : U \to \mathbb{R}^m$ auf einer offenen Teilmenge $U \subseteq \mathbb{R}^n$ können also alle in die Form

$$\partial^{\boldsymbol{\alpha}} f = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} f$$

gebracht werden, wobei die einzelnen Komponenten von $\boldsymbol{\alpha} \in \mathbb{N}_0^n$ angeben wie oft wir nach den einzelnen Koordinatenrichtungen abgeleitet haben (und $\partial_j^0 f = f$ für alle $j = 1, \ldots, n$). Der Satz von Schwarz nimmt in dieser Notation die Form

$$\partial^{\boldsymbol{\alpha}} \partial^{\boldsymbol{\beta}} f = \partial^{\boldsymbol{\beta}} \partial^{\boldsymbol{\alpha}} f = \partial^{\boldsymbol{\alpha}+\boldsymbol{\beta}} f$$

für $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}_0^n$ an, wobei $f$ auf $U$ als $\|\boldsymbol{\alpha} + \boldsymbol{\beta}\|_1$-oft stetig differenzierbar vorausgesetzt wird. Wir bezeichnen in diesem Zusammenhang $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$ als einen **Multiindex**.

### 10.3.2 Mehrdimensionale Taylor-Approximation

**Satz 10.22** (Taylor-Approximation mit Integralrestglied)**.** *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine $(d+1)$-mal stetig differenzierbare Funktion. Sei $x \in U$ und $h \in \mathbb{R}^n$, so dass $x + th \in U$ für alle $t \in [0, 1]$. Dann gilt*

$$f(x+h) = f(x) + \sum_{k=1}^{d} \frac{1}{k!} \big(\partial_h^k f\big)(x) + R_{x,d}^f(h), \tag{10.5}$$

*wobei das Integralrestglied $R_{x,d}^f$ durch*

$$R_{x,d}^f(h) = \int_0^1 \frac{(1-t)^d}{d!} \big(\partial_h^{d+1} f\big)(x + th)\, \mathrm{d}t.$$

*gegeben ist. Insbesondere ist*

$$f(x+h) = f(x) + \sum_{k=1}^{d} \frac{1}{k!} \big(\partial_h^k f\big)(x) + O(\|h\|^{d+1}). \tag{10.6}$$

Dabei bezeichnet $\partial_h^k f$ die $k$-fache Ableitung von $f$ entlang des Vektors $h$. Wir erinnern daran, dass

$$\partial_h f = h_1 \partial_1 f + \ldots + h_n \partial_n f.$$

Auch die höheren Ableitungen $\partial_h^k f$ lassen sich als Linearkombinationen partieller Ableitungen der Ordnung $k$ auffassen, wenn man die Potenz formal ausmultipliziert. Zum Beispiel gilt für den quadratischen Term bei $x \in U$

$$\begin{aligned} \left[\partial_h^2 f\right](x) &= \left[\partial_h (h_1 \partial_1 f + \ldots + h_n \partial_n f)\right](x) \\ &= \sum_{j=1}^n h_j \partial_h (\partial_j f)(x) = \sum_{i,j=1}^n h_i h_j \partial_i \partial_j f(x) = h^t H(x) h \end{aligned} \tag{10.7}$$

für alle $h \in \mathbb{R}^n$, wobei $H(x)$ wieder die Hesse-Matrix der zweiten Ableitungen bei $x$ bezeichnet.

Wie im eindimensionalen Fall wollen wir die Approximation in Gleichung (10.5) (oder auch (10.6)) die **Taylor-Approximation $d$-ter Ordnung** nennen.

*Beweis.* Nach Annahme im Satz gilt $x + th \in U$ für alle $t \in [0,1]$ (oder sogar für $t$ in einem etwas grösseren offenen Intervall). Wir wenden nun die eindimensionale Taylor-Approximation auf die Funktion

$$\varphi : [0,1] \to \mathbb{R},\ t \mapsto f(x + th)$$

an. Nach Theorem 8.58 erhält man für die Taylor-Approximation um 0 bei 1

$$\varphi(1) = P_{0,d}^{\varphi}(1) + \int_0^1 \varphi^{(d+1)}(t) \frac{(1-t)^d}{d!}\, \mathrm{d}t, \tag{10.8}$$

wobei

$$P_{0,d}^{\varphi}(1) = \sum_{k=0}^d \frac{\varphi^{(k)}(0)}{k!}. \tag{10.9}$$

Wenden wir die Kettenregel in Satz 10.13 auf $\varphi$ an, so erhalten wir für $t \in [0,1]$

$$\varphi'(t) = \mathrm{D}_{x+th} f(h) = [\partial_h f](x + th) = h_1 \partial_1 f(x + th) + \ldots + h_n \partial_n f(x + th).$$

Für die zweite Ableitung von $\varphi$ nach $t \in [0,1]$ (für festes $x$ und $h$) ergibt sich ebenso

$$\varphi''(t) = \left[\partial_h^2 f\right](x + th).$$

Per Induktion erhält man schlussendlich

$$\varphi^{(k)}(t) = \left[\partial_h^k f\right](x + th),$$

für alle $k \in \{1, \dots, n\}$ und $t \in [0,1]$. Setzen wir dies in (10.8) und (10.9) ein, so ergibt sich der Satz.

Für die letzte Aussage sei $\varepsilon > 0$ mit $\overline{B_\varepsilon}(x_0) \subseteq U$ so dass alle partiellen Ableitungen der Ordnung $d+1$ auf $\overline{B_\varepsilon}(x_0)$ beschränkt sind. Expandieren wir die Notation $\big(\partial_h^{d+1} f\big)(x+th)$ so erhalten wir eine endliche Linearkombination der $(d+1)$-ten partiellen Ableitungen, die für $\|h\| \leq \varepsilon$ und $t \in [0,1]$ beschränkt sind, wobei die Koeffizienten ein Produkt von $d+1$ Koordinaten von $h = (h_1, \dots, h_n)^t$ sind. Da $|h_j| \leq \|h\|$ für $j = 1, \dots, n$, ergibt sich die behauptete Fehlerabschätzung durch diese endliche Summe und die Dreiecksungleichung für das Riemann-Integral. □

Um zu veranschaulichen, wieso Satz 10.22 gerade die mehrdimensionale Version von Theorem 8.58 ist, wollen wir diesen hier in Multiindexnotation darstellen. Wir setzen für $\boldsymbol{\alpha} \in \mathbb{N}_0^n$ und $h \in \mathbb{R}^n$

$$h^{\boldsymbol{\alpha}} = h_1^{\alpha_1} \cdots h_n^{\alpha_n}$$

sowie $\boldsymbol{\alpha}! = \alpha_1! \cdots \alpha_n!$. Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine $(d+1)$-mal stetig differenzierbare Funktion. Sei $x \in U$ und $h \in \mathbb{R}^n$, so dass $x + th \in U$ für alle $t \in [0,1]$. Dann gilt

$$f(x+h) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}_0^n : \|\boldsymbol{\alpha}\|_1 \leq d} \frac{1}{\boldsymbol{\alpha}!} \partial^{\boldsymbol{\alpha}} f(x) h^{\boldsymbol{\alpha}} + R_{x,d}^f(h) \tag{10.10}$$

wobei

$$R_{x,d}^f(h) = (d+1) \sum_{\boldsymbol{\alpha} \in \mathbb{N}_0^n : \|\boldsymbol{\alpha}\|_1 = d+1} h^{\boldsymbol{\alpha}} \int_0^1 \frac{(1-t)^d}{\boldsymbol{\alpha}!} \partial^{\boldsymbol{\alpha}} f(x+th) \, \mathrm{d}t. \tag{10.11}$$

Spricht man von Taylor-Approximation (insbesondere in der Literatur), so ist meistens die Form in (10.10), (10.11) anstelle von (10.5), (10.6) gemeint.

Wir bemerken an dieser Stelle ebenfalls, dass der Hauptterm auf der rechten Seite von (10.10) genau wie in der eindimensionalen Taylor-Approximation ein Polynom darstellt – diesmal allerdings in $d$ Variablen.

**Wichtige Übung 10.23** (Satz von Taylor in Multiindexnotation)**.** *Zeigen Sie obige Umformulierung der mehrdimensionalen Taylor-Approximation.*

**Applet 10.24** (Taylor-Approximation für Berglandschaft)**.** *Wir sehen anhand der Funktion* $f : (x,y) \in \mathbb{R}^2 \mapsto \sin(x)\cos(y) + 2$ *wie die Taylor-Approximationen erster, zweiter, oder dritter Ordnung die Funktion approximiert.*

Wir werden die mehrdimensionale Taylor-Approximation meist nur für $d = 1, 2$ wie im folgenden Korollar benötigen. VO: 16.03.18

**Korollar 10.25** (Lineare und quadratische Approximation)**.** *Sei* $U \subseteq \mathbb{R}^n$ *offen und* $f : U \to \mathbb{R}$ *eine zweimal stetig differenzierbare Funktion. Dann gilt für alle* $x \in U$

$$f(x+h) = f(x) + \langle \nabla f(x), h \rangle + O(\|h\|^2)$$

*und genauer*

$$f(x+h) = f(x) + \langle \nabla f(x), h\rangle + \tfrac{1}{2} h^t H(x) h + o(\|h\|^2)$$

*für $h \to 0$, wobei $H(x)$ wieder die Hesse-Matrix von $f$ bei $x$ darstellt.*

*Beweis.* Die erste Gleichung folgt direkt aus Satz 10.22. Für die zweite bemerken wir zuerst, dass das Restglied $R^f_{x,1}$ nach Satz 10.22 und (10.7) durch

$$R^f_{x,1}(h) = \int_0^1 (1-t) h^t H(x+th) h \,\mathrm{d}t$$

gegeben ist, wobei $H(x)$ wieder die Hesse-Matrix von $f$ bei $x \in U$ bezeichnet. Für $h \to 0$ unterscheidet sich wegen der Stetigkeit der zweiten Ableitungen $H(x+th)$ um $o(1)$ von $H(x)$ (und die implizite Konstante ist unabhängig von $t \in [0,1]$). Also gilt

$$\begin{aligned} R^f_{x,1}(h) &= \int_0^1 (1-t) h^t H(x+th) h \,\mathrm{d}t = \int_0^1 (1-t) h^t \big(H(x) + o(1)\big) h \,\mathrm{d}t \\ &= h^t H(x) h \int_0^1 (1-t)\,\mathrm{d}t + o(\|h\|^2) = \tfrac{1}{2} h^t H(x) h + o(\|h\|^2) \end{aligned}$$

für $h \to 0$. □

**Beispiel 10.26.** *Wir betrachten die Funktion*

$$f : (x,y)^t \in \mathbb{R}^2 \mapsto (x+1)\sin(x+y) - x$$

*und berechnen die ersten partiellen Ableitungen*

$$\begin{aligned} \partial_1 f(x,y) &= \sin(x+y) + (x+1)\cos(x+y) - 1, \\ \partial_2 f(x,y) &= (x+1)\cos(x+y) \end{aligned}$$

*sowie die zweiten partiellen Ableitungen*

$$\begin{aligned} \partial_1^2 f(x,y) &= 2\cos(x+y) - (x+1)\sin(x+y) \\ \partial_2^2 f(x,y) &= -(x+1)\sin(x+y) \\ \partial_2\partial_1 f(x,y) &= \partial_1\partial_2 f(x,y) = \cos(x+y) - (x+1)\sin(x+y) \end{aligned}$$

*für $(x,y)^t \in \mathbb{R}^2$. Damit gilt beispielsweise für die quadratische Taylor-Approximation von $f$ um $(0,0)^t$*

$$f(x,y) = y + x^2 + xy + o(x^2+y^2)$$

*für $(x,y)^t \to 0$.*

Glücklicherweise ist nicht immer notwendig alle partiellen Ableitungen zu berechnen. Stattdessen kann man auch auf bekannte Reihendarstellungen zurückgreifen. Wir möchten dies an einem Beispiel illustrieren.

**Beispiel 10.27** (Taylor via bekannter Reihendarstellung)**.** *Wir berechnen die Taylor-Approximation zwölfter Ordnung der Funktion*

$$f : (x, y)^t \in \mathbb{R}^2 \mapsto (x + y)\cos(x^2) \in \mathbb{R}$$

*um den Ursprung. Die Taylor-Reihe von* cos *um den Ursprung ist durch*

$$\cos(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$$

*gegeben, womit insbesondere*

$$\cos(x^2) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{4n}}{(2n)!} = 1 - \tfrac{1}{2}x^4 + \tfrac{1}{24}x^8 + O(x^{12})$$

*für alle* $x \in \mathbb{R}$*. Für die gegebene Funktion* $f$ *und* $h = (x, y)^t \to 0$ *ergibt sich also*

$$\begin{aligned} f(h) = (x + y)\cos(x^2) &= (x + y)\Big(1 - \tfrac{1}{2}x^4 + \tfrac{1}{24}x^8 + O(x^{12})\Big) \\ &= x + y - \tfrac{1}{2}x^5 - \tfrac{1}{2}x^4 y + \tfrac{1}{24}x^9 + \tfrac{1}{24}x^8 y + O(x^{12}(x + y)) \\ &= x + y - \tfrac{1}{2}x^5 - \tfrac{1}{2}x^4 y + \tfrac{1}{24}x^9 + \tfrac{1}{24}x^8 y + O(\|h\|^{13}), \end{aligned}$$

*wobei wir verwendet haben, dass* $x^{13} = O(\|h\|^{13})$ *sowie* $x^{12}y = O(\|h\|^{13})$*. Wir empfehlen den Leserinnen und Lesern an dieser Stelle, einige Ableitungen der Funktion* $f$ *zu berechnen und sich davon zu überzeugen, dass das Verwenden der Potenzreihe des Kosinus obige Rechnung erheblich verkürzt.*

## 10.4 Extremwerte

**Definition 10.28** (Extrema)**.** Sei $f$ eine reellwertige Funktion auf einer Menge $X$. Dann sagen wir, dass $f$ in $x_{\max} \in X$ ein **Maximum annimmt**, falls $f(x) \leq f(x_{\max})$ für alle $x \in X$ gilt. Die Funktion $f$ nimmt ein **striktes Maximum** in $x_{\max} \in X$ an, falls $f(x) < f(x_{\max})$ für alle $x \in X \setminus \{x_{\max}\}$ gilt. In beiden Fällen bezeichnen wir $f(x_{\max})$ als das **Maximum** von $f$. Analoge Begriffe definiert man für das **Minimum**. In beiden Fällen sprechen wir von **(globalen) Extremwerten**.

Sei nun $X$ ein metrischer Raum. Dann sagen wir, dass $f$ in $x_{\max} \in X$ ein **lokales Maximum annimmt**, falls es ein $\delta > 0$ gibt, so dass $f(x) \leq f(x_{\max})$ für alle $x \in B_\delta(x_{\max})$. Weiter nimmt $f$ in $x_{\max} \in X$ ein **striktes lokales Maximum** an, falls es ein $\delta > 0$ gibt, so dass $f(x) < f(x_{\max})$ für alle $x \in B_\delta(x_{\max}) \setminus \{x_{\max}\}$. In beiden Fällen wird $f(x_{\max})$ als **lokales Maximum** bezeichnet. Die Definition eines **lokalen Minimum** ist analog und beide werden als **lokale Extremwerte** bezeichnet.

In Satz 9.66 haben wir bereits gesehen, dass stetige Funktionen auf kompakten metrischen Räumen beide Extremwerte (also Maximum und Minimum) besitzen. Des Weiteren wissen wir wegen dem Satz von Heine-Borel (Satz 9.70), dass dies insbesondere für abgeschlossene und beschränkte Teilmengen in $\mathbb{R}^n$ anwendbar ist. Die Methoden dieses Kapitels sind aber eher für offene Teilmengen des $\mathbb{R}^n$ relevant. Für abgeschlossene Teilmengen des $\mathbb{R}^n$ mit „glatten Rändern“ werden wir die hier behandelten Methoden im nächsten Kapitel weiter verfeinern.

**Proposition 10.29** (Notwendige Bedingung für lokale Extrema)**.** *Sei $U \subseteq \mathbb{R}^n$ offen, sei $f : U \to \mathbb{R}$ eine Funktion und sei $x_0 \in U$ ein Punkt. Falls $f$ in $x_0$ ein lokales Extremum annimmt und $f$ in $x_0$ differenzierbar ist, so ist $\mathrm{D}_{x_0} f = 0$.*

Der Beweis dieser Proposition ist weitgehend analog zum Beweis im eindimensionalen Fall (siehe Proposition 7.17).

*Beweis.* Wir nehmen ohne Beschränkung der Allgemeinheit an, dass $f$ in $x_0$ ein lokales Maximum annimmt. Für alle $j \in \{1, \ldots, n\}$ und genügend kleine $h \in \mathbb{R}$ gilt nach Annahme

$$f(x_0 + he_j) - f(x_0) \leq 0.$$

Daher ist

$$\partial_j f(x_0) = \lim_{h \searrow 0} \frac{f(x_0 + he_j) - f(x_0)}{h} \leq 0$$

und gleichzeitig auch

$$\partial_j f(x_0) = \lim_{h \nearrow 0} \frac{f(x_0 + he_j) - f(x_0)}{h} \geq 0.$$

Da $j \in \{1, \ldots, n\}$ beliebig war, folgt $\mathrm{D}_{x_0} f = 0$ aus Proposition 10.6. □

Wie schon für Funktionen auf Intervallen in $\mathbb{R}$ (zum Beispiel für $x \in \mathbb{R} \mapsto x^3 \in \mathbb{R}$) ist das Verschwinden der Ableitung aber keine hinreichende Bedingung für das Vorliegen eines

Extremwerts. Trotzdem ist Proposition 10.29 in der Praxis sehr nützlich, um die Kandidaten für lokale Extrema aufzuspüren. In der Tat sind die Kandidaten jene Punkte, bei denen die Ableitung entweder nicht existiert oder verschwindet.

**Definition 10.30** (Kritische Punkte)**.** Sei $U \subseteq \mathbb{R}^n$ offen und sei $f : U \to \mathbb{R}$ eine differenzierbare Funktion. Ein Punkt $x \in U$ heisst **kritischer Punkt** von $f$, falls $\mathrm{D}_x f = 0$. Ist allgemeiner $f$ eine differenzierbare Abbildung von $U$ nach $\mathbb{R}^m$, so ist $x \in U$ ein kritischer Punkt, falls $\mathrm{D}_x f$ Rang kleiner als $\min(m, n)$ hat.

Weiter nennt man $x \in U$ einen **regulären Punkt** der Abbildung $f : U \to \mathbb{R}^m$, falls $x$ kein kritischer Punkt von $f$ ist. Das Bild eines kritischen Punktes unter $f$ nennt man auch einen **kritischen Wert**; Punkte in $\mathbb{R}^m$ im Komplement der kritischen Werte von $f$ heissen **reguläre Werte**.

Für die Untersuchung, ob bei einem kritischen Punkt ein lokales Extremum angenommen wird, benötigen wir weitere Begriffe aus der Linearen Algebra.

**Definition 10.31.** Sei $A \in \mathrm{Mat}_{n,n}(\mathbb{R})$ eine **symmetrische Matrix** (das heisst, $A^t = A$). Dann nennt man die Abbildung

$$Q_A : v \in \mathbb{R}^n \mapsto v^t A v$$

die **zu $A$ assoziierte quadratische Form** in $n$ Variablen. Die quadratische Form $Q_A$ oder auch die Matrix $A$ heisst

- **positiv definit**, falls $Q_A(v) > 0$ für alle $v \in \mathbb{R}^n \setminus \{0\}$,
- **negativ definit**, falls $Q_A(v) < 0$ für alle $v \in \mathbb{R}^n \setminus \{0\}$,
- **indefinit**, falls $w_-, w_+ \in \mathbb{R}^n$ existieren mit $Q_A(w_+) > 0$ und $Q_A(w_-) < 0$, und
- **nicht-degeneriert**, falls $\det(A) \neq 0$.

*Bemerkung* (Zwei weitere Begriffe zu quadratischen Formen)*.* Sei $A$ eine symmetrische Matrix und $Q_A$ die assoziierte quadratische Form. Nebst den oben eingeführten Begriffen zu $Q_A$ existieren weitere wichtige Begriffe, die wir hier aber nicht verwenden werden. Die quadratische Form $Q_A$ nennt sich positiv semidefinit, falls $Q_A(v) \geq 0$ für alle $v \in \mathbb{R}^n$, und negativ semidefinit, falls $Q_A(v) \leq 0$ für alle $v \in \mathbb{R}^n$.

Der Begriff der Definitheit erlaubt es uns nun wie in Korollar 7.37 zu entscheiden, ob bei einem kritischen Punkt ein lokales Maximum angenommen wird oder nicht.

**Korollar 10.32.** *Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ zweimal stetig differenzierbar, $x_0 \in U$ ein kritischer Punkt und*

$$Q(h) = \sum_{i,j=1}^{n} \partial_i \partial_j f(x_0) h_i h_j$$

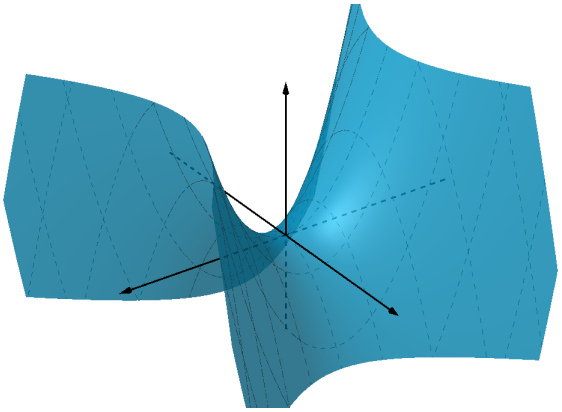
*die quadratische Form assoziiert zur Hesse-Matrix $H(x)$ von $f$ bei $x_0$. Dann gilt*

- *Ist $Q$ positiv definit, so nimmt $f$ bei $x_0$ ein striktes lokales Minimum an.*
- *Ist $Q$ negativ definit, so nimmt $f$ bei $x_0$ ein striktes lokales Maximum an.*
- *Ist $Q$ indefinit, so hat $f$ bei $x_0$ kein lokales Extremum.*

Um sich die obigen Aussagen merken zu können, empfiehlt sich die folgenden einfachen Beispiele im Gedächnis zu behalten.

- $f(x, y) = x^2 + y^2$ hat ein lokales Minimum bei 0.
- $f(x, y) = -x^2 - y^2$ hat ein lokales Maximum bei 0.
- $f(x, y) = x^2 - y^2$ hat kein lokales Extremum bei 0. Allerdings ist 0 ein kritischer Punkt von $f$.

In dem indefiniten Fall spricht man auch von einem **Sattelpunkt**, siehe folgendes Bild.



*Beweis.* Nach Korollar 10.25 gilt

$$f(x_0 + h) - f(x_0) = \tfrac{1}{2}\|h\|^2 \left( Q\left(\frac{h}{\|h\|}\right) + \alpha(x_0, h) \right) \tag{10.12}$$

für $\alpha(x_0, h) = o(1)$ für $h \to 0$. Falls $Q$ positiv definit ist, dann gilt $Q(w) > 0$ für alle $w \in \mathbb{S}^{n-1} = \{v \in \mathbb{R}^n \mid \|v\| = 1\}$. Da $\mathbb{S}^{n-1}$ nach dem Satz von Heine-Borel (Satz 9.70) kompakt ist und $Q$ stetig ist, existiert daher ein $c > 0$ mit $Q(w) \geq c$ für alle $w \in \mathbb{S}^{n-1}$ (siehe Satz 9.66(5)). Es existiert weiter ein $\delta > 0$, so dass der Fehlerterm $\alpha(x_0, h)$ in (10.12) im Absolutbetrag kleiner als $\frac{c}{2}$ ist für $h \in \mathbb{R}^n$ mit $\|h\| < \delta$. Es folgt daher mit (10.12), dass

$$f(x_0 + h) - f(x_0) \geq \tfrac{1}{2}\|h\|^2 \left( Q\left(\frac{h}{\|h\|}\right) - \tfrac{c}{2} \right) \geq \tfrac{c}{4}\|h\|^2 > 0$$

für alle $h \in B_\delta(0)$ gilt, wodurch $f$ in $x_0$ ein striktes lokales Minimum annimmt.

Falls $Q$ negativ definit ist, so ersetzen wir $f$ durch $-f$, womit $Q$ durch $-Q$ ersetzt wird. Die quadratische Form $-Q$ ist aber positiv definit und somit nimmt $-f$ in $x_0$ ein striktes lokales Minimum an, was die Aussage beweist.

Falls $Q$ indefinit ist, so existieren $w_-, w_+ \in \mathbb{S}^{n-1}$, so dass $Q(w_-) < 0$ und $Q(w_+) > 0$. Für hinreichend kleine $s \in \mathbb{R} \setminus \{0\}$ ist dann

$$|\alpha(x_0, sw_-)| < \frac{Q(w_-)}{2}, \quad |\alpha(x_0, sw_+)| < \frac{Q(w_+)}{2}$$

und damit

$$\begin{aligned} f(x_0 + sw_-) - f(x_0) &< \tfrac{1}{2}s^2 \left( Q(w_-) + \frac{|Q(w_-)|}{2} \right) < 0 \\ f(x_0 + sw_+) - f(x_0) &> \tfrac{1}{2}s^2 \left( Q(w_+) - \frac{Q(w_+)}{2} \right) > 0 \end{aligned}$$

Daher nimmt $f$ bei $x_0$ weder ein lokales Minimum noch ein lokales Maximum an. □

Auf Grund von Korollar 10.32 sind wir daran interessiert, für eine gegebene Matrix entscheiden zu können, ob sie positiv definit, negativ definit oder indefinit ist. Folgendes Kriterium aus der Linearen Algebra ist dafür sehr nützlich.

**Satz 10.33** (Charakterisierungen von Definitheit)**.** *Sei $A = (a_{ij})_{ij} \in \mathrm{Mat}_{n,n}(\mathbb{R})$ eine symmetrische Matrix. Dann gilt*

- *$A$ ist genau dann positiv definit, wenn alle der folgenden Determinanten positiv sind:*

$$a_{11}, \quad \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \quad \dots \quad , \quad \det(A).$$

- *$A$ ist genau dann negativ definit, wenn $-A$ positiv definit ist, was genau wechselnden Vorzeichen der Determinanten beginnend mit negativen Vorzeichen entspricht.*
- *Falls $A$ nicht-degeneriert ist und weder positiv noch negativ definit ist, dann ist $A$ indefinit.*

Der Beweis dieses Satzes verwendet nur Methoden der Linearen Algebra und wird deswegen in den nächsten Teilabschnitt ausgelagert. Selbstverständlich existieren weitere, nützliche Charakterisierungen von Definitheit (beispielsweise via der Eigenwerte).

**Beispiel 10.34.** *Seien $a, b \in \mathbb{R}$ feste Parameter. Wir definieren $f : \mathbb{R}^2 \to \mathbb{R}$ durch*

$$f(x, y) = x \sin(y) + ax^2 + by^2$$

*für $(x, y)^t \in \mathbb{R}^2$ und betrachten den kritischen Punkt $(0, 0)^t$. Die Hesse-Matrix von $f$ bei $(0, 0)^t$ ist durch*

$$H = \begin{pmatrix} 2a & 1 \\ 1 & 2b \end{pmatrix}$$

*gegeben. Wir wenden Satz 10.33 und das Kriterium in Korollar 10.32 an und erhalten folgende Fälle.*

- *Falls $a > 0$ ist und $4ab - 1 > 0$ ist, so ist $H$ positiv definit und $f$ hat bei $(0,0)^t$ ein lokales Minimum.*
- *Falls $a < 0$ ist und $4ab - 1 > 0$ ist, so ist $H$ negativ definit und $f$ hat bei $(0,0)^t$ ein lokales Maximum.*
- *Falls $4ab - 1 = 0$ ist, so ist die Hesse-Matrix degeneriert und unsere Kriterien greifen nicht (was nicht heisst, dass man diesen Fall nicht trotzdem entscheiden kann).*
- *Falls $4ab - 1 < 0$ ist, so ist die Hesse-Matrix indefinit und $(0,0)^t$ ist ein Sattelpunkt.*

**Übung 10.35.** *Finden Sie alle kritischen Punkte der Funktion $f : (x,y) \in \mathbb{R}^2 \mapsto x^3 - y^3 + 3\alpha xy$ zu $\alpha \in \mathbb{R}$. Entscheiden Sie jeweils, ob es sich um ein Extremum handelt und wenn ja, ob ein lokales Minimum oder Maximum angenommen wird.*

VO: 19.03.18

### 10.4.1 Beweis des Kriteriums für Definitheit*

Wie angekündigt beweisen wir hier Satz 10.33.

*Beweis von Satz 10.33.* Wir bemerken zuerst, dass für $J \in \mathrm{GL}_n(\mathbb{R})$ die Matrix $A$ genau dann positiv definit (negativ definit oder indefinit) ist, wenn dies für $J^t A J$ der Fall ist.

Der Beweis der ersten Aussage erfolgt per Induktion nach $n$. Für $n = 1$ folgt die Behauptung direkt aus der Definition. Für den Beweis des Induktionsschrittes schreiben wir $A \in \mathrm{Mat}_{n+1,n+1}(\mathbb{R})$ als die Blockmatrix

$$A = \begin{pmatrix} B & v \\ v^t & c \end{pmatrix}$$

für eine symmetrische Matrix $B \in \mathrm{Mat}_{n,n}(\mathbb{R})$, $v \in \mathbb{R}^n$ und $c \in \mathbb{R}$. Falls die Matrix $B$ invertierbar ist, dann gilt mit

$$J = \begin{pmatrix} I_n & -B^{-1}v \\ 0 & 1 \end{pmatrix} \quad \text{und } J^t = \begin{pmatrix} I_n & 0 \\ -v^t B^{-1} & 1 \end{pmatrix},$$

dass

$$\begin{aligned} J^t A J &= \begin{pmatrix} I_n & 0 \\ -v^t B^{-1} & 1 \end{pmatrix} \begin{pmatrix} B & v \\ v^t & c \end{pmatrix} \begin{pmatrix} I_n & -B^{-1}v \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} I_n & 0 \\ -v^t B^{-1} & 1 \end{pmatrix} \begin{pmatrix} B & -BB^{-1}v + v \\ v^t & -v^t B^{-1} v + c \end{pmatrix} = \begin{pmatrix} B & 0 \\ 0 & \tilde{c} \end{pmatrix}, \end{aligned}$$

wobei wir $\tilde{c} = -v^t B^{-1} v + c$ gesetzt haben.

Falls nun $A$ positiv definit ist, dann ist auch $(w^t, 0) A \binom{w}{0} > 0$ für alle $w \in \mathbb{R}^n \setminus \{0\}$. In obiger Notation folgt daraus, dass $B$ ebenfalls positiv definit ist. Gemeinsam mit der Induktionsannahme erhalten wir, dass die ersten $n$ Determinanten positiv sind und insbesondere $B \in \mathrm{GL}_n(\mathbb{R})$ ist. Wenden wir nun obige Rechnung an, so erhalten wir also eine Matrix

$J \in \mathrm{GL}_{n+1}(\mathbb{R})$, so dass

$$J^t A J = \begin{pmatrix} B & 0 \\ 0 & \tilde{c} \end{pmatrix}.$$

Da $A$ als positiv definit vorausgesetzt wurde, gilt $\tilde{c} > 0$ und wegen $\det(J) = 1$ auch

$$\det(A) = \det(J^t A J) = \det(B)\tilde{c} > 0.$$

Dies beweist den Induktionsschritt in der ersten Richtung.

Sei nun $A \in \mathrm{Mat}_{n+1,n+1}(\mathbb{R})$ eine symmetrische Matrix, so dass alle Determinanten wie im Satz positiv sind. Insbesondere hat $B$ eine positive Determinante und ist auf Grund der Induktionssannahme positiv definit. Wir verwenden wieder die oben definierte Matrix $J$ und sehen, dass

$$J^t A J = \begin{pmatrix} B & 0 \\ 0 & \tilde{c} \end{pmatrix} \quad \text{und} \quad \det(A) = \det(B)\tilde{c} > 0$$

und somit $\tilde{c} > 0$. Daraus folgt aber, dass

$$\begin{pmatrix} B & 0 \\ 0 & \tilde{c} \end{pmatrix}$$

und damit auch $A$ positiv definit sind. Dies vollendet den induktiven Beweis der ersten Aussage im Satz.

Für die zweite Aussage verwenden wir, dass $A$ genau dann negativ definit ist, wenn $-A$ positiv definit ist (was direkt aus der Definition folgt). Gemeinsam mit der Multilinearität der Determinante und der ersten Bedingung ergibt sich die gewünschte Charakterisierung mittels der Folge der Determinanten.

Die letzte Behauptung ist keine Charakterisierung, sondern nur eine hinreichende Bedingung. Ihr Beweis ist etwas anders aufgebaut und verwendet folgenden Satz aus der linearen Algebra: Jede symmetrische Matrix $A$ ist diagonalisierbar, wobei es sogar eine orthogonale Matrix $K$ gibt für die $K^{-1}AK$ diagonal ist. Für die orthogonale Matrix $K$ ist aber $K^{-1} = K^t$ und wie schon zuvor haben dadurch $A$ und die Diagonalmatrix $D = K^t A K$ das gleiche Verhalten bezüglich Definitheit. Nach Vorraussetzung ist $A$ nicht-degeneriert, womit alle Eigenwerte von $A$ (also die Diagonaleinträge von $D$) nicht gleich Null sind. Da $A$ nicht positiv definit ist, ist auch $D$ nicht positiv definit und es existiert ein negativer Eintrag in $D$. Dies gilt analog für nicht negativ definit, und zusammen sehen wir, dass sowohl $D$ als auch $A$ indefinit sind. $\square$

VO: 19.03.2018

## 10.5 Parameterintegrale

Seien $a < b$ reelle Zahlen. Ein Integral der Form

$$\int_a^b f(x,t)\,\mathrm{d}t,$$

welches von einer oder mehreren Variablen $x$ abhängt, wird als **Parameterintegral** bezeichnet.

**Satz 10.36** (Differentiation unter dem Integral)**.** *Sei $U \subseteq \mathbb{R}^n$ eine offene Teilmenge, $a < b$ reelle Zahlen und $f : U \times [a,b] \to \mathbb{R}$ stetig. Dann definiert das Parameterintegral*

$$F(x) = \int_a^b f(x,t)\,\mathrm{d}t$$

*für $x \in U$ eine stetige Funktion $F : U \to \mathbb{R}$. Falls zusätzlich die partiellen Ableitungen $\partial_k f$ für $k = 1, \ldots, n$ existieren und auf ganz $U \times [a,b]$ stetig sind, dann ist $F$ stetig differenzierbar und es gilt*

$$\partial_k F(x) = \int_a^b \partial_k f(x,t)\,\mathrm{d}t$$

*für alle $x \in U$ und $k \in \{1, \ldots, n\}$.*

*Beweis.* Man beachte zuerst, dass auf Grund der Stetigkeit von $f$ die Abbildung $t \in [a,b] \mapsto f(x,t)$ für jedes $x \in U$ stetig und somit Riemann-integrierbar ist (Satz 4.42).

Sei nun $x_0 \in U$ und $\eta > 0$, so dass $K = \overline{B_\eta(x_0)} \subseteq U$. Nach dem Satz von Heine-Borel (Satz 9.70) ist $K \times [a,b]$ kompakt und $f|_{K\times[a,b]}$ ist gleichmässig stetig nach Proposition 9.77. Sei also $\varepsilon > 0$. Dann existiert ein $\delta \in (0,\eta)$, so dass für alle $x \in B_\delta(x_0)$ und $t \in [a,b]$ die Abschätzung

$$|f(x,t) - f(x_0,t)| < \varepsilon$$

gilt. Dies impliziert

$$|F(x) - F(x_0)| \leq \int_a^b |f(x,t) - f(x_0,t)|\,\mathrm{d}t < \varepsilon(b-a)$$

für alle $x \in B_\delta(x_0)$ und beweist Stetigkeit von $F$ bei $x_0$.

Sei nun $k \in \{1, \ldots, n\}$ und angenommen $f$ besitzt die stetige partielle Ableitung $\partial_k f$. Seien $x_0 \in U$ und $K = \overline{B_\eta(x_0)} \subseteq U$ wie oben. Für $s \in (-\eta,\eta) \setminus \{0\}$ und $t \in [a,b]$ existiert nach dem Mittelwertsatz (Satz 10.14) ein $\xi_{t,s} \in (0,1)$ mit

$$\frac{f(x_0 + se_k, t) - f(x_0,t)}{s} = \partial_k f(x_0 + \xi_{t,s} s e_k, t).$$

Wir wählen für ein $\varepsilon > 0$ mittels der gleichmässigen Stetigkeit von $\partial_k f$ auf $K \times [a,b]$ ein $\delta \in (0,\eta)$, so dass $x \in B_\delta(x_0)$ die Abschätzung

$$|\partial_k f(x,t) - \partial_k f(x_0,t)| < \varepsilon$$

impliziert. Gemeinsam ergibt sich nun für $s \in (-\delta,\delta) \setminus \{0\}$

$$\begin{aligned}\left|\frac{F(x_0+se_k)-F(x_0)}{s} - \int_a^b \partial_k f(x_0,t)\,\mathrm{d}t\right| &= \left|\int_a^b \left(\frac{f(x_0+se_k,t)-f(x_0,t)}{s} - \partial_k f(x_0,t)\right)\,\mathrm{d}t\right| \\ &= \left|\int_a^b \left(\partial_k f(x_0+\xi_{t,s}se_k,t) - \partial_k f(x_0,t)\right)\,\mathrm{d}t\right| \\ &\leq \varepsilon(b-a).\end{aligned}$$

Da $\varepsilon > 0$ beliebig war, folgt

$$\partial_k F(x_0) = \lim_{s\to 0} \frac{F(x_0+se_k)-F(x_0)}{s} = \int_a^b \partial_k f(x_0,t)\,\mathrm{d}t.$$

Nach dem ersten Teil des Satzes ist $\partial_k F$ stetig und da $k \in \{1,\dots,n\}$ beliebig war, folgt stetige Differenzierbarkeit von $F$ aus Satz 10.10. □

**Beispiel 10.37** (Umfang der Ellipse)**.** *Satz 10.36 erlaubt uns insbesondere, Funktionen zu analysieren, die nur mittels Integralen gegeben sind. Ein Beispiel einer solchen Funktion wollen wir hier finden, indem wir den Umfang der Ellipse berechnen. Seien $a,b > 0$ und (ohne Beschränkung der Allgemeinheit) $a \geq b$. Die Ellipse mit Parametern $a,b$ ist dann gegeben durch die Lösungsmenge der Gleichung*

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

*Eine mögliche Parametrisierung der Ellipse ist somit $\gamma : t \in [0,2\pi] \mapsto (a\cos(t), b\sin(t))$. Der Umfang der Ellipse ist also (siehe Abschnitt 8.3.2)*

$$\begin{aligned}L(\gamma) &= \int_0^{2\pi} \sqrt{\dot\gamma_1(t)^2 + \dot\gamma_2(t)^2}\,\mathrm{d}t = \int_0^{2\pi} \sqrt{a^2\sin^2(t) + b^2\cos^2(t)}\,\mathrm{d}t \\ &= a\int_0^{2\pi} \sqrt{\sin^2(t) + (\tfrac{b}{a})^2\cos^2(t)}\,\mathrm{d}t = a\int_0^{2\pi} \sqrt{1-\varepsilon^2\cos^2(t)}\,\mathrm{d}t \\ &= 4a\int_0^{\pi/2} \sqrt{1-\varepsilon^2\sin^2(t)}\,\mathrm{d}t,\end{aligned}$$

*wobei $\varepsilon = \sqrt{1-\frac{b^2}{a^2}}$ die Exzentrizität der Ellipse bezeichnet, welche gewissermassen die Abweichung der Ellipse von einem Kreis misst. Das Parameterintegral*

$$\varepsilon \in [0,1] \mapsto \int_0^{\pi/2} \sqrt{1-\varepsilon^2\sin^2(t)}\,\mathrm{d}t$$

*nennt sich das vollständige elliptische Integral zweiter Art.*

**Korollar 10.38.** *Sei $U \subseteq \mathbb{R}^n$, seien $a < b$ reelle Zahlen und sei $f : U \times (a,b) \to \mathbb{R}$ stetig mit stetigen partiellen Ableitungen $\partial_k f$ für $k \in \{1, \dots, n\}$. Seien des Weiteren $\alpha, \beta : U \to (a,b)$ stetig differenzierbar. Dann ist das Parameterintegral mit veränderlichen Grenzen*

$$F : x \in U \mapsto \int_{\alpha(x)}^{\beta(x)} f(x,t)\,\mathrm{d}t$$

*stetig differenzierbar und für $k \in \{1, \dots, n\}$ gilt*

$$\partial_k F(x) = f(x, \beta(x))\,\partial_k \beta(x) - f(x, \alpha(x))\,\partial_k \alpha(x) + \int_{\alpha(x)}^{\beta(x)} \partial_k f(x,t)\,\mathrm{d}t$$

*für alle $x \in U$.*

*Beweis.* Wir kombinieren Satz 10.36, den Fundamentalsatz der Integral- und Differentialrechnung (Theorem 8.2) und die mehrdimensionale Kettenregel in Satz 10.13. Dazu definieren wir die Hilfsfunktion

$$\phi : U \times (a,b)^2 \to \mathbb{R}, \quad (x, \alpha, \beta) \mapsto \int_\alpha^\beta f(x,t)\,\mathrm{d}t.$$

Wir zeigen zuerst, dass $\phi$ stetig ist. Sei also $(x_n, \alpha_n, \beta_n) \in U \times (a,b)^2$ eine Folge, die gegen $(x, \alpha, \beta) \in U \times (a,b)^2$ konvergiert. Wir wählen ein $\varepsilon > 0$ so dass $\overline{B_\varepsilon(x)} \subseteq U$ und definieren

$$\begin{aligned} c &= \inf \{\alpha_n, \beta_n \mid n \in \mathbb{N}\} \\ d &= \sup \{\alpha_n, \beta_n \mid n \in \mathbb{N}\}. \end{aligned}$$

Da sowohl $\alpha = \lim_{n\to\infty} \alpha_n \in (a,b)$ und $\beta = \lim_{n\to\infty} \beta_n \in (a,b)$ folgt $c, d \in (a,b)$. (Wieso?) Damit ist $K = \overline{B_\varepsilon(x)} \times [c,d] \subseteq U \times (a,b)^2$ eine kompakte Teilmenge und

$$M = \max_{(x',\alpha',\beta') \in K} |f(x', \alpha', \beta')|$$

existiert. Für alle hinreichend grossen $n$ gilt dann aber $x_n \in B_\varepsilon(x)$ und es folgt

$$\begin{aligned} &\Big|\phi(x_n, \alpha_n, \beta_n) - \phi(x, \alpha, \beta)\Big| \\ &\quad \le \Big|\int_{\alpha_n}^{\beta_n} f(x_n,t)\,\mathrm{d}t - \int_\alpha^\beta f(x_n,t)\,\mathrm{d}t\Big| + \Big|\int_\alpha^\beta f(x_n,t)\,\mathrm{d}t - \int_\alpha^\beta f(x,t)\,\mathrm{d}t\Big| \\ &\quad \le M|\alpha_n - \alpha| + M|\beta_n - \beta| + \Big|\int_\alpha^\beta f(x_n,t)\,\mathrm{d}t - \int_\alpha^\beta f(x,t)\,\mathrm{d}t\Big|, \end{aligned}$$

wobei wir die Dreiecksungleichung für das Integral (Satz 4.24(iii)) über die Teilintervalle zwischen $\alpha_n$ und $\alpha$ (beziehungsweise $\beta_n$ und $\beta$) und die Schranke $M$ für die Funktionswerte von $f$ verwendet haben. Für $n \to \infty$ folgt nun aus Satz 10.36, dass dieser Ausdruck gegen 0 strebt. Da die Folge $(x_n, \alpha_n, \beta_n) \in U \times (a,b)^2$ eine beliebige Folge mit beliebigem Grenzwert $(x, \alpha, \beta) \in U \times (a,b)^2$ war, erhalten wir, dass $\phi$ stetig ist (Proposition 9.37).

Des Weiteren gilt nach Satz 10.36, dass die partiellen Ableitungen $\partial_k \phi$ für $k \in \{1, \ldots, n\}$ existieren und durch

$$\partial_k \phi(x, \alpha, \beta) = \int_\alpha^\beta \partial_k f(x, t) \, \mathrm{d}t$$

für alle $(x, \alpha, \beta) \in U \times (a, b)^2$ gegeben sind. Nach obigem Argument ist $\partial_k \phi$ ebenso stetig. Nach Theorem 8.2 existieren auch die partiellen Ableitungen von $\phi$ nach $\alpha$ und $\beta$ und sind gegeben durch

$$\begin{aligned} \partial_\alpha \phi(x, \alpha, \beta) &= -f(x, \alpha) \\ \partial_\beta \phi(x, \alpha, \beta) &= f(x, \beta) \end{aligned}$$

für $(x, \alpha, \beta) \in U \times (a, b)^2$. Insbesondere sind $\partial_\alpha \phi, \partial_\beta \phi$ wiederum stetig nach Annahme. Nach Satz 10.10 ist $\phi$ also (stetig) differenzierbar.

Wir bemerken nun, dass die Funktion $F$ im Korollar

$$F : x \in U \mapsto \phi \begin{pmatrix} x \\ \alpha(x) \\ \beta(x) \end{pmatrix}$$

erfüllt und ist somit gegeben als Verknüpfung der Funktion

$$\psi : x \in U \mapsto \begin{pmatrix} x \\ \alpha(x) \\ \beta(x) \end{pmatrix}$$

mit der stetig differenzierbaren Funktion $\phi$. Nach Annahme im Korollar ist auch $\psi$ stetig differenzierbar und hat die totale Ableitung

$$\begin{pmatrix} I_n \\ D_x \alpha \\ D_x \beta \end{pmatrix}$$

bei $x \in U$. Wir können also die Kettenregel anwenden und erhalten, dass $F$ stetig differenzierbar ist und bei $x \in U$ und $y = (x, \alpha(x), \beta(x))^t$ gilt

$$\mathrm{D}_x F = \mathrm{D}_y \phi \circ \begin{pmatrix} I_n \\ D_x \alpha \\ D_x \beta \end{pmatrix}$$

beziehungsweise

$$\begin{aligned}\partial_k F = (\mathrm{D}_x F)e_k = \mathrm{D}_y\phi \begin{pmatrix} e_k \\ (\mathrm{D}_x\alpha)e_k \\ (\mathrm{D}_x\beta)e_k \end{pmatrix} &= \mathrm{D}_y\phi \begin{pmatrix} e_k \\ \partial_k\alpha(x) \\ \partial_k\beta(x) \end{pmatrix} \\ &= \partial_k\phi(y) - f(x,\alpha(x))\,\partial_k\alpha(x) + f(x,\beta(x))\,\partial_k\beta(x)\end{aligned}$$

für $k \in \{1,\dots,n\}$ wie gewünscht. □

### 10.5.1 Die Bessel-Differentialgleichung*

Als Anwendung der obigen, allgemeinen Theorie zu Parameterintegralen möchten wir diese hier verwenden, um eine Differentialgleichung zu lösen. Für einen Parameter $n \geq 0$ heisst die Differentialgleichung

$$x^2 J_n''(x) + xJ_n'(x) + (x^2 - n^2)J_n(x) = 0 \tag{10.13}$$

auf $(0,\infty)$ die Bessel-Differentialgleichung. Diese ist linear, homogen von zweiter Ordnung und tritt in mehreren Anwendungen innerhalb und ausserhalb der Mathematik auf.

Aus dem Existenz- und Eindeutigkeitssatz von Picard-Lindelöf (welchen wir gegen Ende des Semesters beweisen werden) folgt, dass (10.13) gemeinsam mit zwei beliebigen Anfangswerten $J_n(x_0) = a$ und $J_n'(x_0) = b$ für $x_0 > 0$ und $a, b \in \mathbb{R}$ eine eindeutig bestimmte Lösung auf $(0,\infty)$ besitzt. Wir wollen hier ein Fundamentalsystem der Lösungen beschreiben oder in anderen Worten zwei linear unabhängige Lösungen angeben, mit denen sich alle weiteren Lösungen als Linearkombinationen ausdrücken lassen. Hierfür nehmen wir $n \in \mathbb{N}_0$ an. Obwohl die Bedeutung dieser Annahmen in (10.13) unklar ist, werden die von uns betrachteten Lösungen diese Annahme benützen.

Das Parameterintegral

$$J_n(x) = \frac{1}{\pi}\int_0^\pi \cos(x\sin(t) - nt)\,\mathrm{d}t \tag{10.14}$$

wird **Bessel-Funktion erster Gattung** genannt und löst die Differentialgleichung (10.13), was wir mit Hilfe von Differentiation unter dem Integral (Satz 10.36) nachrechnen können. In der Tat gilt für beliebige $t \in [0,\pi]$

$$\partial_x(\cos(x\sin(t) - nt)) = -\sin(x\sin(t) - nt)\sin(t)$$

und daher

$$J_n'(x) = \frac{1}{\pi}\int_0^\pi \big(-\sin(x\sin(t) - nt)\sin(t)\big)\,\mathrm{d}t$$

und analog

$$J_n''(x) = \frac{1}{\pi}\int_0^\pi \big(-\cos(x\sin(t)-nt)\sin^2(t)\big)\,\mathrm{d}t.$$

Für den Ausdruck $x^2 J_n''(x) + (x^2-n^2)J_n(x)$ erhalten wir daraus

$$\begin{aligned}
&\frac{1}{\pi}\int_0^\pi \big(-x^2\cos(x\sin(t)-nt)\sin^2(t) + (x^2-n^2)\cos(x\sin(t)-nt)\big)\,\mathrm{d}t \\
&\qquad = \frac{1}{\pi}\int_0^\pi \cos(x\sin(t)-nt)\big(x^2\cos^2(t)-n^2\big)\,\mathrm{d}t \\
&\qquad = \frac{1}{\pi}\int_0^\pi \cos(x\sin(t)-nt)\underbrace{\big(x\cos(t)-n\big)}_{\partial_t(x\sin(t)-nt)}\big(x\cos(t)+n\big)\,\mathrm{d}t \\
&\qquad = \frac{1}{\pi}\sin(x\sin(t)-nt)(x\cos(t)+n)\Big]_0^\pi \\
&\qquad\qquad + \frac{1}{\pi}\int_0^\pi \sin(x\sin(t)-nt)x\sin(t)\,\mathrm{d}t \\
&\qquad = -xJ_n'(x)
\end{aligned}$$

auf Grund von partieller Integration und der Annahme $n \in \mathbb{N}_0$. Daher erfüllt (10.14) die Differentialgleichung (10.13).

Die **Bessel-Funktion zweiter Gattung** ist durch das uneigentliche Integral

$$Y_n(x) = \frac{1}{\pi}\int_0^\pi \sin(x\sin(t)-nt)x\sin(t)\,\mathrm{d}t - \frac{1}{\pi}\int_0^\infty \big(e^t + (-1)^n e^{-nt}\big)\mathrm{e}^{-x\sinh(t)}\,\mathrm{d}t$$

für $x \in (0,\infty)$ definiert. Wir möchten kurz annehmen, dass $Y_n$ die Differentialgleichung (10.13) ebenfalls löst. Dann kann man die gesuchte Lösung eines Anfangswertproblems zu (10.13) in Anwendungen oft bereits mit nur einem Anfangswert $f(x_0) = a$ bestimmen. Denn falls bekannt ist, dass $f$ auf dem Intervall $(0, x_0]$ beschränkt ist, so muss $f$ ein Vielfaches der Bessel-Funktion $J_n$ der ersten Gattung sein. In der Tat gilt

$$\lim_{x\searrow 0} J_n(x) = \frac{1}{\pi}\int_0^\pi \cos(nt)\,\mathrm{d}t$$

nach Satz 10.36 und

$$\lim_{x\searrow 0} Y_n(x) = -\infty. \tag{10.15}$$

Dies schliesst unter den Annahme, dass die gesuchte Lösung auf $(0, x_0]$ beschränkt ist, die Funkion $Y_n$ oder auch Linearkombinationen $\alpha J_n + \beta Y_n$ mit $\beta \neq 0$ aus. Interessanterweise müssen wir aber $Y_n$ wie oben kennen, nur um dann zu sagen, dass die gesuchte Lösung doch ein Vielfaches von $J_n$ sein muss.

**Übung 10.39** (Bessel-Funktionen zweiter Gattung)**.** *Sei* $n \in \mathbb{N}_0$.

(a) *Zeigen Sie, dass die Bessel-Funktion $Y_n$ zweiter Gattung wohl-definiert ist und beweisen Sie die Asymptotik in* (10.15).

(b) *Nehmen Sie eine geeignete Verallgemeinerung der Differentiation unter dem Integral für das uneigentliche Integrale an und beweisen Sie damit, dass $Y_n$ eine Lösung der Bessel-Differentialgleichung* (10.13) *darstellt.*

(c) *Für den Beweis der geeigneten Verallgemeinerung der Differentiation unter dem Integral betrachte man die Funktionen*

$$(x,t) \in (0,\infty) \times [1,\infty) \mapsto f(x,t) = \left(e^t + (-1)^n e^{-nt}\right) \mathrm{e}^{-x\sinh(t)}$$

*und*

$$F : (x,u) \in (0,\infty) \times [0,1] \mapsto \begin{cases} 0 & \textit{falls } u = 0 \\ \frac{f(x,\frac{1}{u})}{u^2} & \textit{falls } u > 0 \end{cases} .$$

*Zeigen Sie, dass*

$$\int_1^\infty f(x,t)\,\mathrm{d}t = \int_0^1 F(x,u)\,\mathrm{d}u$$

*gilt und dass $F$ alle Voraussetzungen von Satz 10.36 erfüllt.*

VO: 21.03.18

## 10.6 Wegintegrale

### 10.6.1 Skalare Wegintegrale

Sei $U \subseteq \mathbb{R}^n$ eine offene Teilmenge und $\gamma : [a, b] \to U$ ein stetig differenzierbarer Weg. Wie wir bereits in Abschnitt 8.3.2 gesehen haben, kann man die Länge von $\gamma$ durch

$$L(\gamma) = \int_a^b \|\gamma'(t)\| \, \mathrm{d}t$$

definieren und berechnen. In Lemma 8.32 haben wir auch gesehen, dass man den Weg oft so reparametrisieren kann, dass $\|\gamma'(s)\| = 1$ für alle $s \in [a, b]$ gilt, womit $s$ bereits die Bedeutung der Bogenlänge entlang des Weges annimmt (was ist die Länge des Weges bis zu Zeitpunkt $t$, falls $\|\gamma'(s)\| = 1$ für alle $s$ gilt?).

Der obige Begriff der Länge eines Weges lässt sich auf eine etwas grössere Klasse erweitern.

**Definition 10.40** (Stückweise differenzierbare Wege und deren Längen)**.** Ein (wie immer stetiger) Weg $\gamma : [a, b] \to \mathbb{R}^n$ heisst **stückweise (stetig) differenzierbar**, falls eine Zerlegung $\mathfrak{Z} = \{a = s_0 < \ldots < s_K = b\}$ von $[a, b]$ existiert, so dass $\gamma|_{[s_{k-1}, s_k]}$ für alle $k \in \{1, \ldots, K\}$ stetig differenzierbar ist. Die Länge des stückweise differenzierbaren Weges $\gamma$ ist definiert durch

$$L(\gamma) = \sum_{k=1}^{K} L(\gamma_k) = \sum_{k=1}^{K} \int_{s_{k-1}}^{s_k} \left\| (\gamma|_{[s_{k-1}, s_k]})'(s) \right\| \, \mathrm{d}s.$$

Eine Zerlegung $\mathfrak{Z}$ wie oben werden wir eine für die stückweise differenzierbare Funktion **erlaubte Zerlegung** nennen.

**Übung 10.41** (Wohldefiniertheit der Länge stückweise differenzierbarer Wege)**.** *Obiger Begriff der Länge eines stückweise differenzierbaren Weges verwendet strenggenommen die gewählte Zerlegung und sollte formal korrekt durch $L(\gamma, \mathfrak{Z})$ bezeichnet werden. Zeigen Sie, dass der Begriff der Länge eines stückweise differenzierbaren Weges nicht von der Wahl einer erlaubten Zerlegung abhängt, also $L(\gamma, \mathfrak{Z}) = L(\gamma, \mathfrak{Z}')$ gilt, falls $\mathfrak{Z}'$ eine weitere erlaubte Zerlegung von $[a, b]$ ist.*

Es gibt auch Situationen, wo das **skalare Wegintegral** einer stetigen reellwertigen Funktion $f : U \to \mathbb{R}$ entlang eines stetigen differenzierbaren Weges $\gamma : [a, b] \to U$

$$\int_a^b f(\gamma(s)) \|\gamma'(s)\| \, \mathrm{d}s$$

von Bedeutung ist. Wie zuvor die Weglänge lässt sich dieser Begriff auf stückweise differenzierbare Wege erweitern.

Zum Beispiel könnte für $f : U \to (0, \infty)$ der Wert $f(x)$ das Inverse der erlaubten Höchstgeschwindigkeit im Punkt $x \in U$ angeben. In der Tat falls $\mathfrak{Z} = \{a = s_0 < s_1 < \ldots < s_K = b\}$ eine bereits sehr feine Zerlegung des Intervalles $[a, b]$ ist, dann gibt für $k \in \{1, \ldots, K\}$ das Produkt $f(\gamma(s_k)) \|\gamma(s_k) - \gamma(s_{k-1})\|$ näherungsweise die Zeit an, die man bei Höchstgeschwindigkeit

braucht, um von $\gamma(s_{k-1})$ nach $\gamma(s_k)$ zu gelangen (wieso?). Dies führt dann zur Interpretation von

$$D_f(\gamma) = \int_a^b f(\gamma(s))\|\gamma'(s)\| \, \mathrm{d}s$$

als die Gesamtdauer der Reise entlang des Weges $\gamma$, wenn man immer mit erlaubter Höchstgeschwindigkeit (von $f$ angegeben) reist. Damit kann man nun die minimale Reisedauer (streng genommen als Infimum definierte, nicht unterschreitbare Reisedauer) als eine natürliche Metrik auf $U$ einführen, die von der gewählten inversen Höchstgeschwindigkeit (Dichte des Verkehrs) $f$ abhängt. Falls $f$ die Dichte eines Mediums beschreibt, so hat dies wiederum eine mögliche physikalische Interpretation wie in Beispiel 7.54.

**Übung 10.42** (Metrik über gewichtete Längen von Wegen)**.** *Sei $U \subseteq \mathbb{R}^n$ eine offene, zusammenhängende Teilmenge und sei $f : U \to (0, \infty)$ stetig. Definieren Sie in Analogie zu Übung 10.18 die Distanz $\mathrm{d}(x, y)$ zweier Punkte $x, y \in U$ durch*

$$\inf \left\{ \int_a^b f(\gamma(t))\|\gamma'(t)\| \, \mathrm{d}t \;\middle|\; \gamma : [a, b] \to U \text{ stückweise differenzierbar mit } \gamma(a) = x,\ \gamma(b) = y \right\}.$$

*Zeigen Sie, dass $\mathrm{d}$ in der Tat eine Metrik auf $U$ definiert und dass diese die Standardtopologie auf $U$ induziert.*

### 10.6.2 Wegintegrale von Vektorfeldern

Für viele weitere Anwendungen ist hingegen ein anderer Begriff des Wegintegrals von Bedeutung.

**Definition 10.43** (Wegintegral eines Vektorfelds)**.** Sei $U \subseteq \mathbb{R}^n$ eine offene Teilmenge und sei $f : U \to \mathbb{R}^n$ ein stetiges Vektorfeld. Wir definieren das **Wegintegral des Vektorfelds** $f$ entlang eines stetig differenzierbaren Weges $\gamma : [a, b] \to U$ durch

$$\int_\gamma f \cdot \mathrm{d}s = \int_a^b \big\langle f(\gamma(s)), \gamma'(s) \big\rangle \, \mathrm{d}s$$

Ist $\gamma : [a, b] \to U$ stückweise differenzierbar und $\mathfrak{Z} = \{a = s_0 < s_1 < \ldots < s_K = b\}$ eine erlaubte Zerlegung von $[a, b]$ für $\gamma$, so setzt man wiederum

$$\int_\gamma f \cdot \mathrm{d}s = \sum_{k=1}^{K} \int_{\gamma|_{[s_{k-1}, s_k]}} f \cdot \mathrm{d}s.$$

Eine von vielen physikalischen Interpretationen ist die Berechnung der Arbeit entlang eines Weges $\gamma$. Angenommen $f(x)$ gibt die Richtung und die Stärke einer Krafteinwirkung auf einen Körper an der Stelle $x \in U$ an. Dann ist $\langle f(\gamma(s_k)), \gamma(s_k) - \gamma(s_{k-1}) \rangle$ näherungsweise die verrichtete Arbeit auf einem Teilintervall $[s_{k-1}, s_k]$ einer Zerlegung $\mathfrak{Z} = \{a = s_0 < \ldots < s_K = b\}$. Die Argumente in Abschnitt 8.3 führen dann zur Interpretation von $\int_\gamma f \cdot \mathrm{d}s$ als die geleistete **Arbeit** für die Reise entlang des Weges $\gamma$ von $\gamma(a)$ nach $\gamma(b)$. Diese Gesamtarbeit hängt im

Allgemeinen vom gewählten Weg und nicht nur vom Anfangsort $\gamma(a)$ und vom Zielort $\gamma(b)$ ab (siehe Beispiel 10.45 unten). Die geleistete Arbeit hängt aber nicht von der gewählten Parametrisierung des Weges ab (vergleiche Lemma 8.32).

**Lemma 10.44** (Reparametrisierungen und Richtungsumkehr eines Weges)**.** *Sei $U \subseteq \mathbb{R}^n$ eine offene Teilmenge, $f : U \to \mathbb{R}^n$ ein stetiges Vektorfeld und sei $\gamma : [a,b] \to \mathbb{R}^d$ ein stetig differenzierbarer Weg für $a < b$. Dann ändert sich der Wert des Wegintegrals $\int_\gamma f \cdot \mathrm{d}s$ nicht unter (orientierungserhaltenden) Reparametrisierungen von $\gamma$.*

*Weiter gilt für den umgekehrten Weg $\tilde{\gamma} : t \in [-b,-a] \mapsto \gamma(-t)$ mit $\tilde{\gamma}(-b) = \gamma(b)$ und $\tilde{\gamma}(-a) = \gamma(a)$*

$$\int_{\tilde{\gamma}} f \cdot \mathrm{d}s = -\int_\gamma f \cdot \mathrm{d}s.$$

*Beweis.* Sei $[\tilde{a}, \tilde{b}]$ ein kompaktes Intervall mit Endpunkten $\tilde{a} < \tilde{b}$ und $\psi : [\tilde{a}, \tilde{b}] \to [a,b]$ eine stetig differenzierbare, monotone wachsende, bijektive Funktion. Dann gilt

$$\begin{aligned}\int_{\gamma\circ\psi} f \cdot \mathrm{d}s &= \int_{\tilde{a}}^{\tilde{b}} \big\langle f(\gamma(\psi(t))), (\gamma\circ\psi)'(t)\big\rangle \,\mathrm{d}t = \int_{\tilde{a}}^{\tilde{b}} \big\langle f(\gamma(\psi(t))), \gamma'(\psi(t))\big\rangle\, \psi'(t)\,\mathrm{d}t \\ &= \int_{\tilde{a}}^{\tilde{b}} \big\langle f(\gamma(s)), \gamma'(s)\big\rangle \,\mathrm{d}s = \int_\gamma f \cdot \mathrm{d}s.\end{aligned}$$

Die zweite Aussage folgt mit selbiger Rechnung und der Funktion $\psi : [-b,-a] \to [a,b], t \mapsto -t$ mit $\psi'(t) = -1$ für alle $t \in [-b,-a]$. Überprüfen Sie dies. □

**Beispiel 10.45** (Wirbelsturm mit Auge)**.** *Wir betrachten das Vektorfeld $f : \mathbb{R}^2 \to \mathbb{R}^2$ definiert durch*

$$f(x,y) = \begin{pmatrix} -y \\ x \end{pmatrix}$$

*für $(x,y)^t \in \mathbb{R}^2$ und betrachten im Folgenden mehrere Wege $\gamma$ von $(0,0)^t$ nach $(1,1)^t$ und berechnen das Wegintegral $\int_\gamma f \cdot \mathrm{d}s$.*

*(a) Sei $\gamma_0 : t \in [0,1] \mapsto (t,t)^t$ der gerade Weg von $(0,0)^t$ nach $(1,1)^t$. Dann gilt*

$$\int_{\gamma_0} f \cdot \mathrm{d}s = \int_0^1 \big\langle f(\gamma_0(t)), \gamma_0'(t)\big\rangle \,\mathrm{d}t = \int_0^1 \left\langle \begin{pmatrix} -t \\ t \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle \mathrm{d}t = 0.$$

*(b) Sei $\gamma_1 : [0,2] \to \mathbb{R}^2$ der Weg definiert durch*

$$\gamma_1(t) = \begin{cases} (t,0)^t & \text{falls } t \in [0,1] \\ (1, t-1)^t & \text{falls } t \in [1,2] \end{cases}$$

*für $t \in [0,2]$, was einen stückweise differenzierbaren Weg von $(0,0)^t$ nach $(1,1)^t$ definiert. Es gilt*

$$\int_{\gamma_1} f \cdot \mathrm{d}s = \int_0^1 \left\langle \begin{pmatrix} 0 \\ t \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle \mathrm{d}t + \int_1^2 \left\langle \begin{pmatrix} 1-t \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle \mathrm{d}t = 1$$

*(c) Sei $\gamma_{-1} : [0,2] \to \mathbb{R}^2$ der Weg definiert durch*

$$\gamma_{-1}(t) = \begin{cases} (0,t)^t & \text{falls } t \in [0,1] \\ (t-1,1)^t & \text{falls } t \in [1,2] \end{cases}$$

*für $t \in [0,2]$, was wiederum einen stückweise differenzierbaren Weg von $(0,0)^t$ nach $(1,1)^t$ definiert. Damit erhalten wir*

$$\int_{\gamma_{-1}} f \cdot \mathrm{d}s = \int_0^1 \left\langle \begin{pmatrix} -t \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle \mathrm{d}t + \int_1^2 \left\langle \begin{pmatrix} -1 \\ t-1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle \mathrm{d}t = -1$$

*Wir sehen also, dass für den „Wirbelsturm“ die geleistete Arbeit $\int_\gamma f \cdot \mathrm{d}s$ vom gewählten Weg $\gamma$ abhängt. Bewegt man sich „senkrecht“ zum Vektorfeld, so wird gar keine Arbeit geleistet (siehe $\gamma_0$); bewegt man sich mit dem Vektorfeld, so wird positive Arbeit geleistet (siehe $\gamma_1$), und bewegt man sich „entgegen dem Vektorfeld“, so wird negative Arbeit geleistet (siehe $\gamma_{-1}$).*

## 10.7 Konservative Vektorfelder

**Definition 10.46.** Sei $U \subseteq \mathbb{R}^n$ ein Gebiet und $f : U \to \mathbb{R}^n$ ein stetiges Vektorfeld. Dann heisst $f$ **konservativ**, falls Wegintegrale des Vektorfelds $f$ nur von Anfangs- und Endpunkt abhängen. Genauer formuliert, falls für alle stückweise stetig differenzierbaren Wege $\gamma : [a, b] \to U$ und $\eta : [a', b'] \to U$ mit $\gamma(a) = \eta(a')$ und $\gamma(b) = \eta(b')$ gilt

$$\int_\gamma f \cdot \mathrm{d}s = \int_\eta f \cdot \mathrm{d}s.$$

**Wichtige Übung 10.47** (Verbindbarkeit)**.** *Zeigen Sie, dass je zwei Punkte in einem Gebiet durch einen stückweise stetig differenzierbaren Weg miteinander verbunden werden können (siehe auch Übung 10.18).*

Eine **Schlaufe** in einer offenen Teilmenge $U \subseteq \mathbb{R}^n$ ist ein Weg mit gleichem Anfangs- und Endpunkt (das heisst, ein Weg $\gamma : [a, b] \to U$ mit $\gamma(a) = \gamma(b)$). Ob ein Vektorfeld konservativ ist oder nicht, lässt sich auch mit Schlaufen charakterisieren.

**Übung 10.48** (Schlaufencharakterisierung)**.** *Zeigen Sie, dass ein stetiges Vektorfeld $f : U \to \mathbb{R}^n$ auf einem Gebiet $U \subseteq \mathbb{R}^n$ genau dann konservativ ist, wenn für jede stückweise stetig differenzierbare Schlaufe $\gamma$ in $U$ gilt $\int_\gamma f \cdot \mathrm{d}s = 0$.*

**Satz 10.49** (Stammfunktion)**.** *Sei $U \subseteq \mathbb{R}^n$ ein Gebiet und $f : U \to \mathbb{R}^n$ ein stetiges Vektorfeld. Dann ist $f$ genau dann konservativ, wenn es eine stetig differenzierbare Funktion $F : U \to \mathbb{R}$ mit $f(x) = \nabla F(x)$ für alle $x \in U$ gibt.*

*Des Weiteren gelten für ein stetig differenzierbares konservatives Vektorfeld $f$ und deren Komponenten $f_1, \ldots, f_n$ die (partiellen) Differentialgleichungen*

$$\partial_j f_k = \partial_k f_j$$

*für alle $j, k \in \{1, \ldots, n\}$.*

Die differenzierbare Funktion $F$ in obigem Satz übernimmt die Rolle der Stammfunktion im Fundamentalsatz der Integral- und Differentialrechnung und wird auch das zum Vektorfeld $f$ assoziierte **Potential** (**Potentialfunktion**) genannt. Diese Funktion existiert aber nicht für alle, sondern nur für gewisse (eben konservative) Vektorfelder.

**Übung 10.50** (Wirbelsturm hat kein Potential)**.** *Zeigen Sie direkt und nochmals unter Verwendung von Satz 10.49), dass das Vektorfeld aus Beispiel 10.45 kein Potential besitzt.*

*Beweis.* Angenommen es gibt eine differenzierbare Funktion $F : U \to \mathbb{R}$ mit $f(x) = \nabla F(x)$ für alle $x \in U$. Sei $\gamma : [a, b] \to U$ ein stetig differenzierbarer Weg. Dann ist für $t \in [a, b]$ also $f(\gamma(t)) = \nabla F(\gamma(t)) = (\mathrm{D}_{\gamma(t)} F)^t$ und somit nach der Kettenregel

$$\begin{aligned}\int_\gamma f \cdot \mathrm{d}s &= \int_a^b \big\langle f(\gamma(t)), \gamma'(t) \big\rangle \,\mathrm{d}t = \int_a^b \mathrm{D}_{\gamma(t)} F \gamma'(t) \,\mathrm{d}t = \int_a^b (F \circ \gamma)'(t) \,\mathrm{d}t \\ &= F(\gamma(b)) - F(\gamma(a)).\end{aligned}$$

Falls $\gamma$ bloss stückweise stetig differenzierbar ist und und $\mathfrak{Z} = \{a = s_0 < s_1 < \ldots < s_K = b\}$ eine erlaubte Zerlegung von $[a, b]$ ist, so kann man obige Rechnung auf die Teilintervalle $[s_{k-1}, s_k]$ anwenden. Dies führt zu einer Teleskopsumme

$$\int_\gamma f \cdot \mathrm{d}s = \sum_{k=1}^{K} \int_{\gamma|_{[s_{k-1},s_k]}} f \cdot \mathrm{d}s = \sum_{k=1}^{K} F(\gamma(s_k)) - F(\gamma(s_{k-1})) = F(\gamma(b)) - F(\gamma(a)).$$

Daher ist $f$ konservativ.

Sei nun $f$ konservativ und $x_0 \in U$ ein fester Punkt. Da $U$ zusammenhängend ist, gibt es nach Übung 10.47 zu jedem $x \in U$ einen stückweise stetig differenzierbaren Weg $\gamma_x$ in $U$ mit Anfangspunkt $x_0$ und Endpunkt $x$. Wir betrachten die Funktion

$$F : x \in U \mapsto \int_{\gamma_x} f \cdot \mathrm{d}s,$$

welche nicht vom gewählten Weg $\gamma_x$ abhängt, da $f$ konservativ ist.

Sei nun $x \in U$, $k \in \{1, \ldots, n\}$ und $h \in \mathbb{R} \setminus \{0\}$ klein genug, so dass $x + the_k \in U$ für alle $t \in [0, 1]$. Wir können dann mit Hilfe eines Weges $\gamma_x : [a, b] \to U$ von $x_0$ nach $x$ einen Weg $\gamma_{x+he_k}$ von $x_0$ nach $x + he_k$ durch

$$t \in [a, b+1] \mapsto \gamma_{x+he_k}(t) = \begin{cases} \gamma_x(t) & \text{falls } t \in [a, b] \\ x + (t-b)he_k & \text{falls } t \in [b, b+1] \end{cases}$$

definieren. Für die partielle Ableitung $\partial_k F$ von $F$ ergibt sich dadurch

$$\begin{aligned} \partial_k F(x) &= \lim_{h\to 0} \frac{F(x + he_k) - F(x)}{h} = \lim_{h\to 0} \tfrac{1}{h} \left( \int_{\gamma_{x+he_k}} f \cdot \mathrm{d}s - \int_{\gamma_x} f \cdot \mathrm{d}s \right) \\ &= \lim_{h\to 0} \tfrac{1}{h} \int_b^{b+1} \langle f(x + (t-b)he_k), he_k \rangle \, \mathrm{d}t = \lim_{h\to 0} \int_0^1 f_k(x + she_k) \, \mathrm{d}s = f_k(x) \end{aligned}$$

auf Grund von Satz 10.36 und der Stetigkeit von $f_k$. Da dies für alle $x \in U$ und $k \in \{1, \ldots, n\}$ gilt und $f_1, \ldots, f_n$ per Annahme stetig sind, folgt aus Satz 10.10, dass die totale Ableitung von $F$ überall existiert und $\nabla F(x) = f(x)$ für alle $x \in U$ gilt.

Sei nun $f$ konservativ und stetig differenzierbar. Dann existiert nach obigem eine Funktion $F$ mit $\nabla F = f$. Für $j, k \in \{1, \ldots, n\}$ gilt dann

$$\partial_j f_k = \partial_j \partial_k F = \partial_k \partial_j F = \partial_k f_j$$

nach dem Satz von Schwarz (Satz 10.20). □

VO: 23.03.2018

### 10.7.1 Integrabilitätsbedingungen

Wie wir in Satz 10.49 gezeigt haben, stellen die partiellen Differentialgleichungen

$$\partial_k f_j = \partial_j f_k \tag{10.16}$$

für $j, k \in \{1, \ldots, n\}$ zu einem stetig differenzierbaren Vektorfeld $f$ auf einem Gebiet eine notwendige Bedingung für die Existenz eines Potentials zu $f$ dar. Wir nennen diese die **Integrabilitätsbedingungen** und wollen hier diskutieren, inwiefern sie auch hinreichend sind.

Wir haben bereits einige Male zuvor gesehen, dass Eigenschaften von Funktionen auf verschiedenen Definitionsgebieten unterschiedlich zusammenhängen, und Konservativität und die Integrabilitätsbedingungen bilden ein weiteres Beispiel dafür: Für gewisse Gebiete reicht es tatsächlich, die Integrabilitätsbedingungen zu überprüfen, um entscheiden zu können, ob ein Potential existiert oder nicht; allerdings nicht für alle Gebiete, wie folgendes Beispiel zeigt.

**Beispiel 10.51** (Wirbelsturm mit Singularität)**.** *Wir betrachten $n = 2$, das Gebiet $U = \mathbb{R}^2 \setminus \{0\}$ und das Vektorfeld $f : U \to \mathbb{R}^2$ gegeben durch*

$$f(x, y) = \begin{pmatrix} \frac{-y}{x^2+y^2} \\ \frac{x}{x^2+y^2} \end{pmatrix}.$$

*für $(x, y)^t \in U$. In diesem Fall ist*

$$\partial_1 f_2(x, y) = \partial_x \left( \frac{x}{x^2 + y^2} \right) = \frac{-x^2 + y^2}{(x^2 + y^2)^2}$$

*und*

$$\partial_2 f_1(x, y) = \partial_y \left( \frac{-y}{x^2 + y^2} \right) = \frac{-(x^2 + y^2) - (-y)2y}{(x^2 + y^2)^2} = \frac{-x^2 + y^2}{(x^2 + y^2)^2},$$

*womit die Integrabilitätsbedingungen in* (10.16) *auf ganz $U$ erfüllt sind. Dennoch ist $f$ nicht konservativ. Sei $\gamma : [0, 2\pi] \to U$ die stetig differenzierbare Schlaufe (der geschlossene Weg) definiert durch*

$$\gamma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}$$

*für $t \in [0, 2\pi]$, die einmal im Gegenuhrzeigersinn um den Einheitskreis läuft. Dann ist*

$$\int_\gamma f \cdot \mathrm{d}s = \int_0^{2\pi} \left\langle \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix}, \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} \right\rangle \mathrm{d}t = 2\pi,$$

*obwohl $\gamma$ ein geschlossener Weg ist mit $\gamma(0) = \gamma(2\pi) = (1, 0)^t$.*

In der Tat misst das Wegintegral in obigen Beispiel die Änderung des Winkels beim Ursprung, doch lässt sich der Winkel als Potential nicht stetig auf ganz $\mathbb{R}^2 \setminus \{0\}$ definieren (was mit der Nicht-Existenz eines Logarithmus auf der komplexen Ebene äquivalent ist). Wie wir jetzt aber zeigen werden, sind für gewisse Gebiete die Integrabilitätsbedingungen dennoch hinreichend.

**Satz 10.52** (Integrabilitätsbedingungen auf sternförmigen Gebieten)**.** *Sei $U \subseteq \mathbb{R}^n$ offen und sternförmig. Ein stetig differenzierbares Vektorfeld $f : U \to \mathbb{R}^n$ ist genau dann konservativ,*

*wenn $f$ den Integrabilitätsbedingungen*

$$\partial_k f_j = \partial_j f_k$$

*für alle $j, k \in \{1, \dots, n\}$ genügt.*

*Beweis.* Die Notwendigkeit der Integrabilitätsbedingungen wurde bereits in Satz 10.49 bewiesen. Für die Umkehrung verwenden wir ein Zentrum $z \in U$ und das Wegintegral von $f$ über die gerade Strecke von $z$ nach $x \in U$, um eine Funktion $F : U \to \mathbb{R}$ durch

$$F(x) = \int_0^1 \langle f(z + t(x - z)), (x - z) \rangle \, \mathrm{d}t$$

für $x \in U$ zu definieren. Entsprechend dem Beweis von Satz 10.49 stellt $F$ gerade einen Kandidaten für ein Potential von $f$ dar.

Wir fixieren ein $j \in \{1, \dots, n\}$ und betrachten als Vorbereitung zur Berechnung von $\partial_j F$ zuerst für $h \in \mathbb{R}^n$

$$\partial_h f_j = \sum_{k=1}^n h_k \partial_k f_j = \sum_{k=1}^n h_k \partial_j f_k \tag{10.17}$$

nach den vorausgesetzten Integrabilitätsbedingungen. Nach Satz 10.36 über Differenzierbarkeit von Parameterintegralen existiert für $j \in \{1, \dots, n\}$ die partielle Ableitung $\partial_j F$ der Funktion $F$, die gleichzeitig als Weg- und als Parameterintegral definiert ist. Des Weiteren gilt für $x \in U$

$$\begin{aligned} \partial_j F(x) &= \partial_{x_j} \int_0^1 \left( \sum_{k=1}^n f_k\Big(z + t(x - z)\Big)(x_k - z_k) \right) \mathrm{d}t \\ &= \int_0^1 \left( \sum_{k=1}^n (\partial_j f_k)\Big(z + t(x - z)\Big) t(x_k - z_k) + f_j\Big(z + t(x - z)\Big) \right) \mathrm{d}t, \end{aligned} \tag{10.18}$$

da einzig der Term mit $k = j$ die Produktregel erfordert und da die partielle Ableitung von $x \in U \mapsto f_k\big(z + t(x - z)\big)$ nach $x_j$ durch $t(\partial_j f_k)\big(z + t(x - z)\big)$ für $x \in U$ gegeben ist. Letzteres folgt aus der mehrdimensionalen Kettenregel oder wie folgt: wir definieren für $t \geq 0$ die Funktion $x \mapsto \psi_t(x) = f_k\big(z + t(x - z)\big)$ und betrachten die partielle Ableitung $\partial_j \psi_t(x)$. Für $t > 0$ berechnet man direkt

$$\begin{aligned} \partial_j \psi_t(x) &= \lim_{s \to 0} \frac{1}{s} \big( f_k\big(z + t(x + se_j - z)\big) - f_k\big(z + t(x - z)\big) \big) \\ &= t \lim_{s \to 0} \frac{1}{st} \big( f_k\big(z + t(x - z) + ste_j\big) - f_k\big(z + t(x - z)\big) \big) \\ &= t \partial_j f_k(z + t(x - z)). \end{aligned}$$

Für $t = 0$ hängt $\psi_0 = f_k(z)$ ja nicht von $x$ ab, womit die partielle Ableitung $\partial_j \psi_0$ verschwindet.

Wir setzen nun $h = x - z$, verwenden (10.17) in (10.18) und erhalten mit partieller Integration

$$\begin{aligned}\partial_j F(x) &= \int_0^1 t \underbrace{\partial_h f_j(z+th)}_{=\frac{\mathrm{d}}{\mathrm{d}t}(t\mapsto f_j(z+th))} \,\mathrm{d}t + \int_0^1 f_j(z+th)\,\mathrm{d}t \\ &= [t f_j(z+th)]_0^1 - \int_0^1 f_j(z+th)\,\mathrm{d}t + \int_0^1 f_j(z+th)\,\mathrm{d}t \\ &= f_j(z+h) = f_j(x).\end{aligned}$$

Daher ist $f = \nabla F$, $F$ ist stetig differenzierbar nach Satz 10.10 und der Satz folgt aus der Charakterisierung der Konservativität in Satz 10.49. □

**Übung 10.53.** *Für welchen Wert von $\lambda \in \mathbb{R}$ ist das Vektorfeld $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ definiert durch*

$$f(x,y) = \begin{pmatrix} \lambda x \mathrm{e}^y \\ (y+1+x^2)\mathrm{e}^y \end{pmatrix}$$

*für $(x,y)^t \in \mathbb{R}^2$ konservativ? Bestimmen Sie für diesen Wert ein Potential von $f$.*

**Applet 10.54** (Integrabilitätsbedingungen)**.** *Welche verschiedenen Werte für das Wegintegral (dargestellt unten rechts) können Sie erzielen wenn Sie geschlossene Wege betrachten? Warum ändert sich der Wert des Wegintegral meist nicht aber manchmal schon wenn Sie die mittleren drei Punkte bewegen?*

## 10.8 Weitere Lernmaterialien

### 10.8.1 Verwendung des Kapitels

Dieses Kapitel stellt die Grundlagen für die restlichen Themen des Semesters zur Verfügung. In der Tat ist der Begriff der totalen Ableitung für alle weiteren Diskussionen wie zum Beispiel rund um Teilmannigfaltigkeiten im nächsten Kapitel oder auch der mehrdimensionalen Substitutionsregel von fundamentaler Bedeutung. Für die mehrdimensionalen Integralsätze werden wir allerdings noch weitere Ableitungsbegriffe kennenlernen, die allerdings ohne ein gutes Verständnis der Ableitung im Sinne von Definition 10.3 schwer verständlich sein werden. Der Begriff der partiellen Ableitung ist für die Theorie aber vor allem für alle praktischen Berechnungen unabdingbar und stellt einen direkten Zusammenhang zu allen Sätzen und Regeln der eindimensionalen Differentialrechnung her.

Des Weiteren sind Wegintegrale für die Anwendungen (zum Beispiel in der Physik) aber auch für den weiteren Aufbau der mehrdimensionalen Analysis notwendig. Dies haben wir bereits bei der Besprechung der konservativen Vektorfelder gesehen. Die Sätze rund um diesen Begriff zeigen ebenso, dass die mehrdimensionale Analysis deutlich komplexer als die eindimensionale Analysis ist: Im Fundamentalsatz der Differential- und Integralrechnung haben wir gesehen, dass in der eindimensionalen Analysis jede stetige Funktion auf einem Intervall als Ableitung einer Funktion (Stammfunktion) auftritt. Doch ein Vektorfeld auf einer offenen Teilmengen in $\mathbb{R}^n$ ist nicht immer eine Ableitung einer Potentialfunktion. Selbst die notwendigen Integrabilitätsbedingungen (welche sich aus dem Satz von Schwarz ergeben) stellen keine Charakterisierung der Konservativität dar, da die Gestalt des Definitionsbereiches einen weiteren Einfluss auf den Begriff hat. Wir werden diesen Problemen etwas allgemeiner im Zusammenhang der mehrdimensionalen Integralsätze nochmals begegnen.

Das Berechnen von höheren Ableitungen von reellwertigen Funktionen auf offenen Teilmengen von $\mathbb{R}^n$ ist zwar prinzipiell nicht schwierig, ist aber ab der dritten Ableitung kaum mehr praktisch. In der Tat hat bereits die dritte totale Ableitung $\binom{n}{3} + n(n-1) + n$ verschiedene Komponenten (da manche der $n^3$ iterierten partiellen Ableitungen nach dem Satz von Schwarz übereinstimmen). Für $n = 3$ sind dies $1 + 6 + 3 = 10$ und für $n = 4$ bereits $4 + 12 + 4 = 18$ Komponenten. Aus diesem Grund beschränkt man sich meist auf die erste und zweite Ableitung, welche durch den Gradienten und die Hesse-Matrix gegeben sind. Sollte die Berechnung einer höheren Taylor-Approximation gewünscht sein, so ist es einfacher diese mittels der Definition der gegebenen Funktion aus eindimensionalen Taylor-Approximationen zu berechnen. Die Hesse-Matrix und die Begriffe der positiv und negativ definiten Matrizen sind hingegen sehr nützlich, da wir diese zur Bestimmung von lokalen Extremwerten benötigen.

### 10.8.2 Übungen

**Übung** (Konvexität für Funktionen und Teilmengen)**.** *Sei $I$ ein Intervall und $f : I \to \mathbb{R}$ eine Funktion. Zeigen Sie, dass $f$ genau dann konvex ist, wenn die Teilmenge*

$$\{(x, y) \in I \times \mathbb{R} \mid y \geq f(x)\} \subseteq \mathbb{R}^2$$

*der Punkte in der Ebene oberhalb des Graphen von $f$ konvex ist.*

**Übung.** *Sei $f : \mathbb{R}^2 \to \mathbb{R}$ definiert durch*

$$f(x,y) = \begin{cases} 0 & \text{falls } (x,y) = (0,0), \\ (x^2+y^2)\sin(\frac{1}{x^2+y^2}) & \text{andernfalls.} \end{cases}$$

*(i) Zeigen Sie, dass $f$ auf ganz $\mathbb{R}^2$ differenzierbar ist.*

*(ii) Zeigen Sie, dass die partiellen Ableitungen von $f$ nicht überall stetig sind.*

**Übung** (Notwendigkeit der Annahmen im Satz von Schwarz)**.** *In dieser Übung möchten wir eine zweimal differenzierbare Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ konstruieren, deren partielle Ableitungen $\partial_1\partial_2 f$, $\partial_2\partial_1 f$ beim Punkt $(0,0)$ nicht stetig sind und dort verschiedene Werte haben. Dies zeigt, dass die Annahme der zweifachen stetigen Differenzierbarkeit im Satz von Schwarz (Satz 10.20) notwendig war.*

*Sei $f : \mathbb{R}^2 \to \mathbb{R}$ definiert für alle für $(x,y)^t \in \mathbb{R}^2$ durch*

$$f(x,y) = \begin{cases} xy\frac{x^2-y^2}{x^2+y^2} & \text{falls } (x,y) \neq (0,0), \\ 0 & \text{falls } (x,y) = (0,0). \end{cases}$$

*(i) Zeigen Sie, dass $f$ zweimal differenzierbar ist. Überzeugen Sie sich dazu zuerst davon, dass $f$ stetig ist.*

*(ii) Zeigen Sie, dass $\partial_{xy} f$ und $\partial_{yx} f$ auf $\mathbb{R}^2 \setminus \{(0,0)\}$ stetig sind.*

*(iii) Zeigen Sie, dass $\partial_{xy} f(0,0) = -\partial_{yx} f(0,0) = 1$ gilt.*

**Übung.** *Zeigen Sie, dass das vollständige elliptische Integral zweiter Art (siehe Beispiel 10.37)*

$$E : x \in [0,1] \mapsto \int_0^{\pi/2} \sqrt{1 - x^2\sin^2(t)}\,\mathrm{d}t$$

*der Differentialgleichung*

$$(x^2-1)xE''(x) + (x^2-1)E'(x) - xE(x) = 0$$

*genügt.*

**Übung** (Unabhängigkeit von Parametrisierung)**.** *Sei $U \subseteq \mathbb{R}^n$ eine offene, zusammenhängende Teilmenge und sei $f : U \to \mathbb{R}$ stetig. Sei $\gamma : [a,b] \to U$ ein stetig differenzierbarer Weg. Zeigen Sie, dass der Wert*

$$\int_a^b f(\gamma(s))\|\gamma'(s)\|\,\mathrm{d}s$$

*gleich bleibt unter Reparametrisierungen des Weges $\gamma$.*

**Übung** (Differenzierbare Reparametrisierungen)**.** *Sei $\gamma : [a, b] \to \mathbb{R}^n$ ein stückweise differenzierbarer Weg. Zeigen Sie, dass eine Reparametrisierung von $\gamma$ existiert, die stetig differenzierbar ist.*

**Übung** (Eindeutigkeit des Potentials)**.** *Sei $U \subseteq \mathbb{R}^n$ ein Gebiet und $f : U \to \mathbb{R}^n$ ein konservatives Vektorfeld. Zeigen Sie, dass Potentiale von $f$ sich um Konstanten unterscheiden.*

### 10.8.3 Lernkarten

Sie können wiederum die Lernkarten für Ihre Wiederholung der Themen des Kapitels verwenden.

# Kapitel 11

# Anfänge der Differentialgeometrie

Wie wir bereits am Anfang von Kapitel 10 erwähnt haben, kann eine Abbildung der Form $f : U \to \mathbb{R}^m$ auf einer offenen Menge $U \subseteq \mathbb{R}^n$ verschiedene Bedeutungen haben. In diesem Kapitel werden wir folgende Blickwinkel in den Vordergrund stellen:

- Für $m = n$ könnte $f$ einen glatten Koordinatenwechsel („Diffeomorphismus“) darstellen. Polar-, Zylinder-, und Kugelkoordinaten werden drei einfache Beispiele dafür darstellen.
- Für $m < n$ könnten wir uns für die Lösungsmenge eines (meist nicht-linearen) Gleichungssystems $f(x) = 0$ interessieren. Unter geeigneten Annahmen wird die Lösungsmenge eine „Teilmannigfaltigkeit“ von $\mathbb{R}^n$ darstellen. Zum Beispiel ist die Lösungsmenge der Gleichung $x^2 + y^2 + z^2 - 1 = 0$ die Teilmannigfaltigkeit $\mathbb{S}^2 = \{v \in \mathbb{R}^3 \mid \|v\| = 1\}$ des dreidimensionalen euklidschen Raumes $\mathbb{R}^3$.
- Für $n < m$ könnte $f : U \to \mathbb{R}^m$ eine Parametrisierung einer derartigen „$n$-dimensionalen Teilmannigfaltigkeit“ im $\mathbb{R}^m$ darstellen. Zum Beispiel könnten wir $U = (0, \pi) \times (-\pi, \pi)$ gemeinsam mit Breiten- und Längengraden verwenden, um $\mathbb{S}^2$ fast vollständig mittels

  $$f(\theta, \varphi) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$$

  für $(\theta, \varphi) \in U$ zu parametrisieren. Die Abbildung $f$ liefert also gewissermassen Koordinaten auf der Sphäre $\mathbb{S}^2$ (und stellt, wenn man will, einen Teil eines Atlas zu $\mathbb{S}^2$ dar).

Wir beginnen unsere Diskussion mit dem zweiten Blickwinkel, doch deuten obige Beispiele schon an, dass die drei Themen eng miteinander verwandt sind. Die Sätze und Begriffe dieses Kapitels stellen die Anfänge der Differentialgeometrie dar, welche nebst grosser mathematischer Relevanz auch in anderen Fachgebieten wie zum Beispiel der allgemeinen Relativitätstheorie der Physik eine zentrale Rolle einnimmt.

## 11.1 Sätze zur impliziten Funktion und zur inversen Abbildung

Innerhalb der Mathematik oder auch in Anwendungen treten oft Gleichungen der Form $F(x, y) = 0$ auf, die man dann gerne als Definition einer „implizit definierten Funktion“ $y =$

$y(x)$ auffassen würde. Zum Beispiel könnte $F$ das sehr konkrete Polynom

$$F(x,y) = y^5 + x^2y^4 + (1+x)y^3 - x^5y + 2$$

sein, womit wir bereits ein sehr schwieriges Problem vor uns haben: Wie können wir für ein vorgegebenes $x \in \mathbb{R}$ die Lösung $y = y(x)$ von $F(x,y) = 0$ durch eine Formel angeben? Definiert dies überhaupt eine Funktion?

Da im Allgemeinen ein Polynom fünften Grades (hier $y \in \mathbb{R} \mapsto F(x,y)$ für ein fest gewähltes $x \in \mathbb{R}$) bis zu fünf reelle Nullstellen besitzen kann, definiert $F(x,y) = 0$ keine eindeutig bestimmte Funktion $x \mapsto y(x)$. Wir werden dieses Problem im Folgenden umgehen, indem wir von einem bekannten Lösungspaar $(x_0, y_0)$ mit $F(x_0, y_0) = 0$ ausgehen und für $x$ in der Nähe von $x_0$ eine Lösung $y = y(x)$ für $F(x,y) = 0$ ebenfalls nur in der Nähe von $y_0$ suchen. Unter geeigneten Annahmen wird dies $y$ eindeutig als Funktion von $x$ bestimmen.

Die Antwort auf die erste Frage ist vielleicht etwas überraschender. Selbst für eine polynomielle Gleichung der Form $F(x,y) = 0$ wie oben gibt es im Allgemeinen keine Formel, die die gesuchte Lösung mittels Wurzelausdrücken bestimmt. (Der Beweis dieser Aussage verwendet Galois-Theorie aus der Vorlesung Algebra II im zweiten Studienjahr des Mathematikstudiums.)

Für gewisse nicht-polynomielle Gleichungen wie zum Beispiel

$$F(x,y) = x - y - \mathrm{e}^y = 0 \tag{11.1}$$

ist es etwas weniger überraschend, dass wir keine konkrete Lösungsformel angeben können, selbst wenn für die Gleichung (11.1) zu jedem $x \in \mathbb{R}$ ein eindeutig bestimmtes $y = y(x) \in \mathbb{R}$ mit $x = y + \mathrm{e}^y$ existiert (wieso?). Dennoch wollen wir die Funktion auch ohne Angabe einer konkreten Formel weiter untersuchen.

Wir geben uns also im Zusammenhang mit einer impliziten Gleichung $F(x,y) = 0$ die Aufgabe, ausgehend von einer Lösung $(x_0, y_0)$ mit $F(x_0, y_0) = 0$, zu zeigen, dass es für jedes $x$ nahe an $x_0$ ein eindeutig bestimmtes $y$ nahe an $y_0$ gibt mit $F(x,y) = 0$. Genauer werden wir offene Mengen $U \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}^m$, eine Funktion $F : U \times V \to \mathbb{R}^m$, eine Lösung $(x_0, y_0) \in U \times V$ der Gleichung $F(x_0, y_0) = 0$ betrachten und zeigen, dass unter geeigneten Annahmen an $F$ bei $(x_0, y_0)$ eine lokale Lösungsfunktion $y = f(x)$ existiert, die (mehrmals) stetig differenzierbar ist.

Auf der Suche nach einer inversen Funktion einer injektiven Funktion $f$ von $U \subseteq \mathbb{R}^n$ nach $\mathbb{R}^n$ können wir durch $F : U \times \mathbb{R}^n \to \mathbb{R}^n$ mit $F(x,y) = f(x) - y$ eine Funktion definieren, so dass für gegebenes $y$ die Lösungen $x = x(y)$ der impliziten Gleichung $F(x,y) = 0$ gerade die inverse Funktion beschreiben.

### 11.1.1 Satz zur impliziten Funktion

**Satz 11.1** (Stetige lokale Lösungsfunktion)**.** *Sei $r > 0$ ein Radius und seien $x_0 \in \mathbb{R}^n$, $y_0 \in \mathbb{R}^m$ Punkte. Wir betrachten die offene Teilmenge*

$$B_r^{\mathbb{R}^n}(x_0) \times B_r^{\mathbb{R}^m}(y_0) = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid \|x - x_0\|_2 < r \text{ und } \|y - y_0\|_2 < r\}$$

*von $\mathbb{R}^n \times \mathbb{R}^m$. Sei $F : B_r^{\mathbb{R}^n}(x_0) \times B_r^{\mathbb{R}^m}(y_0) \to \mathbb{R}^m$ eine stetige Funktion, die die folgenden drei Bedingungen erfüllt:*

- $F(x_0, y_0) = 0$.
- *Die partiellen Ableitungen*

  $$\partial_{y_k} F : B_r^{\mathbb{R}^n}(x_0) \times B_r^{\mathbb{R}^m}(y_0) \to \mathbb{R}^m$$

  *existieren für alle $k \in \{1, \ldots, m\}$ und sind auf $B_r^{\mathbb{R}^n}(x_0) \times B_r^{\mathbb{R}^m}(y_0)$ stetig.*
- *Die totale Ableitung $A$ bei $y_0$ der Abbildung $y \in B_r(y_0) \mapsto F(x_0, y)$ ist invertierbar, das heisst, die Matrix $A = (\partial_{y_k} F_j(x_0, y_0))_{j,k} \in \mathrm{Mat}_{m,m}(\mathbb{R})$ hat nicht-verschwindende Determinante.*

*Dann existiert ein offener Ball $U_0 = B_\alpha(x_0)$ um $x_0$ und ein offener Ball $V_0 = B_\beta(y_0)$ um $y_0$ mit $\alpha, \beta \in (0, r)$ und eine stetige Funktion $f : U_0 \to V_0$, so dass für alle $(x, y) \in U_0 \times V_0$ die Gleichung $F(x, y) = 0$ genau dann gilt, wenn $y = f(x)$ gilt. Insbesondere ist $f(x_0) = y_0$.*

Wie wir sehen werden, besteht unserer Beweis des obigen Satzes aus einem (leicht veränderten und auf höhere Dimensionen übertragenen) Newton-Verfahren (siehe Beispiel 8.65) und dem Banachschen Fixpunktsatz, wobei wir $\alpha, \beta > 0$ geeignet wählen werden, um eine Lipschitz-Kontraktion auf einem vollständigen metrischen Raum zu erhalten.

*Beweis.* Da wir zu einem jeweils fest gewählten $x$ ein $y$ mit $F(x, y) = 0$ suchen wollen, wird die Notation $F_x(y) = F(x, y)$ für $(x, y) \in B_r(x_0) \times B_r(y_0)$ nützlich sein. Wir verwenden diese bereits, um für ein festes $x \in B_r(x_0)$ die Hilfsfunktion

$$T_x : y \in B_r(y_0) \mapsto y - A^{-1} F_x(y) \in \mathbb{R}^m$$

zu definieren. Diese entspricht gerade der Iterationsvorschrift im Newton-Verfahren, wobei wir allerdings die Ableitung von $F_x$ bei $y$ schlicht durch $A = \mathrm{D}_{y_0} F_{x_0}$ ersetzt haben. Trotz dieser Änderungen bemerken wir, dass für $(x, y) \in B_r(x_0) \times B_r(y_0)$ die Gleichung $F(x, y) = 0$ zur Fixpunktgleichung $T_x(y) = y$ äquivalent ist (wieso?).

Die Abbildung $T_x$ als Lipschitz-Kontraktion: Sei $x \in B_r(x_0)$. Nach Annahme und Satz 10.10 ist $F_x$ auf $B_r(y_0)$ eine stetig differenzierbare Funktion, womit nach der Kettenregel $T_x$ ebenso stetig differenzierbar ist mit Ableitung

$$\mathrm{D}_y T_x = I_m - A^{-1} \mathrm{D}_y F_x = A^{-1}(A - \mathrm{D}_y F_x)$$

für $y \in B_r(y_0)$, wobei $I_m \in \mathrm{Mat}_{m,m}(\mathbb{R})$ die Identitätsmatrix bezeichnet. Für $x = x_0$ und $y = y_0$ ergibt sich damit

$$\mathrm{D}_{y_0} T_{x_0} = A^{-1}(A - \mathrm{D}_{y_0} F_{x_0}) = 0.$$

Auf Grund der angenommenen Stetigkeit von $(x, y) \in B_r(x_0) \times B_r(y_0) \mapsto \mathrm{D}_y F_x$ existiert also ein $\delta \in (0, r)$, so dass für alle $(x, y) \in B_\delta(x_0) \times B_\delta(y_0)$ die Abschätzung $\|\mathrm{D}_y T_x\|_{\mathrm{op}} \leq \frac{1}{2}$ gilt.

Wir zeigen, dass dies die Lipschitz-Kontraktionseigenschaft impliziert. In der Tat folgt für $x \in B_\delta(x_0)$ und $y_1, y_2 \in B_\delta(y_0)$ mit Hilfe des geraden Weges $\gamma : t \in [0, 1] \mapsto (1 - t)y_1 + ty_2$ von $y_1$ nach $y_2$ (mit Ableitung $y_2 - y_1$)

$$\begin{aligned}\|T_x(y_1) - T_x(y_2)\| &= \|T_x \circ \gamma(1) - T_x \circ \gamma(0)\| = \left\| \int_0^1 (T_x \circ \gamma)'(t)\,\mathrm{d}t \right\| \\ &\leq \int_0^1 \|\mathrm{D}_{\gamma(t)} T_x(y_2 - y_1)\|\,\mathrm{d}t \leq \int_0^1 \|\mathrm{D}_{\gamma(t)} T_x\|_{\mathrm{op}} \|(y_2 - y_1)\|\,\mathrm{d}t \leq \tfrac{1}{2}\|y_2 - y_1\|.\end{aligned}$$

Sei nun $\beta = \frac{\delta}{2}$, $V_0 = B_\beta(y_0)$ und $Y = \overline{B_\beta(y_0)}$. Wir erhalten also, dass für jedes fest gewählte $x \in B_\delta(x_0)$ die eingeschränkte Abbildung (wieder mit $T_x$ bezeichnet)

$$\begin{aligned} T_x : &Y \to \mathbb{R}^m \\ &y \mapsto T_x(y) = y - A^{-1} F_x(y), \end{aligned}$$

die eine Lipschitz-Kontraktion mit Lipschitz-Konstante $\frac{1}{2}$ ist. Um den Banachschen Fixpunktsatz anwenden zu können, müssen wir noch $T_x(Y) \subseteq Y$ zeigen.

VO: 26.03.18

Eine Selbstabbildung: Nach Stetigkeit von $F$ und wegen $F(x_0, y_0) = 0$ existiert ein $\alpha \in (0, \delta)$, so dass für alle $x \in \overline{B_\alpha(x_0)}$ die Abschätzung

$$\|T_x(y_0) - y_0\| = \|A^{-1} F(x, y_0)\| < \tfrac{\beta}{3}$$

gilt. Falls nun $x \in \overline{B_\alpha(x_0)}$ und $y \in Y = \overline{B_\beta(y_0)}$ sind, dann folgt

$$\begin{aligned}\|T_x(y) - y_0\| &= \|T_x(y) - T_x(y_0) + T_x(y_0) - y_0\| \\ &\leq \|T_x(y) - T_x(y_0)\| + \|T_x(y_0) - y_0\| \\ &\leq \tfrac{1}{2}\|y - y_0\| + \tfrac{\beta}{3} \leq \tfrac{5}{6}\beta < \beta. \end{aligned} \tag{11.2}$$

Für $x \in \overline{B_\alpha(x_0)}$ und den oben definierten und nach Proposition 9.52 vollständigen metrischen Raum $Y = \overline{B_\beta(y_0)}$ gilt daher $T_x(Y) \subseteq Y$. Aus dem Banachschen Fixpunktsatz (Satz 9.54) folgt, dass es einen eindeutig bestimmten Punkt $y \in Y$ mit $T_x(y) = y$ gibt. Auf Grund von (11.2) gilt des Weiteren $\|y - y_0\| < \beta$.

Definition der Lösungsfunktion

Zusammenfassend haben wir also gezeigt, dass es für jedes $x \in \overline{B_\alpha(x_0)}$ ein eindeutig bestimmtes $y = y(x) \in \overline{B_\beta(y_0)}$ mit $F(x, y) = 0$ gibt, welches zusätzlich auch $y \in B_\beta(y_0)$ erfüllt.

Dies definiert somit eine Funktion $f : \overline{B_\alpha(x_0)} \to B_\beta(y_0)$ mit der Eigenschaft $F(x, f(x)) = 0$ für alle $x \in \overline{B_\alpha(x_0)}$.

Stetigkeit der Lösungsfunktion: Um die Stetigkeit von $f$ zu zeigen, wiederholen wir obiges Argument für alle $x \in \overline{B_\alpha(x_0)}$ „gleichzeitig“. Wir betrachten dazu die Teilmenge

$$\tilde{Y} = \left\{ g : \overline{B_\alpha(x_0)} \to Y \mid g \text{ ist stetig} \right\} \subseteq C(\overline{B_\alpha(x_0)}, \mathbb{R}^m).$$

Nach Satz 9.86 ist $C(\overline{B_\alpha(x_0)}, \mathbb{R}^m)$ ausgestattet mit der Supremumsnorm

$$\|g\|_\infty = \sup_{x \in \overline{B_\alpha(x_0)}} \|g(x)\|$$

ein vollständiger metrischer Raum. Da $Y$ abgeschlossen ist, folgt des Weiteren, dass $\tilde{Y}$ als abgeschlossene Teilmenge ebenso vollständig ist. (Wieso?)

Zu einer Funktion $g \in \tilde{Y}$ definieren wir nun die Funktion

$$\tilde{T}g : x \in \overline{B_\alpha(x_0)} \mapsto T_x(g(x)) = g(x) - A^{-1}F(x, g(x)),$$

welche auf Grund der Stetigkeit von $g \in \tilde{Y}$ und $F$ wiederum stetig ist. Für $g \in \tilde{Y}$ gilt nach Definition $\|g(x) - y_0\| \leq \beta$ für alle $x \in \overline{B_\alpha(x_0)}$ und somit gilt auf Grund von (11.2) auch $\|\tilde{T}g(x) - y_0\| < \beta$ für alle $x \in \overline{B_\alpha(x_0)}$, wodurch $\tilde{T}g$ ebenso in $\tilde{Y}$ liegt. Schlussendlich gilt für $g_1, g_2 \in \tilde{Y}$ und $x \in \overline{B_\alpha(x_0)}$

$$\|(\tilde{T}g_1 - \tilde{T}g_2)(x)\| = \|T_x(g_1(x)) - T_x(g_2(x))\| \leq \tfrac{1}{2}\|g_1(x) - g_2(x)\| \leq \tfrac{1}{2}\|g_1 - g_2\|_\infty,$$

da $T_x$ Lipschitz-stetig ist mit Lipschitz-Konstante $\frac{1}{2}$.

Dies zeigt, dass $\tilde{T} : \tilde{Y} \to \tilde{Y}$ eine Lipschitz-Kontraktion auf einem vollständigen metrischen Raum darstellt und damit nach dem Banachschen Fixpunktsatz einen eindeutig bestimmten Fixpunkt besitzt. Sei $y = \tilde{T}y$ dieser Fixpunkt. Dann ist $y$ stetig und für alle $x \in \overline{B_\alpha(x_0)}$ gilt $y(x) = \tilde{T}y(x) = T_x(y(x)) \in B_\beta(y_0)$, wodurch $y(x) = f(x)$ die eindeutige Lösung der Gleichung $F(x, y) = 0$ mit $y \in \overline{B_\beta(y_0)}$ sein muss. □

Wie wir in obigem Satz gesehen haben, „erbt“ die Lösungsfunktion $f$ die Stetigkeitseigenschaft von der Funktion $F$, die die implizite Gleichung definiert. Wir möchten nun zeigen, dass dasselbe auch für Differenzierbarkeit zutrifft.

**Satz 11.2** (Differenzierbarkeit der lokalen Lösungsfunktion)**.** *Seien $r > 0$, $x_0 \in \mathbb{R}^n$, $y_0 \in \mathbb{R}^m$ und $F : B_r(x_0) \times B_r(y_0) \to \mathbb{R}^m$ mit den Eigenschaften aus Satz 11.1 und sei $f : U_0 \to V_0$ die stetige lokale Lösungsfunktion aus Satz 11.1. Angenommen $F$ ist $d$-mal stetig differenzierbar für $d \geq 1$. Dann ist die stetige Lösungsfunktion $f$ ebenso $d$-mal stetig differenzierbar und die Ableitung von $f$ bei $x \in U_0$ ist durch*

$$\mathrm{D}_x f = -((\boldsymbol{\partial}_y F)(x, f(x)))^{-1}(\boldsymbol{\partial}_x F)(x, f(x)) \tag{11.3}$$

*gegeben.*

Dabei bezeichnet $\boldsymbol{\partial}_x F(x,y)$ die Matrix der partiellen Ableitungen von $F$ in den Koordinatenrichtungen der Variable $x \in B_r(x_0)$ ausgewertet an der Stelle $(x,y)$, das heisst,

$$\boldsymbol{\partial}_x F(x,y) = \big(\partial_{x_j} F_i(x,y)\big)_{ij} \in \mathrm{Mat}_{m,n}(\mathbb{R}).$$

Alternativ könnte man, konform mit unserer bisherigen Notation, auch schreiben

$$\boldsymbol{\partial}_x F(x,y) = \mathrm{D}_x(F(\cdot,y)),$$

womit also $\boldsymbol{\partial}_x F(x,y)$ die Ableitung nach $x$ unter Festhalten von $y$ an der Stelle $(x,y)$ bezeichnet. Analog ist $\boldsymbol{\partial}_y F(x,y)$ die Ableitung nach $y$ unter Festhalten von $x$ an der Stelle $(x,y)$

$$\boldsymbol{\partial}_y F(x,y) = \big(\partial_{y_j} F_i(x,y)\big)_{ij} \in \mathrm{Mat}_{m,m}(\mathbb{R}).$$

Wir bemerken noch, dass a priori die Existenz der Inversen dieser Matrix an der Stelle $(x, f(x))$ nicht klar ist und im Beweis unten nachgewiesen wird.

Falls bereits bekannt wäre, dass $f$ in der Tat differenzierbar ist, so würde (11.3) auch aus der Kettenregel folgen. Um dies zu sehen, nehmen Sie zuerst $m = n = 1$ an und wenden die Kettenregel auf $F(x, f(x)) = 0$ an. Der allgemeine Fall ist bloss in der Notation schwieriger.

**Beispiel 11.3** (Zum Differential der impliziten Funktion)**.**

*(i) Der Einheitskreis $\mathbb{S}^1$ in $\mathbb{R}^2$ ist definiert durch die Gleichung $x^2 + y^2 = 1$. Nahe jedem Punkt $(x_0, y_0) \in \mathbb{S}^1$ mit $x_0 \neq 0$ ist diese Gleichung nach der Variablen $x$ auflösbar und für die so definierte implizite Funktion $x = x(y)$ erhalten wir durch Ableiten der definierenden Gleichung nach $y$ und Auswerten im Punkt $(x_0, y_0)$*

$$2x_0 x'(y_0) + 2y_0 = 0,$$

*also $x'(y_0) = -\frac{y_0}{x_0}$.*

*(ii) Die Funktion $F\colon \mathbb{R}^4 \to \mathbb{R}^2$ sei definiert durch*

$$F(x,y,z,w) = \begin{pmatrix} x^2 + y + z - 2 \\ y - 2z + w^3 \end{pmatrix}$$

*für $(x,y,z,w)^t \in \mathbb{R}^4$. Wir wollen die Auflösbarkeit der Gleichung $F(x,y,z,w) = 0$ in einer Umgebung des Punktes $(x_0, y_0, z_0, w_0) = (0,1,1,1)$ untersuchen. Dazu betrachten wir das Differential*

$$\mathrm{D}_{(x_0,y_0,z_0,w_0)} F = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 1 & -2 & 3 \end{pmatrix}.$$

*Der Satz über implizite Funktionen besagt nun, dass wir die Gleichung $F(x,y,z,w) = 0$ nahe $(x_0, y_0, z_0, w_0)$ nach zwei der Variablen $x, y, z, w$ auflösen können, sofern die $2 \times 2$-Teilmatrix obigen Differentials bestehend aus den zu diesen Variablen gehörigen partiellen*

*Ableitungen invertierbar ist. Zum Beispiel ist dies der Fall für das Variablenpaar $y$ und $w$, da die Teilmatrix $\left(\begin{smallmatrix}1&0\\1&3\end{smallmatrix}\right)$ bestehend aus der 2. und 4. Spalte invertierbar ist. Die Ableitung der so definierten impliziten Funktionen $y = y(x,z)$ und $w = w(x,z)$ in $(x_0,z_0) = (0,1)$ ist gemäss der Formel in Satz 11.2 dann gegeben durch*

$$\mathrm{D}_{(x_0,z_0)}\begin{pmatrix}y\\w\end{pmatrix} = -\begin{pmatrix}1&0\\1&3\end{pmatrix}^{-1}\begin{pmatrix}0&1\\0&-2\end{pmatrix} = \begin{pmatrix}0&-1\\0&1\end{pmatrix}.$$

*Beweis von Satz 11.2.* Nach Annahme in Satz 11.1 ist $A = \mathrm{D}_{y_0}F_{x_0}$ invertierbar, wobei wir wieder die Notation $F_{x_0}(y) = F(x_0,y)$ für $y \in B_r(y_0)$ aus dem Beweis von Satz 11.1 verwenden. Seien weiter $T_x$, $\delta > 0$ mit $\|\mathrm{D}_yT_x\| \leq \frac{1}{2}$ für alle $(x,y) \in B_\delta(x_0) \times B_\delta(y_0)$ sowie $\alpha$ und $\beta$ ebenso wie im letzten Beweis. Dann gilt $\mathrm{D}_yT_x = I_m - A^{-1}(\mathrm{D}_yF_x)$ und dadurch auch

$$A^{-1}(\mathrm{D}_yF_x) = I_m - \mathrm{D}_yT_x$$

für alle $(x,y) \in B_\delta(x_0)\times B_\delta(y_0)$. Daraus folgt (siehe Übung 11.4), dass $A^{-1}(\mathrm{D}_yF_x)$ invertierbar ist und

$$\big(A^{-1}(\mathrm{D}_yF_x)\big)^{-1} = \big(I_m - \mathrm{D}_yT_x\big)^{-1} = \sum_{j=0}^{\infty}(\mathrm{D}_yT_x)^j.$$

Insbesondere ist $\boldsymbol{\partial}_yF(x,y) = \mathrm{D}_yF_x = A(A^{-1}(\mathrm{D}_yF_x))$ für alle $(x,y) \in B_\delta(x_0) \times B_\delta(y_0)$ als Produkt invertierbarer Matrizen invertierbar. Da $\alpha,\beta \in (0,\delta)$ sind und $f(B_\alpha(x_0)) \subseteq B_\beta(y_0)$ ebenso in $B_\delta(y_0)$ enthalten ist, folgt insbesondere, dass der Ausdruck

$$\big(\boldsymbol{\partial}_yF(x,f(x))\big)^{-1} = \big(\mathrm{D}_{f(x)}F_x\big)^{-1}$$

in (11.3) für alle $x \in B_\alpha(x_0)$ definiert ist.

Um Differenzierbarkeit von $f$ und (11.3) zu beweisen, fixieren wir $x \in B_\alpha(x_0)$ und definieren $A_x = \boldsymbol{\partial}_yF(x,f(x))$ sowie $L_x = -A_x^{-1}\boldsymbol{\partial}_xF(x,f(x))$. Für genügend kleine $h \in \mathbb{R}^n$ mit $x+h$ in $B_\alpha(x_0)$ erhalten wir damit

$$\begin{aligned}\|f(x+h) - f(x) - L_x(h)\| &= \|f(x+h) - f(x) + A_x^{-1}\boldsymbol{\partial}_xF(x,f(x))h\|\\ &\leq \|A_x^{-1}\|_{\mathrm{op}}\|\boldsymbol{\partial}_xF(x,f(x))h + \boldsymbol{\partial}_yF(x,f(x))(f(x+h)-f(x))\|,\end{aligned}$$

wobei wir $A_x^{-1}$ herausgehoben haben und die Definition der Matrixnorm verwendeten. Wir bemerken nun, dass der Ausdruck $\boldsymbol{\partial}_xF(x,f(x))h + \boldsymbol{\partial}_yF(x,f(x)(f(x+h)-f(x))$ genau die totale Ableitung von $F$ in $(x,f(x))$ angewendet auf den Vektor $\begin{pmatrix}h\\f(x+h)-f(x)\end{pmatrix}$ ist. Als nächstes ändern wir das Vorzeichen dieses Terms innerhalb der Norm und fügen die Funktionswerte

von $F$ bei $\binom{x}{f(x)}$ beziehungsweise $\binom{x+h}{f(x+h)}$ hinzu. Dies führt zu

$$\begin{aligned}
&\|f(x+h)-f(x)-L_x(h)\| \\
&\qquad \le \|A_x^{-1}\|_{\mathrm{op}} \Big\| \underbrace{F\Big(\binom{x+h}{f(x+h)}\Big)}_{=0} - \underbrace{F\Big(\binom{x}{f(x)}\Big)}_{=0} - \mathrm{D}_{(x,f(x))}F\Big(\binom{h}{f(x+h)-f(x)}\Big)\Big\| \\
&\qquad \le \|A_x^{-1}\|_{\mathrm{op}}\ \alpha(h)\ \Big\|\binom{h}{f(x+h)-f(x)}\Big\| \\
&\qquad \le \|A_x^{-1}\|_{\mathrm{op}}\ \alpha(h)\ \big(\|h\| + \|f(x+h)-f(x)\|\big) \qquad (11.4)
\end{aligned}$$

mit $\alpha(h) = o(1)$ für $h \to 0$, da $F$ bei $(x, f(x))$ differenzierbar ist und $f$ stetig ist. Dies wäre die gewünschte Abschätzung für die Differenzierbarkeit von $f$, wenn rechts nicht der Term $\|f(x+h)-f(x)\|$ auftauchen würde. Wir können diesen jedoch mit einer kleinen Rechnung loswerden, wozu wir $a = \|A_x^{-1}\|_{\mathrm{op}}$ und $b = \|L_x\|_{\mathrm{op}}$ setzen, um den Fokus auf das Wesentliche zu lenken. Es gilt

$$\begin{aligned}
\|f(x+h)-f(x)\| &= \|f(x+h)-f(x)-L_x(h)+L_x(h)\| \\
&\le \|f(x+h)-f(x)-L_x(h)\| + \|L_x(h)\| \\
&\le a\alpha(h)\big(\|h\| + \|f(x+h)-f(x)\|\big) + b\|h\|
\end{aligned}$$

und daher auch

$$\big(1 - a\alpha(h)\big)\|f(x+h)-f(x)\| \le \big(a\alpha(h)+b\big)\|h\|.$$

Da $\alpha(h) = o(1)$ für $h \to 0$, gibt es nun ein $\rho \in (0,1)$ klein genug, so dass $a\alpha(h) \le \frac{1}{2}$ für alle $h \in B_\rho(0)$. Damit ergibt sich $1 - a\alpha(h) \ge \frac{1}{2}$ und

$$\|f(x+h)-f(x)\| \le 2\big(a\alpha(h)+b\big)\|h\| \le c\|h\|$$

für $c = 2(\frac{1}{2}+b)$. Wir setzen dies in (11.4) ein und erhalten

$$\|f(x+h)-f(x)-L_x(h)\| \le a\alpha(h)(1+c)\|h\|$$

für $h \in B_\rho(0)$. Da $\alpha(h) = o(1)$ für $h \to 0$, ergibt dies die Differenzierbarkeit von $f$ sowie die Gleichung (11.3) für jedes $x \in B_\alpha(x_0)$.

Sei nun $F$ zweimal stetig differenzierbar. Nach Gleichung (11.3) gilt, dass die partiellen Ableitungen von $f$ sich durch die Inversenbildung einer invertierbaren Matrix und den ersten partiellen Ableitungen von $F$ ausgewertet bei $(x, f(x))$ ausdrücken lassen. Wir bemerken, dass die Einträge der Inverse einer Matrix sich wiederum durch ein Polynom in den Einträgen der Matrix und der Inversen der Determinante ausdrücken lassen (nach der Cramer'schen Regel). Das heisst, die partiellen Ableitungen lassen sich als Summen von Produkten darstellen, wobei jeder Faktor entweder der Kehrwert von $h(x) = \det\big(\boldsymbol{\partial}_y F(x, f(x))\big)$ ist oder eine erste partiellen Ableitung von $F$ ausgewertet bei $(x, f(x))$ ist. Da wir bereits wissen, dass $f$ differenzierbar ist

mit stetiger Ableitung, können wir nun diese partiellen Ableitungen nochmals ableiten: Summen von Produkten werden von einer partiellen Ableitung wieder auf Summen von Produkten abgebildet, wobei die Faktoren entweder unverändert bleiben oder diese partiell abgeleitet werden. Nach der Kettenregel hat aber jede partielle Ableitung von $F$ ausgewertet bei $(x, f(x))$ wieder eine stetige partielle Ableitung (da ja $f$ stetig differenzierbar ist). Dieses Argument gilt ebenso für $h(x)$ und es folgt, dass die partielle Ableitung von $h(x)^{-1}$ gleich $-h(x)^{-2}$ mal der partiellen Ableitung von $h(x)$ ist. Es folgt also, dass jede zweite partielle Ableitung existiert und als eine Summe von Produkten dargestellt werden kann, wobei die Faktoren entweder eine Potenz von $h^{-1}$, erste oder zweite partielle Ableitungen von $F$ ausgewertet bei $(x, f(x))$ sind.

Mit Induktion nach $d$ ergibt sich dieselbe Aussage für alle partiellen Ableitungen bis zur Ordnung $d$, da $F$ als $d$-mal stetig differenzierbar vorausgesetzt wurde. □

Die obigen beiden Sätze zur impliziten Funktion (Satz 11.1 und Satz 11.2) hatten eher lange Beweise. Doch werden wir sehen, dass im restlichen Kapitel diese als wichtige Grundlage dienen und alle weitere Diskussionen dadurch viel einfacher werden.

**Übung 11.4** (Geometrische Reihe von Matrizen)**.** *Es bezeichne $\|\cdot\|_{\mathrm{op}}$ die Matrixnorm (siehe Definition 9.75).*

*(i) Zeigen Sie, dass für $B \in \mathrm{Mat}_{m,m}(\mathbb{R})$ mit $\|B\|_{\mathrm{op}} < 1$ die Matrix $I_m - B$ invertierbar ist mit*

$$(I_m - B)^{-1} = \sum_{j=0}^{\infty} B^j.$$

*Die Reihe $\sum_{j=0}^{\infty} B^j$ ist dabei als Grenzwert der Folge $\left(\sum_{j=0}^{n} B^j\right)_n$ aufzufassen; ein Teil der Aussage ist also, dass diese Folge konvergiert (unter den getroffenen Annahmen).*

*(ii) Zeigen Sie, dass die Menge der invertierbaren Matrizen in $\mathrm{Mat}_{m,m}(\mathbb{R})$ offen ist und dass zu $B \in \mathrm{GL}_m(\mathbb{R})$ jedes $C \in \mathrm{Mat}_{m,m}(\mathbb{R})$ mit $\|C - B\|_{\mathrm{op}} < \|B^{-1}\|_{\mathrm{op}}^{-1}$ ebenfalls invertierbar ist.*

VO: 28.03.18

Ostern

### 11.1.2 Satz zur inversen Abbildung

Wir möchten nun die Differenzierbarkeit der inversen Abbildung vom eindimensionalen Fall in Satz 7.14 auf Abbildungen $f : U \subseteq \mathbb{R}^n \to f(U) \subseteq \mathbb{R}^n$ verallgemeinern. Wie zu Beginn dieses Abschnitts erwähnt wurde, werden wir dazu den Satz der impliziten Funktion (Satz 11.2) verwenden.

**Satz 11.5** (Satz zur inversen Abbildung)**.** *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^n$ eine $d$-mal stetig differenzierbare Funktion mit $d \geq 1$. Sei $x_0 \in U$ mit invertierbarer totaler Ableitung $\mathrm{D}_{x_0} f \in \mathrm{Mat}_{n,n}(\mathbb{R})$ (das heisst, $x_0$ ist ein regulärer Punkt von $f$). Dann gibt es eine offene Umgebung $U_0 \subseteq U$ und eine offene Umgebung $V_0 \subseteq f(U)$ von $y_0 = f(x_0)$, so dass $f|_{U_0} : U_0 \to V_0$ bijektiv*

*ist und die Umkehrabbildung ebenso d-mal stetig differenzierbar ist. Des Weiteren gilt*

$$\mathrm{D}_y(f^{-1}) = (\mathrm{D}_x f)^{-1}$$

*für alle* $x \in U_0$ *und* $y = f(x) \in V_0$.

Die zentrale Bedingung in obigem Satz ist wohlgemerkt die Invertierbarkeit der totalen Ableitung bei $x_0$. Diese verallgemeinert die zentrale Annahme in Satz 7.14, wo verlangt wurde, dass die Ableitung nicht verschwindet.

Satz 11.5 wird es uns in gewissen Fällen auch erlauben, eine globale Inverse zu finden.

**Definition 11.6** (Diffeomorphismus)**.** Seien $U, V \subseteq \mathbb{R}^n$ offen. Eine bijektive, glatte Funktion $f : U \to V$ mit glatter Inversen $f^{-1} : V \to U$ wird ein (glatter) **Diffeomorphismus** genannt. Sind $f$ und $f^{-1}$ jeweils bloss $d$-mal stetig differenzierbar für $d \geq 1$, so nennen wir $f$ einen $C^d$**-Diffeomorphismus**.

Satz 11.5 besagt also auch, dass eine glatte Funktion bei einem regulären Punkt lokal ein Diffeomorphismus auf eine offene Teilmenge im Bild ist.

*Beweis von Satz 11.5.* Sei $r > 0$ ein Radius mit $B_r(x_0) \subseteq U$. Wir definieren $F : B_r(x_0) \times B_r(y_0) \to \mathbb{R}^n$ durch

$$F(x, y) = f(x) - y$$

für $(x, y) \in B_r(x_0) \times B_r(y_0)$ und wir wollen die Gleichung $F(x, y) = 0$ nach $x$ auflösen. Hierzu bemerken wir, dass $F$ $d$-mal stetig differenzierbar ist, dass $F(x_0, y_0) = 0$ ist und dass $\boldsymbol{\partial}_x F(x_0, y_0) = \mathrm{D}_{x_0} f$ per Annahme invertierbar ist. Daher erfüllt $F$ alle Voraussetzungen des Satzes über implizite Funktionen (Satz 11.1 und Satz 11.2, wobei $x$ und $y$ vertauschte Rollen einnehmen). Es folgt, dass es Radien $\alpha, \beta \in (0, r)$ und eine $d$-mal stetig differenzierbare Funktion $g : V_0 = B_\beta(y_0) \to \tilde{U} = B_\alpha(x_0)$ gibt, sodass für alle $(x, y) \in \tilde{U} \times V_0$ die Äquivalenzen

$$y = f(x) \iff F(x, y) = 0 \iff x = g(y) \tag{11.5}$$

gelten. Wir definieren $U_0 = \tilde{U} \cap f^{-1}(V_0)$, welche als Durchschnitt zweier offener Mengen (siehe Proposition 9.37) wieder offen ist. Aus (11.5) folgt nun, dass $f|_{U_0} : U_0 \to V_0$ invertierbar ist (und insbesondere $V_0 \subseteq f(U)$ ist) und dass $g = (f|_{U_0})^{-1} : V_0 \to U_0$ die inverse Abbildung ist.

Aus Satz 11.2 folgt weiters, dass $g$ $d$-mal stetig differenzierbar ist und dass für $x \in U_0$ und $y = f(x) \in V_0$

$$\mathrm{D}_y g = (\mathrm{D}_x f)^{-1}$$

(was auch aus der Kettenregel und $g \circ f|_{U_0} = I_{U_0}$ folgt). Dies beendet den Beweis des Satzes. □

**Übung 11.7.** *Verwenden Sie den Satz zur inversen Abbildung (Satz 11.5), um den Satz zur impliziten Abbildung (Satz 11.1 und Satz 11.2) zu beweisen.*

Wir wollen hier noch betonen, dass obiger Existenzsatz für inverse Abbildungen wirklich nur eine lokale Aussage liefert. In mehr als einer Dimension gibt es kein lokales Kriterium für die Injektivität einer Abbildung $f : U \to V$ (im Gegensatz zur Kombination von Korollar 7.35 und Satz 3.65). Beispielsweise ist die komplexe Exponentialabbildung $\exp : \mathbb{C} \to \mathbb{C}$ lokal injektiv, aber nicht injektiv (selbiges gilt auch für die Polarkoordinatenabbildung $(r, \varphi) \in (0, \infty) \times \mathbb{R} \mapsto (r \cos \varphi, r \sin \varphi) \in \mathbb{R}^2$). Können wir Injektivität aus anderen Überlegungen erhalten, so kann der Satz der inversen Abbildung (Satz 11.5) trotzdem global nützlich sein.

**Korollar 11.8** (Kriterium für Diffeomorphie)**.** *Sei $U \subseteq \mathbb{R}^n$ offen und sei $f : U \to \mathbb{R}^n$ eine $d$-mal stetig differenzierbare, injektive Funktion mit $d \geq 1$. Angenommen jeder Punkt $x \in U$ hat die Eigenschaft, dass $\mathrm{D}_x f$ invertierbar ist (oder in anderen Worten: jeder Punkt in $U$ ist regulär). Dann ist $V = f(U) \subseteq \mathbb{R}^n$ offen und $f : U \to V$ ist ein $C^d$-Diffeomorphismus mit*

$$\mathrm{D}_y(f^{-1}) = (\mathrm{D}_x f)^{-1}$$

*für alle $x \in U$ und $y = f(x) \in V$.*

*Beweis.* Wir zeigen zuerst, dass $V = f(U)$ offen ist. Sei also $y_0 = f(x_0)$ für ein $x_0 \in U$. Da $\mathrm{D}_{x_0} f$ invertierbar ist, können wir den Satz der inversen Abbildung (Satz 11.5) anwenden und erhalten offene Umgebungen $U_0$ von $x_0$ und $V_0$ von $y_0$, so dass $f|_{U_0} : U_0 \to V_0$ ein Diffeomorphismus ist. Insbesondere ist $V_0 = f(U_0) \subseteq f(U) = V$, womit $V$ offen ist, da $y_0 \in V$ beliebig war (eine beliebige Vereinigung offener Mengen ist offen). Des Weiteren stimmt $(f|_{U_0})^{-1} : V_0 \to U_0$ mit $f^{-1}|_{V_0} : V_0 \to U_0$ überein, wobei $f^{-1} : V \to U$ auf Grund der vorausgesetzten Injektivität existiert.

Wir erhalten, dass $f^{-1}$ auf $V_0$ also $d$-mal stetig differenzierbar ist. Da $y_0 \in V$ allerdings beliebig war und da stetige Differenzierbarkeit eine lokale Eigenschaft ist, zeigt dies, dass $f^{-1}$ $d$-mal stetig differenzierbar ist. Somit ist $f : U \to V$ ein $C^d$-Diffeomorphismus. □

**Beispiel 11.9.** *Wir betrachten hier die Abbildung*

$$f : \mathbb{R}^2 \to \mathbb{R}^2, \quad (x, y)^t \mapsto (x + y^2, y + x^2)^t$$

*und wollen geeignete offene Mengen $U \subseteq \mathbb{R}^2$, $V \subseteq \mathbb{R}^2$ finden, so dass $f|_U : U \to V$ ein Diffeomorphismus ist. Die Jacobi-Matrix von $f$ bei $(x, y)^t \in \mathbb{R}^2$ ist durch*

$$\mathrm{D}_{(x,y)} f = \begin{pmatrix} 1 & 2y \\ 2x & 1 \end{pmatrix}$$

*gegeben und deren Determinante (auch Jacobi-Determinante genannt) ist durch $1 - 4xy$ gegeben. Das heisst, für alle Punkte $(x, y)^t \in \mathbb{R}^2$, die nicht auf der Hyperbel $xy = \frac{1}{4}$ liegen, existiert nach dem Satz über die inverse Abbildung (Satz 11.5) eine Umgebung, so dass $f$ eingeschränkt auf diese Umgebung ein Diffeomorphismus auf eine andere offene Menge ist. Wir wollen hier aber das Quadrat $U = (-\frac{1}{2}, \frac{1}{2}) \times (-\frac{1}{2}, \frac{1}{2})$ zwischen den beiden Punkten $(-\frac{1}{2}, -\frac{1}{2})$ und $(\frac{1}{2}, \frac{1}{2})$ auf der Hyperbel betrachten.*

*Wir behaupten, dass $f|_U$ injektiv ist. Aus dieser Behauptung und Korollar 11.8 folgt dann, dass $f|_U : U \to f(U)$ ein Diffeomorphismus ist, da die Jacobi-Matrix auf $U$ nicht-verschwindende Determinante besitzt.*

*Seien also $(x_1, y_1), (x_2, y_2) \in U$ mit $f(x_1, y_1) = f(x_2, y_2)$. Dann gilt*

$$\left.\begin{aligned} x_1 + y_1^2 &= x_2 + y_2^2 \\ y_1 + x_1^2 &= y_2 + x_2^2 \end{aligned}\right\} \Longrightarrow \begin{aligned} |x_1 - x_2| &= |y_2^2 - y_1^2| = |y_1 - y_2||y_1 + y_2| \\ |y_1 - y_2| &= |x_2^2 - x_1^2| = |x_1 - x_2||x_1 + x_2| \end{aligned}$$

*Da aber $|x_1 + x_2| < 1$ und $|y_1 + y_2| < 1$ gelten, ist dies nur möglich, wenn $x_1 = x_2$ und $y_1 = y_2$ (wieso?).*

*Da die Jacobi-Matrix nur auf der Hyperbel $xy = \frac{1}{4}$ verschwindet, könnte man vermuten, dass obige Wahl von $U$ als Definitionsbereichs des Diffeomorphismus nicht optimal ist und man diesen deutlich vergrössern könnte. Doch ebenso wie bei den Polarkoordinaten gibt es keine natürliche maximale Wahl des Definitionsbereichs, da zum Beispiel $f(1,0) = (1,1)^t = f(0,1)$ und die Punkte $(1,0)^t$, $(0,1)^t$ in derselben Zusammenhangskomponente des Komplements der Hyperbel liegen.*

### 11.1.3 Polar- und Zylinderkoordinaten

Die **Polarkoordinatenabbildung** ist durch

$$\begin{aligned} f : (0,\infty) \times (-\pi,\pi) &\to \mathbb{R}^2 \setminus ((-\infty,0] \times \{0\}) \\ (r,\varphi) &\mapsto \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \end{pmatrix} \end{aligned}$$

definiert und bildet einen Diffeomorphismus. In der Tat ist $f$ bijektiv (nach Abschnitt 6.6.4) und hat bei $(r,\varphi) \in (0,\infty) \times (-\pi,\pi)$ die Jacobi-Matrix

$$\begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}$$

mit Determinante $r \neq 0$, womit nach Korollar 11.8 die Abbildung $f$ ein Diffeomorphismus ist.

Des Weiteren kann man natürlich $f$ auch auf den Punkten in $\{0\} \times \mathbb{R}$ durch $(0,0)^t$ definieren, doch wäre $f$ bei diesen nicht lokal invertierbar.

Eine dreidimensionale Verallgemeinerung der Polarkoordinaten stellen die **Zylinderkoordinaten** dar. Die entsprechende Abbildung ist der Diffeomorphismus

$$\begin{aligned} f : (0,\infty) \times (-\pi,\pi) \times \mathbb{R} &\to \mathbb{R}^3 \setminus ((-\infty,0] \times \{0\} \times \mathbb{R}) \\ (r,\varphi,z) &\mapsto \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \\ z \end{pmatrix} \end{aligned}$$

mit Jacobi-Matrix

$$\begin{pmatrix} \cos\varphi & -r\sin\varphi & 0 \\ \sin\varphi & r\cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

und Jacobi-Determinante $r$ bei $(r, \varphi, z) \in (0, \infty) \times (-\pi, \pi) \times \mathbb{R}$.

### 11.1.4 Kugelkoordinaten

Eine weitere Verallgemeinerung der Polarkoordinaten für den dreidimensionalen Raum stellen die **Kugelkoordinaten** dar. Hier werden ein Radius $r \in (0, \infty)$ und zwei Winkel $\theta \in (0, \pi)$, $\varphi \in (-\pi, \pi)$ verwendet, wobei $\theta$ den Winkel eines Punktes relativ zum Nordpol der Sphäre $\mathbb{S}^2$ und $\varphi$ den Winkel der Projektion auf die $xy$-Ebene relativ zu $(1, 0, 0)^t$ angeben soll. Der entsprechende Diffeomorphismus ist also gegeben durch

$$\begin{aligned} f : (0, \infty) \times (0, \pi) \times (-\pi, \pi) &\to \mathbb{R}^3 \setminus ((-\infty, 0] \times \{0\} \times \mathbb{R}) \\ (r, \theta, \varphi) &\mapsto \begin{pmatrix} r\sin\theta\cos\varphi \\ r\sin\theta\sin\varphi \\ r\cos\theta \end{pmatrix}. \end{aligned}$$

Für $(r, \theta, \varphi) \in (0, \infty) \times (0, \pi) \times (-\pi, \pi)$ ist die Jacobi-Matrix durch

$$\mathrm{D}_{(r,\theta,\varphi)} f = \begin{pmatrix} \sin\theta\cos\varphi & r\cos\theta\cos\varphi & -r\sin\theta\sin\varphi \\ \sin\theta\sin\varphi & r\cos\theta\sin\varphi & r\sin\theta\cos\varphi \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix}$$

gegeben und die Jacobi-Determinante (nach der Regel von Sarrus) durch

$$\begin{aligned} \det\left(\mathrm{D}_{(r,\theta,\varphi)} f\right) &= r^2 \sin\theta\cos^2\theta\cos^2\varphi + r^2\sin^3\theta\sin^2\varphi \\ &\quad + r^2\sin\theta\cos^2\theta\sin^2\varphi + r^2\sin^3\theta\cos^2\varphi \\ &= r^2(\sin\theta\cos^2\theta + \sin^3\theta) = r^2\sin\theta \neq 0 \end{aligned}$$

gegeben.

VO: 09.04.18

## 11.2 Teilmannigfaltigkeiten des Euklidschen Raumes

Falls wir die Extremwerte einer stetig differenzierbaren, reellwertigen Funktion $f$ eingeschränkt auf die abgeschlossene Einheitskugel

$$K = \{v \in \mathbb{R}^n \mid \|v\| \leq 1\}$$

finden wollen, dann existieren diese zwar wegen Kompaktheit von $K$ (nach Satz 9.66(5)), aber die Methoden von Kapitel 10 können nur die Extremwerte von $f$ im offenen Ball $B_1(0) = \{v \in \mathbb{R}^n \mid \|v\| < 1\}$ finden (welche vielleicht nicht existieren). Dieses Problem drängt die Untersuchung des Verhaltens der Funktion $f$ auf der Sphäre $\mathbb{S}^2 = K \setminus B_1(0)$ auf.

Ähnliche und auch andere Probleme führen dazu, dass man Teilmannigfaltigkeiten von $\mathbb{R}^n$ wie zum Beispiel $\mathbb{S}^2$ (analytisch) studieren möchte. Eine Teilmannigfaltigkeit des $\mathbb{R}^n$ sollte man sich als Teilmenge vorstellen, die lokal jeweils wie eine offene Teilmenge eines Euklidischen Raumes (mit meist niegriger Dimension) aussieht und eine gewisse „Glattheit“ erfüllt. Beispielsweise kann man sich intuitiv gut vorstellen, dass man um jeden Punkt auf der Sphäre eine Umgebung findet, die sich durch „Flachdrücken“ zu einer offenen Kreisscheibe im $\mathbb{R}^2$ formen lässt.

Wir werden in unseren Diskussionen keine „Kanten“ oder „Ecken“ einer Teilmannigfaltigkeit erlauben, da sich diese nicht auf „glatte Weise“ flach machen lassen. Zum Beispiel ist der Rand $\big\{(x, y, z) \in \mathbb{R}^3 \mid \max\{|x|, |y|, |z|\} = 1\big\}$ des Quaders $[-1, 1]^3$ keine Teilmannigfaltigkeit des $\mathbb{R}^3$.

Allgemeinere Mannigfaltigkeiten sind von fundamentaler geometrischer Bedeutung und erscheinen deswegen zum einen in sehr vielen fortführenden Vorlesungen des Mathematikstudiums wie zum Beispiel der Differentialgeometrie im dritten oder vierten Studienjahr, aber auch in Physik-Vorlesungen wie zum Beispiel der allgemeinen Relativitätstheorie. Wir begnügen uns hier mit der Untersuchung von Teilmannigfaltigkeiten des Euklidischen Raumes.

### 11.2.1 Definition und Beispiele

**Definition 11.10.** Sei $0 \leq k \leq n$ für $n \geq 1$. Eine Teilmenge $M \subseteq \mathbb{R}^n$ ist eine $k$-dimensionale (glatte) **Teilmannigfaltigkeit**, falls für jeden Punkt $p \in M$ eine offene Umgebung $U_p$ in $\mathbb{R}^n$ von $p$ und ein Diffeomorphismus $\varphi_p : U_p \to V_p = \varphi_p(U_p)$ auf eine weitere offene Teilmenge $V_p \subseteq \mathbb{R}^n$ existiert, so dass

$$\varphi_p(U_p \cap M) = \{y \in V_p \mid y_i = 0 \text{ für alle } i > k\}\,.$$

In anderen Worten sieht eine Teilmannigfaltigkeit $M$ – in der Nähe eines jeden Punktes in $M$ und bis auf einen Diffeomorphismus – wie ein Teil des Teilraumes $\mathbb{R}^k \times \{0\}^{n-k}$ von $\mathbb{R}^n$ aus. Wenn wir so wollen, können wir nach Anpassen des obigen Diffeomorphismus weitere Bedingungen an $U_p$ oder $V_p$ stellen. Zum Beispiel könnten wir verlangen, dass $U_p = B_\varepsilon(p)$ für ein geeignetes $\varepsilon > 0$ (durch Einschränken von $\varphi_p$) oder dass $V_p = (-\varepsilon, \varepsilon)^n$ für ein geeignetes $\varepsilon > 0$ (durch Verschieben von $\varphi_p$, um $\varphi_p(p) = 0$ zu erreichen, und durch Einschränken). Letzteres legt den Vergleich von $\varphi_p$ zu einer Kartenabbildung nahe, wobei $(-\varepsilon, \varepsilon)^k$ eine „quadratische

Karte“ und

$$(y_1, \ldots, y_k) \in (-\varepsilon, \varepsilon)^k \mapsto \varphi_p^{-1}(y_1, \ldots, y_k, 0, \ldots, 0) \in M$$

die Kartenabbildung darstellt, deren Bild eine Umgebung von $p \in M$ ist.

Alternativ lässt sich $\varphi_p$ als eine Abbildung auffassen, die, wie zu Beginn dieses Abschnitts erklärt wurde, die Umgebung $U_p$ von $p$ „verbiegt“, so dass $U_p \cap M$ „flach“ wird. Hierbei ist eine offene Teilmenge des Teilraumes $\mathbb{R}^k \times \{0\}^{n-k}$ sozusagen flach. Des Weiteren können wir uns die Umkehrabbildung $\psi = \varphi^{-1}$ nach Einschränkung auf $V_p \cap (\mathbb{R}^k \times \{0\}^{n-k})$ als eine Parametrisierung von $M \cap U_p$ vorstellen, und ohne Einschränkung als eine Parametrisierung der (im $\mathbb{R}^n$ offenen) Umgebung $U_p$ der Punkte in $M \cap U_p$.

**Beispiel 11.11** (Erste Beispiele von Teilmannigfaltigkeiten)**.**

*(a) Jede offene Teilmenge $U$ in $\mathbb{R}^n$ ist eine $n$-dimensionale Teilmannigfaltigkeit. In der Tat lässt sich der Diffeomorphismus $\varphi_p$ zu $p \in U$ als die Identitätsabbildung auf $U$ wählen (wieso?). Weiter ist jede $n$-dimensionale Teilmannigfaltigkeit von $\mathbb{R}^n$ eine offene Teilmenge von $\mathbb{R}^n$ (siehe Übung 11.15).*

*(b) Jede endliche (oder diskrete) Teilmenge $M \subseteq \mathbb{R}^n$ ist eine nulldimensionale Teilmannigfaltigkeit. Dabei lässt sich $\varphi_p$ zu $p \in M$ als eine Verschiebung $B_\varepsilon(p) \to B_\varepsilon(0),\ x \mapsto x - p$ für $\varepsilon > 0$ klein genug wählen (wieso?). Hierbei heisst eine Menge* diskret, *falls es zu jedem $p \in M$ ein $\varepsilon > 0$ gibt mit $M \cap B_\varepsilon(p) = \{p\}$.*

*(c) Wir wollen nun zu ersten interessanten Beispiel einer Teilmannigfaltigkeit – Graphen von glatten Funktionen – kommen, welche in einem gewissen Sinn fast den allgemeinen Fall darstellen (siehe Proposition 11.12). Sei also $k < n$ und $U \subseteq \mathbb{R}^k$ offen. Falls $f : U \to \mathbb{R}^{n-k}$ eine glatte Abbildung ist, dann ist*

$$M = \mathrm{Graph}(f) = \{(x, f(x)) \mid x \in U\} \subseteq \mathbb{R}^n$$

*eine $k$-dimensionale Teilmannigfaltigkeit. Denn für $p = (x_0, f(x_0)) \in M$ zu $x_0 \in U$ können wir $U_p = V_p = U \times \mathbb{R}^{n-k}$ setzen und die glatte Abbildung*

$$\varphi = \varphi_p : U_p \to V_p, \quad (x, y) \mapsto (x, y - f(x))$$

*betrachten. Diese definiert in der Tat einen Diffeomorphismus und erfüllt für alle $(x, y)$ in $U \times \mathbb{R}^{n-k}$*

$$\varphi(x, y) \in \mathbb{R}^k \times \{0\}^{n-k} \Longleftrightarrow (x, y) \in U_p \cap M = M.$$

*(d) Die $(n-1)$-dimensionale (Einheits-) Sphäre*

$$\mathbb{S}^{n-1} = \{p \in \mathbb{R}^n \mid \|p\| = 1\}$$

*ist eine $(n-1)$-dimensionale Teilmannigfaltigkeit von $\mathbb{R}^n$.*

*Um die Idee des Beweises (auch graphisch) klarer darstellen zu können, beschränken wir uns hier auf den Fall $n = 3$ und verweisen auf Theorem 11.16 unten für den allgemeinen Fall. Sei $p_0 = (x_0, y_0, z_0)^t \in \mathbb{S}^2$.*

- *Falls $z_0 > 0$ ist, verwenden wir die Abbildung*

$$\begin{aligned}\varphi_{p_0} = \varphi_+ : U_{p_0} = U_+ = B_1^{\mathbb{R}^2}(0) \times (0, \infty) &\to \varphi_+(U_+) \\ (x, y, z)^t &\mapsto (x, y, z - \sqrt{1 - x^2 - y^2})^t\end{aligned}$$

  *wie oben im Teil (c) dieses Beispiels.*

- *Falls $z_0 < 0$ ist, verwenden wir analog*

$$\begin{aligned}\varphi_- : U_- = B_1^{\mathbb{R}^2}(0) \times (-\infty, 0) &\to \varphi_-(U_-) \\ (x, y, z)^t &\mapsto (x, y, z + \sqrt{1 - x^2 - y^2})^t.\end{aligned}$$

- *Falls $z_0 = 0$ und $y_0 > 0$ ist, verwenden wir die Menge*

$$\tilde{U}_+ = \left\{(x, y, z) \in \mathbb{R}^3 \mid y > 0 \text{ und } x^2 + z^2 < 1\right\}$$

  *(was obigem $U_+$ nach Vertauschung der Koordinaten $y$ und $z$ entspricht) gemeinsam mit*

$$\tilde{\varphi}_+ : (x, y, z)^t \in \tilde{U}_+ \mapsto (x, z, y - \sqrt{1 - x^2 - z^2})^t.$$

- *Die verbleibenden Fälle (das wären $z_0 = 0$, $y_0 < 0$ beziehungsweise $y_0 = z_0 = 0$, $x_0 = 1$ beziehungsweise $y_0 = z_0 = 0$, $x_0 = -1$) sind analog.*

Die Diskussion in Teil (d) von Beispiel 11.11 ist gewissermassen typisch für Teilmannigfaltigkeiten. Obwohl wir uns die Teilmannigfaltigkeit $\mathbb{S}^2$ lokal als Graphen einer Funktion vorstellen (siehe Proposition 11.12 unten), benötigen wir oft mehrere Funktionen (oben wegen der Vorzeichenwechsel der Koordinaten) und müssen die Auswahl an Koordinaten, mit denen sich die restlichen Koordinaten darstellen lassen, entsprechend dem Punkt $p \in M$ wählen (damit Punkte am Äquator auch erlaubt sind). Dies ergibt zusammen eine Kollektion von Karten, das heisst, einen „Atlas“ von $M$: lokal sieht $M$ in jedem Punkt wie die einzelnen Karten (zum Beispiel $(-\eta, \eta)^k$ für $\eta > 0$) aus, doch gibt es zwischen den Karten Überlappungen, die zusammen $M$ komplett beschreiben. Dieser Gesichtspunkt erlaubt es, Definition 11.10 zum Begriff von $k$-dimensionalen Mannigfaltigkeiten zu erweitern, die dann als metrische Räume, aber nicht zwingend als Teilmenge von $\mathbb{R}^n$ gegeben sind. Wie bereits erwähnt, begnügen wir uns hier aber mit dem Begriff einer Teilmannigfaltigkeit des Euklidschen Raumes.

Wir möchten nun zeigen, dass sich jede $k$-dimensionale Teilmannigfaltigkeit lokal als Graph von Funktionen $B \subseteq \mathbb{R}^k \to \mathbb{R}^{n-k}$ darstellen lässt. Dazu beachtet man aber, dass diese Darstellung nicht zwingend über den ersten $k$ Koordinaten möglich sein muss – siehe den Spezialfall

$z_0 = 0$ in Beispiel 11.11(d). Wir lassen also Koordinatenvertauschungen zu. Zu $\sigma \in S_n$ sei

$$P_\sigma : (x_1, \ldots, x_n)^t \in \mathbb{R}^n \mapsto (x_{\sigma(1)}, \ldots, x_{\sigma(n)})^t \in \mathbb{R}^n$$

die induzierte Permutation der Koordinaten.

**Proposition 11.12** (Lokale Darstellbarkeit durch Graphen)**.** *Eine Teilmenge $M \subseteq \mathbb{R}^n$ ist genau dann eine $k$-dimensionale Teilmannigfaltigkeit, wenn es zu jedem Punkt $p \in M$ eine offene Umgebung $U_p$ von $p$ in $\mathbb{R}^n$, eine glatte Funktion $f_p : \tilde{U}_p \to \mathbb{R}^{n-k}$ auf einer offenen Teilmenge $\tilde{U}_p \subseteq \mathbb{R}^k$ und ein $\sigma \in S_n$ gibt, so dass*

$$M \cap U_p = P_\sigma(\mathrm{Graph}(f_p)).$$

*Beweis.* Angenommen $M$ ist eine $k$-dimensionale Teilmannigfaltigkeit. Sei $p \in M$, $U_p$ eine offene Umgebung von $p \in \mathbb{R}^n$ und $\varphi_p : U_p \to V_p \subseteq \mathbb{R}^n$ ein Diffeomorphismus wie in Definition 11.10 mit $\varphi_p(p) = 0$. Sei

$$\psi : (y_1, \ldots, y_k) \in (-\varepsilon, \varepsilon)^k \mapsto \varphi_p^{-1}(y_1, \ldots, y_k, 0, \ldots, 0) \in M$$

für ein $\varepsilon > 0$ klein genug. Dann hat das Differential $\mathrm{D}_0\psi$ Rang $k$, womit $k$ linear unabhängige Zeilen in $\mathrm{D}_0\psi$ existieren. Nach Koordinatenvertauschung (von hier stammt $\sigma$ in der Aussage) können wir annehmen, dass diese Zeilen die ersten $k$ sind. Damit hat die Abbildung

$$g : y \in (-\varepsilon, \varepsilon)^k \mapsto (\psi_1(y), \ldots, \psi_k(y))^t$$

ein invertierbares Differential bei 0. Also existiert nach dem Satz zur inversen Abbildung (Satz 11.5) eine (nicht-leere) offene Menge $U \subseteq (-\varepsilon, \varepsilon)^k$, so dass die Einschränkung von $g$ auf $U$ ein Diffeomorphismus ist. Wir betrachten nun die Abbildung $f = \psi \circ (g|_U)^{-1} : g(U) \to M$. Für $i \leq k$ und alle $y \in g(U)$ gilt $f_i(y) = \psi_i(g|_U^{-1}(y)) = y_i$ nach Konstruktion. Die Teilmannigfaltigkeit $M$ ist also lokal der Graph der Abbildung $y \in g(U) \mapsto (f_{k+1}(y), \ldots, f_n(y))$ (nach Permutation der Koordinaten), womit der erste Teil der Aussage bewiesen ist.

Für die Umkehrung kann man analog vorgehen wie in Beispiel 11.11(c). □

**Applet 11.13** (Illustration zur Darstellbarkeit durch Graphen)**.** *In diesem Applet illustrieren wir Proposition 11.12 anhand einer eindimensionalen Teilmannigfaltigkeit von $\mathbb{R}^2$, wobei Sie rechts auswählen können ob für den ausgewählten Punkt die Vertauschung der Koordinaten angewendet werden sollte.*

**Übung 11.14** (Ein Überschneidungspunkt)**.** *Sei $M = \{(x, y) \in \mathbb{R}^2 \mid x(y^2 - x) = 0\}$. Zeigen Sie unter Verwendung von Proposition 11.12, dass $M \setminus \{0\}$ eine Teilmannigfaltigkeit ist, aber $M$ nicht.*

**Übung 11.15** (Topologische Eigenschaften von Teilmannigfaltigkeiten)**.** *Zeigen Sie, dass jede $n$-dimensionale Teilmannigfaltigkeit von $\mathbb{R}^n$ eine offene Teilmenge von $\mathbb{R}^n$ ist. Zeigen Sie ebenso, dass jede nulldimensionale Teilmannigfaltigkeit von $\mathbb{R}^n$ eine diskrete Teilmenge von $\mathbb{R}^n$ ist.*

*Bemerkung* ($C^\ell$-Teilmannigfaltigkeiten). Der Begriff der glatten Teilmannigfaltigkeit in Definition 11.10 lässt sich zum Begriff von $C^\ell$-Teilmannigfaltigkeiten für ein beliebiges $\ell \in \mathbb{N} \cup \{\infty\}$ verallgemeinern. Dazu betrachtet man in Analogie zu Definition 11.10 lokale $C^\ell$-Diffeomorphismen $\varphi_p$. Da wir in Zukunft aber sowieso meist an glatten Teilmannigfaltigkeiten interessiert sein werden und die Notation einfach halten möchten, haben wir auf diese Verallgemeinerung verzichtet.

### 11.2.2 Niveaumengen als Teilmannigfaltigkeiten

Wie schon die Sphäre, sind viele Teilmannigfaltigkeiten über Gleichungen definiert, und folgender Satz gibt ein allgemeines Resultat in diese Richtung.

**Theorem 11.16** (Satz über den konstanten Rang)**.** *Sei $U \subseteq \mathbb{R}^n$ offen, $1 \leq m < n$ und $F : U \to \mathbb{R}^m$ eine glatte Funktion, so dass $F$ keine kritischen Punkte in $M = \{p \in U \mid F(p) = 0\}$ besitzt (das heisst, $\mathrm{D}_pF$ hat Rang $m$ für alle $p \in M$ beziehungsweise $0$ ist ein regulärer Wert von $F$). Dann ist $M$ eine $(n-m)$-dimensionale Teilmannigfaltigkeit von $\mathbb{R}^n$.*

Insbesondere zeigt der Satz, dass für jedes $n$ die Sphäre $\mathbb{S}^{n-1} \subseteq \mathbb{R}^n$ eine Teilmannigfaltigkeit von $\mathbb{R}^n$ ist. In der Tat ist $\mathbb{S}^{n-1} = \left\{(x_1, \ldots, x_n)^t \in \mathbb{R}^n \mid F(x_1, \ldots, x_n) = 0\right\}$ für $F : (x_1, \ldots, x_n)^t \in \mathbb{R}^n \mapsto x_1^2 + \ldots + x_n^2 - 1$ und der einzige kritische Punkt von $F$ ist der Ursprung $0 \notin \mathbb{S}^{n-1}$ (wieso?).

*Beweis.* Sei $p \in M$. Nach Voraussetzung ist der Rang der totalen Ableitung $\mathrm{D}_pF$ gleich $\min(m, n) = m$. Wie zuvor identifizieren wir $\mathrm{D}_pF$ mit der Jacobi-Matrix $\mathrm{D}_pF \in \mathrm{Mat}_{m,n}(\mathbb{R})$ (mit $m$ Zeilen und $n$ Spalten). Nach Annahme existieren $m$ dieser Spalten, welche gemeinsam eine invertierbare Matrix bilden. Wenn nötig, vertauschen wir die Koordinaten im $\mathbb{R}^n$ (diese Abbildung wird Teil des Diffeomorphismus) und bezeichnen die ersten $k = n - m$ Koordinaten mit $x$ und die letzten $m$ Koordinaten mit $y$. Durch diese Vertauschung können wir annehmen, dass $\boldsymbol{\partial}_y F(p)$ invertierbar ist. In anderen Worten erfüllt die implizite Gleichung $F(x, y) = 0$ bei dem Punkt $p = (x_0, y_0)$ und für ein geeignetes $r > 0$ mit $B_r(x_0) \times B_r(y_0) \subseteq U$ die Voraussetzung des Satzes über implizite Funktionen (Satz 11.1 und Satz 11.2) und wir erhalten offene Umgebungen $U_0 \subseteq B_r(x_0)$ von $x_0$ und $V_0 \subseteq B_r(y_0)$ von $y_0$, so dass

$$M \cap (U_0 \times V_0) = \mathrm{Graph}(f)$$

für eine glatte Funktion $f : U_0 \to V_0$. Wir können also $U_p = U_0 \times V_0$ und den Diffeomorphismus (genau wie in Beispiel 11.11(1)) durch

$$\begin{aligned} \varphi_0 : U_0 \times V_0 &\to \varphi_0(U_0 \times V_0) \\ (x, y) &\mapsto (x, y - f(x)) \end{aligned}$$

definieren.

Da $p \in M$ beliebig war, beweist dies den Satz. □

**Beispiel 11.17.** *Der einzige kritische Punkt der Funktion* $F : (x, y, z)^t \mapsto x^2 + y^2 - z^2$ *auf* $\mathbb{R}^3$ *ist wiederum der Ursprung (wieso?). Aus diesem Grund und wegen Theorem 11.16 sind das sogenannte* ***einschalige Hyperboloid***

$$H_1 = \left\{(x, y, z)^t \in \mathbb{R}^3 \mid x^2 + y^2 - z^2 - 1 = 0\right\}$$

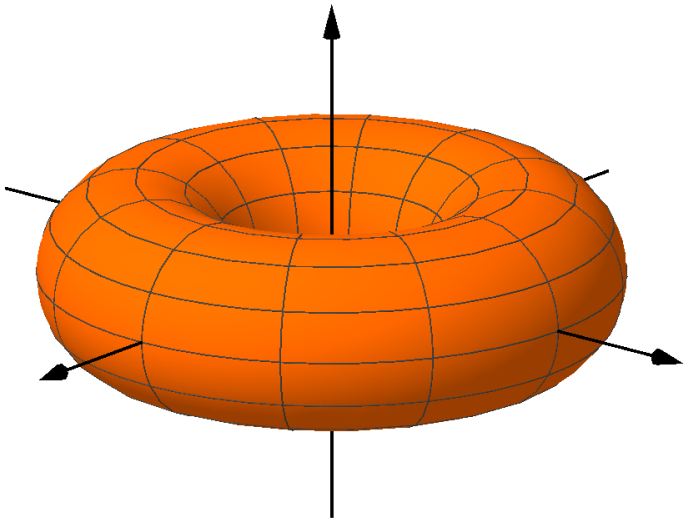
*und das sogenannte* ***zweischalige Hyperboloid***

$$H_2 = \left\{(x, y, z)^t \in \mathbb{R}^3 \mid x^2 + y^2 - z^2 + 1 = 0\right\}$$

*zweidimensionale Teilmannigfaltigkeiten von* $\mathbb{R}^3$.

s Obiges Beispiel lässt sich mit derselben Methodik auf beliebige quadratische Formen erweitern (siehe Abschnitt 11.4.2).

**Beispiel 11.18** (2-Torus)**.** *Seien* $0 < r < R$ *Radien. Der* 2*-Torus* $\mathbb{T}^2$ *mit Parametern* $R, r$ *ist der Rotationskörpers, der entsteht, wenn man den Kreis in der* $x$*-*$z$*-Ebene mit Radius* $r$ *um den Punkt* $(0, R)$ *um die* $z$*-Achse rotiert.*



*Wir möchten hier zeigen, dass der* 2*-Torus eine zweidimensionale Teilmannigfaltigkeit des* $\mathbb{R}^3$ *ist. Passend zu der geometrischen Beschreibung können wir den* 2*-Torus auch durch eine Gleichung definieren, nämlich*

$$\mathbb{T}^2 = \left\{(x, y, z)^t \in \mathbb{R}^3 \mid \left(\sqrt{x^2 + y^2} - R\right)^2 + z^2 = r^2\right\}.$$

*Wir betrachten also die glatte Funktion*

$$\begin{aligned} F : (x, y, z)^t \in (\mathbb{R}^2 \setminus \{0\}) \times \mathbb{R} \mapsto & \left(\sqrt{x^2 + y^2} - R\right)^2 + z^2 - r^2 \\ & = x^2 + y^2 + z^2 - 2R\sqrt{x^2 + y^2} + R^2 - r^2. \end{aligned}$$

*und zeigen, dass* $0$ *ein regulärer Wert von* $F$ *ist. Zuerst bemerken wir, dass*

$$F(0,0,z) = z^2 + R^2 - r^2 > 0$$

*für alle* $z \in \mathbb{R}$. *Für* $p = (x,y,z)^t \in M$ *gilt also* $p \in (\mathbb{R}^2 \setminus \{0\}) \times \mathbb{R}$ *und*

$$\begin{aligned} \mathrm{D}_p F &= \left(2x - 2R\frac{x}{\sqrt{x^2+y^2}}\ , 2y - 2R\frac{y}{\sqrt{x^2+y^2}}\ , 2z\right) \\ &= \left(2x\left(1 - \frac{R}{\sqrt{x^2+y^2}}\right)\ , 2y\left(1 - \frac{R}{\sqrt{x^2+y^2}}\right)\ , 2z\right). \end{aligned}$$

*Wenn nun obiges* $p$ *ein kritischer Punkt (das heisst,* $\mathrm{D}_p F = 0$*) ist, so gilt* $z = 0$ *und* $R = \sqrt{x^2+y^2}$ *wegen* $(x,y) \neq 0$. *Keiner dieser Punkte liegt aber auf* $\mathbb{T}^2$ *und somit ist* $0$ *ein regulärer Wert von* $F$. *Nach Theorem 11.16 ist der* $2$*-Torus eine zweidimensionale Teilmannigfaltigkeit des* $\mathbb{R}^3$.

**Übung 11.19** (Parametrisierung des 2-Torus)**.** *Wir möchten hier eine Parametrisierung des* $2$*-Torus mit Parametern* $R, r$ *besprechen.*

*Dazu betrachten wir als erstes zwei Winkel* $\varphi, \psi$. *Dabei stellt der erste Winkel* $\varphi$ *den Winkel in der* $xy$*-Ebene dar und der zweite Winkel* $\psi$ *stellt den Winkel in der Ebene dar, die von der* $z$*-Richtung und der von* $\varphi$ *fixierten Richtung in der* $xy$*-Ebene gegeben ist.*

*(i) Sei*

$$\bar{f} : [0,2\pi] \times [0,2\pi] \to \mathbb{T}^2, \quad (\varphi,\psi) \mapsto ((R + r\sin\psi)\cos\varphi, (R + r\sin\psi)\sin\varphi, r\cos\psi)^t.$$

*Zeigen Sie, dass* $\bar{f}$ *tatsächlich eine stetige Abbildung ist, die auf* $\mathbb{T}^2$ *abbildet. Überzeugen Sie sich auch davon, dass* $\bar{f}$ *surjektiv, aber nicht injektiv ist.*

*(ii) Die Abbildung* $\bar{f}$ *lässt sich auch nicht so einschränken, dass* $\bar{f}$ *eine stetige Bijektion mit stetiger Inverser wird (was wir wollen). Überzeugen Sie sich als Beispiel davon, dass* $\bar{f}|_{[0,2\pi)^2}$ *bijektiv und stetig ist, aber keine stetige Inverse besitzt.*

*Nach (ii) ist* $\bar{f}$ *offenbar nicht die optimale Weise, den* $2$*-Torus topologisch zu verstehen. Wir wechseln nun deswegen vom Intervall* $[0,2\pi]$ *auf den Einheitskreis, wo die Winkel* $0$ *und* $2\pi$ *automatisch identifiziert werden.*

*(iii) Zeigen Sie, dass die Abbildung*

$$f : \mathbb{S}^1 \times \mathbb{S}^1 \to \mathbb{T}^2, \quad ((x,y)^t, (v,w)^t) \mapsto ((R + rw)x, (R + rw)y, rv)^t$$

*eine stetige Bijektion mit stetiger Inversen definiert.*

*Tatsächlich ist, wenn man weiss, was differenzierbare Abbildung zwischen Teilmannigfaltigkeiten sind, die obige Abbildung* $f$ *glatt mit glatter Inversen und somit die richtige Art und Weise* $\mathbb{T}^2$ *zu verstehen.*

### 11.2.3 Tangentialraum und Tangentialbündel

Wir wollen hier zu einer gegebenen Teilmannigfaltigkeit $M \subseteq \mathbb{R}^n$ den zugehörigen „Phasenraum“ definieren, der analog zum einfacheren Fall $M = U$ für eine offene Teilmenge $U$ von $\mathbb{R}^n$ (siehe Abschnitt 10.2.2), aus allen Paaren der möglichen Orte $p \in M$ und den bei $p$ möglichen Geschwindigkeitsvektoren bestehen soll. Dies legt die folgende Definition nahe.

**Definition 11.20** (Tangentialraum)**.** Sei $M \subseteq \mathbb{R}^n$ eine $k$-dimensionale Teilmannigfaltigkeit. Der **Tangentialraum** (oder **Phasenraum**) von $M$ bei $p \in M$ ist durch

$$\begin{aligned}\mathrm{T}_pM &= \big\{(p, \gamma'(0)) \mid \exists \varepsilon > 0,\ \gamma : (-\varepsilon, \varepsilon) \to M \text{ differenzierbar mit } \gamma(0) = p\big\}\\ &\subseteq \mathrm{T}_p\mathbb{R}^n = \{p\} \times \mathbb{R}^n\end{aligned}$$

und das **Tangentialbündel** von $M$ durch

$$\mathrm{T}M = \bigsqcup_{p \in M} \mathrm{T}_pM \subseteq \mathrm{T}\mathbb{R}^n = \mathbb{R}^n \times \mathbb{R}^n$$

definiert.

In Worten ausgedrückt ist $\mathrm{T}_pM$ also die Menge der Geschwindigkeitsvektoren von kurzen Wegen durch $p$ in $M$ wie oben schon erklärt. Als erstes kann man sich nun ähnliche Fragen wie in Abschnitt 10.2.2 stellen. Beispielsweise: Ist $\mathrm{T}_pM$ ein Unterraum von $\mathrm{T}_p\mathbb{R}^n$?

Die Identität $\mathrm{T}\mathbb{R}^n = \mathbb{R}^n \times \mathbb{R}^n$ drängt weiter die Frage auf, ob vielleicht $\mathrm{T}M$ in einem geeigneten Sinne isomorph zu $M \times \mathbb{R}^k$ ist, wenn $M$ eine $k$-dimensionale Teilmannigfaltigkeit ist. Diese, im Allgemeinen sehr schwierige Frage werden wir in dieser Vorlesung nur für sehr wenige Teilmannigfaltigkeiten vollständig (je nach Fall positiv oder negativ) beantworten können.

Stattdessen wollen wir hier vorerst diese Frage nur lokal (wie zuvor heisst das für eine kleine Umgebung eines vorgegebenen Punktes) beantworten. Zur Vereinfachung der Notation werden wir im Folgenden $\mathbb{R}^k$ mit dem Teilraum $\mathbb{R}^k \times \{0\}^{n-k} \subseteq \mathbb{R}^n$ identifizieren.

**Satz 11.21** (Lokale Beschreibung des Tangentialbündels)**.** *Sei $M \subseteq \mathbb{R}^n$ eine $k$-dimensionale Teilmannigfaltigkeit.*

- *Sei $U_0 \subseteq \mathbb{R}^n$ eine offene Umgebung und sei $\varphi : U_0 \to V_0 \subseteq \mathbb{R}^n$ ein Diffeomorphismus mit*

  $$\varphi(U_0 \cap M) = \{y \in V_0 \mid y_{k+1} = \ldots = y_n = 0\} = V_0 \cap \mathbb{R}^k$$

  *wie in Definition 11.10. Wir definieren $\psi = \varphi^{-1} : V_0 \to U_0$ und die Ableitung*

  $$\begin{aligned}\mathrm{D}\psi : \mathrm{T}V_0 &\to \mathrm{T}U_0\\ (y, h) &\mapsto (\psi(y), \mathrm{D}_y\psi(h)).\end{aligned}$$

  *Dann ist die Einschränkung von $D\psi$ eine Bijektion von $\mathrm{T}(V_0 \cap \mathbb{R}^k) = (V_0 \cap \mathbb{R}^k) \times \mathbb{R}^k$ nach $\mathrm{T}(U_0 \cap M)$. Insbesondere ist $\mathrm{T}_pM$ ein $k$-dimensionaler Unterraum von $\mathrm{T}_p\mathbb{R}^n$ für alle $p \in M$.*

- *Angenommen $M = \{x \in U \mid F(x) = 0\}$ ist gegeben als Niveaumenge einer glatten Funktion $F : U \to \mathbb{R}^{n-k}$ auf einer offenen Teilmenge $U \subseteq \mathbb{R}^n$, so dass $0$ ein regulärer Wert von $F$ ist (wie im Theorem 11.16 über den konstanten Rang). Dann ist*

$$\mathrm{T}M = \{(p, v) \in M \times \mathbb{R}^n \mid \mathrm{D}_p F(v) = 0\}.$$

Insbesondere zweitere Aussage erlaubt in spezifischen Fällen eine effiziente Berechnung des Tangentialraums an einem Punkt, welcher dann im Wesentlichen durch den Kern des Differentials der Funktion $F$ an diesem Punkt gegeben ist.

**Beispiel 11.22.** *Wir betrachten die Sphäre $\mathbb{S}^{n-1} = \big\{(x_1, \ldots, x_n)^t \in \mathbb{R}^n \mid F(x_1, \ldots, x_n) = 0\big\}$ wobei $F : (x_1, \ldots, x_n)^t \in \mathbb{R}^n \mapsto x_1^2 + \ldots + x_n^2 - 1$. Dann ist für $x \in \mathbb{S}^{n-1}$*

$$\mathrm{D}_x F = (2x_1, \ldots, 2x_n) = 2x^t.$$

*Somit ist der Tangentialraum bei $x \in \mathbb{S}^{n-1}$ gegeben durch*

$$\mathrm{T}_x\mathbb{S}^{n-1} = \{x\} \times \big\{v \in \mathbb{R}^n \mid 2x^t v = 0\big\} = \langle x \rangle^\perp,$$

*was gerade die Menge der Vektoren ist, die senkrecht auf $x$ stehen (das orthogonale Komplement von $x$), und der Anschauung entspricht.*

Wir wenden uns nun dem Beweis von Satz 11.21 zu. Die wesentliche Idee für die erste Aussage ist dabei, dass sich ein Weg in der „flachen lokalen Kopie“ $V_0 \cap \mathbb{R}^k$ via des Diffeomorphismus $\psi$ auf $M$ schieben lässt und umgekehrt.

*Beweis von Satz 11.21.* Wir beweisen die erste Aussage zuerst. Da $V_0 \cap \mathbb{R}^k$ offen ist in $\mathbb{R}^k$, gibt es zu jedem $y \in V_0 \cap \mathbb{R}^k$ und $h \in \mathbb{R}^k$ einen differenzierbaren Weg $\gamma : (-\varepsilon, \varepsilon) \to V_0 \cap \mathbb{R}^k$ mit $\gamma(0) = y$ und $\gamma'(0) = h$ (zum Beispiel den geraden Weg $\gamma(t) = y + th$). Damit ist $\psi \circ \gamma : (-\varepsilon, \varepsilon) \to U_0 \cap M$ ein differenzierbarer Weg mit $\psi \circ \gamma(0) = \psi(y)$ und $(\psi \circ \gamma)'(0) = \mathrm{D}_y\psi(h)$. Also gilt

$$\mathrm{D}\psi(y, h) = (\psi(y), \mathrm{D}_y\psi(h)) \in \mathrm{T}_{\psi(y)}M \subseteq \mathrm{T}(U_0 \cap M).$$

Da $(y, h) \in (V_0 \cap \mathbb{R}^k) \times \mathbb{R}^k$ beliebig war, ergibt sich daraus

$$\mathrm{D}\psi((V_0 \cap \mathbb{R}^k) \times \mathbb{R}^k) \subseteq \mathrm{T}(U_0 \cap M).$$

Für die andere Inklusion nehmen wir an, dass $\gamma : (-\varepsilon, \varepsilon) \to M$ ein differenzierbarer Weg mit $\gamma(0) = p = \psi(y) \in U_0 \cap M$ ist. Da $U_0$ offen ist, existiert ein $\delta \in (0, \varepsilon)$ mit $\gamma((-\delta, \delta)) \subseteq U_0 \cap M$. Wir betrachten nun den Weg $\varphi \circ \gamma : (-\delta, \delta) \to V_0 \cap \mathbb{R}^k$, welcher bei 0 die Ableitung $h = \mathrm{D}_p\varphi(\gamma'(0)) \in \mathbb{R}^k$ hat. Gemeinsam mit $\psi \circ \varphi = \mathrm{id}$ auf $U_0$ und der Kettenregel in der Form $\mathrm{D}_y\psi \circ \mathrm{D}_p\varphi = I_n$ erhalten wir

$$(p, \gamma'(0)) = \big(\psi(y), \mathrm{D}_y\psi(\mathrm{D}_p\varphi(\gamma'(0)))\big) = \big(\psi(y), \mathrm{D}_y\psi(h)\big) \in \mathrm{D}\psi((V_0 \cap \mathbb{R}^k) \times \mathbb{R}^k)).$$

Dies zeigt, dass $\mathrm{D}\psi((V_0 \cap \mathbb{R}^k) \times \mathbb{R}^k) = \mathrm{T}(U_0 \cap M)$.

Da es nach Definition einer Teilmannigfaltigkeit für jeden Punkt $p \in M$ einen derartigen Diffeomorphismus mit $p \in U_0$ gibt, folgt für jeden Punkt $p \in M$, dass $\mathrm{T}_pM = \mathrm{D}\psi(\{p\} \times \mathbb{R}^k)$ ein $k$-dimensionaler Unterraum von $\mathrm{T}_p\mathbb{R}^n = \{p\} \times \mathbb{R}^n$ ist.

Für die zweite Behauptung sei $p \in M$ beliebig. Da nach obigem $\mathrm{T}_pM$ $k$-dimensional ist, aber auch der Kern von $\mathrm{D}_pF$ $k$-dimensional ist (wieso?), reicht es, eine Inklusion zu zeigen. Sei $\gamma : (-\varepsilon, \varepsilon) \to M$ ein differenzierbarer Weg mit $\gamma(0) = p$. Dann gilt $F \circ \gamma(t) = 0$ für alle $t \in (-\varepsilon, \varepsilon)$ nach Definition von $M$ und daher ist

$$\mathrm{D}_pF(\gamma'(0)) = (F \circ \gamma)'(0) = 0.$$

Somit ist $\gamma'(0) \in \ker(\mathrm{D}_pF)$, was den Beweis des Satzes abschliesst. □

VO: 11.04.18

Wir verwenden nun obigen Satz, um einen spezifischen Fall zu diskutieren, wo die Isomorphie $\mathrm{T}M \cong M \times \mathbb{R}^k$ gilt. Mit Satz 11.21 sind wir nun auch in der Lage zu erklären, was wir hier mit Isomorphie meinen. Es soll eine Bijektion $f : \mathrm{T}M \cong M \times \mathbb{R}^k$ (stetig mit stetiger Inversen) existieren, so dass für alle $p \in M$ die Einschränkung $f|_{\mathrm{T}_pM} : \mathrm{T}_pM \to \{p\} \times \mathbb{R}^k$ wohldefiniert ist und ein Isomorphismus von Vektorräumen ist. Wir verlangen insbesondere von einem solchen Isomorphismus, dass er den Fusspunkt $p$ erhält.

**Beispiel 11.23** (Tangentialbündel des Einheitskreises und des 2-Torus*)**.**

*(a) Wir behaupten als erstes, dass* $\mathrm{T}\mathbb{S}^1 \cong \mathbb{S}^1 \times \mathbb{R}$.

*Sei* $\iota : \mathbb{R}^2 \to \mathbb{R}^2$ *die Rotation um 90 Grad im Gegenuhrzeigersinn, das heisst,* $\iota(x_1, x_2) = (-x_2, x_1)^t$ *für alle* $(x_1, x_2)^t \in \mathbb{R}^2$. *Wir betrachten nun die Abbildung*

$$f : \mathrm{T}\mathbb{S}^1 \to \mathbb{S}^1 \times \mathbb{R}, \quad (x, v) \mapsto (x, \langle v, \iota(x)\rangle$$

*und behaupten, dass diese alle gewünschten Eigenschaft hat. Grund dafür ist im Wesentlichen, dass es bei einem fixierten Punkt* $x \in \mathbb{S}^1$ *nur eine, zu diesem orthogonale Richtung gibt, welche gerade von* $\iota(x)$ *aufgespannt wird. In der Tat definiert deswegen*

$$g : \mathbb{S}^1 \times \mathbb{R} \to \mathrm{T}\mathbb{S}^1, \quad (x, \alpha) \mapsto (x, \alpha\iota(x))$$

*eine beidseitige Inverse von* $f$. *Weiter erhält* $f$ *per Definition den Fusspunkt und die Abbildung* $f|_{\mathrm{T}_x\mathbb{S}^1}$ *ist linear und von Null verschieden, also ein Isomorphismus (da beide Räume eindimensional sind).*

*(b) Wir möchten nun auch* $\mathrm{T}\mathbb{T}^2 \cong \mathbb{T}^2 \times \mathbb{R}^2$ *zeigen. Dazu verwenden wir, dass sich* $\mathbb{T}^2$ *mit* $\mathbb{S}^1 \times \mathbb{S}^1$ *via* $f : \mathbb{S}^1 \times \mathbb{S}^1 \to \mathbb{T}^2, ((x, y)^t, (v, w)^t) \mapsto ((R + rw)x, (R + rw)y, rv)^t$ *eineindeutig parametrisieren lässt (siehe Übung 11.19). Wir erweitern diese Abbildung zu einer*

*Abbildung*

$$\begin{aligned} \tilde{f} : \mathbb{S}^1 \times \mathbb{S}^1 \times \mathbb{R}^2 &\to \mathrm{T}\mathbb{T}^2 \\ \left( \underbrace{\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} v \\ w \end{pmatrix}}_{=p}, (s,t) \right) &\mapsto \left( f(p), s \begin{pmatrix} -(R+rw)y \\ (R+rw)x \\ 0 \end{pmatrix} + t \begin{pmatrix} rvx \\ rvy \\ -rw \end{pmatrix} \right). \end{aligned}$$

*Man kann nun überprüfen, dass $\tilde{f}$ einen Isomorphismus zwischen $\mathbb{S}^1 \times \mathbb{S}^1 \times \mathbb{R}^2$ und $\mathrm{T}\mathbb{T}^2$ definiert.*

**Übung 11.24** (Tangentialbündel von $\mathbb{S}^1 \times \mathbb{S}^1$*)**.**

*(a) Zeigen Sie, dass $\mathbb{S}^1 \times \mathbb{S}^1$ eine zweidimensionale Teilmannigfaltigkeit von $\mathbb{R}^4$ ist und dass $\mathrm{T}(\mathbb{S}^1 \times \mathbb{S}^1) \cong (\mathbb{S}^1 \times \mathbb{S}^1) \times \mathbb{R}^2$.*

*(b) Ergänzen Sie die unterlassenen Details in Beispiel 11.23(b) und konstruieren Sie einen Isomorphismus $\mathrm{T}\mathbb{T}^2 \cong \mathbb{T}^2 \times \mathbb{R}^2$.*

## 11.3 Extremwertprobleme

Die letzten beiden Abschnitte enthielten, wenn man so will, sehr wenige konkrete Rechnungen und zeigten stattdessen bloss Existenz gewisser unbekannter Funktionen oder führten neue geometrische Begriffe ein. In diesem Sinne mag es überraschend sein, dass diese für praktische Rechnungen trotzdem relevant sind. Genau dies wollen wir hier nun demonstrieren, indem wir Extremwertaufgaben allgemeiner als zuvor besprechen.

### 11.3.1 Extrema auf kompakten Teilmengen

Jede stetige, reellwertige Funktion $f$ auf einer kompakten Teilmenge $K \subseteq \mathbb{R}^n$ hat ein globales Maximum und ein globales Minimum (siehe Satz 9.66(5)), die beide in Anwendungen von Interesse sein könnten. Allerdings ist diese Aussage für die Berechnung dieser Extrema (siehe Definition 10.28) nicht sonderlich hilfreich. Wir wollen hier erläutern, wie wir mit Hilfe der bereits entwickelten Theorie und der sogenannten Methode der Lagrange-Multiplikatoren diese Extremwerte finden können.

Wir werden im Folgenden die topologischen Begriffe des Inneren, des Abschlusses und des Randes verwenden – siehe Definition 9.23.

**Beispiel 11.25** (Extrema im Inneren und auf dem Rand)**.** *Wir betrachten im Folgenden mehrere Situationen, in welchen jeweils eine Teilmenge $B \subseteq \mathbb{R}^n$ und eine stetig differenzierbare, reellwertige Funktion $f$ auf $B$ gegeben ist. (Wir wollen des Weiteren immer annehmen, dass $f$ auf einer grösseren offenen Menge $U \supseteq B$ definiert und dort stetig differenzierbar ist.)*

*(a) Falls $B = I = [a, b]$ ein kompaktes Intervall mit $a < b$ (das heisst, mit nicht-leerem Inneren $(a, b)$) ist, so verwenden wir die Methoden aus Abschnitt 7.1.3. Das heisst, wir suchen die kritischen Punkte von $f$ in $(a, b)$ und betrachten getrennt noch die Randpunkte $\partial I = \{a, b\}$.*

*(b) Falls $B = P$ ein kompaktes Polygon im $\mathbb{R}^2$ wie zum Beispiel*

$$P = \big\{(x, y)^t \in \mathbb{R}^2 \mid x \geq 0,\ y \geq 0,\ x + y \leq 1\big\}$$

*ist, so können wir die Extremwerte auf $B$ finden, indem wir zuerst nach kritischen Punkten von $f$ im Innern $P^\circ$ suchen (mit den Methoden aus Abschnitt 10.4). Anschliessend schränken wir $f$ auf die einzelnen Kanten von $P$ (in obigem Beispiel $[0, 1] \times \{0\}$, $\{0\} \times [0, 1]$ und $\big\{(x, 1 - x)^t \mid x \in [0, 1]\big\}$) ein und verwenden dort die Methoden von (a). Insbesondere muss man die Werte von $f$ an den Ecken $(0, 0)^t$, $(1, 0)^t$ und $(0, 1)^t$ auch in Betracht ziehen.*

*(c) Falls $B = \{v \in \mathbb{R}^n \mid \|v\|_2 \leq 1\}$ der abgeschlossene Einheitsball im $\mathbb{R}^n$ ist, so können wir die kritischen Punkte von $f$ im Inneren $B^\circ = \{v \in \mathbb{R}^n \mid \|v\| < 1\}$ finden und des Weiteren die Funktion $f$ auf der Teilmannigfaltigkeit $\partial B = \mathbb{S}^{n-1}$ untersuchen. Letzteres werden wir im nächsten Unterabschnitt in Angriff nehmen.*

*(d) Falls*

$$B = \left\{(x, y, z)^t \in \mathbb{R}^3 \mid x^2 + y^2 \leq z \leq 1\right\}$$

*ist, so können wir die Funktion $f$*

- *im Inneren $B^\circ$ von $B$,*
- *auf der Teilmannigfaltigkeit $\left\{(x, y, z)^t \in \mathbb{R}^3 \mid x^2 + y^2 = z < 1\right\}$,*
- *auf der Kreisscheibe $\left\{(x, y, z)^t \in \mathbb{R}^3 \mid x^2 + y^2 < z = 1\right\}$ und*
- *auf dem Kreis $\left\{(x, y, z)^t \in \mathbb{R}^3 \mid x^2 + y^2 = z = 1\right\}$*

*getrennt betrachten. Wir finden dann die Extremwerte, indem wir alle kritischen Punkte der Einschränkungen von $f$ auf diese Teilmengen finden.*

In den obigen Beispielen ist es einfach eine Parametrisierung der Ränder zu finden, welche man verwenden könnte, um die Suche der Extremwerte auf die Diskussion von glatten Funktionen auf offenen Teilmengen (den Karten) zurückzuspielen. Allerdings ist es, wie bereits besprochen, oft sehr schwer oder sogar unmöglich die Parametrisierung der Ränder explizit durch Formeln und uns wohlbekannten Funktionen zu beschreiben. Aus diesem Grund wäre es von Vorteil eine Methode zur Verfügung zu haben, welche ohne Verwendung der Kartenabbildungen die lokalen Extremwerte berechnen könnte.

### 11.3.2 Extrema mit Nebenbedingungen und Lagrange-Multiplikatoren

Wie im letzten Teilabschnitt angedeutet wurde, wollen wir nun Funktionen auf Teilmannigfaltigkeiten untersuchen und dort notwendige Bedingungen für lokale Extrema angeben.

**Proposition 11.26** (Notwendige Bedingungen für Extrema mit Nebenbedingungen)**.** *Sei $U \subseteq \mathbb{R}^n$ offen und $M \subseteq U$ eine Teilmannigfaltigkeit von $\mathbb{R}^n$. Weiter sei $f : U \to \mathbb{R}$ eine differenzierbare Funktion. Angenommen $f|_M$ nimmt in $p \in M$ ein lokales Extremum an. Dann ist $\nabla f(p)$ ein **Normalenvektor** an $M$ bei $p$, das heisst, es gilt $\langle \nabla f(p), v\rangle = 0$ für alle $(p, v) \in \mathrm{T}_pM$.*

Die Menge der Normalenvektoren an $M$ bei $p$ werden wir mit

$$(\mathrm{T}_pM)^\perp = \{(p, w) \mid \langle w, v\rangle = 0 \text{ für alle } (p, v) \in \mathrm{T}_pM\}$$

bezeichnen. Genauso wie $\mathrm{T}_pM$ bildet $(\mathrm{T}_pM)^\perp$ einen Unterraum von $\mathrm{T}_p\mathbb{R}^n$. Wenn $M$ eine $k$-dimensionale Teilmannigfaltigkeit ist, dann hat $(\mathrm{T}_pM)^\perp$ die Dimension $(n - k)$ (wieso?).

*Beweis.* Wir betrachten einen differenzierbaren Weg $\gamma : (-\varepsilon, \varepsilon) \to M$ mit $\gamma(0) = p$ und $\varepsilon > 0$. Da $f$ in $p$ ein lokales Extremum annimmt, nimmt $f \circ \gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}$ in 0 ein lokales Extremum an. Daher gilt nach Proposition 7.17 und der Kettenregel

$$0 = (f \circ \gamma)'(0) = \left\langle \nabla f(p), \gamma'(0)\right\rangle .$$

Da $\varepsilon > 0$ beliebig und $\gamma : (-\varepsilon, \varepsilon) \to M$ ein beliebiger differenzierbarer Weg mit $\gamma(0) = p$ war, folgt daraus mit der Definition des Tangentialraums $\mathrm{T}_p M$ die Proposition. $\square$

Damit wir obiges Resultat in die Praxis umsetzen können, benötigen wir eine weitere Definition. Dazu nehmen wir im Folgenden an, dass $U \subseteq \mathbb{R}^n$ offen ist und die Teilmannigfaltigkeit $M$ durch $M = \{x \in U \mid F(x) = 0\}$ für eine glatte Funktion $F : U \to \mathbb{R}^{n-k}$ mit konstantem Rang wie in Theorem 11.16 gegeben ist. Des Weiteren soll $f : U \to \mathbb{R}$ differenzierbar sein.

**Definition 11.27.** Die **Lagrange-Funktion** $L : U \times \mathbb{R}^{n-k} \to \mathbb{R}$ für eine Funktion $f : U \to \mathbb{R}$ ist durch

$$L : (x, \lambda) \in U \times \mathbb{R}^{n-k} \mapsto L(x, \lambda) = f(x) - \sum_{j=1}^{n-k} \lambda_j F_j(x)$$

definiert. Die Komponenten von $\lambda$ werden auch **Lagrange-Multiplikatoren** genannt.

Wir wollen noch betonen, dass wir für den Beweis der im folgenden Korollar vorgestellten praktische Methode eigentlich alle Themen dieses Kapitel (direkt oder indirekt) verwendet werden.

**Korollar 11.28** (Notwendige Bedingungen mit Lagrange-Multiplikatoren)**.** *Sei $U \subseteq \mathbb{R}^n$ eine offene Teilmenge und $M = \{x \in U \mid F(x) = 0\}$ eine $k$-dimensionale Teilmannigfaltigkeit gegeben als Niveaumenge durch eine glatte Funktion $F : U \to \mathbb{R}^{n-k}$ mit regulärem Wert $0$ (siehe Theorem 11.16).*

*Sei $f : U \to \mathbb{R}$ eine differenzierbare Funktion, für die $f|_M$ in $p \in M$ ein lokales Extremum annimmt, und sei $L$ die zu $M$ und $f$ gehörige Lagrange-Funktion. Dann existieren Lagrange-Multiplikatoren $\lambda \in \mathbb{R}^{n-k}$, so dass die Gleichungen*

$$\partial_{x_i} L(p, \lambda) = 0, \quad \partial_{\lambda_j} L(p, \lambda) = 0$$

*für alle $i \in \{1, \dots, n\}$ und $j \in \{1, \dots, n-k\}$ erfüllt sind. Dabei ist zu $(x, \lambda) \in U \times \mathbb{R}^{n-k}$*

$$\partial_{x_i} L(x, \lambda) = \partial_i f(x) - \sum_{j=1}^{n-k} \lambda_j \partial_i F_j(x), \quad \partial_{\lambda_j} L(x, \lambda) = -F_j(x) \tag{11.6}$$

*für $i \in \{1, \dots, n\}$ und $j \in \{1, \dots, n-k\}$.*

VO: 13.04.18

*Beweis.* Die Gleichungen $\partial_{\lambda_j} L(x, \lambda) = -F_j(x)$ für $j = 1, \dots, n-k$ folgen direkt aus der Definition von $L$, womit die Gleichungen $\partial_{\lambda_j} L(p, \lambda) = 0$ für $j = 1, \dots, n-k$ auf Grund der Definition von $M$ gelten.

Die Beschreibung von $\partial_{x_i} L(x, \lambda)$ für $i = 1, \dots, n$ in (11.6) folgen ebenso direkt aus der Definition der Lagrange Funktion. Betrachten wir diese gemeinsam, so erhalten wir eine Umformulierung der zu beweisenden Behauptung: wir wollen zeigen, dass $\nabla f(p)$ eine Linearkombination der Vektoren $\nabla F_1(p), \dots, \nabla F_{n-k}(p)$ ist, wobei die Lagrange Multiplikatoren die Koeffizienten der Linearkombination darstellen.

Wir beobachten zuerst, dass $\mathrm{D}_pF_1 = (\nabla F_1(p))^t, \ldots, \mathrm{D}_pF_{n-k} = (\nabla F_{n-k}(p))^t$ die Zeilen der Matrix $\mathrm{D}_pF$ sind, welche nach Annahme an $F$ linear unabhängig sind. Nach Satz 11.21 gilt somit

$$\begin{aligned}\mathrm{T}_pM &= \{(p,v) \in \mathrm{T}_p\mathbb{R}^n \mid \mathrm{D}_pF(v) = 0\} \\ &= \{(p,v) \in \mathrm{T}_p\mathbb{R}^n \mid \langle \nabla F_j(p), v\rangle = 0 \text{ für alle } j = 1,\ldots,n-k\}\,.\end{aligned}$$

Also sind die Vektoren $\nabla F_1(p), \ldots, \nabla F_{n-k}(p)$ Normalenvektoren an $M$ bei $p$. Wegen linearer Unabhängigkeit und $\dim((\mathrm{T}_pM)^\perp) = n-k$ lässt sich nun jeder Normalenvektor als Linearkombination dieser Vektoren schreiben. Nach Proposition 11.26 ist $\nabla f(p)$ ein Normalenvektor und das Korollar folgt. □

**Applet 11.29** (Lagrange-Multiplikatoren und Normalenvektoren)**.** *In diesem Applet illustrieren wir Proposition 11.26 und Korollar 11.28 anhand einer eindimensionalen Teilmannigfaltigkeit. Unter dieser Annahme liegt nur ein Gradientenvektor $\nabla F$ vor, womit Proposition 11.26 besagt, dass $\nabla F$ und $\nabla f$ parallel sein sollen.*

Für die Praxis ist es wichtig, dass wir zwar für den Beweis von Korollar 11.28 den Satz über implizite Funktionen (Satz 11.1 und Satz 11.2) indirekt (wie genau?) verwendet haben, aber die implizite Funktion (welche wir eher selten berechnen können oder wollen) in der Methode von Lagrange nicht vorkommt.

Wir wollen diese Methode an einem Beispiel erproben. Dieses wird auch zeigen, dass der Methode der Lagrange-Multiplikatoren (Korollar 11.28) Voraussetzungen zugrunde liegen, die, wenn sie ignoriert werden, zum falschen Ergebnis führen.

**Beispiel 11.30.** *Wir betrachten die kompakte Menge*

$$K = \left\{(x,y)^t \in [-1,1]^2 \mid y^3 - x^2 = 0\right\},$$

*welche durch die Funktion $F : (x,y)^t \in \mathbb{R}^2 \mapsto y^3 - x^2$ definiert wird. Wir beobachten zuerst, dass $K$ die Punkte $(1,1)^t$ und $(-1,1)^t$ enthält, welche eine getrennte Behandlung erfordern (wieso?). Wir betrachten die Funktion $f(x,y) = 4y - 3x$. Die Punkte $(1,1)^t$ und $(-1,1)^t$ erfüllen*

$$f(1,1) = 1, \quad f(-1,1) = 7.$$

*Wir betrachten also $M = K \setminus \left\{(1,1)^t, (-1,1)^t\right\}$ und wollen alle möglichen lokalen Extremwerte von $f|_M$ auffinden. Dazu verwenden wir (vorerst ohne gross zu überlegen) die Methode der Lagrange-Multiplikatoren (Korollar 11.28) an. Anschliessend sollten wir das globale Minimum und das globale Maximum von $f|_K$ sehr leicht unter den gefundenen Funktionswerten bestimmen können. Die zu $f$ und $M$ gehörige Lagrange-Funktion ist gegeben durch*

$$L(x,y,\lambda) = 4y - 3x - \lambda(y^3 - x^2).$$

*Man berechnet*

$$\begin{aligned}\partial_x L(x,y,\lambda) &= -3+2\lambda x\\ \partial_y L(x,y,\lambda) &= 4-3\lambda y^2\\ \partial_\lambda L(x,y,\lambda) &= -(y^3-x^2).\end{aligned}$$

*Aus* $-3+2\lambda x=0$ *folgt* $\lambda\neq 0$ *und* $x\neq 0$ *sowie* $\lambda=\frac{3}{2x}$. *Genauso folgt aus* $4-3\lambda y^2=0$, *dass* $y\neq 0$ *und* $\lambda=\frac{4}{3y^2}$. *Damit gilt* $\frac{3}{2x}=\frac{4}{3y^2}$ *oder äquivalent dazu* $x=\frac{9}{8}y^2$. *Andererseits gilt des Weiteren* $\partial_\lambda L(x,y,\lambda)=-(y^3-x^2)=0$. *Setzen wir nun* $x=\frac{9}{8}y^2$ *ein, so erhalten wir*

$$0=y^3-\left(\tfrac{9}{8}y^2\right)^2=y^3-\tfrac{9^2}{8^2}y^4=y^3\left(1-\tfrac{9^2}{8^2}y\right).$$

*Da* $y\neq 0$ *ist, ergibt dies* $y=\frac{8^2}{9^2}$ *und* $x=\frac{9}{8}y^2=\frac{8^3}{9^3}$. *Wir erhalten also mit der Methode der Lagrange-Multiplikatoren einen einzigen weiteren Kandidaten nebst den Randpunkten für die Extremwerte, nämlich*

$$f(\tfrac{8^3}{9^3},\tfrac{8^2}{9^2})=4\tfrac{8^2}{9^2}-3\tfrac{8^3}{9^3}=1.053\ldots.$$

*Haben wir jetzt alle lokalen Extremwerte gefunden? Nein, denn* $f$ *nimmt das globale Minimum auf* $K$ *im Punkt* $(0,0)$ *an. Wir haben diesen Punkt mit der Methode der Lagrange-Multiplikatoren nicht gefunden, da* $M$ *keine Teilmannigfaltigkeit von* $\mathbb{R}^2$ *ist und* $(0,0)$ *ein kritischer Punkt von* $F$ *ist.*

### 11.3.3 Diagonalisierbarkeit symmetrischer Matrizen*

Obwohl wir hier ja eigentlich Analysis betreiben wollen, können wir mit der Methode der Lagrange-Multiplikatoren relativ einfach folgenden wichtigen Satz aus der Linearen Algebra beweisen (welcher für uns bereits von Relevanz war beim Beweis der Charakterisierung von Indefinitheit in Satz 10.33).

**Satz 11.31.** *Jede symmetrische Matrix* $A\in\mathrm{Mat}_{n,n}(\mathbb{R})$ *ist über* $\mathbb{R}$ *diagonalisierbar. Des Weiteren existiert sogar eine Orthonormalbasis von* $\mathbb{R}^n$ *bestehend aus Eigenvektoren von* $A$.

Der schwierigste Schritt für den Beweis des Satzes ist folgendes Lemma, welches wir mit der Methode von Lagrange beweisen wollen. Schön an diesem Argument ist insbesondere, dass der Fundamentalsatz der Algebra (Theorem 9.81) nicht verwendet wird (siehe auch Übung 11.33).

**Lemma 11.32.** *Sei* $n\geq 1$ *und* $A\in\mathrm{Mat}_{n,n}(\mathbb{R})$ *eine symmetrische Matrix. Dann besitzt* $A$ *einen reellen Eigenvektor.*

*Beweis.* Wir betrachten die Sphäre $\mathbb{S}^{n-1}$, die als Niveaumenge der Funktion $F: x\in\mathbb{R}^3\mapsto \|x\|^2-1\in\mathbb{R}$ gegeben ist, und die reellwertige Funktion (quadratische Form)

$$f:\mathbb{R}^n\to\mathbb{R},\ x\mapsto x^tAx.$$

Da $\mathbb{S}^{n-1}$ kompakt ist, nimmt $f|_{\mathbb{S}^{n-1}}$ ein Maximum und ein Minimum an. Also angenommen $f$ nimmt in $p \in \mathbb{S}^{n-1}$ ein Extremum an. Sei

$$L : (x, \lambda) \in \mathbb{R}^n \times \mathbb{R} \mapsto f(x) - \lambda F(x) = x^t Ax - \lambda(\|x\|^2 - 1)$$

die zu $\mathbb{S}^{n-1}$ und $f$ gehörige Lagrange-Funktion. Nach Korollar 11.28 existiert also ein $\lambda \in \mathbb{R}$ mit

$$\partial_{x_j} L(p, \lambda) = \partial_j f(p) - \lambda \partial_j F(p) = 0 \tag{11.7}$$

für alle $j = 1, \ldots, n$ sowie $\partial_\lambda L(p, \lambda) = F(p) = 0$. Letzteres besagt bloss, dass $\|p\| = 1$ oder anders ausgedrückt $p \in \mathbb{S}^{n-1}$, wie bereits bekannt ist.

Wir berechnen nun die partiellen Ableitungen von $F$ und $f$. Es gilt für alle $x \in \mathbb{R}^n$

$$\begin{aligned}
\partial_j F(x) &= \partial_j \left( \sum_{k=1}^n x_k^2 \right) = 2x_j \\
\partial_j f(x) &= \partial_j \left( \sum_{k,\ell=1}^n a_{k\ell} x_k x_\ell \right) = \sum_{\ell=1}^n a_{j\ell} x_\ell + \sum_{k=1}^n a_{kj} x_k
\end{aligned}$$

nach der Produktregel und da $\partial_j(x_k)$ genau dann Null ist, wenn $k \neq j$ ist. Da $A$ per Annahme aber symmetrisch ist, erhalten wir

$$\partial_j f(x) = 2 \sum_{\ell=1}^n a_{j\ell} x_\ell = 2(Ax)_j.$$

Somit gilt $(Ap)_j = \lambda p_j$ für alle $j = 1, \ldots, n$ oder äquivalent dazu $Ap = \lambda p$. □

*Beweis von Satz 11.31.* Wir benötigen zusätzlich zu Lemma 11.32 etwas mehr Lineare Algebra für den Beweis, den wir jetzt mit Induktion nach $n$ durchführen werden. Für $n = 1$ gibt es nichts zu beweisen. Sei also $A \in \mathrm{Mat}_{n,n}(\mathbb{R})$ eine symmetrische Matrix. Nach Lemma 11.32 existiert ein reeller Eigenvektor $v_1 \in \mathbb{S}^{n-1}$ zu einem Eigenwert $\lambda_1 \in \mathbb{R}$. Wir betrachten das orthogonale Komplement

$$W = v_1^\perp = \{w \in \mathbb{R}^n \mid \langle w, v_1 \rangle = 0\}$$

von $v_1$. Für $w \in W$ gilt auf Grund der Symmetrie von $A$

$$\langle Aw, v_1 \rangle = \langle w, Av_1 \rangle = \lambda_1 \langle w, v_1 \rangle = 0$$

und es folgt, dass $A(W) \subseteq W$. Sei $w_1, \ldots, w_{n-1}$ eine Orthonormalbasis von $W$ bezüglich $\langle \cdot, \cdot \rangle$ (welche wegen des Gram-Schmidt-Orthonormalisierungsverfahrens existiert). Für $i, j \in \{1, \ldots, n-1\}$ gilt nun

$$\langle Aw_j, w_i \rangle = \langle w_j, Aw_i \rangle = \langle Aw_i, w_j \rangle .$$

In anderen Worten ist die Basisdarstellung $B$ von $A|_W : W \to W$ bezüglich der Basis $w_1, \ldots, w_{n-1}$ wieder symmetrisch. Nach Induktionsvoraussetzung existiert für $B$ eine Orthonormalbasis bestehend aus Eigenvektoren von $B$. Da aber $B$ (gemeinsam mit der Standardbasis von $\mathbb{R}^{n-1}$) genau $A|_W$ (gemeinsam mit der Orthonormalbasis $w_1, \ldots, w_{n-1}$) entspricht, existiert also auch für $W$ eine Orthonormalbasis $v_2, \ldots, v_n$ aus Eigenvektoren von $A$.

Wir erhalten damit, dass $v_1, \ldots, v_n$ eine Orthonormalbasis von $\mathbb{R}^n$ bildet und aus Eigenvektoren von $A$ besteht. □

**Übung 11.33.** *Der Vollständigkeit halber möchten wir hier ein elementares Argument zum Beweis von Lemma 11.32 unter Verwendung des Fundamentalsatzes angeben. Sei $n \geq 1$ und $A \in \mathrm{Mat}_{n,n}(\mathbb{R})$ eine symmetrische Matrix.*

*(i) Zeigen Sie, dass alle komplexen Eigenwerte von $A$ reell sind.*

*(ii) Beweisen Sie Lemma 11.32, indem Sie zeigen, dass $A$ genau dann einen komplexen Eigenvektor besitzt, wenn $A$ einen reellen Eigenvektor besitzt.*

*Das in unserem Beweis gewonnene geometrische Verständnis der Eigenwerte von $A$ kann aber auch anders genutzt werden. Als Beispiel davon kann man einen Spezialfall des Satzes von Courant-Fischer beweisen.*

*(iii) Zeigen Sie, dass die Werte*

$$\min_{x \in \mathbb{S}^{n-1}} x^t A x, \qquad \max_{x \in \mathbb{S}^{n-1}} x^t A x$$

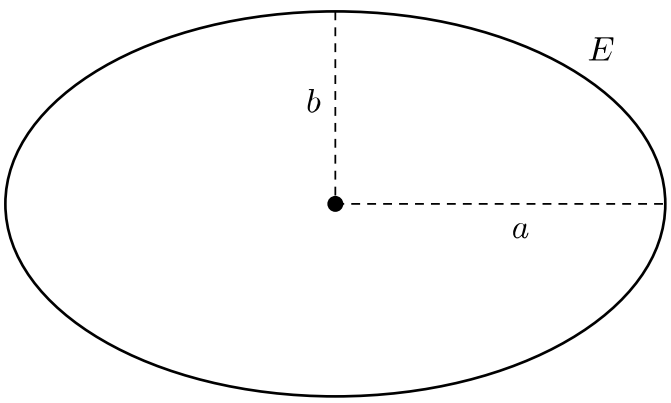
*den kleinsten respektive den grössten Eigenwert von $A$ darstellen.*

### 11.3.4 Eine hinreichende Bedingung für Lagrange-Multiplikatoren*

**Beispiel 11.34.** *Wir betrachten für $a > b > 0$ die Teilmannigfaltigkeit*

$$E = \left\{ (x, y)^t \in \mathbb{R}^2 \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0 \right\},$$

*welche eine Ellipse in Hauptachsenlage darstellt, siehe folgendes Bild.*



*Des Weiteren betrachten wir die Funktion*

$$f : \mathbb{R}^2 \to \mathbb{R},\ (x, y)^t \mapsto x^2 + y^2,$$

*welche eingeschränkt auf $E$ globale Minima bei $(0, -b)^t$ und $(0, b)^t$ und globale Maxima bei $(-a, 0)^t$ und $(a, 0)^t$ annimmt, was aus dem Bild folgt (wieso?).*

*Falls wir aber mit Hilfe unseren bisherigen Methoden entscheiden wollen, bei welchen der vier Punkte $(0, -b)^t$, $(0, b)^t$, $(-a, 0)^t$ und $(a, 0)^t$ lokale Minima oder lokale Maxima vorliegen, so ist es vorerst unklar, wie wir dies machen können. In der Tat greift die Diskussion des letzten Abschnitts diese Thematik überhaupt nicht auf und auch das Kriterium aus 10.32 greift nicht, denn die kritischen Punkte von $f$ als Funktion auf $\mathbb{R}^2$ liegen nicht auf $E$. Versuchen wir diese beiden Methoden ohne gross darüber nachzudenken zu verbinden, so würden wir fälschlicherweise zu dem Schluss kommen, dass alle vier Punkte ein Minimum der Funktion darstellen da die Hesse-Matrix von $f$ bei jedem Punkt positiv definit ist. Aber wo wäre dann das Maximum von $f|_E$?*

**Korollar 11.35** (Hesse-Matrix der Lagrange-Funktion)**.** *Sei $U \subseteq \mathbb{R}^n$ eine offene Teilmenge, sei $M = \{x \in U \mid F(x) = 0\}$ eine $k$-dimensionale Teilmannigfaltigkeit gegeben als Niveaumenge durch eine glatte Funktion $F : U \to \mathbb{R}^{n-k}$ mit regulärem Wert $0$ und sei $f : U \to \mathbb{R}$ eine zweimal stetig differenzierbare Funktion. Sei des Weiteren $p \in M$ ein kritischer Punkt für $f|_M$ mit Lagrange-Parameter $\lambda \in \mathbb{R}^{n-k}$, das heisst, $(p, \lambda)$ erfüllt die Lagrange-Gleichungen*

$$\partial_{x_i} L(p, \lambda) = 0$$

*für alle $j \in \{1, \ldots, n\}$. Wir definieren eine quadratische Form $Q : \mathrm{T}_p M \to \mathbb{R}$ durch*

$$Q(v) = \sum_{i,j=1}^{n} \partial_{x_i} \partial_{x_j} L(p, \lambda) v_i v_j$$

*für $(p, v) \in \mathrm{T}_p M$ (wobei $v_1, \ldots, v_n$ die Koordinaten von $v$ bezüglich der Standardbasis in $\mathbb{R}^n$ darstellen). Dann gelten folgende Aussagen:*

- *Falls $Q$ positiv definit ist, dann nimmt $f|_M$ bei $p$ ein striktes lokales Minimum an.*
- *Falls $Q$ negativ definit ist, dann nimmt $f|_M$ bei $p$ ein striktes lokales Maximum an.*
- *Falls $Q$ indefinit ist, dann nimmt $f|_M$ bei $p$ kein lokales Extremum an.*

Mit Korollar 11.35 lässt sich unter anderem auch die in Beispiel 11.34 festgestellte Problematik beheben.

**Übung 11.36.**

*(i) Überprüfen Sie, dass in Beispiel 11.34 die Methode der Lagrange-Multiplikatoren genau die vier Punkte $(0, -b)^t$, $(0, b)^t$, $(-a, 0)^t$ und $(a, 0)^t$ aufspürt. Berechnen Sie dabei jeweils die dazugehörigen Lagrange-Multiplikatoren.*

*(ii) Verwenden Sie Korollar 11.35, um bei jedem der vier Punkte zu entscheiden, ob und, wenn ja, was für ein lokales Extremum angenommen wird.*

*Beweis von Korollar 11.35.* Wir bemerken, dass es für die Einschränkung auf $M$ keine Rolle spielt, ob wir $f|_M$ oder $L|_M$ betrachten. Dies liegt daran, dass die Funktion $L$ durch $L(x,\lambda) = f(x) - \sum_{j=1}^{n-k} \lambda_j F_j(x)$ für $(x,\lambda) \in U \times \mathbb{R}^{n-k}$ gegeben ist und $F$ auf $M$ identisch verschwindet. Der entscheidende Vorteil von $L$ ist, dass nach geeigneter Wahl von $\lambda$ der Punkt $(p,\lambda)$ sogar ein kritischer Punkt von $L$ auf $U$ ist und wir dann die uns bekannten Methoden (siehe Korollar 10.32) anwenden können. Die quadratische (Taylor-) Approximation von $L$ hat nach Korollar 10.25 die Form

$$L(p+h) - L(p) = \sum_{i,j=1}^{n} \partial_{x_i}\partial_{x_j} L(p,\lambda) h_i h_j + o(\|h\|^2) \tag{11.8}$$

für $h \to 0$ (was für $f$ im Allgemeinen falsch wäre).

Sei nun $\psi : B_\varepsilon^{\mathbb{R}^k}(0) \to M$ eine glatte Parametrisierung einer Umgebung von $p \in M$, welche die Eigenschaften

- $\psi(0) = p$,
- $\|\psi(y) - p\| \leq C\|y\|$ für alle $y \in B_\varepsilon^{\mathbb{R}^k}(0)$ und
- $\|\psi(y) - (p + \mathrm{D}_0\psi(y)\| = o(\|y\|)$ für $y \to 0$

erfüllt (wieso existiert eine solche Parametrisierung?). Setzen wir $h = \psi(y) - p$ für $y \in B_\varepsilon^{\mathbb{R}^k}(0)$ in Gleichung (11.8) ein, so ergibt sich (wieso?)

$$L(\psi(y)) - L(p) = \sum_{i,j=1}^{n} \partial_{x_i}\partial_{x_j} L(p,\lambda)(\mathrm{D}_0\psi(y))_i(\mathrm{D}_0\psi(y))_j + o(\|y\|^2)$$

Wir bemerken nun, dass nach Satz 11.21

$$\mathrm{T}_pM = \Big\{(p,v) \in \mathrm{T}_p\mathbb{R}^n \mid v \in \mathrm{D}_0\psi(\mathbb{R}^k)\Big\} = \{p\} \times \mathrm{D}_0\psi(\mathbb{R}^k)$$

gilt. Das heisst, dass unsere Annahmen an die quadratische Form $Q$ sich direkt auf die quadratische Form

$$y \in \mathbb{R}^k \mapsto \sum_{i,j=1}^{n} \partial_{x_i}\partial_{x_j} L(p,\lambda)(\mathrm{D}_0\psi(y))_i(\mathrm{D}_0\psi(y))_j$$

übertragen. Wir können das Korollar also aus demselben Argument wie im Beweis der analogen Aussage auf offenen Teilmengen (Korollar 10.32) erhalten. □

## 11.4 Weitere Lernmaterialien

### 11.4.1 Verwendung des Kapitels

Wie bereits erwähnt ist der Satz über die implizite Funktionen grundlegend für das Verständnis von Gleichungen der Form $F(x, y) = 0$ und deren lokalen Lösungenfunktionen. Der Satz über implizite Funktionen hat auch zum Satz über die inverse Abbildung (einem Kriterium für Diffeomorphie) und einer Möglichkeit, um zu zeigen, dass eine gegebene Teilmenge von $\mathbb{R}^n$ eine Teilmannigfaltigkeit ist (das Theorem über den konstanten Rang), geführt.

Sie sollten sich einen Diffeomorphismus als einen glatten (statt linearen) Koordinatenwechsel vorstellen. Hierzu sind eben Polarkoordinaten, Zylinderkoordinaten und Kugelkoordinaten wichtige Beispiele. Wir werden Diffeomorphismen wieder im Zusammenhang der mehrdimensionalen Substitutionsregel im nächsten Kapitel begegnen.

Der Begriff der Teilmannigfaltigkeit ist in dieser Vorlesung vor allem für Flächen im $\mathbb{R}^3$ später von Bedeutung, auf welchen wir auch ein Integral definieren werden. Zum Beispiel werden die mehrdimensionalen Integralsätze ein dreidimensionales Integral über einen Bereich im $\mathbb{R}^3$ mit einem zweidimensionalen Integral über den Rand des Bereichs (oft eine Teilmannigfaltigkeit) in Verbindung bringen wird.

Schlussendlich erwähnen wir noch, dass die Methode der Lagrange-Multiplikatoren eine wichtige praktische Methode zur Bestimmung von Extremwerten darstellt. Mitunter muss aber diese Methode mit einer Untersuchung des Definitionsbereichs (zum Beispiel bestehend aus einer offenen Menge und einer glatten Teilmannigfaltigkeit als Rand) kombiniert werden.

### 11.4.2 Übungen

**Übung** (Umkehrabbildung zu Kugelkoordinaten)**.** *Wir betrachten den Diffeomorphismus*

$$\begin{aligned} f : (0, \infty) \times (0, \pi) \times (-\pi, \pi) &\to \mathbb{R}^3 \setminus ((-\infty, 0] \times \{0\} \times \mathbb{R}) \\ (r, \theta, \varphi) &\mapsto \begin{pmatrix} r \sin\theta \cos\varphi \\ r \sin\theta \sin\varphi \\ r \cos\theta \end{pmatrix}. \end{aligned}$$

*wie in Abschnitt 11.1.4.*

*(i) Begründen Sie, warum die offene Menge $f((0, \infty) \times (0, \pi) \times (-\pi, \pi))$ in der Tat Komplement $((-\infty, 0] \times \{0\} \times \mathbb{R})$ hat.*

*(ii) Bestimmen Sie die Umkehrabbildung von $f$.*

**Übung** (Elliptische Kurven)**.** *Sei $a \in \mathbb{R}$ und $M_a = \{(x, y) \in \mathbb{R}^2 \mid y^2 = x^3 + a\}$. Für welche $a$ ist $M_a$ eine Teilmannigfaltigkeit?*

**Übung** (Quadratische Hyperflächen)**.** *Sei $A \in \mathrm{Mat}_{n,n}(\mathbb{R})$ eine symmetrische Matrix, so dass die assoziierte quadratische Form $Q_A : x \mapsto x^t A x$ nicht-degeneriert ist. Zeigen Sie, dass $\{x \in \mathbb{R}^n \mid Q(x) = 1\}$ eine $(n-1)$-dimensionale Teilmannigfaltigkeit von $\mathbb{R}^n$ definiert.*

**Übung** ($\mathrm{SL}_2(\mathbb{R})$)**.** *In dieser Übung möchten wir zeigen, dass die Gruppe*

$$\mathrm{SL}_2(\mathbb{R}) = \left\{ \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathrm{Mat}_{2,2}(\mathbb{R}) \mid a_{11}a_{22} - a_{12}a_{21} = 1 \right\}$$

*eine Teilmannigfaltigkeit von* $\mathrm{Mat}_{2,2}(\mathbb{R}) \cong \mathbb{R}^4$ *darstellt.*

*(i) Zeigen Sie, dass* $I_2 \in \mathrm{SL}_2(\mathbb{R})$ *ein regulärer Punkt von* $\det$ *ist und dass* $\mathrm{D}_{I_2} \det = \mathrm{Tr}$ *ist, wobei* $\mathrm{Tr} : (a_{ij})_{ij} \in \mathrm{Mat}_{2,2}(\mathbb{R}) \mapsto a_{11} + a_{22}$ *die Spur bezeichnet.*

*(ii) Sei* $g \in \mathrm{SL}_2(\mathbb{R})$ *und* $v \in \mathrm{Mat}_{2,2}(\mathbb{R})$*. Zeigen Sie, dass die Richtungableitung* $\partial_v \det(g)$ *existiert und durch* $\mathrm{Tr}(g^{-1}v)$ *gegeben ist.*

*(iii) Zeigen Sie, dass* $\mathrm{SL}_2(\mathbb{R})$ *eine Teilmannigfaltigkeit ist.*

**Übung** (Tangentialbündel ist auch eine Teilmannigfaltigkeit)**.** *Sei* $M \subseteq \mathbb{R}^n$ *eine* $k$*-dimensionale Teilmannigfaltigkeit. Zeigen Sie, dass* $\mathrm{T}M \subseteq \mathrm{T}\mathbb{R}^n = \mathbb{R}^n \times \mathbb{R}^n$ *eine* $2k$*-dimensionale Teilmannigfaltigkeit ist.*

**Übung** (Tangentialbündel eines Graphen)**.** *Sei* $M$ *eine* $k$*-dimensionale Teilmannigfaltigkeit gegeben durch den Graphen einer glatten Funktion wie in Beispiel 11.11(c) (oder Proposition 11.12). Charakterisieren Sie in dieser Situation des Tangentialbündel von* $M$ *in Analogie zum zweiten Teil der Aussage von Satz 11.21.*

**Übung** (Challenge – Parametrisierung eindimensionaler Teilmannigfaltigkeiten)**.** *Sei* $M$ *eine zusammenhängende,* 1*-dimensionale, kompakte Teilmannigfaltigkeit von* $\mathbb{R}^n$*. Zeigen Sie, dass ein regulärer, einfacher, glatter Weg* $\gamma : [0,1] \to M$ *existiert mit* $\gamma([0,1]) = M$*.*

**Übung.** *Sei* $n \geq 2$*. Zeigen Sie, dass zwei Punkte* $x, y \in \mathbb{S}^{n-1}$ *genau dann (maximalen) Abstand* 2 *haben, wenn* $x = -y$ *ist. Betrachten Sie hierzu die Funktion* $(x,y) \mapsto \|x-y\|^2$ *auf* $\mathbb{S}^{n-1} \times \mathbb{S}^{n-1} \subseteq \mathbb{R}^{2n}$*.*

### 11.4.3 Lernkarten

Sie können wiederum die Lernkarten oder den Graphen für Ihre Wiederholung der Themen des Kapitels verwenden.

# Kapitel 12

# Mehrdimensionale Integralrechnung

Wir wollen in diesem Kapitel das Riemann-Integral über ein Intervall $[a, b] \subseteq \mathbb{R}$ zu einem mehrdimensionalen Riemann-Integral über Quader im $\mathbb{R}^n$ oder gewisse andere Teilmengen von $\mathbb{R}^n$ („Jordan-messbare Teilmengen") ausbauen. Es genügt hier wiederum, reellwertige Funktionen zu betrachten, da die Verallgemeinerung zu komplex- oder vektorwertigen Funktionen komponentenweise erfolgt (siehe Abschnitt 5.7.6).

Wie wir sehen werden, sind die Definitionen und die anfängliche Theorie in Abschnitt 12.1 weitgehend analog zum eindimensionalen Fall in Kapitel 4. Doch wollen wir den Begriff der Riemann-Integrierbarkeit tiefgründiger verstehen. Des Weiteren werden wir bei der Berechnung der Integrale in Abschnitt 12.5 und der Substitutionsregel in Abschnitt 12.6 noch neue Phänomene antreffen, die schon in Abschnitt 4.9 angedeutet wurden und für die wir hier die formal korrekte mathematische Grundlage liefern wollen.

Im nächsten Kapitel werden wir die hier entwickelte Theorie der (gewissermassen flachen) Integration über Teilmengen im $\mathbb{R}^n$ zu einer Integralrechnung über $k$-dimensionale Teilmannigfaltigkeiten im $\mathbb{R}^n$ erweitern.

## 12.1 Das Riemann-Integral für Quader

Sei $n \in \mathbb{N}$. Wie bereits erwähnt, wollen wir im Folgenden Intervalle als Integrationsbereiche durch Teilmengen im $\mathbb{R}^n$ ersetzen. Hierfür ist es von Vorteil mit dem Fall eines Quaders zu beginnen.

### 12.1.1 Definition

Unter einem **Quader im** $\mathbb{R}^n$ (oder einen $n$**-dimensionalen Quader**) verstehen wir eine Teilmenge $Q$, die gegeben ist als Produkt von Intervallen, also

$$Q = I_1 \times \ldots \times I_n$$

für beschränkte nichtleere Intervalle $I_1, \ldots, I_n \subseteq \mathbb{R}$. Falls wir die Definition der Quader als direktes Produkt betonen wollen, werden wir von „achsenparallelen Quadern" sprechen. Falls die Intervalle $I_k = [a_k, b_k]$ für $a_k \leq b_k$ und $k = 1, \ldots, n$ abgeschlossen sind, so ist auch der

Quader im $\mathbb{R}^n$ abgeschlossen. Falls die Intervalle $I_k = (a_k, b_k)$ für $a_k < b_k$ und $k = 1, \ldots, n$ offen sind, so ist auch der Quader im $\mathbb{R}^n$ offen. Wir werden häufig bloss von abgeschlossenen und offenen Quadern sprechen wenn die Dimension $n \in \mathbb{N}$ aus dem Zusammenhang ersichtlich ist. Falls die Länge der Intervalle $I_1, \ldots, I_n$ alle übereinstimmen, so nennen wir $Q$ auch einen ($n$-dimensionalen) **Würfel**. Für $n = 2$ sprechen wir auch von **Rechtecken**.

**Definition 12.1** (Volumen von Quadern)**.** Für $n \geq 1$ und beschränkte nichtleere Intervalle $I_1, \ldots, I_n$ ist das **Volumen** des $n$-dimensionalen Quaders $Q = I_1 \times \cdots \times I_n$ durch

$$\mathrm{vol}(Q) = \mathrm{vol}(\overline{Q}) = \mathrm{vol}(Q^\circ) = (b_1 - a_1) \cdots (b_n - a_n)$$

definiert, wobei $a_k = \inf I_k$ und $b_k = \sup I_k$ für $k = 1, \ldots, n$ ist.

Wir werden im restlichen Abschnitt die Dimension $n \geq 1$ und die reellen Zahlen

$$a_1 < b_1, \ldots, a_n < b_n,$$

als Endpunkte der Intervalle $I_1, \ldots, I_n$ als gegeben annehmen und den Quader durch

$$Q = I_1 \times \cdots \times I_n$$

definieren.

Grundlage unserer Diskussion ist neben dem Volumen auch eine mehrdimensionale Verallgemeinerung von Zerlegungen.

**Definition 12.2** (Zerlegung von $Q$)**.** Sei der Quader $Q$ wie oben. Für jedes $k \in \{1, \ldots, n\}$ sei

$$\mathfrak{Z}_k = \left\{a_k = x_{k,0} < x_{k,1} < \ldots < x_{k,A(k)} = b_k\right\}$$

eine Zerlegung von $[a_k, b_k]$. Dann bezeichnen wir

$$\mathfrak{Z} = (\mathfrak{Z}_1, \ldots, \mathfrak{Z}_n)$$

als eine **Zerlegung des Quaders** $Q$ und die offenen Quader

$$Q_{\boldsymbol{\alpha}} = (x_{1,\alpha_1-1}, x_{1,\alpha_1}) \times \ldots \times (x_{n,\alpha_n-1}, x_{n,\alpha_n})$$

für $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$ mit $\alpha_k \in \{1, \ldots, A(k)\}$ für alle $k \in \{1, \ldots, n\}$ als die **der Zerlegung $\mathfrak{Z}$ entsprechenden (offenen) Teilquader**. Wir werden für die der Zerlegung $\mathfrak{Z}$ entsprechenden offenen Teilquader kurz auch $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$ schreiben (für $\boldsymbol{\alpha}$ implizit wie oben).

Die obige Definition mag zuerst sehr formal wirken, doch stellt diese bloss die Formalisierung von einem Raster (wie bei Häuschenpapier) in der Ebene und einer Zerlegung eines dreidimensionalen Würfels in kleinere Teilwürfel (wie bei Lego oder Minecraft) dar. Dabei schränken wir uns weder in der Dimension $n$ noch in der Grösse und Form der kleineren Quader ein, doch wollen wir die Schnitte durch $Q$ immer achsenparallel und vollständig durch $Q$

ziehen, um die Zerlegung zu erhalten. Des Weiteren soll der Vektor $\boldsymbol{\alpha} \in \prod_{k=1}^{n} \{1, \ldots, A(k)\}$ in diesem Zusammenhang als die Adresse des Teilquaders $Q_{\boldsymbol{\alpha}}$ verstanden werden.

Ist $\mathfrak{Z}$ eine Zerlegung eines Quaders $Q$, so gilt die **Additionsformel**

$$\operatorname{vol}(Q) = \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} \operatorname{vol}(Q_{\boldsymbol{\alpha}}) \tag{12.1}$$

Den Beweis davon überlassen wir als Übung. Aussagen wie diese, die geometrisch anschaulich sind und elementare aber vielleicht etwas langwierige Beweise haben, werden wir mitunter ohne formalen Beweis verwenden. Wir wollen auch noch erwähnen, dass die Summe $\sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}}$ in (12.1) und vielen weiteren Ausdrücken als (deutlich hübschere) Kurzschreibweise für die Summe $\sum_{\alpha_1=1}^{A(1)} \sum_{\alpha_2=1}^{A(2)} \cdots \sum_{\alpha_n=1}^{A(n)}$ zu verstehen ist.

**Wichtige Übung 12.3** (Additivität des Volumens bezüglich Zerlegungen)**.** *Beweisen Sie die Additionsformel* (12.1)*. Betrachten Sie zuerst $n = 2$; der allgemeine Fall ist nur in der Notation schwieriger.*

**Definition 12.4** (Riemann-Integral auf Quadern)**.** Für eine beschränkte Funktion $f : Q \to \mathbb{R}$ auf einem abgeschlossenen Quader $Q$ und eine Zerlegung $\mathfrak{Z}$ von $Q$ definieren wir die **Untersumme**

$$U(f, \mathfrak{Z}) = \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} \inf(f(Q_{\boldsymbol{\alpha}})) \operatorname{vol}(Q_{\boldsymbol{\alpha}})$$

und die **Obersumme**

$$O(f, \mathfrak{Z}) = \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} \sup(f(Q_{\boldsymbol{\alpha}})) \operatorname{vol}(Q_{\boldsymbol{\alpha}}).$$

Das **untere Integral** von $f$ wird durch

$$\underline{I}(f) = \sup \{U(f, \mathfrak{Z}) \mid \mathfrak{Z} \text{ ist eine Zerlegung von } Q\}$$

und das **obere Integral** von $f$ durch

$$\overline{I}(f) = \inf \{O(f, \mathfrak{Z}) \mid \mathfrak{Z} \text{ ist eine Zerlegung von } Q\}$$

definiert. Die Funktion $f$ heisst **Riemann-integrierbar**, falls $\underline{I}(f) = \overline{I}(f)$ gilt. Der gemeinsame Wert wird in diesem Fall als das **Riemann-Integral**

$$\int_Q f \, \mathrm{dvol} = \underline{I}(f) = \overline{I}(f)$$

bezeichnet.

Wir bemerken an dieser Stelle, dass obige Definition des Riemann-Integrals (zumindest auf den ersten Blick) von der Definition des eindimensionalen Riemann-Integrals abweicht. In der Tat beschränken wir uns hier auf spezifische Treppenfunktionen, die die Funktion von unten

respektive oben beschränken. Für eine Funktion $g : Q \to \mathbb{R}$, die auf den einer Zerlegung $\mathfrak{Z}$ entsprechenden Teilquadern konstant ist und die Ungleichung $g \leq f$ erfüllt, gilt aber

$$\sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} g(x_{\boldsymbol{\alpha}}) \operatorname{vol}(Q_{\boldsymbol{\alpha}}) \leq \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} \inf(f(Q_{\boldsymbol{\alpha}})) \operatorname{vol}(Q_{\boldsymbol{\alpha}}) = U(f, \mathfrak{Z}),$$

wobei $x_{\boldsymbol{\alpha}} \in Q_{\boldsymbol{\alpha}}$ jeweils eine beliebige Wahl eines Punktes bezeichnet. In anderen Worten, wir verwenden oben bereits die für die Zerlegung optimale untere Treppenfunktion in der Definition der Untersumme $U(f, \mathfrak{Z})$. Dies (und die analoge Rechnung für obere Treppenfunktionen) begründet, warum die Definition 12.4 in der Tat Definition 4.11 verallgemeinert.

Sie sollten sich zum Beispiel ein Raster auf einer rechteckigen topographischen Karte vorstellen und über jedem Teilrechteck einen massstabsgetreuen Quader anbringen, womit Sie eine approximative (digitale) Landschaft erhalten. Bei der Höhe der Teilquader hat man die Wahl entweder die minimale Höhe oder die maximale Höhe innerhalb des Teilrechtecks zu verwenden, und erhält dadurch die Untersumme oder die Obersumme zur Zerlegung. Sie sollten versuchen sich diesbezüglich eine mehrdimensionale Verallgemeinerung von Applet 4.13 vorzustellen und bei den folgenden Lemmas die Aussagen und Beweise mittels diesem Bild zu interpretieren.

**Lemma 12.5.** *Seien $\mathfrak{Z} = (\mathfrak{Z}_1, \ldots, \mathfrak{Z}_n)$ und $\mathfrak{Z}' = (\mathfrak{Z}'_1, \ldots, \mathfrak{Z}'_n)$ zwei Zerlegungen eines abgeschlossenen Quaders $Q$. Dann gilt $U(f, \mathfrak{Z}) \leq O(f, \mathfrak{Z}')$ für jede beschränkte Funktion $f : Q \to \mathbb{R}$. Falls des Weiteren $\mathfrak{Z}'$ eine Verfeinerung von $\mathfrak{Z}$ ist (das heisst, für jedes $k \in \{1, \ldots, n\}$ die Zerlegung $\mathfrak{Z}'_k$ eine Verfeinerung von $\mathfrak{Z}_k$ ist), dann gilt*

$$U(f, \mathfrak{Z}) \leq U(f, \mathfrak{Z}') \leq O(f, \mathfrak{Z}') \leq O(f, \mathfrak{Z}).$$

*Insbesondere gilt für eine beschränkte Funktion $f : Q \to \mathbb{R}$ die Ungleichung $\underline{I}(f) \leq \overline{I}(f)$.*

*Beweis.* Angenommen $\mathfrak{Z}'$ ist eine Verfeinerung von $\mathfrak{Z}$. Dann ist jeder offene Teilquader $Q'_{\boldsymbol{\beta}}$ entsprechend der Zerlegung $\mathfrak{Z}'$ in einem eindeutig bestimmten Teilquader $Q_{\boldsymbol{\alpha}}$ entsprechend der Zerlegung $\mathfrak{Z}$ enthalten und es gilt $\inf f(Q'_{\boldsymbol{\beta}}) \geq \inf f(Q_{\boldsymbol{\alpha}})$ (wegen $f(Q'_{\boldsymbol{\beta}}) \subseteq f(Q_{\boldsymbol{\alpha}})$ und den Eigenschaften des Infimums). Gemeinsam mit der Additionsformel (12.1) erhält man

$$\begin{aligned} U(f, \mathfrak{Z}) &= \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} \inf(f(Q_{\boldsymbol{\alpha}})) \operatorname{vol}(Q_{\boldsymbol{\alpha}}) \\ &= \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} \inf(f(Q_{\boldsymbol{\alpha}})) \sum_{Q'_{\boldsymbol{\beta}} \subseteq Q_{\boldsymbol{\alpha}}} \operatorname{vol}(Q'_{\boldsymbol{\beta}}) \\ &\leq \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} \sum_{Q'_{\boldsymbol{\beta}} \subseteq Q_{\boldsymbol{\alpha}}} \inf(f(Q'_{\boldsymbol{\beta}})) \operatorname{vol}(Q'_{\boldsymbol{\beta}}) = U(f, \mathfrak{Z}'). \end{aligned}$$

Der Beweis von $O(f, \mathfrak{Z}') \leq O(f, \mathfrak{Z})$ ist analog.

Falls nun $\mathfrak{Z}$ und $\mathfrak{Z}'$ zwei beliebige Zerlegungen sind, dann können wir eine gemeinsame Verfeinerung $\mathfrak{Z}''$ definieren (in dem wir $\mathfrak{Z}''_k = \mathfrak{Z}_k \cup \mathfrak{Z}'_k$ für alle $k \in \{1, \ldots, n\}$ als die Vereinigung

definieren) und erhalten aus obigem Argument

$$U(f,\mathfrak{Z}) \leq U(f,\mathfrak{Z}'') \leq O(f,\mathfrak{Z}'') \leq O(f,\mathfrak{Z}')$$

wie gewünscht. □

Folgende Proposition stellt eine Verallgemeinerung von Proposition 4.12 dar.

**Proposition 12.6** (Eine erste Charakterisierung von Riemann-Integrierbarkeit)**.** *Sei $f$ eine beschränkte reellwertige Funktion auf einem abgeschlossenen Quader $Q \subseteq \mathbb{R}^n$. Die Funktion $f$ ist genau dann Riemann-integrierbar, wenn zu jedem $\varepsilon > 0$ eine Zerlegung $\mathfrak{Z}$ von $Q$ mit $O(f,\mathfrak{Z}) - U(f,\mathfrak{Z}) < \varepsilon$ existiert.*

*Beweis.* Angenommen $f$ ist Riemann-integrierbar. Ist $\varepsilon > 0$, so existieren per Definition des unteren Integrals eine Zerlegung $\underline{\mathfrak{Z}}$ mit $\underline{I}(f) - \frac{\varepsilon}{2} < U(f,\underline{\mathfrak{Z}})$ und analog eine Zerlegung $\overline{\mathfrak{Z}}$ von $Q$ mit $O(f,\overline{\mathfrak{Z}}) < \overline{I}(f) + \frac{\varepsilon}{2}$. Nach Riemann-Integrierbarkeit von $f$ gilt $\underline{I}(f) = \overline{I}(f)$ und nach Lemma 12.5 hat die gemeinsame Verfeinerung $\mathfrak{Z}$ von $\underline{\mathfrak{Z}}$ und $\overline{\mathfrak{Z}}$ die gewünschte Eigenschaft.

Die Umkehrung folgt direkt aus der Tatsache, dass für jede Zerlegung $\mathfrak{Z}$ von $Q$ die Ungleichung $\overline{I}(f) - \underline{I}(f) \leq O(f,\mathfrak{Z}) - U(f,\mathfrak{Z})$ gilt. □

### 12.1.2 Erste Eigenschaften

**Proposition 12.7** (Linearität des Riemann-Integrals)**.** *Seien $f_1, f_2$ Riemann-integrierbare reellwertige Funktionen auf einem abgeschlossenen Quader $Q \subseteq \mathbb{R}^n$ und $s_1, s_2 \in \mathbb{R}$. Dann ist auch die Linearkombination $s_1 f_1 + s_2 f_2$ Riemann-integrierbar und es gilt*

$$\int_Q (s_1 f_1 + s_2 f_2)\,\mathrm{dvol} = s_1 \int_Q f_1\,\mathrm{dvol} + s_2 \int_Q f_2\,\mathrm{dvol}.$$

*Beweis.* Sei $f : Q \to \mathbb{R}$ Riemann-integrierbar und $s > 0$. Weiter sei $\mathfrak{Z}$ eine Zerlegung von $Q$. Dann gilt

$$\begin{aligned} U(sf,\mathfrak{Z}) &= \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} \inf(sf(Q_{\boldsymbol{\alpha}}))\,\mathrm{vol}(Q_{\boldsymbol{\alpha}}) \\ &= \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} s\inf(f(Q_{\boldsymbol{\alpha}}))\,\mathrm{vol}(Q_{\boldsymbol{\alpha}}) = sU(f,\mathfrak{Z}) \end{aligned}$$

und analog $sO(f,\mathfrak{Z}) = O(sf,\mathfrak{Z})$. Insbesondere ist

$$\underline{I}(sf) = s\underline{I}(f) = s\overline{I}(f) = \overline{I}(sf).$$

Somit ist $sf$ Riemann-integrierbar und es gilt $\int_Q sf\,\mathrm{dvol} = s\int_Q f\,\mathrm{dvol}$.

Für $s = 0$ gibt es nicht zu zeigen und für $s < 0$ gilt

$$U(sf,\mathfrak{Z}) = sO(f,\mathfrak{Z}), \quad O(sf,\mathfrak{Z}) = sU(f,\mathfrak{Z})$$

für jede Zerlegung $\mathfrak{Z}$ von $Q$, womit der Rest analog folgt.

Wir zeigen nun Additivität des Integrals. Seien also $f_1, f_2 : Q \to \mathbb{R}$ zwei Riemann-integrierbare Funktionen und $\mathfrak{Z}$ eine Zerlegung von $Q$. Wir bemerken nun zuerst, dass wegen $(f_1 + f_2)(Q_{\boldsymbol{\alpha}}) \subseteq f_1(Q_{\boldsymbol{\alpha}}) + f_2(Q_{\boldsymbol{\alpha}})$ auch

$$\begin{aligned} U(f_1 + f_2, \mathfrak{Z}) &= \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} \inf((f_1 + f_2)(Q_{\boldsymbol{\alpha}})) \operatorname{vol}(Q_{\boldsymbol{\alpha}}) \\ &\geq \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} \big( \inf(f_1(Q_{\boldsymbol{\alpha}})) + \inf(f_2(Q_{\boldsymbol{\alpha}})) \big) \operatorname{vol}(Q_{\boldsymbol{\alpha}}) \\ &= U(f_1, \mathfrak{Z}) + U(f_2, \mathfrak{Z}) \end{aligned}$$

gilt. Analog zeigt man $O(f_1 + f_2, \mathfrak{Z}) \leq O(f_1, \mathfrak{Z}) + O(f_2, \mathfrak{Z})$. Nach Lemma 12.5 existiert zu $\varepsilon > 0$ nach Verfeinerung eine Zerlegung $\mathfrak{Z}$ mit $\underline{I}(f_i) - \varepsilon \leq U(f_i, \mathfrak{Z}) \leq O(f_i, \mathfrak{Z}) \leq \overline{I}(f_i) + \varepsilon$ für $i = 1, 2$. Dies ergibt

$$\begin{aligned} \underline{I}(f_1) + \underline{I}(f_2) - 2\varepsilon &\leq U(f_1, \mathfrak{Z}) + U(f_2, \mathfrak{Z}) \\ &\leq U(f_1 + f_2, \mathfrak{Z}) \\ &\leq O(f_1 + f_2, \mathfrak{Z}) \\ &\leq O(f_1, \mathfrak{Z}) + O(f_2, \mathfrak{Z}) \leq \overline{I}(f_1) + \overline{I}(f_2) + 2\varepsilon. \end{aligned}$$

Somit gilt nach Riemann-Integrierbarkeit von $f_1$ und $f_2$

$$\begin{aligned} \int_Q f_1 \operatorname{dvol} + \int_Q f_2 \operatorname{dvol} - 2\varepsilon &\leq \underline{I}(f_1 + f_2) \\ &\leq \overline{I}(f_1 + f_2) \leq \int_Q f_1 \operatorname{dvol} + \int_Q f_2 \operatorname{dvol} + 2\varepsilon \end{aligned}$$

und da $\varepsilon > 0$ beliebig war, folgt die Aussage. □

**Proposition 12.8** (Monotonie und die Dreiecksungleichung des Riemann-Integrals)**.** *Sei $Q$ ein abgeschlossener Quader. Für zwei Riemann-integrierbare Funktionen $f_1, f_2 : Q \to \mathbb{R}$ gelten folgende Monotonie-Eigenschaften:*

*(i) Falls $f_1 \geq 0$ ist, so gilt $\int_Q f_1 \operatorname{dvol} \geq 0$.*

*(ii) Falls $f_1 \leq f_2$ ist, so gilt $\int_Q f_1 \operatorname{dvol} \leq \int_Q f_2 \operatorname{dvol}$.*

*(iii) Die Funktion $|f_1| : Q \to \mathbb{R}$ ist Riemann-integrierbar und es gilt die Dreiecksungleichung*

$$\left| \int_Q f_1 \operatorname{dvol} \right| \leq \int_Q |f_1| \operatorname{dvol}.$$

*Beweis.* Die erste Aussage (i) folgt direkt aus

$$0 \leq \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} \inf(f_1(Q_{\boldsymbol{\alpha}})) \operatorname{vol}(Q_{\boldsymbol{\alpha}}) = U(f_1, \mathfrak{Z})$$

für eine beliebige Zerlegung $\mathfrak{Z}$ von $Q$. Weiter folgt (ii) nach Betrachtung von $f_2 - f_1$ und Anwendung von (i) und Linearität des Riemann-Integrals (Proposition 12.7).

Für (iii) zeigen wir, dass $f_1^+ = \max(f_1, 0)$ Riemann-integrierbar ist. Daraus folgt dann, dass $|f_1| = 2f_1^+ - f_1$ nach Linearität ebenfalls Riemann-integrierbar ist und

$$\int_Q f_1 \,\mathrm{dvol} \leq \int_Q |f_1| \,\mathrm{dvol},$$

was mit der analogen Aussage für $-f_1$ die Proposition beweist.

Die Eigenschaften der Abbildung $s \in \mathbb{R} \mapsto s^+ = \max(s, 0)$ (Monotonie und Stetigkeit würden hierfür ausreichen) implizieren, dass $\sup\{x^+ \mid x \in A\} = (\sup A)^+$ und $\inf\{x^+ \mid x \in A\} = (\inf A)^+$ für jede nichtleere beschränkte Teilmenge $A \subseteq \mathbb{R}$. Dies impliziert des Weiteren $O(f_1^+, \mathfrak{Z}) - U(f_1^+, \mathfrak{Z}) \leq O(f_1, \mathfrak{Z}) - U(f_1, \mathfrak{Z})$ für jede Zerlegung $\mathfrak{Z}$ von $Q$, womit Riemann-Integrierbarkeit von $f_1^+$ aus Proposition 12.6 folgt (siehe auch den Beweis von Satz 4.24(iii)). □

Wir wollen hier nochmals das erste Applet aus Abschnitt 4.9 betrachten, da dieses eben die grundlegenden Definitionen dieses Abschnittes anschaulich machen sollte.

**Applet 12.9** (Zelt)**.** *Wir sehen, dass wir das Volumen des Zeltes von unten und von oben abschätzen können, wodurch wir immer genauere Annäherungen für das Volumen erhalten können. Das zwei-dimensionale Integral* $\int_{[-1,1]^2}(2 - x^2 - y^2)\,\mathrm{dvol}$ *gibt das Volumen fehlerfrei an.*

**Übung 12.10** (Direkte Berechnung des Zeltes)**.** *Berechnen Sie das zwei-dimensionale Integral* $\int_{[-1,1]^2}(2 - x^2 - y^2)\,\mathrm{dvol}$ *unter Verwendung der Resultate dieses Abschnittes und der eindimensionalen Integration. Betrachten Sie dazu für reelle Zahlen* $a < b$ *beziehungsweise* $c < d$ *und eine reellwertige Riemann-integrierbare Funktion* $f$ *auf dem Intervall* $[a, b]$ *die Funktion*

$$f \circ \pi_1 : (x, y) \in [a, b] \times [c, d] \mapsto f(x).$$

*Zeigen Sie, dass* $f \circ \pi_1$ *auf dem Rechteck* $Q = [a, b] \times [c, d]$ *Riemann-integrierbar ist und*

$$\int_Q f \circ \pi_1 \,\mathrm{dvol} = (d - c) \int_a^b f(x)\,\mathrm{d}x$$

*erfüllt.*

## 12.2 Riemann-Integrierbarkeit und Stetigkeit

Stetige Funktionen sind in vielerlei Hinsicht einfacher zu verstehen als Riemann-integrierbare Funktionen. In diesem Abschnitt möchten wir (unter anderem) nachweisen, dass sich Riemann-integrierbare Funktionen zwischen stetigen Funktionen „einklemmen“ lassen. Dies wird uns später erlauben, einige Aussagen über Riemann-integrierbare Funktionen auf stetige Funktionen zu reduzieren. Eine solche Aussage ist die Approximierbarkeit des Riemann-Integrals mit Riemann-Summen (man vergleiche hierzu Abschnitt 5.5).

### 12.2.1 Stetige Funktionen sind Riemann-integrierbar

Als erstes wollen wir hier Satz 4.42 wie folgt verallgemeinern.

**Proposition 12.11** (Stetigkeit als Vorraussetzung)**.** *Sei $Q$ ein abgeschlossener Quader mit nicht-leerem Inneren. Dann ist jede stetige Funktion $f : Q \to \mathbb{R}$ auch Riemann-integrierbar.*

Im Beweis werden wir eine Zerlegung $\mathfrak{Z}$ suchen, der Teilquader in allen Koordinatenrichtungen „kleine“ Seitenlängen haben. Genauer kann man folgenden Begriff verwenden.

**Definition 12.12** (Maschenweite)**.** Für eine Zerlegung $\mathfrak{Z}$ des Quaders $Q$ definieren wir die **Maschenweite der Zerlegung** $\mathfrak{Z}$ als

$$\begin{aligned} |\mathfrak{Z}| &= \max\{\text{grösste Kantenlänge von } Q_{\boldsymbol{\alpha}} \mid Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}\} \\ &= \max_{k=1,\dots,n} \max_{\alpha_k=1,\dots,A(k)} (x_{k,\alpha_k} - x_{k,\alpha_k-1}). \end{aligned}$$

Wir erwähnen noch, dass es in diesem Kapitel oft einfacher ist, die Maximumsnorm

$$\|v\|_\infty = \max_{j=1,\dots,n} |v_j|$$

für Vektoren $v \in \mathbb{R}^n$ zu verwenden. Denn der Ball $B_r^\infty(v)$ mit Radius $r > 0$ in dieser Norm ist selbst der Würfel $(v_1 - r, v_1 + r) \times \ldots \times (v_n - r, v_n + r)$ und damit mit den bereits auftretenden Quadern besser kompatibel als zum Beispiel die Bälle in der 2-Norm.

*Beweis von Proposition 12.11.* Nach dem Satz von Heine-Borel (Satz 9.70) sind abgeschlossene Quader kompakt (sie sind abgeschlossen und beschränkt). Nach Proposition 9.77 ist die stetige Funktion $f$ auf $Q$ damit sogar gleichmässig stetig. Sei $\varepsilon > 0$. Dann existiert also ein $\delta > 0$, so dass für alle $x_1, x_2 \in Q$ gilt

$$\|x_1 - x_2\|_\infty < \delta \implies |f(x_1) - f(x_2)| < \varepsilon. \tag{12.2}$$

Sei nun $\mathfrak{Z}$ eine Zerlegung von $Q$ mit Maschenweite kleiner als $\delta$. Dann gilt

$$\begin{aligned} O(f, \mathfrak{Z}) - U(f, \mathfrak{Z}) &= \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} \big(\sup(f(Q_{\boldsymbol{\alpha}})) - \inf(f(Q_{\boldsymbol{\alpha}}))\big) \operatorname{vol}(Q_{\boldsymbol{\alpha}}) \\ &\le \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} \varepsilon \operatorname{vol}(Q_{\boldsymbol{\alpha}}) = \varepsilon \operatorname{vol}(Q), \end{aligned}$$

da $\sup(f(Q_{\boldsymbol{\alpha}})) - \inf(f(Q_{\boldsymbol{\alpha}})) \leq \varepsilon$ für alle $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$ wegen (12.2). Aus Proposition 12.6 folgt nun, dass $f$ Riemann-integrierbar ist. □

### 12.2.2 Sandwich-Charakterisierung von Riemann-Integrierbarkeit

**Proposition 12.13** (Sandwich-Charakterisierung mit stetigen Funktionen)**.** *Sei $Q$ ein abgeschlossener Quader mit nichtleerem Inneren und $f : Q \to \mathbb{R}$ beschränkt. Die Funktion $f$ ist genau dann Riemann-integrierbar, wenn es für jedes $\varepsilon > 0$ zwei stetige Funktionen $f_-, f_+ : Q \to \mathbb{R}$ gibt, die*

$$f_- \leq f \leq f_+ \quad \textit{und} \quad \int_Q (f_+ - f_-)\,\mathrm{dvol} < \varepsilon$$

*erfüllen.*

In Worten ausgedrückt ist eine beschränkte Funktionen auf einem Quader also genau dann Riemann-integrierbar, wenn sie sich beliebig genau zwischen zwei stetigen Funktionen „einquetschen“ lässt.

VO: 16.04.18

*Beweis.* Angenommen $f$ erfüllt die Sandwich-Bedingung. Dann existieren zu $\varepsilon > 0$ zwei Funktionen $f_-, f_+ \in C(Q)$ mit $f_- \leq f \leq f_+$ und $\int_Q (f_+ - f_-)\,\mathrm{dvol} < \varepsilon$. Da $f_-, f_+$ nach Proposition 12.11 Riemann-integrierbar sind, gibt es nach Definition der Riemann-Integrierbarkeit Zerlegungen $\mathfrak{Z}_-, \mathfrak{Z}_+$, so dass

$$\int_Q f_-\,\mathrm{dvol} - \varepsilon < U(f_-, \mathfrak{Z}_-)$$
$$O(f_+, \mathfrak{Z}_+) < \int_Q f_+\,\mathrm{dvol} + \varepsilon.$$

Für die gemeinsame Verfeinerung $\mathfrak{Z}$ von $\mathfrak{Z}_-$ und $\mathfrak{Z}_+$ gilt nach Lemma 12.5 und der Voraussetzung $f_- \leq f \leq f_+$, dass

$$\begin{aligned} O(f, \mathfrak{Z}) - U(f, \mathfrak{Z}) &\leq O(f_+, \mathfrak{Z}_+) - U(f_-, \mathfrak{Z}_-) \\ &\leq \int_Q (f_+ - f_-)\,\mathrm{dvol} + 2\varepsilon < 3\varepsilon \end{aligned}$$

ist. Da $\varepsilon > 0$ aber beliebig war, folgt die Riemann-Integrierbarkeit von $f$ aus der Charakterisierung in Proposition 12.6.

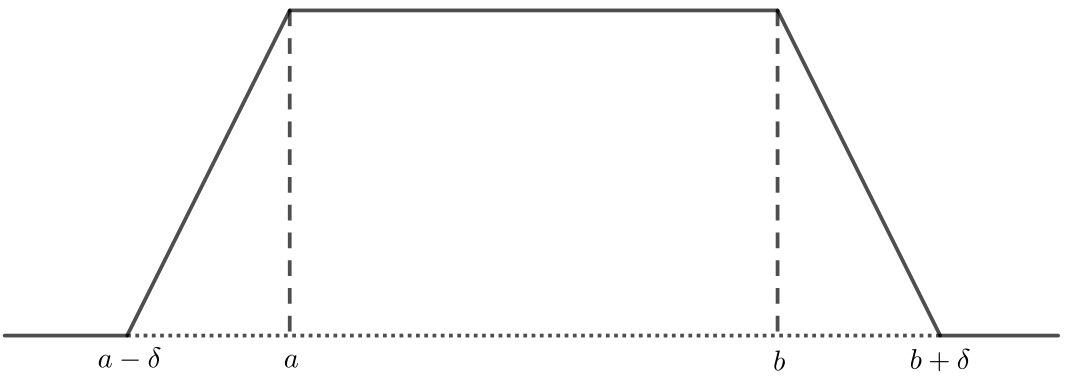
Für die Umkehrung werden wir nun schrittweise vorgehen, indem wir immer schwächere Annahmen an die Funktionen $f$ treffen.

Charakteristische Funktionen von Quadern: Angenommen $f$ ist von der Form $f = \mathbb{1}_{Q'}$ für einen abgeschlossenen Quader $Q' \subseteq Q$. Wir schreiben $Q' = [a_1', b_1'] \times \ldots \times [a_n', b_n']$ und möchten $f$ gewissermassen in allen Koordinaten durch eine stetige Funktion approximieren. Dazu definieren wir für $a < b$ in $\mathbb{R}$ und ein (noch zu bestimmendes) $\delta > 0$ eine stetige

Funktion

$$\psi_{a,b} : \mathbb{R} \to [0,1],\ x \mapsto \begin{cases} 0 & \text{falls } x < a - \delta \\ \frac{1}{\delta}(x - a + \delta) & \text{falls } a - \delta \leq x < a \\ 1 & \text{falls } x \in [a,b] \\ \frac{1}{\delta}(b + \delta - x) & \text{falls } b < x \leq b + \delta \\ 0 & \text{falls } x > b + \delta \end{cases}$$

Wir bemerken, dass $\psi_{a,b}$ gewissermassen eine stetige Approximation der charakteristischen Funktion des Intervalls $[a,b]$ darstellt.




Figur 12.1: Darstellung der Abbildung $\psi_{a,b}$.

Unter Verwendung von Funktionen der Form $\psi_{a,b}$ für $a < b$ definieren wir nun

$$h_+ : x \in Q \mapsto \prod_{k=1}^{n} \psi_{a'_k, b'_k}(x_k) \in [0,1].$$

Dann ist $h_+(x) = 1$ für alle $x \in Q'$ und damit $h_+ \geq f = \mathbb{1}_{Q'}$. Des Weiteren ist $h_+(x) = 0$ falls $x_k \notin [a'_k - \delta, b'_k + \delta]$ für ein $k \in \{1, \ldots, n\}$. Also lässt sich das Integral $\int_Q (h_+ - f)\,\mathrm{dvol}$ durch eine endliche Summe der Volumen von $2n$ Quadern beschränkt, die jeweils eine Kantenlänge kleiner gleich $\delta$ haben. Also ist $\int_Q (h_+ - f)\,\mathrm{dvol} \leq C\delta$ für ein $C > 0$ (welches von $n$ und den Längen $b_1 - a_1, \ldots, b_n - a_n$ abhängt aber von $\delta$ unabhängig ist). Die Wahl $\delta = \frac{\varepsilon}{2C}$ beweist damit den ersten Teil der Aussage.

Um eine geeignete stetige Funktion unter $f = \mathbb{1}_{Q'}$ zu finden, stellen wir zuerst fest, dass $h_- = 0$ alle gewünschten Eigenschaften erfüllt, falls $Q'$ leeres Inneres hat. Wir nehmen nun an, dass $a'_k < b'_k$ für alle $k \in \{1, \ldots, n\}$ gilt. Wir betrachten ein $\delta > 0$ wie zuvor, wobei wir nach Verkleinerung von $\delta$ annehmen können, dass

$$\delta < \tfrac{1}{3} \max_{k=1,\ldots,n} (b'_k - a'_k)$$

gilt. Sei

$$h_- : x \in Q \mapsto \prod_{k=1}^{n} \psi_{a'_k+\delta, b'_k-\delta}(x_k) \in [0,1].$$

Analoge Argumente wie für $h_+$ beweisen dann $h_- \leq f$ und $\int_Q (f - h_-)\,\mathrm{dvol} < \varepsilon$ wie gewünscht.

Treppenfunktionen: Sei nun $\mathfrak{Z}$ eine Zerlegung von $Q$ und $f : Q \to \mathbb{R}$ eine beschränkte Funktion mit der Eigenschaft, dass $f|_{Q_{\boldsymbol{\alpha}}}$ konstant ist für alle $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$. Sei $c_{\boldsymbol{\alpha}} \in \mathbb{R}$ der auf $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$ angenommene Wert und sei $M = \sup(f(Q))$. Wir bemerken, dass per Definition gilt

$$\begin{aligned}\int_Q f\,\mathrm{dvol} &= \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} c_{\boldsymbol{\alpha}} \operatorname{vol}(Q_{\boldsymbol{\alpha}}) \\ &= \int_Q \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} c_{\boldsymbol{\alpha}} \mathbb{1}_{Q_{\boldsymbol{\alpha}}}\,\mathrm{dvol}\end{aligned}$$

Wir wollen vorerst annehmen, dass $f \geq 0$ ist. Um die Werte von $f$ an den Rändern der $\mathfrak{Z}$ entsprechenden Teilquadern zu kontrollieren, bezeichnen wir mit $Q_1, \ldots, Q_N$ jene Quader, die man erhält, wenn man die Vereinigung der Ränder $\partial Q_{\boldsymbol{\alpha}}$ für alle $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$ in Quader (mit leerem Inneren) zerlegt. Seien nun $h_{+,\boldsymbol{\alpha}}, h_{-,\boldsymbol{\alpha}}$ respektive $h_{+,i}, h_{-,i}$ die aus dem vorherigen Schritt erhaltenen, stetigen Funktionen für die charakteristischen Funktionen von $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$ respektive $Q_i$. Wir definieren

$$\begin{aligned} f_-(x) &= \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} c_{\boldsymbol{\alpha}} h_{-,\boldsymbol{\alpha}} \\ f_+(x) &= \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} c_{\boldsymbol{\alpha}} h_{+,\boldsymbol{\alpha}} + \sum_{i=1}^{N} M h_{+,i}. \end{aligned}$$

Dann gilt $f_- \leq f \leq f_+$ sowie

$$\begin{aligned}\int_Q (f_+ - f_-)\,\mathrm{dvol} &= \int_Q \Big( \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} c_{\boldsymbol{\alpha}} (h_{+,\boldsymbol{\alpha}} - h_{-,\boldsymbol{\alpha}}) \Big)\,\mathrm{dvol} + M \sum_{i=1}^{N} \int_Q h_{+,i}\,\mathrm{dvol} \\ &\leq 2M |\{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}\}| \varepsilon + MN\varepsilon,\end{aligned}$$

wobei $M, |\{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}\}|$ und $N$ von $\varepsilon$ unabhängige Konstanten sind. Ist $f$ nicht zwingend nicht-negativ, so kann man $f$ zerlegen als $f_2 - f_1$ für nicht-negative Funktionen $f_1, f_2$ und die gerade bewiesene Aussage für $f_1, f_2$ verwenden.

Beliebige Riemann-integrierbare Funktionen: Sei $f : Q \to \mathbb{R}$ Riemann-integrierbar, sei $\varepsilon > 0$ und sei $\mathfrak{Z}$ eine Zerlegung mit $O(f, \mathfrak{Z}) - U(f, \mathfrak{Z}) < \varepsilon$ (siehe Proposition 12.6).

Wir setzen $M = \sup |f(Q)|$ sowie

$$g_+ : x \in Q \mapsto \begin{cases} M & \text{falls } x \in \partial Q_{\boldsymbol{\alpha}} \text{ für ein } Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z} \\ \sup(f(Q_{\boldsymbol{\alpha}})) & \text{falls } x \in Q_{\boldsymbol{\alpha}} \text{ für ein } Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z} \end{cases}$$
$$g_- : x \in Q \mapsto \begin{cases} -M & \text{falls } x \in \partial Q_{\boldsymbol{\alpha}} \text{ für ein } Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z} \\ \inf(f(Q_{\boldsymbol{\alpha}})) & \text{falls } x \in Q_{\boldsymbol{\alpha}} \text{ für ein } Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}. \end{cases}$$

Nach Annahme an $\mathfrak{Z}$ gilt

$$\int_Q (g_+ - g_-) \,\mathrm{dvol} = O(f, \mathfrak{Z}) - U(f, \mathfrak{Z}) < \varepsilon$$

sowie $g_- \leq f \leq g_+$. Wir wenden nun den vorherigen Schritt auf $g_+, g_-$ an und erhalten stetige Funktionen $f_+, f_- : Q \to \mathbb{R}$ mit $g_+ \leq f_+$ und $f_- \leq g_-$ und $\int_Q (f_+ - g_+) \,\mathrm{dvol} < \varepsilon$ und $\int_Q (g_- - f_-) \,\mathrm{dvol} < \varepsilon$. Zusammenfassend haben wir $f_- \leq f \leq f_+$ sowie

$$\int_Q (f_+ - f_-) \,\mathrm{dvol} = \int_Q \big((f_+ - g_+) + (g_+ - g_-) + (g_- - f_-)\big) \,\mathrm{dvol} < 3\varepsilon,$$

was die Aussage beweist, da $\varepsilon > 0$ beliebig war. □

### 12.2.3 Integrierbarkeit über Riemann-Summen

**Definition 12.14** (Riemann-Summen)**.** Für eine Zerlegung $\mathfrak{Z}$ eines abgeschlossenen Quaders $Q$ bezeichnen wir $\mathbf{z} = (z_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha}}$ als eine **erlaubte Wahl von Zwischenpunkten** der Zerlegung $\mathfrak{Z}$, falls $z_{\boldsymbol{\alpha}} \in Q_{\boldsymbol{\alpha}} \cup \partial Q_{\boldsymbol{\alpha}}$ für alle $\boldsymbol{\alpha}$. Für eine reellwertige Funktion $f$ auf $Q$, eine Zerlegung $\mathfrak{Z}$ von $Q$ und eine erlaubte Wahl von Zwischenpunkten $\mathbf{z}$ definieren wir die **Riemann-Summe** durch

$$R(f, \mathfrak{Z}, \mathbf{z}) = \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} f(z_{\boldsymbol{\alpha}}) \,\mathrm{vol}(Q_{\boldsymbol{\alpha}}).$$

**Satz 12.15** (Charakterisierung mittels Riemann-Summen)**.** *Sei $f : Q \to \mathbb{R}$ eine beschränkte Funktion auf einem abgeschlossenen Quader $Q \subseteq \mathbb{R}^n$. Dann ist $f$ genau dann Riemann-integrierbar, wenn die Riemann-Summen für gegen Null strebende Maschenweiten gegen eine reelle Zahl $I$ (gegeben durch $I = \int_Q f \,\mathrm{dvol}$) konvergieren. Genauer formuliert besagt letzteres, dass für jedes $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass für jede Zerlegung $\mathfrak{Z}$ von $Q$ mit Maschenweite kleiner als $\delta$ und jede erlaubte Wahl $\mathbf{z}$ von Zwischenpunkten der Zerlegung $\mathfrak{Z}$ die Abschätzung*

$$|R(f, \mathfrak{Z}, \mathbf{z}) - I| < \varepsilon$$

*gilt.*

Bevor wir mit dem Beweis beginnen, möchten wir kurz anmerken, dass für die Konvergenz der Riemann-Summen einer Riemann-integrierbaren Funktion bis auf kleinere Anpassungen der gleiche Beweis wie für Satz 5.79 durchgeführt werden kann. Die Sandwich-Charakterisierung der Riemann-Integrierbarkeit aus Proposition 12.13 wird es uns aber erlauben, einen weniger technischen Beweis zu geben.

*Beweis.* Wir nehmen zuerst an, dass die Riemann-Summen gegen $I \in \mathbb{R}$ konvergieren. Sei $\varepsilon > 0$. Dann existiert ein $\delta > 0$, so dass für jede Zerlegung $\mathfrak{Z}$ mit Maschenweite kleiner als $\delta$ und jede erlaubte Wahl von Zwischenpunkten $\mathbf{z}$ die Abschätzung

$$I - \varepsilon < R(f, \mathfrak{Z}, \mathbf{z}) < I + \varepsilon$$

erfüllt ist. Wir wählen eine Zerlegung $\mathfrak{Z}$ mit Maschenweite kleiner als $\delta$ und variieren die Zwischenpunkte, um

$$\begin{aligned} U(f, \mathfrak{Z}) &= \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} \inf(f(Q_{\boldsymbol{\alpha}})) \operatorname{vol}(Q_{\boldsymbol{\alpha}}) \\ &= \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} \inf \{ f(z_{\boldsymbol{\alpha}}) \operatorname{vol}(Q_{\boldsymbol{\alpha}}) \mid z_{\boldsymbol{\alpha}} \in Q_{\boldsymbol{\alpha}} \} \\ &\geq \inf \{ R(f, \mathfrak{Z}, \mathbf{z}) \mid \mathbf{z} \text{ erlaubt} \} \\ &\geq I - \varepsilon \end{aligned}$$

und analog ebenso

$$\begin{aligned} O(f, \mathfrak{Z}) &= \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} \sup(f(Q_{\boldsymbol{\alpha}})) \operatorname{vol}(Q_{\boldsymbol{\alpha}}) \\ &\leq \sup \{ R(f, \mathfrak{Z}, \mathbf{z}) \mid \mathbf{z} \text{ erlaubt} \} \leq I + \varepsilon \end{aligned}$$

zu erhalten. Wir bemerken, dass die ersten beiden Ungleichungen keine Gleichungen sind, da erlaubte Zwischenpunkte auch am Rand $\partial Q_{\boldsymbol{\alpha}}$ der jeweiligen Teilquader $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$ liegen dürfen. Daraus folgt $I - \varepsilon \leq \underline{I}(f) \leq \overline{I}(f) \leq I + \varepsilon$. Da $\varepsilon > 0$ aber beliebig war, ist $f$ also Riemann-integrierbar mit Riemann-Integral $\int_Q f \operatorname{dvol} = I$.

Die Umkehrung ist wie schon erwähnt die Verallgemeinerung von Satz 5.79 für mehrdimensionale Integrale. Wir betrachten hierzu zuerst eine stetige Funktion $f : Q \to \mathbb{R}$. Sei $\varepsilon > 0$. Wie bereits im Beweis von Proposition 12.11 wählen wir mittels Proposition 9.77 ein $\delta > 0$ so dass für alle $x, y \in Q$ gilt

$$\|x - y\|_\infty < \delta \implies |f(x) - f(y)| < \varepsilon.$$

Sei nun $\mathfrak{Z}$ eine Zerlegung von $Q$ mit Maschenweite kleiner als $\delta$. Dann folgt für $x, y \in \overline{Q_{\boldsymbol{\alpha}}}$ und $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$, dass $\|x - y\|_\infty < \delta$ und damit $f(y) - \varepsilon < f(x) < f(y) + \varepsilon$. Variieren wir $y \in Q_{\boldsymbol{\alpha}}$ und setzen wir $x = z_{\boldsymbol{\alpha}}$ so impliziert dies

$$\sup(f(Q_{\boldsymbol{\alpha}})) - \varepsilon \leq f(z_{\boldsymbol{\alpha}}) \leq \inf(f(Q_{\boldsymbol{\alpha}})) + \varepsilon$$

für alle $z_{\boldsymbol{\alpha}} \in \overline{Q_{\boldsymbol{\alpha}}}$. Wir multiplizieren dies mit $\operatorname{vol}(Q_{\boldsymbol{\alpha}})$, summieren über alle $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$ und erhalten dadurch

$$\begin{aligned}\int_Q f \operatorname{dvol} - \varepsilon \operatorname{vol}(Q) &\leq O(f, \mathfrak{Z}) - \varepsilon \operatorname{vol}(Q) \\ &\leq R(f, \mathfrak{Z}, \mathbf{z}) \\ &\leq U(f, \mathfrak{Z}) + \varepsilon \operatorname{vol}(Q) \leq \int_Q f \operatorname{dvol} + \varepsilon \operatorname{vol}(Q)\end{aligned}$$

für jede Zerlegung $\mathfrak{Z}$ mit Maschenweite kleiner als $\delta$ und jede erlaubte Wahl an Zwischenpunkten $\mathbf{z}$. Da $\varepsilon > 0$ beliebig war, beweist dies die gewünschte Konvergenz der Riemannsummen.

Sei nun $f : Q \to \mathbb{R}$ bloss Riemann-integrierbar und $\varepsilon > 0$. Dann existieren nach Proposition 12.13 stetige Funktionen $f_-, f_+ \in C(Q)$ mit $f_- \leq f \leq f_+$ und $\int_Q (f_+ - f_-) \operatorname{dvol} < \varepsilon$. Nach obigem existiert des Weiteren ein $\delta > 0$ so dass für alle Zerlegungen $\mathfrak{Z}$ mit Maschenweite kleiner als $\delta$ und jede erlaubte Wahl $\mathbf{z}$ von Zwischenpunkten die Abschätzungen

$$\begin{aligned}\left| R(f_-, \mathfrak{Z}, \mathbf{z}) - \int_Q f_- \operatorname{dvol} \right| &< \varepsilon \\ \left| R(f_+, \mathfrak{Z}, \mathbf{z}) - \int_Q f_+ \operatorname{dvol} \right| &< \varepsilon\end{aligned}$$

gelten. Daraus folgt nun

$$\begin{aligned}\int_Q f \operatorname{dvol} - 2\varepsilon &\leq \int_Q f_- \operatorname{dvol} - \varepsilon \\ &\leq R(f_-, \mathfrak{Z}, \mathbf{z}) \\ &\leq R(f, \mathfrak{Z}, \mathbf{z}) \\ &\leq R(f+, \mathfrak{Z}, \mathbf{z}) \\ &\leq \int_Q f_+ \operatorname{dvol} + \varepsilon \leq \int_Q f \operatorname{dvol} + 2\varepsilon\end{aligned}$$

wiederum für jede Zerlegung $\mathfrak{Z}$ mit Maschenweite kleiner als $\delta$ und jede erlaubte Wahl an Zwischenpunkten. Dies beweist die gewünschte Konvergenz für die Riemann-integrierbare Funktion $f$. $\square$

## 12.3 Das Lebesgue-Kriterium

In diesem Abschnitt wollen wir den Zusammenhang zwischen Riemann-Integrierbarkeit und Stetigkeit vollständig erklären. Wie in den letzten beiden Abschnitten wird $Q$ immer einen abgeschlossenen Quader wie in Definition 12.1 mit nichtleerem Inneren bezeichnen. VO: 18.04.18

### 12.3.1 Nullmengen

Für die genaue Charakterisierung (nach Lebesgue) der Riemann-Integrierbarkeit mittels Stetigkeit benötigen wir den folgenden Begriff.

**Definition 12.16** (Nullmengen)**.** Eine Teilmenge $N \subseteq \mathbb{R}^n$ wird eine **Nullmenge** (genauer eine **Lebesgue-Nullmenge** im $\mathbb{R}^n$) genannt, falls es zu jedem $\varepsilon > 0$ eine Folge $(Q_\ell)_\ell$ von offenen Quadern im $\mathbb{R}^n$ gibt, so dass

$$N \subseteq \bigcup_{\ell=1}^{\infty} Q_\ell, \quad \sum_{\ell=1}^{\infty} \operatorname{vol}(Q_\ell) < \varepsilon \tag{12.3}$$

erfüllt sind.

Wie bereits im Namen „Nullmenge“ enthalten ist, interpretieren wir die Aussagen in (12.3) so, dass 0 der einzig vernünftige Wert für das $n$-dimensionale Volumen von $N$ ist. In der Tat lässt sich (12.3) so lesen, dass $N$ in einer Menge $\bigcup_{\ell=1}^{\infty} Q_\ell$ enthalten ist, deren Volumen klein ist. Dabei wurde das Volumen von $\bigcup_{\ell=1}^{\infty} Q_\ell$ noch nicht definiert, aber $\sum_{\ell=1}^{\infty} \operatorname{vol}(Q_\ell)$ lässt sich als eine obere Schranke von diesem auffassen.

Beispielsweise ist eine Punktmenge $\{x\} \subseteq \mathbb{R}^n$ eine Nullmenge, da $x$ in beliebig kleinen Quadern enthalten ist.

**Lemma 12.17** (Eigenschaften von Nullmengen)**.** *Eine Teilmenge einer Nullmenge ist eine Nullmenge. Eine abzählbare Vereinigung von Nullmengen ist wiederum eine Nullmenge.*

*Beweis.* Die erste Aussage folgt direkt aus (12.3). Sei nun $N_1, N_2, \ldots, N_j, \ldots$ eine Folge von Nullmengen im $\mathbb{R}^n$ und sei $N$ ihre Vereinigung. Sei $\varepsilon > 0$. Dann existiert per Definition der Nullmengen für jedes $N_j$ eine Folge von offenen Quadern $(Q_{j,\ell})_\ell$ mit

$$N_j \subseteq \bigcup_{\ell=1}^{\infty} Q_{j,\ell}, \quad \sum_{\ell=1}^{\infty} \operatorname{vol}(Q_{j,\ell}) < \frac{\varepsilon}{2^j}.$$

Dies impliziert aber

$$N = \bigcup_{j=1}^{\infty} N_j \subseteq \bigcup_{j=1}^{\infty} \bigcup_{\ell=1}^{\infty} Q_{j,\ell},$$
$$\sum_{j=1}^{\infty} \sum_{\ell=1}^{\infty} \operatorname{vol}(Q_{j,\ell}) < \sum_{j=1}^{\infty} \frac{\varepsilon}{2^j} = \varepsilon.$$

Da $\mathbb{N} \times \mathbb{N}$ abzählbar ist und $\varepsilon > 0$ beliebig war, erhalten wir, dass $N$ eine Nullmenge ist. $\square$

**Beispiel 12.18** (Erste Beispiele von Nullmengen)**.** *Mit Lemma 12.17 ist es nun relativ leicht, viele Beispiele von Nullmengen anzugeben.*

*(a) Jede abzählbare Teilmenge im $\mathbb{R}^n$ ist eine Nullmenge. Insbesondere ist zum Beispiel $\mathbb{Q}^n$ eine Nullmenge, da $\mathbb{Q}^n$ abzählbar ist.*

*(b) Jeder Quader mit leerem Inneren ist eine Nullmenge. Insbesondere ist jeder Unterraum von Dimension kleiner $n$ eine Nullmenge im $\mathbb{R}^n$.*

*(c) Die Cantormenge ist eine Nullmenge im $\mathbb{R}^n$.*

**Wichtige Übung 12.19** (Einige Nullmengen)**.** *Beweisen Sie die Aussagen in obigem Beispiel.*

Es wäre eine unangenehme Überraschung, wenn sich nun herausstellen würde, dass sogar $\mathbb{R}^n$ eine Nullmenge im $\mathbb{R}^n$ ist. Dass dies nicht so ist, ist eine Konsequenz der Vollständigkeit von $\mathbb{R}^n$ (oder genauer formuliert der Kompaktheit von Quadern).

**Proposition 12.20** (Nicht-Nullmengen)**.** *Ein Quader $Q \subseteq \mathbb{R}^n$ mit nicht-leerem Inneren ist keine Nullmenge im $\mathbb{R}^n$.*

Nach Lemma 12.17 ist also auch $\mathbb{R}^n$ keine Nullmenge.

*Beweis.* Angenommen $Q = [a_1, b_1] \times \cdots \times [a_n, b_n]$ ist ein abgeschlossener Quader definiert durch $a_1 < b_1, \ldots, a_n < b_n$. Dann ist $Q$ nach dem Satz von Heine Borel (Satz 9.70) kompakt. Falls nun $Q$ eine Nullmenge ist, dann existieren offene Quader $O_1, O_2, \ldots$ im $\mathbb{R}^n$ mit

$$Q \subseteq \bigcup_{\ell=1}^{\infty} O_\ell, \quad \sum_{\ell=1}^{\infty} \operatorname{vol}(O_\ell) < \frac{\operatorname{vol}(Q)}{2}.$$

Nach Satz 9.66(3) (Kompaktheit ist äquivalent zu abzählbarer Überdeckungskompaktheit) existiert dann aber auch ein $m \in \mathbb{N}$ mit

$$Q \subseteq \bigcup_{\ell=1}^{m} O_\ell, \quad \sum_{\ell=1}^{m} \operatorname{vol}(O_\ell) < \frac{\operatorname{vol}(Q)}{2}. \tag{12.4}$$

Wir definieren die Quader $Q_\ell = \overline{O_\ell} \cap Q$ für alle $\ell \in \{1, \ldots, m\}$ und schreiben

$$Q_\ell = [a_{\ell,1}, b_{\ell,1}] \times \ldots \times [a_{\ell,n}, b_{\ell,n}] \subseteq Q.$$

Insbesondere gilt $Q = \bigcup_{\ell=1}^{m} Q_\ell$. Wir zerlegen nun die Quader $Q_\ell$ für alle $\ell \in \{1, \ldots, m\}$ weiter, um eine Zerlegung von $Q$ zu erhalten. Genauer lässt sich für jedes $k \in \{1, \ldots, n\}$ eine Zerlegung $\mathfrak{Z}_k$ von $[a_k, b_k]$ durch Anordnen der Punkte $\{a_{1,k}, b_{1,k}, \ldots, a_{m,k}, b_{m,k}\}$ definieren. Damit ist $\mathfrak{Z} = (\mathfrak{Z}_1, \ldots, \mathfrak{Z}_n)$ eine Zerlegung von $Q$. Nach Gleichung (12.1) gilt

$$\operatorname{vol}(Q) = \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} \operatorname{vol}(Q_{\boldsymbol{\alpha}}).$$

Per Definition der Zerlegung $\mathfrak{Z}$ und (12.4) ist jeder der $\mathfrak{Z}$ entsprechenden Quader $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$ in einem (oder mehreren) der Quader $Q_\ell$ enthalten. Sammeln wir nun alle Quader $Q_{\boldsymbol{\alpha}}$ zusammen, die in einem Quader $Q_\ell$ enthalten sind, so ergibt sich nach der Additionsformel (12.1)

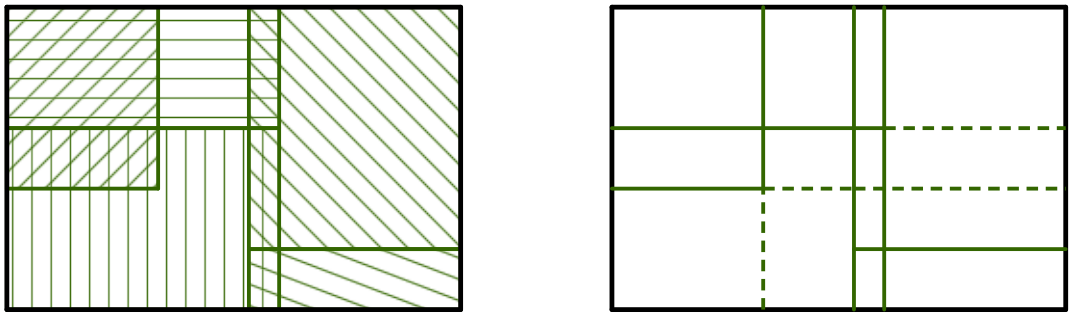
$$\sum_{\substack{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z} \\ Q_{\boldsymbol{\alpha}} \subseteq Q_\ell}} \operatorname{vol}(Q_{\boldsymbol{\alpha}}) = \operatorname{vol}(Q_\ell).$$

Summiert man nun über $\ell = 1, \ldots, m$, so ergibt sich

$$\operatorname{vol}(Q) \leq \sum_{\ell=1}^{m} \operatorname{vol}(Q_\ell),$$

wobei im Allgemeinen nur eine Ungleichung gilt, da ein $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$ möglicherweise in mehreren der Quader $Q_\ell$ liegen kann. Dies widerspricht aber (12.4).

Falls nun $Q$ ein Quader im $\mathbb{R}^n$ mit nichtleerem Inneren ist, so gibt es einen abgeschlossenen Quader $Q' \subseteq Q$ mit nichtleerem Inneren. Nach Obigem ist dann $Q'$ keine Nullmenge, und Lemma 12.17 zeigt, dass auch $Q$ keine Nullmenge sein kann. □



Figur 12.2: Nach Reduktion zu einer endlichen Überdeckung des Quaders $Q$ ist es wesentlich einfacher, zu zeigen, dass das Gesamtvolumen der überdeckenden Quader mindestens so gross ist wie das Volumen von $Q$. In der Tat, für letzteres fabriziert man aus $Q_1, \ldots, Q_m$ eine Zerlegung, wie in obigem Bild illustriert ist, und wendet dann die Additionsformel (12.1) an.

**Wichtige Übung 12.21** (Über die Definition von Nullmengen)**.** *Zeigen Sie, dass eine Teilmenge $N \subseteq \mathbb{R}^n$ genau dann eine Nullmenge ist, wenn es für jedes $\varepsilon > 0$ eine Folge $(Q_\ell)_\ell$ abgeschlossener (!) Quader im $\mathbb{R}^n$ gibt mit*

$$N \subseteq \bigcup_{\ell=1}^{\infty} Q_\ell, \quad \sum_{\ell=1}^{\infty} \operatorname{vol}(Q_\ell) < \varepsilon.$$

**Proposition 12.22.** *Sei $Q \subseteq \mathbb{R}^{n-1}$ ein abgeschlossener Quader und $f : Q \to \mathbb{R}$ Riemann-integrierbar. Dann ist der Graph*

$$\operatorname{graph}(f) = \{(x, f(x)) \mid x \in Q\}$$

*von $f$ eine Nullmenge in $\mathbb{R}^n$.*

*Beweis.* Sei $M = \sup|f(Q)|$. Da $f$ als Riemann-integrierbar vorrausgesetzt wurd, existiert für jedes $\varepsilon > 0$ eine Zerlegung $\mathfrak{Z}$ von $Q$ so dass $O(f,\mathfrak{Z}) - U(f,\mathfrak{Z}) < \varepsilon$. Zu jedem $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$ definieren wir den Quader $P_{\boldsymbol{\alpha}} = Q_{\boldsymbol{\alpha}} \times [\inf(f(Q_{\boldsymbol{\alpha}})), \sup(f(Q_{\boldsymbol{\alpha}}))]$ und erhalten

$$\operatorname{graph}(f) \subseteq \bigcup_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} P_{\boldsymbol{\alpha}} \cup \underbrace{\bigcup_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} \big(\partial Q_{\boldsymbol{\alpha}} \times [-M, M]\big)}_{=R_{\mathfrak{Z}} \times [-M,M]},$$

wobei die zweite Vereinigung sozusagen dem Raster $R_{\mathfrak{Z}} = \bigcup_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} \partial Q_{\boldsymbol{\alpha}} \subseteq \mathbb{R}^{n-1}$ der Zerlegung $\mathfrak{Z}$ von $Q$ entspricht. Nun bemerken wir, dass

$$\sum_{Q_{\boldsymbol{\alpha}} \sqsubseteq \mathfrak{Z}} \operatorname{vol}(P_{\boldsymbol{\alpha}}) = O(f,\mathfrak{Z}) - U(f,\mathfrak{Z}) < \varepsilon.$$

Des Weiteren ist $R_{\mathfrak{Z}}$ als Teilmenge einer endlichen Vereinigung von Hyperebenen eine Nullmenge in $\mathbb{R}^{n-1}$. Daher können wir $R_{\mathfrak{Z}} \times [-M, M]$ mit einer (ebenso endlichen) Vereinigung von weiteren Quadern überdecken, deren Volumina in Summe auch kleiner als $\varepsilon$ ist. Wir erhalten also, dass $\operatorname{graph}(f)$ von endlich vielen Quadern mit Summe der Volumina kleiner als $2\varepsilon$ überdeckt werden kann. Da $\varepsilon > 0$ beliebig war, folgt die Proposition (aus Übung 12.21). □

### 12.3.2 Charakterisierung mittels Stetigkeit

Mit dem Begriff der Nullmenge sind wir nun in der Lage, ein genaues Kriterium anzugeben, welches Riemann-Integrierbarkeit und Stetigkeit verknüpft.

**Satz 12.23** (Lebesgue-Kriterium für Riemann-Integrierbarkeit)**.** *Sei $f$ eine beschränkte reellwertige Funktion auf einem abgeschlossenen Quader $Q$. Die Funktion $f$ ist genau dann Riemann-integrierbar, wenn die Menge*

$$N = \{x \in Q \mid f \text{ ist unstetig in } x\}$$

*eine Nullmenge im $\mathbb{R}^n$ ist.*

Die zweite Bedingung wird auch kurz als „$f$ ist fast überall stetig“ bezeichnet. Allgemeiner sagt man, dass eine Aussage $A(x)$ über Elemente $x \in \mathbb{R}^n$ (oder $x$ in einer Teilmenge von $\mathbb{R}^n$) *fast überall* gilt falls

$$N = \{x \mid A(x) \text{ gilt nicht}\}$$

eine Nullmenge ist.

Für den Beweis werden wir den Begriff der Oszillation (siehe Abschnitt 9.5.6) verwenden, den wir nun kurz rekapitulieren wollen.

### 12.3.3 Oszillation einer Funktion

Wir betrachten eine beschränkte Funktion $f : Q \to \mathbb{R}$ auf einem Quader $Q$. Für solche Funktionen wurde in Definition 9.78 die Oszillation eingeführt, welche im Wesentlichen die

Schwankung der Funktion um jeden Punkt beschreibt. Wir wollen hier nochmals die Maximumsnorm auf $\mathbb{R}^n$ verwenden und den $\delta$-Ball in der Unendlich-Norm um $x$ mit $B_\delta^\infty(x)$ bezeichnen. Dann ist die Oszillation bei $x \in Q$ durch

$$\omega(f,x) = \lim_{\delta \searrow 0} \omega(f,x,\delta),$$

definiert, wobei für $\delta > 0$

$$\omega(f,x,\delta) = \sup(f(Q \cap B_\delta^\infty(x))) - \inf(f(Q \cap B_\delta^\infty(x))).$$

Ist $X \subseteq Q$ abgeschlossen (und damit nach dem Satz von Heine-Borel – Satz 9.70 – kompakt) und die Oszillation von $f$ erfüllt $\omega(f,x) \leq \eta$ für alle $x \in X$, so lassen sich auch die Schwankungen $\omega(f,x,\delta)$ für ein geeignetes und gleichmässiges $\delta$ durch $\eta + \varepsilon$ kontrollieren. Für die genaue Aussage siehe Proposition 9.79.

Im Beweis des Lebesgue-Kriteriums werden wir eine weitere Aussage über die Oszillation benötigen, die wir hier im allgemeinen Kontext formulieren und beweisen.

**Lemma 12.24.** *Sei $f : X \to \mathbb{R}$ eine beschränkte Funktion auf einem metrischen Raum $X$. Für jedes $\eta \geq 0$ ist die Teilmenge $N_\eta = \{x \in X \mid \omega(f,x) \geq \eta\} \subseteq X$ abgeschlossen.*

*Beweis.* Sei $\eta \geq 0$ und sei $(x_k)_k$ eine konvergente Folge in $Q$ mit $\omega(f,x_k) \geq \eta$ für alle $k \in \mathbb{N}$ und mit Grenzwert $x \in X$. Sei $\delta > 0$ beliebig. Dann existiert ein $k$ mit $x_k \in B_\delta(x)$. Da $B_\delta(x)$ offen ist, gibt es des Weiteren ein $\delta_k > 0$ mit $B_{\delta_k}(x_k) \subseteq B_\delta(x)$. Nun ist

$$\begin{aligned} \sup f(B_\delta(x)) &\geq \sup(f(B_{\delta_k}(x_k))), \\ \inf f(B_\delta(x)) &\leq \inf(f(B_{\delta_k}(x_k))), \end{aligned}$$

und daher $\omega(f,x,\delta) \geq \omega(f,x_k,\delta_k)$. Da aber $\omega(f,x_k,\delta_k) \geq \omega(f,x_k) \geq \eta$ gilt, erhalten wir $\omega(f,x,\delta) \geq \eta$. Da $\delta > 0$ beliebig war, folgt das Lemma. □

**Übung 12.25.** *Zeigt Obiges, dass die Menge $\{x \in X \mid f$ ist in $x$ unstetig$\}$ abgeschlossen ist? Begründen Sie Ihre Antwort mit einem Beweis oder einem Gegenbeispiel.*

### 12.3.4 Das Lebesgue-Kriterium ist notwendig

*Beweis von Satz 12.23, erste Hälfte.* Wir nehmen zuerst an, dass die beschränkte reellwertige Funktion $f$ auf $Q$ Riemann-integrierbar ist. Seien $\eta > 0$ und $\varepsilon > 0$ beliebig. Dann existiert nach Proposition 12.6 eine Zerlegung $\mathfrak{Z}$ von $Q$, so dass $O(f,\mathfrak{Z}) - U(f,\mathfrak{Z}) < \varepsilon\eta$ oder äquivalenterweise

$$\sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} \big(\sup f(Q_{\boldsymbol{\alpha}}) - \inf f(Q_{\boldsymbol{\alpha}})\big) \operatorname{vol}(Q_{\boldsymbol{\alpha}}) < \varepsilon\eta.$$

Wir definieren des Weiteren

$$\operatorname{Bad}(\eta) = \{\boldsymbol{\alpha} \mid \sup f(Q_{\boldsymbol{\alpha}}) - \inf f(Q_{\boldsymbol{\alpha}}) \geq \eta\}$$

und erhalten

$$\begin{aligned}\eta \sum_{\boldsymbol{\alpha}\in\mathrm{Bad}(\eta)} \mathrm{vol}(Q_{\boldsymbol{\alpha}}) &\leq \sum_{\boldsymbol{\alpha}\in\mathrm{Bad}(\eta)} \big(\sup f(Q_{\boldsymbol{\alpha}}) - \inf f(Q_{\boldsymbol{\alpha}})\big)\,\mathrm{vol}(Q_{\boldsymbol{\alpha}}) \\ &< \varepsilon\eta\end{aligned}$$

und nach Division mit $\eta$ damit

$$\sum_{\boldsymbol{\alpha}\in\mathrm{Bad}(\eta)} \mathrm{vol}(Q_{\boldsymbol{\alpha}}) < \varepsilon.$$

Nun betrachten wir die Menge

$$N_\eta = \{x \in Q \mid \omega(f,x) \geq \eta\}$$

aus Lemma 12.24. Für $x \in Q_{\boldsymbol{\alpha}}$ mit $\boldsymbol{\alpha} \notin \mathrm{Bad}(\eta)$ gibt es ein $\delta > 0$ mit $B_\delta^\infty(x) \subseteq Q_{\boldsymbol{\alpha}}$ (da der Teilquader $Q_{\boldsymbol{\alpha}}$ nach Definition offen ist), womit

$$\omega(f,x) \leq \omega(f,x,\delta) \leq \sup(f(Q_{\boldsymbol{\alpha}})) - \inf(f(Q_{\boldsymbol{\alpha}})) < \eta$$

folgt. Damit gilt also für jedes $x \in N_\eta$, dass entweder $x$ ein Element eines Randes eines der Teilquader ist oder dass $x \in Q_{\boldsymbol{\alpha}}$ für ein $\boldsymbol{\alpha} \in \mathrm{Bad}(\eta)$, in Symbolen

$$\begin{aligned}N_\eta \subseteq \bigcup_{Q_{\boldsymbol{\alpha}}\sqsubset 3} \partial Q_{\boldsymbol{\alpha}} \cup \bigcup_{\boldsymbol{\alpha}\in\mathrm{Bad}(\eta)} Q_{\boldsymbol{\alpha}}, \\ \sum_{\boldsymbol{\alpha}\in\mathrm{Bad}(\eta)} \mathrm{vol}(Q_{\boldsymbol{\alpha}}) < \varepsilon.\end{aligned}$$

Da aber $\partial Q_{\boldsymbol{\alpha}}$ durch endlich viele offene Quadern mit beliebig kleinem Gesamtvolumen überdeckt werden kann, können wir auch eine endliche Überdeckung von $N_\eta$ durch offene Quader finden, so dass das Gesamtvolumen kleiner als $\varepsilon$ ist. Dies zeigt, dass $N_\eta$ eine Nullmenge ist, da $\varepsilon > 0$ beliebig war.

Wir wenden nun obiges für $\eta = \frac{1}{k}$ zu $k \in \mathbb{N}$ an und erhalten aus Lemma 12.17, dass

$$\begin{aligned}N &= \{x \in Q \mid f \text{ ist bei } x \text{ unstetig}\} \\ &= \{x \in Q \mid \omega(f,x) > 0\} \\ &= \bigcup_{k=1}^{\infty} N_{\frac{1}{k}}\end{aligned}$$

eine Nullmenge ist. □

### 12.3.5 Das Lebesgue-Kriterium ist hinreichend

*Beweis von Satz 12.23, zweite Hälfte.* Wir nehmen nun an, dass $f : Q \to \mathbb{R}$ beschränkt ist und $N = \{x \in Q \mid f \text{ ist bei } x \text{ unstetig}\}$ eine Nullmenge ist. Des Weiteren definieren wir die

obere Schranke $M = \sup |f(Q)|$.

Sei $\varepsilon > 0$. Dann ist $N_\varepsilon = \{x \in Q \mid \omega(f,x) \geq \varepsilon\} \subseteq N$ nach Lemma 12.17 eine Nullmenge, nach Lemma 12.24 abgeschlossen und nach dem Satz von Heine-Borel (Satz 9.70) kompakt. Per Definition von Nullmengen und abzählbarer Überdeckungskompaktheit in Satz 9.66(3) folgt daher, dass es eine endliche Überdeckung

$$N_\varepsilon \subseteq \bigcup_{\ell=1}^{m} O_\ell$$

mit offenen Quadern $O_1, \ldots, O_m$ gibt, so dass

$$\sum_{\ell=1}^{m} \operatorname{vol}(O_\ell) < \varepsilon \tag{12.5}$$

gilt.

Wir betrachten nun die Teilmenge $K = Q \setminus \bigcup_{\ell=1}^{m} O_\ell$. Wiederum nach dem Satz von Heine-Borel ist $K$ kompakt. Nach Konstruktion gilt $K \subseteq Q \setminus N_\varepsilon$ und damit $\omega(f,x) < \varepsilon$ für alle $x \in K$. Nach Proposition 9.79 angewendet auf $\eta = \varepsilon$ und $\varepsilon$ gibt es also ein $\delta > 0$, so dass für alle $x \in K$ die Abschätzung

$$\omega\big(f|_K, x, \delta\big) < 2\varepsilon \tag{12.6}$$

gilt.

Wir wählen nun eine Zerlegung $\mathfrak{Z}$, so dass die Maschenweite kleiner als $\delta$ ist und jeder der Quader $\overline{O_1} \cap Q, \ldots, \overline{O_m} \cap Q$ eine Vereinigung von der Zerlegung $\mathfrak{Z}$ entsprechenden abgeschlossenen Quadern ist. Wir betrachten schlussendlich die Differenz

$$O(f,\mathfrak{Z}) - U(f,\mathfrak{Z}) = \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} \big(\sup(f(Q_{\boldsymbol{\alpha}})) - \inf(f(Q_{\boldsymbol{\alpha}}))\big) \operatorname{vol}(Q_{\boldsymbol{\alpha}})$$

und möchten zeigen, dass diese klein ist. Dazu trennen wir die Summe in zwei Teile auf. Im ersten summieren wir gerade über jene Teilquader von $\mathfrak{Z}$, die Teil eines Quaders $\overline{O_\ell}$ sind und im zweiten Teil summieren wir über die restlichen Quader. Auf die erste Summe wenden wir nun (12.5) und auf die zweite (12.6) an. Dies ergibt

VO: 20.04.18

$$\begin{aligned} O(f,\mathfrak{Z}) - U(f,\mathfrak{Z}) &\leq \sum_{\ell=1}^{m} \sum_{\substack{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z} \\ Q_{\boldsymbol{\alpha}} \subseteq O_\ell}} 2M \operatorname{vol}(Q_{\boldsymbol{\alpha}}) + \sum_{\substack{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z} \\ Q_{\boldsymbol{\alpha}} \subseteq K}} 2\varepsilon \operatorname{vol}(Q_{\boldsymbol{\alpha}}) \\ &\leq 2M\varepsilon + 2\varepsilon \operatorname{vol}(Q) = (2M + 2\operatorname{vol}(Q))\varepsilon. \end{aligned}$$

Da $\varepsilon > 0$ beliebig war, folgt nun die Riemann-Integrierbarkeit von $f$. $\square$

## 12.4 Das Riemann-Integral über Jordan-messbare Mengen

Wir wenden uns nun von der Integration über achsenparallele Quader ab und möchten allgemeinere Mengen wie zum Beispiel die Kreisscheibe im $\mathbb{R}^2$ betrachten können.

**Definition 12.26** (Jordan-Messbarkeit)**.** Eine Teilmenge $B$ von $\mathbb{R}^n$ heisst **Jordan-messbar**, falls es einen abgeschlossenen Quader $Q$ in $\mathbb{R}^n$ mit $Q \supseteq B$ gibt, so dass die charakteristische Funktion $\mathbb{1}_B$ auf $Q$ Riemann-integrierbar ist. Das **Volumen** (der **Inhalt** oder das **Jordan-Mass**) ist in diesem Fall durch

$$\mathrm{vol}(B) = \int_Q \mathbb{1}_B \,\mathrm{dvol}$$

definiert.

Mit dem Lebesgue-Kriterium aus Satz 12.23 erhalten wir das folgende Resultat, welches insbesondere zeigt, dass Jordan-Messbarkeit einer Teilmenge $B \subseteq \mathbb{R}^n$ nicht von der Wahl des Quaders $Q$ mit $Q \supseteq B$ abhängt (siehe auch Lemma 12.30 für die analoge Aussage zum Volumen).

**Korollar 12.27.** *Eine Teilmenge $B \subseteq \mathbb{R}^n$ ist genau dann Jordan-messbar, wenn $B$ beschränkt ist und der Rand $\partial B$ eine Nullmenge ist. Falls $B_1, B_2 \subseteq \mathbb{R}^n$ Jordan-messbar sind, so sind auch $B_1 \cup B_2$, $B_1 \cap B_2$ und $B_1 \setminus B_2$ Jordan-messbar.*

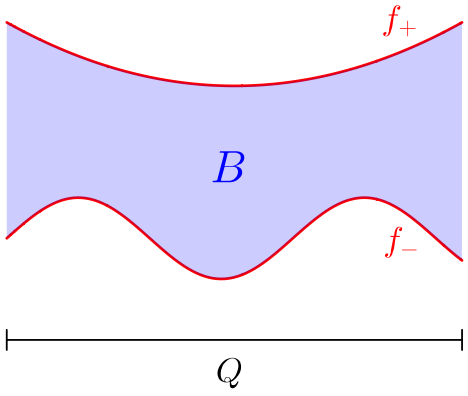
*Schlussendlich gilt für einen abgeschlossenen Quader $Q \subseteq \mathbb{R}^{n-1}$ und zwei stetige Funktionen $f_-, f_+ : Q \to \mathbb{R}$ mit $f_- \leq f_+$, dass die Menge*

$$B = \{(x, y) \in \mathbb{R}^n \mid x \in Q,\ f_-(x) \leq y \leq f_+(x)\}$$

*zwischen den Graphen von $f_-$ und $f_+$ (siehe das folgende Bild) Jordan-messbar ist. Das gleiche gilt, wenn man zwei Riemann-integrierbare Funktionen $f_-, f_+$ auf $Q$ mit $f_- \leq f_+$ und eine Jordan-messbare Teilmenge $D \subseteq Q$ verwendet um*

$$B = \{(x, y) \in \mathbb{R}^n \mid x \in D,\ f_-(x) \leq y \leq f_+(x)\}$$

*zu definieren.*



*Beweis.* Wir zeigen zuerst die erste Aussage. Sei $B \subseteq \mathbb{R}^n$ eine Teilmenge.

Falls es für ein $x \in B$ ein $\varepsilon > 0$ mit $B_\varepsilon(x) \subseteq B$ gibt, so ist $\mathbb{1}_B$ in $x$ stetig (da konstant 1 in einer Umgebung). Falls es für ein $x \in \mathbb{R}^n \setminus B$ ein $\varepsilon > 0$ mit $B_\varepsilon(x) \subseteq \mathbb{R}^n \setminus B$ gibt, so ist $\mathbb{1}_B$ in $x$ stetig (da konstant 0 in einer Umgebung). Für $x \in \partial B$ hingegen gibt es für jedes $\varepsilon > 0$ Punkte $y_-, y_+ \in B_\varepsilon(x)$ mit $\mathbb{1}_B(y_-) = 0$ und $\mathbb{1}_B(y_+) = 1$, womit also $\mathbb{1}_B$ in $x$ unstetig ist. In anderen Worten gilt

$$\{x \in \mathbb{R}^n \mid \mathbb{1}_B \text{ ist in } x \text{ unstetig}\} = \partial B. \tag{12.7}$$

Ist $\partial B$ nun eine Nullmenge, so ist $\mathbb{1}_B$ auf jedem abgeschlossenen Quader $Q \supseteq B$ Riemann-integrierbar auf Grund des Lebesgue-Kriteriums (Satz 12.23). Ist umgekehrt $\mathbb{1}_B$ auf $Q$ Riemann-integrierbar, so ist $\mathbb{1}_B|_Q$ ebenso wegen dem Lebesgue-Kriterium (Satz 12.23) fast überall auf $Q$ stetig. Die Stetigkeit von $\mathbb{1}_B|_Q$ bei $x \in Q$ unterscheidet sich von der Stetigkeit von $\mathbb{1}_B$ bei $x \in Q$ höchstens wenn $x \in \partial Q$. (Dieser Unterschied entsteht, da zum Beispiel $\mathbb{1}_Q|_Q$ als konstante Funktion auf ganz $Q$ stetig ist.) Da aber $\partial Q$ eine Nullmenge ist, erhalten wir trotzdem, dass $\partial B$ eine Nullmenge ist.

Seien nun $B_1, B_2 \subseteq Q$ Jordan-messbar. Wir bemerken zuerst, dass die Formeln

$$\begin{aligned}
\mathbb{1}_{B_1 \cap B_2} &= \mathbb{1}_{B_1} \cdot \mathbb{1}_{B_2} \\
\mathbb{1}_{B_1 \cup B_2} &= \mathbb{1}_{B_1} + \mathbb{1}_{B_2} - \mathbb{1}_{B_1 \cap B_2} \\
\mathbb{1}_{B_1 \setminus B_2} &= \mathbb{1}_{B_1} - \mathbb{1}_{B_1} \cdot \mathbb{1}_{B_2}
\end{aligned}$$

gelten, was aus direktem Nachrechnen durch Fallunterscheidungen folgt. Nach Gleichung (12.7) sind die Ränder $\partial(B_1 \cup B_2)$, $\partial(B_1 \cap B_2)$ und $\partial(B_1 \setminus B_2)$ in der Vereinigung der Unstetigkeitsstellen von $\mathbb{1}_{B_1}$ und $\mathbb{1}_{B_2}$ oder also in $\partial B_1 \cup \partial B_2$ enthalten. Nach obigem und Lemma 12.17 ist aber $\partial B_1 \cup \partial B_2$, und damit sind $\partial(B_1 \cup B_2)$, $\partial(B_1 \cap B_2)$ und $\partial(B_1 \setminus B_2)$ Nullmengen. Wiederum nach dem Lebesgue-Kriterium sind also $B_1 \cup B_2$, $B_1 \cap B_2$ und $B_1 \setminus B_2$ Jordan-messbar.

Sei nun $Q \subseteq \mathbb{R}^{n-1}$ ein abgeschlossener Quader, seien $f_-, f_+ : Q \to \mathbb{R}$ stetige Funktionen mit $f_- \leq f_+$ und sei $B$ wie im Korollar. Dann ist $B$ beschränkt (auf Grund der Kompaktheit von $Q$) und der Rand von $B$ ist in der Vereinigung

$$(\partial Q \times \mathbb{R}) \cup \operatorname{graph}(f_-) \cup \operatorname{graph}(f_+)$$

enthalten (wieso?). Nach Proposition 12.22 sind auch $\operatorname{graph}(f_-)$ und $\operatorname{graph}(f_+)$ Nullmengen. Da die Teilmenge $\partial Q \times \mathbb{R}$ in einer endlichen Vereinigung von Hyperebenen enthalten ist, ist sie ebenso eine Nullmenge. Inbesondere ist der Rand von $B$ als Teilmenge einer Vereinigung von Nullmengen eine Nullmenge, womit $B$ Jordan-messbar ist.

Für den allgemeinen Fall begnügen wir uns mit einer Skizze und nehmen nochmals an, dass $D = Q \subseteq \mathbb{R}^{n-1}$ ein Quader und $f_-, f_+ : D \to \mathbb{R}$ Riemann-integrierbare Funktionen mit $f_- \leq f_+$ sind. Wir definieren $M = \max\{\sup|f_-(Q)|, \sup|f_+(Q)|\}$, die Menge $B$ wie im

Korollar und die Nullmenge $N_Q = \{x \in Q \mid f_- \text{ oder } f_+ \text{ ist in } x \text{ unstetig}\}$. Dann gilt

$$\begin{aligned} N &= \{(x,y) \in Q \times (-M,M) \mid \mathbb{1}_B \text{ ist in } (x,y) \text{ unstetig}\} \\ &\subseteq (N_Q \times [-M,M]) \cup \operatorname{graph}(f_-) \cup \operatorname{graph}(f_+) \end{aligned} \tag{12.8}$$

(siehe Übung 12.28). Da $N_Q$ eine Nullmenge im $\mathbb{R}^{n-1}$ ist, folgt, dass $N_Q \times [-M,M]$ eine Nullmenge im $\mathbb{R}^n$ ist (siehe Übung 12.28). Des Weiteren sind $\operatorname{graph}(f_-)$ und $\operatorname{graph}(f_+)$ nach Proposition 12.22 Nullmengen im $\mathbb{R}^n$.

Falls $D \subseteq Q$ eine Jordan-messbare Teilmenge ist, so ist auch $D \times [-M,M]$ Jordan-messbar (siehe Übung 12.28) und wir können den Bereich der Punkte $(x,y) \in D \times \mathbb{R}$ zwischen zwei Riemann-integrierbare Funktionen $f_-, f_+$ als Durchschnitt der Jordan-messbaren Mengen $D \times [-M,M]$ und der Menge, die durch die Bedingung $f_-(x) \leq y \leq f_+(x)$ für $x \in Q$ definiert wird, erhalten. □

**Übung 12.28** (Jordan-Messbarkeit von Teilmengen zwischen Graphen)**.** *In dieser Übung möchten wir nun die notwendigen Details für die letzte Aussage in Korollar 12.27 nachliefern.*

*(i) Zeigen Sie* (12.8)*.*

*(ii) Zeigen Sie, dass für eine Nullmenge $N \subseteq \mathbb{R}^{n-1}$ auch $N \times \mathbb{R}$ eine Nullmenge in $\mathbb{R}^n$ ist. Zeigen Sie, dass für eine Jordan messbare Menge $D \subseteq \mathbb{R}^{n-1}$ auch $D \times [a,b]$ (für beliebige $a < b$) eine Jordan-messbare Teilmenge von $\mathbb{R}^n$ ist.*

*(iii) Schliessen Sie auf die letzte Aussage im Korollar.*

**Definition 12.29** (Das Riemann-Integral auf $B$)**.** Sei $B \subseteq \mathbb{R}^n$ eine Jordan-messbare Teilmenge und sei $f$ eine reellwertige Funktion auf $B$. Dann heisst $f$ **Riemann-integrierbar**, falls es einen abgeschlossenen Quader $Q \subseteq \mathbb{R}^n$ mit $B \subseteq Q$ gibt, so dass die Funktion

$$x \in Q \mapsto (\mathbb{1}_B f)(x) = \begin{cases} f(x) & \text{falls } x \in B \\ 0 & \text{falls } x \in Q \setminus B \end{cases}$$

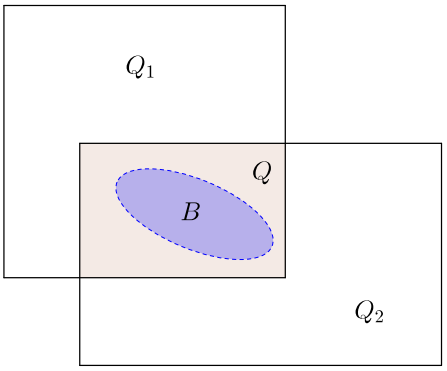
auf $Q$ Riemann-integrierbar ist.Wir schreiben in diesem Fall $f \in \mathcal{R}(B)$ und nennen

$$\int_B f \,\mathrm{dvol} = \int_Q \mathbb{1}_B f \,\mathrm{dvol}$$

das **Riemann-Integral** von $f$ über $B$.

**Lemma 12.30** (Wohldefiniertheit des Riemann-Integrals)**.** *Die Wahl des abgeschlossenen Quaders $Q \supseteq B$ in Definition 12.29 beinflusst die Riemann-Integrierbarkeit einer Funktion $f : B \to \mathbb{R}$ und den Wert des Riemann-Integrals $\int_B f \,\mathrm{dvol}$ nicht.*

*Beweis.* Angenommen $Q_1, Q_2 \subseteq \mathbb{R}^n$ sind zwei abgeschlossene Quader mit $B \subseteq Q_1$ und $B \subseteq Q_2$. Dann erfüllt der Quader $Q = Q_1 \cap Q_2$ ebenso $B \subseteq Q$; siehe folgendes Bild.



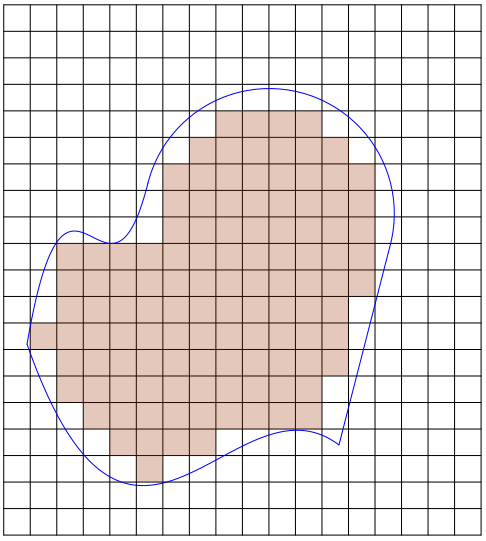

Nun wählen wir eine Zerlegung $\mathfrak{Z}_1$ von $Q_1$, so dass die der Zerlegung $\mathfrak{Z}_1$ entsprechenden Quader entweder in $Q$ oder komplett ausserhalb liegen. Wir schränken diese Zerlegung auf $Q$ ein und erhalten eine Zerlegung $\mathfrak{Z}$ von $Q$. Die Funktion $\mathbb{1}_B f$ auf $Q_1$ (beziehungsweise $Q_2$, $Q$) wie in Definition 12.29 bezeichnen wir mit $(\mathbb{1}_B f)_{Q_1}$ (beziehungsweise $(\mathbb{1}_B f)_{Q_2}$, $(\mathbb{1}_B f)_Q$). Mit dieser Notation gilt

$$U((\mathbb{1}_B f)_{Q_1}, \mathfrak{Z}) = U((\mathbb{1}_B f)_Q, \mathfrak{Z}),$$
$$O((\mathbb{1}_B f)_{Q_1}, \mathfrak{Z}) = O((\mathbb{1}_B f)_Q, \mathfrak{Z}),$$

da die Funktion $(\mathbb{1}_B f)_{Q_1}$ auf allen $\mathfrak{Z}_1$ entsprechenden offenen Teilquadern, die nicht in $Q$ enthalten sind, verschwindet.

Ist also $(\mathbb{1}_B f)_{Q_1}$ auf $Q_1$ Riemann-integrierbar, so ist nach Proposition 12.6 die Funktion $(\mathbb{1}_B f)_Q$ auf $Q$ Riemann-integrierbar auf $Q$ und $\int_{Q_1} \mathbb{1}_B f \,\mathrm{dvol} = \int_Q \mathbb{1}_B f \,\mathrm{dvol}$. Aus demselben Argument folgt, dass $(\mathbb{1}_B f)_{Q_2}$ auf $Q_2$ Riemann-integrierbar ist und dass $\int_{Q_2} \mathbb{1}_B f \,\mathrm{dvol} = \int_Q \mathbb{1}_B f \,\mathrm{dvol}$ gilt. Damit folgt das Lemma. $\square$

Insbesondere zeigt Lemma 12.30, dass Jordan-Messbarkeit einer Teilmenge $B \subseteq \mathbb{R}^n$ und das Volumen $\mathrm{vol}(B) = \int_Q \mathbb{1}_B \,\mathrm{dvol}$ in Definition 12.26 nicht von der Wahl eines Quaders $Q \supseteq B$ abhängt. Geometrisch sollte man sich diese Definition wie folgt vorstellen. Die Teilmenge $B \subseteq \mathbb{R}^n$ ist genau dann Jordan-messbar, wenn das Volumen von $B$ durch ein genügend feines Raster beliebig genau bestimmt werden kann, siehe folgendes Bild.



Figur 12.3: Wir unterlegen $B$ mit einem genügend feinen Raster und zählen alle Teilquader ab, die komplett in $B$ enthalten sind. Multipliziert man das Resultat mit dem Volumen eines Teilquaders, so erhält man die Untersumme zur Definition von $\mathrm{vol}(B)$. Des Weiteren zählen wir alle Teilquader ab, die notwendig sind (abgesehen vom Raster), um $B$ zu überdecken. Das Produkt davon mit dem Volumen eines Teilquaders ist die Obersumme zu $\mathrm{vol}(B)$. Das Volumen (das Jordan-Mass) von $B$ ist definiert, wenn diese untere und obere Approximationen einander beliebig nahe gebracht werden können (nach Proposition 12.6).

**Übung 12.31** (Translations- und Streckungsverhalten des Volumens)**.** *Sei $a \in \mathbb{R}^n$. Zeigen Sie, dass für jede Jordan-messbare Teilmenge $B \subseteq \mathbb{R}^n$ auch die Teilmenge $a + B$ Jordan-messbar ist mit Volumen*

$$\mathrm{vol}(a + B) = \mathrm{vol}(B).$$

*Verifizieren Sie auch, dass für $\alpha \in \mathbb{R}$ und eine Jordan-messbare Teilmenge $B \subseteq \mathbb{R}^n$ die Teilmenge $\alpha B$ ebenfalls Jordan-messbar ist mit*

$$\mathrm{vol}(\alpha B) = |\alpha|^n \,\mathrm{vol}(B).$$

Das Lebesgue-Kriterium (Satz 12.23) lässt sich relativ schnell zu einem Kriterium für Riemann-Integrierbarkeit auf beliebigen Jordan-messbaren Mengen verallgemeinern.

**Korollar 12.32** (Lebesgue-Kriterium)**.** *Sei $B \subseteq \mathbb{R}^n$ Jordan-messbar und sei $f : B \to \mathbb{R}$ beschränkt. Dann ist $f$ genau dann Riemann-integrierbar, wenn $f$ auf $B$ fast überall stetig ist, das heisst, wenn die Menge der Unstetigkeitsstellen von $f$ eine Nullmenge ist. Insbesondere ist jede beschränkte stetige Funktion auf einer Jordan-messbaren Menge Riemann-integrierbar.*

*Beweis.* Sei $Q$ ein abgeschlossener Quader mit $B \subseteq Q$. Da nach Annahme $B$ Jordan-messbar ist, ist $\mathbb{1}_B$ auf $Q$ Riemann-integrierbar (nach Lemma 12.30) und die Menge der Unstetigkeitsstellen von $\mathbb{1}_B$ (welche nach Gleichung (12.7) gerade durch $\partial B$ gegeben ist) ist eine Nullmenge nach dem Lebesgue-Kriterium in Satz 12.23. Wie schon im Beweis von Korollar 12.27 ändern etwaige Probleme auf der Nullmenge $\partial Q$ bei Einschränkung von (12.7) auf $Q$ diese Aussage

nicht. Es gilt

$$\begin{aligned}\{x \in B \mid f \text{ ist unstetig in } x\} \\ &\subseteq \{x \in Q \mid \mathbb{1}_B f \text{ ist unstetig in } x\} \\ &\subseteq \{x \in Q \mid \mathbb{1}_B \text{ ist unstetig in } x\} \cup \{x \in B \mid f \text{ ist unstetig in } x\}\,.\end{aligned}$$

Das Korollar folgt somit aus dem Lebesgue-Kriterium und aus Lemma 12.17. □

**Korollar 12.33.** *Sei $B \subseteq \mathbb{R}^n$ eine Jordan-messbare Teilmenge. Die Linearität des Riemann-Integrals (Proposition 12.7), die Monotonie und die Dreiecksungleichung (Proposition 12.8) gelten analog für das Riemann-Integral von Riemann-integrierbaren Funktionen auf $B$.*

*Beweis.* Alle Aussagen folgen direkt durch Anwendung der Propositionen in Abschnitt 12.1.2 auf $\mathbb{1}_B f$ wie in Definition 12.29. □

**Proposition 12.34** (Additivität des Riemann-Integrals)**.** *Seien $B_1, B_2$ zwei Jordan-messbare Teilmengen des $\mathbb{R}^n$ und sei $f : B_1 \cup B_2 \to \mathbb{R}$ eine Riemann-integrierbare Funktion. Dann sind $f|_{B_1} \in \mathcal{R}(B_1)$, $f|_{B_2} \in \mathcal{R}(B_2)$ und $f|_{B_1 \cap B_2} \in \mathcal{R}(B_1 \cap B_2)$ ebenso Riemann-integrierbar und es gilt*

$$\int_{B_1 \cup B_2} f \,\mathrm{dvol} = \int_{B_1} f \,\mathrm{dvol} + \int_{B_2} f \,\mathrm{dvol} - \int_{B_1 \cap B_2} f \,\mathrm{dvol}.$$

*Beweis.* Es gilt

$$\mathbb{1}_{B_1 \cup B_2} = \mathbb{1}_{B_1} + \mathbb{1}_{B_2} - \mathbb{1}_{B_1 \cap B_2}.$$

Die Proposition folgt damit nach Multiplikation mit $f$ aus dem Lebesgue-Kriterium (um die Integrierbarkeit der einzelnen Funktionen zu überprüfen) und der Linearität des Riemann-Integrals in Proposition 12.7. □

Manche (aber eben nicht alle) der obigen Aussage hätten wir auch mit folgendem einfacheren (aber eben weniger nützlichen) Nullmengenbegriff treffen können.

**Definition 12.35.** Eine Teilmenge $\mathcal{J} \subseteq \mathbb{R}^n$ heisst eine **Jordan-Nullmenge**, falls es zu jedem $\varepsilon > 0$ eine endliche Liste $Q_1, \ldots, Q_L \subseteq \mathbb{R}^n$ offener Quader gibt, so dass

$$\mathcal{J} \subseteq \bigcup_{\ell=1}^{L} Q_\ell, \quad \sum_{\ell=1}^{L} \mathrm{vol}(Q_\ell) < \varepsilon$$

gilt.

**Übung 12.36.** *Sei $\mathcal{J} \subseteq \mathbb{R}^n$ eine Teilmenge. Zeigen Sie die Äquivalenz folgender Aussagen.*

*(i) $\mathcal{J}$ ist eine Jordan-Nullmenge.*

*(ii) $\mathcal{J}$ ist Jordan-messbar und* $\mathrm{vol}(\mathcal{J}) = 0$.

*(iii)* $\overline{\mathcal{J}} = J \cup \partial J$ *ist eine beschränkte Lebesgue-Nullmenge.*

Beispielsweise ist also $[0,1] \cap \mathbb{Q} \subseteq \mathbb{R}$ eine Lebesgue-Nullmenge (da abzählbar), aber keine Jordan-Nullmenge (da $\overline{[0,1] \cap \mathbb{Q}} = [0,1]$ keine Lebesgue-Nullmenge ist). Allerdings ist jede Jordan-Nullmenge $\mathcal{J} \subseteq \mathbb{R}^n$ eine Lebesgue-Nullmenge. Andererseits ist dafür jede Jordan-Nullmenge Jordan-messbar, was für Lebesgue-Nullmengen nicht stimmt (nach Beispiel 4.17 ist $[0,1] \cap \mathbb{Q} \subseteq \mathbb{R}$ nicht Jordan-messbar).

**Übung 12.37** (Integrale über Jordan-Nullmengen)**.** *Sei $\mathcal{J}$ eine Jordan-Nullmenge und $f : \mathcal{J} \to \mathbb{R}$ eine beschränkte Funktion. Zeigen Sie, dass $f$ auf $\mathcal{J}$ Riemann-integrierbar ist und $\int_{\mathcal{J}} f \,\mathrm{dvol} = 0$ erfüllt.*

**Übung 12.38.** *Sei $Q$ ein abgeschlossener Quader und sei $f : Q \to \mathbb{R}$ eine Riemann-integrierbare Funktion mit $f \geq 0$ und $\int_Q f \,\mathrm{dvol} = 0$. Zeigen Sie (in Analogie zu Übung 9.12), dass dann $f = 0$ fast überall gilt.*

*Formulieren und beweisen Sie die analoge Aussage für Funktionen auf Jordan-messbaren Teilmengen.*

*Bemerkung.* Sei $Q$ ein abgeschlossener Quader. Betrachten wir den Vektorraum $\mathcal{R}(Q)$ der Riemann-integrierbaren Funktionen auf $Q$, so definiert

$$\|\cdot\|_1 : f \in \mathcal{R}(Q) \mapsto \int_Q |f| \,\mathrm{dvol}$$

„fast“ eine Norm auf $\mathcal{R}(Q)$. In der Tat erfüllt $\|\cdot\|_1$ alle Eigenschaften aus Definition 5.84 ausser die Definitheit. Stattdessen gilt für $f \in \mathcal{R}(Q)$ mit $\|f\|_1 = 0$ nach Übung 12.38, dass $f$ fast überall gleich Null ist. Die Menge

$$\mathcal{N}(Q) = \{f \in \mathcal{R}(Q) \mid \|f\|_1 = 0\}$$

bildet einen Teilraum von $\mathcal{R}(Q)$ und die Abbildung $\|\cdot\|_1$ kann als eine Norm auf dem Quotientenraum $\mathcal{R}(Q)/\mathcal{N}(Q)$ interpretiert werden. Leider ist der so erhaltene normierte Vektorraum nicht vollständig und damit für die weiterführende Analysis nicht so nützlich. Dieses Manko des Riemann-Integrals wird im Zuge der Vorlesung „Mass und Integral“ im zweiten Studienjahr des Mathematikstudiums mit dem Lebesgue-Integral behoben.

## 12.5 Der Satz von Fubini

Wir haben in den letzten drei Abschnitten bereits die meisten theoretischen Grundlagen der mehrdimensionalen Integralrechnung kennengelernt. Bisher haben wir jedoch noch keine allgemeine Methode zur expliziten Berechnung von Integralen über mehrdimensionale Teilmengen des Euklidschen Raumes gesehen. Wir holen dies in diesem Abschnitt mit der Diskussion des Satzes von Fubini nach.

Wir betrachten zuerst den einfachsten Fall eines Quaders $Q = [a, b] \times [c, d] \subseteq \mathbb{R}^2$ und einer stetigen Funktion $f : Q \to \mathbb{R}$. Hier ist die Idee, dass man (zum Beispiel) zuerst die Funktion $f(x, y)$ bezüglich der Variable $y$ integriert, wodurch wir das stetige Parameterintegral

$$x \in [a, b] \mapsto \int_c^d f(x, y)\,\mathrm{d}y$$

erhalten, siehe Satz 10.36. Integrieren wir dieses nun nach $x$ über $[a, b]$, so erhalten wir eine reelle Zahl

$$\int_a^b \left[\int_c^d f(x, y)\,\mathrm{d}y\right]\,\mathrm{d}x.$$

Wir werden nun zeigen, dass wir zum selben Ergebnis gelangen, wenn wir zuerst nach $x$ über $[a, b]$ und anschliessend über $y$ in $[c, d]$ integriert hätten. Genauer gilt sogar

$$\int_Q f\,\mathrm{dvol} = \int_c^d \left[\int_a^b f(x, y)\,\mathrm{d}x\right]\,\mathrm{d}y = \int_a^b \left[\int_c^d f(x, y)\,\mathrm{d}y\right]\,\mathrm{d}x.$$

Da wir bereits viele Methoden zur Berechnung zur Integration von Funktionen in einer Variablen kennen, ist diese Formel sehr nützlich. Aus diesem Grund wollen wir des öfteren auch $\mathrm{dvol}(x, y) = \mathrm{d}x\,\mathrm{d}y = \mathrm{d}y\,\mathrm{d}x$ und analog $\mathrm{dvol}(x_1, \ldots, x_n) = \mathrm{d}x_1 \ldots \mathrm{d}x_n$ für Integration im $\mathbb{R}^n$ schreiben. Für nicht notwendigerweise stetige Funktionen gibt es in der genauen Formulierung allerdings kleine Probleme:

**Theorem 12.39** (Fubini)**.** *Seien $X \subseteq \mathbb{R}^n$ und $Y \subseteq \mathbb{R}^m$ zwei abgeschlossene Quader und sei $f : X \times Y \to \mathbb{R}$ eine Riemann-integrierbare Funktion. Dann existiert das* ***Parameterintegral***

$$x \in X \mapsto \int_Y f(x, y)\,\mathrm{dvol}(y)$$

*für fast alle $x \in X$ und es gilt*

$$\int_{X \times Y} f(x, y)\,\mathrm{dvol}((x, y)) = \int_X \left[\int_Y f(x, y)\,\mathrm{dvol}(y)\right]\,\mathrm{dvol}(x).$$

*Genauer formuliert, definieren wir die Funktionen $f_x : y \in Y \mapsto f(x, y)$ sowie*

$$F : x \in X \mapsto \underline{I}(f_x) = \begin{cases} \int_Y f_x(y)\,\mathrm{dvol}(y) & \textit{falls } f_x \textit{ Riemann-integrierbar ist} \\ \underline{I}(f_x) & \textit{sonst.} \end{cases}$$

*Dann ist $F$ auf $X$ Riemann-integrierbar und es gilt*

$$\int_{X\times Y} f(x,y)\,\mathrm{dvol}((x,y)) = \int_X F(x)\,\mathrm{dvol}(x).$$

*Selbiges gilt, wenn man in der Definition von $F$ das untere Integral durch das obere Integral ersetzt.*

Die Idee für den Beweis ist relativ einfach: der Satz stimmt für charakteristische Funktionen von abgeschlossenen Quadern (wieso?) und Riemann-integrierbare Funktionen sind mittels Linearkombinationen solcher Funktionen „approximierbar“.

*Beweis.* Jede Zerlegung von $X\times Y$ hat die Form $\mathfrak{Z} = (\mathfrak{Z}_X, \mathfrak{Z}_Y)$ für Zerlegungen $\mathfrak{Z}_X$ von $X$ und $\mathfrak{Z}_Y$ von $Y$ und die $\mathfrak{Z}$ entsprechenden offenen Quader $Q_{\boldsymbol{\alpha}}$ haben genau die Form $Q_{\boldsymbol{\alpha}} = O_{\boldsymbol{\beta}} \times P_{\boldsymbol{\gamma}}$ für offene Quader $O_{\boldsymbol{\beta}} \sqsubset \mathfrak{Z}_X$ und $P_{\boldsymbol{\gamma}} \sqsubset \mathfrak{Z}_Y$. Wir werden diese Korrespondenz im Beweis mehrmals implizit verwenden. Nach Definition des Volumens von Quadern gilt des Weiteren

$$\mathrm{vol}_{m+n}(Q_{\boldsymbol{\alpha}}) = \mathrm{vol}_n(O_{\boldsymbol{\beta}}) \cdot \mathrm{vol}_m(P_{\boldsymbol{\gamma}}),$$

wobei wir (in diesem Beweis) $\mathrm{vol}_n$ (beziehungsweise $\mathrm{vol}_m$, $\mathrm{vol}_{m+n}$) für das Volumen von Quadern im $\mathbb{R}^n$ (beziehungsweise im $\mathbb{R}^m$, $\mathbb{R}^{n+m}$) schreiben.

Mit diesen Vorbereitungen betrachten wir nun eine Funktion $F$ mit $\underline{I}(f_x) \le F(x) \le \overline{I}(f_x)$ für alle $x \in X$. Es folgt mit einigen weiteren Überlegungen

$$\begin{aligned}
U(f,\mathfrak{Z}) &= \sum_{Q_{\boldsymbol{\alpha}} = O_{\boldsymbol{\beta}}\times P_{\boldsymbol{\gamma}} \sqsubset \mathfrak{Z}} \inf(f(Q_{\boldsymbol{\alpha}}))\,\mathrm{vol}_n(O_{\boldsymbol{\beta}}) \cdot \mathrm{vol}_m(P_{\boldsymbol{\gamma}}) \\
&\le \sum_{O_{\boldsymbol{\beta}} \sqsubset \mathfrak{Z}_X} \inf\left\{U(f_x, \mathfrak{Z}_Y) \mid x \in O_{\boldsymbol{\beta}}\right\} \mathrm{vol}_n(O_{\boldsymbol{\beta}}) && (12.9a)\\
&\le \sum_{O_{\boldsymbol{\beta}} \sqsubset \mathfrak{Z}_X} \inf\left\{\underline{I}(f_x) \mid x \in O_{\boldsymbol{\beta}}\right\} \mathrm{vol}_n(O_{\boldsymbol{\beta}}) && (12.9b)\\
&\le \sum_{O_{\boldsymbol{\beta}} \sqsubset \mathfrak{Z}_X} \inf(F(O_{\boldsymbol{\beta}}))\,\mathrm{vol}_n(O_{\boldsymbol{\beta}}) && (12.9c)\\
&= U(F, \mathfrak{Z}_X)
\end{aligned}$$

für eine Zerlegung $\mathfrak{Z} = (\mathfrak{Z}_X, \mathfrak{Z}_Y)$ von $X \times Y$.

VO: 23.04.18

Um die Ungleichung (12.9a) zu beweisen, bemerken wir zuerst, dass

$$\begin{aligned}
\inf f(O_{\boldsymbol{\beta}} \times P_{\boldsymbol{\gamma}}) &= \inf\left\{f(x,y) \mid x \in O_{\boldsymbol{\beta}} \text{ und } y \in P_{\boldsymbol{\gamma}}\right\} \\
&\le \inf\left\{f(x_0, y) \mid y \in P_{\boldsymbol{\gamma}}\right\}
\end{aligned}$$

für alle $x_0 \in O_{\boldsymbol{\beta}} \sqsubset \mathfrak{Z}_X$ und $P_{\boldsymbol{\gamma}} \sqsubset \mathfrak{Z}_Y$ gilt. Wir multiplizieren dies mit $\operatorname{vol}_m(P_{\boldsymbol{\gamma}})$ und summieren über $P_{\boldsymbol{\gamma}} \sqsubset \mathfrak{Z}_Y$, was

$$\begin{aligned}\sum_{P_{\boldsymbol{\gamma}} \sqsubset \mathfrak{Z}_Y} \inf f(O_{\boldsymbol{\beta}} \times P_{\boldsymbol{\gamma}}) \operatorname{vol}_m(P_{\boldsymbol{\gamma}}) &\leq \sum_{P_{\boldsymbol{\gamma}} \sqsubset \mathfrak{Z}_Y} \inf \{f(x_0, y) \mid y \in P_{\boldsymbol{\gamma}}\} \operatorname{vol}_m(P_{\boldsymbol{\gamma}}) \\ &= U(f_{x_0}, \mathfrak{Z}_Y)\end{aligned}$$

für alle $x_0 \in O_{\boldsymbol{\beta}} \sqsubset \mathfrak{Z}_X$ impliziert. Da $x_0 \in O_{\boldsymbol{\beta}}$ aber beliebig ist, erhalten wir daher

$$\sum_{P_{\boldsymbol{\gamma}} \sqsubset \mathfrak{Z}_Y} \inf f(O_{\boldsymbol{\beta}} \times P_{\boldsymbol{\gamma}}) \operatorname{vol}_m(P_{\boldsymbol{\gamma}}) \leq \inf \left\{U(f_{x_0}, \mathfrak{Z}_Y) \mid x_0 \in O_{\boldsymbol{\beta}}\right\}$$

und damit (12.9a) nach Multiplikation mit $\operatorname{vol}_n(O_{\boldsymbol{\beta}})$ und Summation über $O_{\boldsymbol{\beta}} \sqsubset \mathfrak{Z}_X$.

Die Ungleichung $U(f_x, \mathfrak{Z}_Y) \leq \underline{I}(f_x)$ für $x \in X$ folgt direkt aus der Definition des unteren Integrals $\underline{I}(f_x)$ als Supremum aller Untersummen und impliziert (12.9b).

Zuletzt gilt die Ungleichung (12.9c), da $\underline{I}(f_x) \leq F(x)$ vorausgesetzt wurde.

Zusammenfassend haben wir also gezeigt, dass für eine beliebige Zerlegung $\mathfrak{Z} = (\mathfrak{Z}_X, \mathfrak{Z}_Y)$ von $X \times Y$ die Ungleichung

$$U(f, \mathfrak{Z}) \leq U(F, \mathfrak{Z}_X)$$

von Untersummen gilt. Analog zeigt man $O(F, \mathfrak{Z}_X) \leq O(f, \mathfrak{Z})$ und es gilt

$$U(f, \mathfrak{Z}) \leq U(F, \mathfrak{Z}_X) \leq O(F, \mathfrak{Z}_X) \leq O(f, \mathfrak{Z}).$$

Variieren wir die Zerlegung $\mathfrak{Z} = (\mathfrak{Z}_X, \mathfrak{Z}_Y)$, so ergibt sich

$$\underline{I}(f) \leq \underline{I}(F) \leq \overline{I}(F) \leq \overline{I}(f).$$

Da aber $f$ per Annahme Riemann-integrierbar ist, gilt $\underline{I}(f) = \overline{I}(f)$. Also ist $F$ ebenfalls Riemann-integrierbar und es gilt

$$\int_{X \times Y} f(x, y) \operatorname{dvol}_{n+m}((x, y)) = \int_X F(x) \operatorname{dvol}_n(x).$$

Wir zeigen nun, dass die Funktion $f_x : y \in Y \mapsto f(x, y)$ für fast alle $x \in X$ Riemann-integrierbar ist. Dabei verwenden wir obiges sowohl für die Funktion $F$ definiert durch $x \in X \mapsto \underline{I}(f_x)$ als auch für die Funktion $x \in X \mapsto \overline{I}(f_x)$. Obiges zeigt, dass beide Funktionen auf $X$ Riemann-integrierbar und Integral $\int_{X \times Y} f \operatorname{dvol}_{n+m}$ besitzen. Wir schliessen daraus, dass die Funktion

$$h : x \in X \mapsto \overline{I}(f_x) - \underline{I}(f_x) \in [0, \infty)$$

Riemann-integrierbar ist und

$$\int_X h(x)\,\mathrm{dvol}_n(x) = 0$$

genügt. Nach dem Lebesgue-Kriterium (Satz 12.23) ist $h$ in fast allen Punkten $x \in X$ stetig.

Wir behaupten nun, dass $h(x_0) = 0$ für $x_0 \in X$ gelten muss, wenn $h$ in $x_0$ stetig ist (man vergleiche dies zu Übung 12.38). Daraus folgt, dass $\overline{I}(f_x) = \underline{I}(f_x)$ für fast alle $x \in X$ gilt, woraus schliesslich der Satz folgt. Angenommen $x_0 \in X$ ist ein Stetigkeitspunkt von $h$, der $h(x_0) > 0$ erfüllt. Dann gibt es nach Stetigkeit ein $\delta > 0$ mit $h(x) > \frac{h(x_0)}{2} = \varepsilon$ für alle $x \in B_\delta(x_0)$. Nun kann man aber einen offenen Quader $Q_0 \subseteq B_\delta(x_0)$ mit $x_0 \in Q_0$ finden, woraus $h \geq \varepsilon \mathbb{1}_{Q_0}$ und somit $\int_X h(x)\,\mathrm{dvol}_n(x) \geq \varepsilon\,\mathrm{vol}_n(Q_0) > 0$ folgt. Dies widerspricht aber $\int_X h(x)\,\mathrm{dvol}_n(x) = 0$. □

Wir betrachten nochmals ein Applet aus Abschnitt 4.9, welches den Satz von Fubini veranschaulichen sollte.

**Applet 12.40** (Volumen des Zeltes)**.** *Wir können den Satz von Fubini und das Vorgehen der Berechnung des Volumens auch geometrisch veranschaulichen. Dabei bestimmt die $x$-Koordinate einen ebenen Querschnitt durch das Zelt, und die $y$-Koordinate animiert die Berechnung des Flächeninhaltes des Querschnittes. Versuchen Sie mit den Schiebern die iterierte Integration nachzustellen.*

**Korollar 12.41.** *Sei $Q = [a_1, b_1] \times \ldots \times [a_n, b_n]$ ein $n$-dimensionaler abgeschlossener Quader und sei $f : Q \to \mathbb{R}$ Riemann-integrierbar. Dann gilt*

$$\int_Q f\,\mathrm{dvol} = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x_1, \ldots, x_n)\,\mathrm{d}x_n \ldots \mathrm{d}x_1,$$

*wobei dieselben formalen Komplikationen wie in Theorem 12.39 auftreten können.*

Für stetige Funktionen treten wohlgemerkt (da nach Satz 10.36 Parameterintegrale von stetigen Funktionen für alle Parameter existieren) keinerlei Komplikationen auf.

*Beweis.* Dies folgt durch mehrmaliges Anwenden des Satzes von Fubini (Theorem 12.39). □

**Übung 12.42.** *Präzisieren Sie die in Korollar 12.41 auftretenden Komplikationen und erklären Sie den Beweis genauer.*

**Korollar 12.43** (Fubini für Bereiche zwischen Graphen)**.** *Sei $D \subseteq \mathbb{R}^{n-1}$ eine Jordan-messbare Menge, seien $\varphi_-, \varphi_+ : D \to \mathbb{R}$ beschränkt und stetig mit $\varphi_- \leq \varphi_+$ und sei $B \subseteq \mathbb{R}^n$ die Jordan-messbare Teilmenge*

$$B = \{(x, y) \in D \times \mathbb{R} \mid \varphi_-(x) \leq y \leq \varphi_+(x)\}\,.$$

*Für eine Riemann-integrierbare Funktion $f$ auf $B$ gilt*

$$\int_B f(x, y)\,\mathrm{dvol}((x, y)) = \int_D \int_{\varphi_-(x)}^{\varphi_+(x)} f(x, y)\,\mathrm{d}y\,\mathrm{dvol}(x),$$

*wobei wieder die selben Komplikationen wie in Theorem 12.39 auftreten können.*

*Beweis.* Wir erinnern zuerst daran, dass $B$ nach Korollar 12.27 tatsächlich Jordan-messbar ist. Wenden wir nun den Satz von Fubini für einen Quader $X \subseteq \mathbb{R}^{n-1}$ mit $D \subseteq X$ und den Quader $Y = [\inf \varphi_-(D), \sup \varphi_+(D)]$ an, so ergibt sich

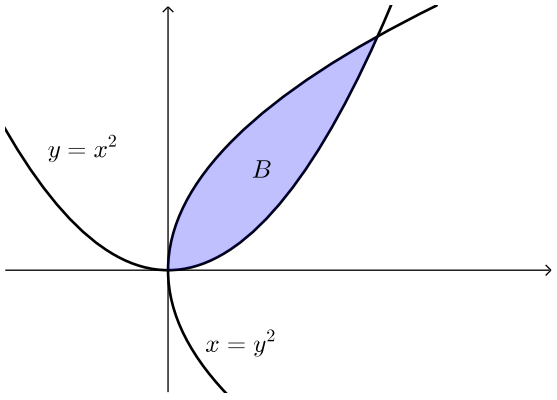
$$\begin{aligned}
\int_B f(x,y)\,\mathrm{dvol}((x,y)) &= \int_{X\times Y} (\mathbb{1}_B f)(x,y)\,\mathrm{dvol}((x,y)) \\
&= \int_X \int_{\inf \varphi_-(D)}^{\sup \varphi_+(D)} \mathbb{1}_B(x,y) f(x,y)\,\mathrm{d}y\,\mathrm{dvol}(x) \\
&= \int_X \mathbb{1}_D(x) \int_{\varphi_-(x)}^{\varphi_+(x)} f(x,y)\,\mathrm{d}y\,\mathrm{dvol}(x) \\
&= \int_D \int_{\varphi_-(x)}^{\varphi_+(x)} f(x,y)\,\mathrm{d}y\,\mathrm{dvol}(x)
\end{aligned}$$

wie behauptet. □

**Beispiel 12.44** (Schwerpunkt einer Teilmenge)**.** *Wir betrachten die Teilmenge*

$$B = \left\{(x,y) \in \mathbb{R}^2 \mid x \geq 0,\ x^2 \leq y \leq \sqrt{x}\right\}$$

*zwischen den Parabeln $y = x^2$ und $x = y^2$.*




*Wir möchten nun den Schwerpunkt dieser Menge berechnen. Dabei nehmen wir an, dass die Masse gleichmässig auf $B$ verteilt ist. Also ist diese bis auf einen Faktor gleich dem Flächeninhalt. Aus Abschnitt 1.1 wissen wir bereits, dass der Bereich definiert durch $0 \leq y \leq x^2$ und $0 \leq x \leq 1$ Flächeninhalt $\frac{1}{3}$ hat. Aus Symmetriegründen folgt daraus, dass der Flächeninhalt von $B$ ebenfalls $\frac{1}{3}$ ist.*

*Die $x$-Koordinate $x_S$ des Schwerpunkts $S$ ist per Definition durch $x_S = \frac{1}{\mathrm{vol}(B)} \int_B x \,\mathrm{dvol}$ gegeben und berechnet sich mit Korollar 12.43 also zu*

$$\begin{aligned} x_S &= \tfrac{1}{\mathrm{vol}(B)} \int_B x \,\mathrm{dvol} = 3 \int_0^1 \int_{x^2}^{\sqrt{x}} x \,\mathrm{d}y \,\mathrm{d}x \\ &= 3 \int_0^1 x(\sqrt{x} - x^2) \,\mathrm{d}x = 3 \int_0^1 (x^{\frac{3}{2}} - x^3) \,\mathrm{d}x \\ &= 3(\tfrac{2}{5} x^{\frac{5}{2}} - \tfrac{1}{4} x^4)\Big]_0^1 = \frac{9}{20}. \end{aligned}$$

*Genauso ist $y_S = \frac{1}{\mathrm{vol}(B)} \int_B y \,\mathrm{dvol}$ durch*

$$\begin{aligned} y_S &= 3 \int_0^1 \int_{x^2}^{\sqrt{x}} y \,\mathrm{d}y \,\mathrm{d}x = 3 \int_0^1 \left[\tfrac{1}{2} y^2\right]_{y=x^2}^{y=\sqrt{x}} \,\mathrm{d}x \\ &= 3 \int_0^1 (\tfrac{1}{2} x - \tfrac{1}{2} x^4) \,\mathrm{d}x = 3(\tfrac{1}{4} x^2 - \tfrac{1}{10} x^5)\Big]_0^1 \\ &= 3(\tfrac{1}{4} - \tfrac{1}{10}) = \frac{9}{20} \end{aligned}$$

*gegeben, was man auf Grund der Symmetrie von $B$ um die Gerade $y = x$ hätte erwarten können.*

**Übung 12.45** (Fubini)**.** *Berechnen Sie*

$$\begin{aligned} &\int_{B_1} xyz \,\mathrm{dvol} \text{ für } B_1 = \left\{(x,y,z) \in \mathbb{R}^3 \mid 0 \le x \le y \le z \le 1\right\}, \\ &\int_{B_2} xyz \,\mathrm{dvol} \text{ für } B_2 = \left\{(x,y,z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 \le 1\right\} \end{aligned}$$

*und das Volumen $\mathrm{vol}(B_3)$ für die Menge $B_3$, die zwischen den Kurven $x^2 + y^2 = 8$ und $4y = x^2 + 4$ eingeschlossen wird.*

**Korollar 12.46** (Prinzip von Cavalieri)**.** *Falls $B \subseteq [a,b] \times \mathbb{R}^{n-1}$ beschränkt und Jordan-messbar ist, dann ist*

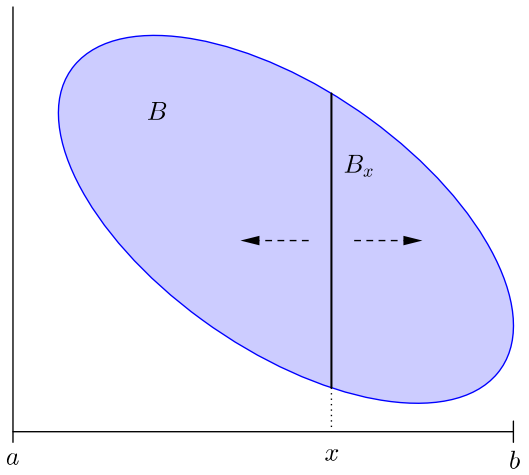
$$\mathrm{vol}(B) = \int_a^b \mathrm{vol}(B_x) \,\mathrm{d}x,$$

*wobei für $x \in [a,b]$ die Teilmenge $B_x \subseteq \mathbb{R}^{n-1}$ durch*

$$B_x = \left\{y \in \mathbb{R}^{n-1} \mid (x,y) \in B\right\}$$

*gegeben ist und für fast alle $x \in [a,b]$ Jordan-messbar ist.*

*Beweis.* Folgt direkt aus Theorem 12.39 mit $f = \mathbb{1}_B$. □




**Übung 12.47** (Volumen von Rotationskörpern)**.** *In Abschnitt 8.3.4 wurde das Volumen von Rotationskörpern definiert. Wir möchten hier verifizieren, dass das im letzten Abschnitt eingeführte Volumen Jordan-messbarer Teilmengen damit kompatibel ist. Sei also* $f : [a, b] \to \mathbb{R}_{\geq 0}$ *stetig und sei*

$$K_f = \left\{(x, y, z)^t \in \mathbb{R}^3 \mid a \leq x \leq b,\ 0 \leq \sqrt{y^2 + z^2} \leq f(x)\right\}$$

*der durch* $f$ *gegebene Rotationskörper. Zeigen Sie, dass* $K_f$ *Jordan-messbar ist und dass das das Volumen von* $K_f$ *(im Sinne von Definition 12.26) durch* $\pi \int_a^b f(x)^2 \,\mathrm{d}x$ *gegeben ist.*

In vielen konkreten Rechnungen zum Beispiel in der Ebene hat man die Wahl ob man zuerst nach der $y$-Variable und anschliessend nach der $x$-Variable oder umgekehrt integriert. Meist funktionieren beide Möglichkeiten, obwohl es vorkommt, dass eine Möglichkeit deutlich weniger Arbeit bereitet. Es empfiehlt sich daher diese Entscheidung bewusst zu tätigen und zu üben (siehe Übung 12.48). Zum Beispiel bei Integration über die Menge

$$B = \left\{(x, y) \in [0, 1] \times [0, 2] \mid 0 \leq y \leq x(2 + \sin(\tfrac{1}{x})) \text{ für alle } x > 0\right\}$$

wäre eine Integration über $x$ bei fest gehaltenem $y$ nur eine theoretische Möglichkeit aber praktisch unmöglich. (Wieso?) Derartige Komplikationen treten auch im nächsten Kapitel bei den Beweisen der mehrdimensionalen Integralsätze auf und erfordern mitunter geschicktere Argumentationen als man auf dem ersten Blick vermuten würde.

**Übung 12.48** (Reihenfolge)**.** *Berechnen Sie* $\int_0^1 \int_x^1 \mathrm{e}^{y^2} \,\mathrm{d}y \,\mathrm{d}x$.

*Bemerkung.* Theorem 12.39 und Korollar 12.46 sind unter anderem, wie vielleicht nicht auf Anhieb klar ist, sogar für sehr elementare Volumsberechnungen nötig. Beispielsweise könnte man sich fragen, ob sich das Volumen einer Pyramide im $\mathbb{R}^3$ elementar geometrisch berechnen lässt, in dem man die Pyramide durch endlich viele Operationen von Schneiden und Kleben zu Quadern zerlegt. In der Tat, im $\mathbb{R}^2$ funktioniert diese Methode ausgezeichnet für Polygone. Bereits für Pyramiden im $\mathbb{R}^3$ kann man jedoch beweisen (mit Hilfe der sogenannten Dehn-Invariante), dass diese Methode tatsächlich im Allgemeinen nicht funktioniert - man muss

also bereits für die Volumenberechnung von Polyedern im $\mathbb{R}^3$ auf Integrationsmethoden der Analysis zurückgreifen.

## 12.6 Mehrdimensionale Substitutionsregel

Wir wollen in diesem Abschnitt die Substitutionsregel für mehrdimensionale Integrale beweisen und werden hierbei wieder den Begriff „Diffeomorphismus“ aus Definition 11.6 benötigen. Diffeomorphismen kann man sich als glatte Koordinatenwechsel vorstellen. In diesem Sinne wird die mehrdimensionale Substitutionsregel erklären, inwiefern ein solcher Koordinatenwechsel das Volumen einer Menge oder allgemeiner das Integral einer Funktion beeinflusst.

### 12.6.1 Lineare Substitution

Wir betrachten zuerst den Fall eines linearen Koordinatenwechsels, da dies einen wichtigen Spezialfall darstellt, der auch im Beweis des allgemeinen Falles von Bedeutung sein wird.

**Proposition 12.49** (Lineare Substitutionsregel)**.** *Sei $Q \subseteq \mathbb{R}^n$ ein abgeschlossener Quader und sei $L$ in $\mathrm{GL}_n(\mathbb{R})$. Dann ist das Bild $L(Q)$ von $Q$ unter $L$ Jordan-messbar und es gilt*

$$\mathrm{vol}(L(Q)) = |\det(L)|\,\mathrm{vol}(Q).$$

Eine gute geometrische Erklärung für den obigen Satz wird durch die Singulärwertzerlegung (SVD-Zerlegung oder Cartan-Zerlegung) gegeben, nach der man die Matrix $L$ wie aus dem Satz als $L = k_1 A k_2$ schreiben kann für $k_1, k_2 \in \mathrm{SO}(n, \mathbb{R})$ und eine Diagonalmatrix $A \in \mathrm{Mat}_{n,n}(\mathbb{R})$.

**Übung 12.50.** *Zeigen Sie die Existenz einer solchen Singulärwertzerlegung.*

Für die Diagonalmatrix $A$ und einen achsenparallelen Quadern $Q$ ist es anschaulich (und ebenso formal) klar, dass $\mathrm{vol}(A(Q)) = |\det(A)|\,\mathrm{vol}(Q)$ gilt. In der Tat ist $A(Q)$ wieder ein achsenparalleler Quader, der in jeder Richtung entsprechend den Diagonaleinträgen von $A$ gestreckt oder gestaucht wurde. Dies lässt sich dann auf jede andere Jordan-messbare Menge verallgemeinern (siehe den ersten Teil des Beweises der Proposition unten). Für die (orthogonalen) Drehungsmatrizen $k_1, k_2$ ist es geometrisch einleuchtend, dass diese das $n$-dimensionale Volumen nicht ändern. Gleichzeitig gilt aber $\det(k_1) = \det(k_2) = 1$ per Definition. Zusammen ergibt sich gerade die Aussage in obiger Proposition. Dies stellt aber auf Grund des Falls der Drehungsmatrizen keinen formalen Beweis dar, weswegen wir ein anderes Argument geben werden.

*Beweis.* Wir beginnen zuerst mit dem bereits erwähnten Spezialfall der Diagonalmatrizen, den wir später auch verwenden werden.

Der Satz für Diagonalmatrizen: Angenommen $L = A \in \mathrm{GL}_n(\mathbb{R})$ ist eine diagonale Matrix. Für einen achsenparallelen Quader $Q \subseteq \mathbb{R}^n$ ist $A(Q)$ wieder ein achsenparalleler Quader mit Volumen

$$\mathrm{vol}(A(Q)) = |\det(A)|\,\mathrm{vol}(Q)$$

und insbesondere Jordan-messbar.

Ist nun $B \subseteq \mathbb{R}^n$ Jordan-messbar und $Q$ ein Quader mit $B \subseteq Q$, dann ist per Definition $\mathbb{1}_B$ auf $Q$ Riemann-integrierbar. Falls $\mathfrak{Z}_Q$ eine Zerlegung von $Q$ ist, dann können wir die

Eigenwerte von $A$ verwenden, um eine Zerlegung $\mathfrak{Z}'$ von $Q' = A(Q)$ zu definieren, so dass die zu $\mathfrak{Z}'$ gehörigen Quader genau von der Form $A(Q_{\boldsymbol{\alpha}})$ für $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$ sind. Des Weiteren ist jede Zerlegung von $Q'$ von dieser Form. Wegen $\mathbb{1}_{A(B)} \circ A = \mathbb{1}_B$ und $\operatorname{vol}(A(Q_{\boldsymbol{\alpha}})) = |\det(A)| \operatorname{vol}(Q_{\boldsymbol{\alpha}})$ für alle $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$ gilt nun

$$\begin{aligned} U(\mathbb{1}_{A(B)}, \mathfrak{Z}') &= \sum_{A(Q_{\boldsymbol{\alpha}}) \sqsubset \mathfrak{Z}'} \inf(\mathbb{1}_{A(B)}(A(Q_{\boldsymbol{\alpha}}))) \operatorname{vol}(A(Q_{\boldsymbol{\alpha}})) \\ &= \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} \inf(\mathbb{1}_B(Q_{\boldsymbol{\alpha}})) |\det(A)| \operatorname{vol}(Q_{\boldsymbol{\alpha}}) \\ &= |\det(A)| U(\mathbb{1}_B, \mathfrak{Z}). \end{aligned}$$

Genau gleich zeigt man die analoge Aussage für obere Summen. Damit ist das untere (beziehungsweise das obere) Integral von $\mathbb{1}_{A(B)}$ auf $Q' = A(Q)$ gleich dem unteren (beziehungsweise dem oberen) Integral von $\mathbb{1}_B$ auf $Q$ multipliziert mit $|\det(A)|$. Also ist $\mathbb{1}_{A(B)}$ Riemann-integrierbar, $A(B)$ ist Jordan-messbar und es gilt

$$\begin{aligned} \operatorname{vol}(A(B)) &= \int_{A(Q)} \mathbb{1}_{A(B)} \operatorname{dvol} \\ &= |\det(A)| \int_Q \mathbb{1}_B \operatorname{dvol} \\ &= |\det(A)| \operatorname{vol}(B). \end{aligned}$$

Zerlegung in Elementarmatrizen: Um nun die Proposition für eine allgemeine Matrix $L \in \mathrm{GL}_n(\mathbb{R})$ zu beweisen, zerlegen wir $L$ in Elementarmatrizen. Das heisst, wir betrachten folgende drei Typen von Matrizen:

- *Diagonalmatrizen.* Für diese gilt die Proposition bereits auf Grund obiger Diskussion.
- *Permutationsmatrizen.* Eine Permutationsmatrix $P = P_\sigma \in \mathrm{GL}_n(\mathbb{R})$ ist durch eine Permutation $\sigma \in \mathrm{Sym}(n)$ definiert und erfüllt $Pe_j = e_{\sigma(j)}$ für alle $j \in \{1, \ldots, n\}$. Der Beweis für Diagonalmatrizen oben lässt sich auf Permutationsmatrizen übertragen, da wiederum achsenparallele Quader auf achsenparallele Quader abgebildet werden.
- *Scherungsmatrizen.* Hier bezeichnen wir eine Matrix als eine Scherungsmatrix falls sie für ein $k \neq \ell \in \{1, \ldots, n\}$ und $s \in \mathbb{R}$ durch

  $$(U^s_{k\ell})_{ij} = \begin{cases} 1 & \text{falls } i = j \\ s & \text{falls } k = i \text{ und } \ell = j \\ 0 & \text{sonst} \end{cases}$$

  für $i, j \in \{1, \ldots, n\}$ gegeben ist.

Wir bemerken, dass wir obige Scherungsmatrizen durch Permutationsmatrizen konjugieren können und damit erreichen können, dass $U^s_{k\ell}$ zu

$$U^s = U^s_{n1} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ s & 0 & \cdots & 1 \end{pmatrix} \in \mathrm{SL}_n(\mathbb{R})$$

konjugiert ist. Aus dieser Bemerkung und dem Gauss-Algorithmus der Linearen Algebra folgt nun, dass sich $L$ als Produkt

$$L = L_1 \dots L_M$$

schreiben lässt, wobei für jedes $m \in \{1, \dots, M\}$ die Matrix $L_m$ eine Diagonalmatrix, eine Scherungsmatrix der Form $U^s$ für $s \in \mathbb{R}$ oder eine Permutationsmatrix ist.

Reduktion auf Scherungsmatrizen: Für Diagonalmatrizen und Permutationsmatrizen wurde die Aussage der Proposition oben bereits allgemeiner gezeigt. Wir werden unten beweisen, dass die Formel

$$\mathrm{vol}(U(B)) = |\det U| \,\mathrm{vol}(B) = \mathrm{vol}(B) \tag{12.10}$$

auch für eine Scherungsmatrix der Form $U^s = U^s_{n1}$ gilt, wenn $B = L'(Q)$ das Bild eines Quaders unter einer linearen Abbildung $L' \in \mathrm{GL}_n(\mathbb{R})$ ist. Für Mengen dieser Form ist es einfach zu sehen, dass diese Jordan messbar sind: denn der Rand $\partial B = L'(\partial Q)$ ist eine endliche Vereinigung von beschränkten Teilmengen von Hyperebenen und damit wieder eine (sogar Jordan) Nullmenge. Damit ist dann die Proposition bereits für $L_1, \dots, L_M$ getrennt gültig. Für einen Quader $Q \subseteq \mathbb{R}^n$ folgt dann mittels Induktion die Formel

$$\begin{aligned} \mathrm{vol}(L(Q)) &= \mathrm{vol}(L_1 \underbrace{L_2 \cdots L_M(Q)}_{B=L'(Q)}) \\ &= |\det(L_1)| \,\mathrm{vol}(L_2 \cdots L_M(Q)) \\ &= |\det(L_1)||\det(L_2 \cdots L_M)| \,\mathrm{vol}(Q) \\ &= |\det(L)| \,\mathrm{vol}(Q). \end{aligned}$$

Der Fall der Scherungsmatrizen: Sei $s \in \mathbb{R}$ und sei $B = L'(Q)$ für einen Quader $Q$. Sei des Weiteren $B \subseteq [a,b] \times \mathbb{R}^{n-1}$ für gewisse $a, b \in \mathbb{R}$. Für $x \in [a,b]$ und $y \in \mathbb{R}^{n-1}$ gilt

$$\begin{aligned} U^s \begin{pmatrix} x \\ y_1 \\ \vdots \\ y_{n-1} \end{pmatrix} &= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ s & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} x \\ y_1 \\ \vdots \\ y_{n-1} \end{pmatrix} \\ &= \begin{pmatrix} x \\ y_1 \\ \vdots \\ y_{n-1} + xs \end{pmatrix}. \end{aligned}$$

Insbesondere ist ebenso $U^s(B) \subseteq [a,b] \times \mathbb{R}^{n-1}$. Nach dem Satz von Fubini in der Form des Prinzips von Cavalieri (Korollar 12.46) und der Translationsinvarianz des Volumens in Übung 12.31 gilt

$$\begin{aligned} \operatorname{vol}(U^s(B)) &= \int_a^b \operatorname{vol}\left(\left\{y \in \mathbb{R}^{n-1} \mid (x, y_1, \ldots, y_{n-1})^t \in U^s(B)\right\}\right) \mathrm{d}x \\ &= \int_a^b \operatorname{vol}\left(\left\{y \in \mathbb{R}^{n-1} \mid (x, y_1, \ldots, y_{n-1} - sx)^t \in B\right\}\right) \mathrm{d}x \\ &= \int_a^b \operatorname{vol}\left(\left\{y \in \mathbb{R}^{n-1} \mid (x, y_1, \ldots, y_{n-1})^t \in B\right\}\right) \mathrm{d}x = \operatorname{vol}(B), \end{aligned}$$

was zu zeigen war. □

VO: 25.04.18

Die Annahme in obiger Proposition, dass $L \in \mathrm{Mat}_{n,n}(\mathbb{R})$ invertierbar ist, ist streng genommen nicht notwendig.

**Übung 12.51** (Proposition 12.49 für nicht-invertierbare Matrizen)**.** *In dieser Übung möchten wir Proposition 12.49 für eine Matrix $L \in \mathrm{Mat}_{n,n}(\mathbb{R})$ mit $\det(L) = 0$ zeigen. Verifizieren Sie also, dass für jede Jordan-messbare Teilmenge $B \subseteq \mathbb{R}^n$ das Bild $L(B)$ eine Jordan-messbare Menge mit Volumen $0$ (also eine Jordan-Nullmenge) ist.*

**Übung 12.52** (Ellipsoide)**.** *Berechnen Sie für eine symmetrische positiv definite Matrix $A \in \mathrm{Mat}_{n,n}(\mathbb{R})$ das Volumen der Menge $\{x \in \mathbb{R}^n \mid \langle Ax, x\rangle \leq 1\}$ (wobei Sie das Volumen $\omega_n = \operatorname{vol}(B_1(0))$ der Einheitskugel als bekannt vorraussetzen dürfen).*

Im Zusammenhang mit linearen Abbildungen wollen wir noch folgenden Begriff aus der linearen Algebra wiederholen.

**Definition 12.53.** Zu Vektoren $v_1, \ldots, v_n \in \mathbb{R}^n$ wird

$$\operatorname{gram}(v_1, \ldots, v_n) = \det(\langle v_i, v_j\rangle)_{ij}$$

die **Gramsche Determinante** von $v_1, \ldots, v_n$ genannt.

**Korollar 12.54** (Volumen von Parallelotopen)**.** *Sei* $L \in \mathrm{Mat}_{n,n}(\mathbb{R})$ *mit Spalten* $v_1, \ldots, v_n \in \mathbb{R}^n$*. Dann ist das* ***Parallelotop***

$$P = L([0,1]^n) = \left\{ \sum_{i=1}^n s_i v_i \mid s_1, \ldots, s_n \in [0,1] \right\}$$

*Jordan-messbar und es gilt*

$$\begin{aligned} \mathrm{vol}(P) &= \mathrm{vol}(L([0,1]^n)) \\ &= |\det(L)| = \sqrt{\mathrm{gram}(v_1, \ldots, v_n)}. \end{aligned}$$

Der Vorteil der Volumenformel in obigem Korollar ist, dass diese nur von den inneren Produkten der Vektoren und nicht von der Basis selbst Gebrauch macht. Für $n = 2$ wird das Parallelotop auch Parallelogram genannt und für $n = 3$ auch Parallelepiped.

*Beweis.* Nach Proposition 12.49 ist $L([0,1]^n)$ Jordan-messbar und hat Volumen $|\det(L)|$ (da $\mathrm{vol}([0,1]^n) = 1$). Für die letzte Gleichheit bemerken wir, dass

$$\begin{aligned} \det(L)^2 &= \det(L^t)\det(L) \\ &= \det(L^t L) \\ &= \det(\langle v_i, v_j \rangle)_{ij} = \mathrm{gram}(v_1, \ldots, v_n) \end{aligned}$$

gilt. Für den Fall $\det(L) = 0$ verweisen wir auf Übung 12.51. □

### 12.6.2 Substitution mit Diffeomorphismen

Wir möchten uns nun von linearen Substitutionen abwenden und allgemeinere Diffeomorphismen zulassen, wobei wir allerdings vorerst folgende zusätzliche Annahme an die Funktion treffen wollen.

**Definition 12.55.** Sei $n \geq 1$ und $X \subseteq \mathbb{R}^n$ offen. Für eine Funktion $f : X \to \mathbb{R}$ ist der **Träger** $\mathrm{supp}(f)$ durch

$$\mathrm{supp}(f) = \overline{\{x \in X \mid f(x) \neq 0\}}$$

definiert. Wir sagen, dass $f$ **kompakten Träger** hat, falls $\mathrm{supp}(f)$ eine kompakte Teilmenge von $X$ ist. Für eine Funktion $f : X \to \mathbb{R}$ mit kompakten Träger $K = \mathrm{supp}(f)$ definieren wir das Riemann-Integral

$$\int_X f \,\mathrm{dvol} = \int_Q \mathbb{1}_X f \,\mathrm{dvol},$$

wobei $Q$ ein Quader mit $K \subseteq Q$ ist und $\mathbb{1}_X f$ für alle $x \in \mathbb{R}^n$ durch

$$\mathbb{1}_X f(x) = \begin{cases} f(x) & \text{für } x \in X, \\ 0 & \text{sonst} \end{cases}$$

definiert ist.

Wir bemerken, dass wir für den Begriff einer Funktion mit kompakten Träger den Abschluss sowohl in $X$ oder auch in $\mathbb{R}^n$ in obiger Definition nehmen können. Denn falls zum Beispiel der Abschluss von $\{x \in X \mid f(x) \neq 0\}$ in $X$ (anstatt in $\mathbb{R}^n$) eine kompakte Teilmenge $K$ von $X$ ist, so ist der Abstand von $K$ zu $\mathbb{R}^n \setminus X$ positiv (da dieser in einem Punkt von $K$ angenommen wird und $X$ offen ist). Hieraus folgt nun, dass $K$ ebenso der Abschluss in $\mathbb{R}^n$ ist. Des Weiteren wollen wir noch bemerken, dass wir ebenso das Integral der Funktion $\mathbb{1}_K f$ über den Quader $Q$ in obiger Definition betrachten könnten, denn $\mathbb{1}_K f = \mathbb{1}_X f$ gilt auf ganz $\mathbb{R}^n$.

**Satz 12.56** (Substitutionsregel mit kompaktem Träger)**.** *Seien $X, Y \subseteq \mathbb{R}^n$ offene Teilmengen und sei $\Phi : X \to Y$ ein $C^1$-Diffeomorphismus. Des Weiteren sei $f : Y \to \mathbb{R}$ eine Riemann-integrierbare Funktion mit kompakten Träger in $Y$. Dann hat die Funktion*

$$x \in X \mapsto (f \circ \Phi(x))|\det(\mathrm{D}_x\Phi)|$$

*kompakten Träger in $X$, ist Riemann-integrierbar und es gilt*

$$\int_Y f(y)\,\mathrm{dvol}(y) = \int_X (f \circ \Phi(x))|\det(\mathrm{D}_x\Phi)|\,\mathrm{dvol}(x),$$

*wobei die Funktion $x \in X \mapsto \det(\mathrm{D}_x\Phi)$ als die **Jacobi-Determinante** von $\Phi$ bezeichnet wird.*

Wir bemerken, dass für lineare Abbildungen Satz 12.56 sehr schnell aus Proposition 12.49 folgen würde, aber der allgemeine Fall für uns viel interessanter ist. Des Weiteren bemerken wir, dass $X$ oder $Y$ nicht Jordan-messbar oder beschränkt sein müssen, was einen der Gründe für unsere Annahme über den kompakten Träger darstellt. Bevor wir uns dem Beweis zuwenden, möchten wir zwei konkrete Beispiele betrachten (siehe auch Abschnitt 4.9).

**Beispiel 12.57** (Substitution für Polarkoordinaten)**.** *Sei $f : B_R(0) \to \mathbb{R}$ eine Riemann-integrierbare Funktion auf der offenen Kreisscheibe $B_R(0) = \big\{x \in \mathbb{R}^2 \mid \|x\| < R\big\}$. Wir können nun (informell - siehe unten) Satz 12.56 wie folgt anwenden:*

$$\begin{aligned}\int_{B_R(0)} f(y)\,\mathrm{dvol}(y) &= \int_U f(y)\,\mathrm{dvol}(y)\\ &= \int_0^R \int_0^{2\pi} f(r\cos\varphi, r\sin\varphi) r\,\mathrm{d}\varphi\,\mathrm{d}r, \end{aligned} \tag{12.11}$$

*wobei $U = B_R(0) \setminus \{(x,y) \mid y = 0,\ x \geq 0\}$. Dabei haben wir mehrere Tatsachen ausgenützt. In der ersten Gleichung haben wir unter Verwendung von Proposition 12.34 auf die Integration über die Jordan-Nullmenge $\{(x,y) \mid y = 0,\ x \geq 0\}$ verzichtet, da die Polarkoordinatenabbildung (siehe Abschnitt 11.1.3)*

$$\Phi : \begin{pmatrix} r \\ \varphi \end{pmatrix} \in (0,R) \times (0,2\pi) \mapsto \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \end{pmatrix} \in U.$$

*damit einen Diffeomorphismus darstellt. Der Faktor $r$ im Integral rechts in Gleichung* (12.11) *ist die Jacobi-Determinante der Polarkoordinaten-Abbildung. In der Tat ist die Jacobi-Matrix*

*bei* $(r,\varphi)^t \in (0,R)\times(0,2\pi)$ *durch*

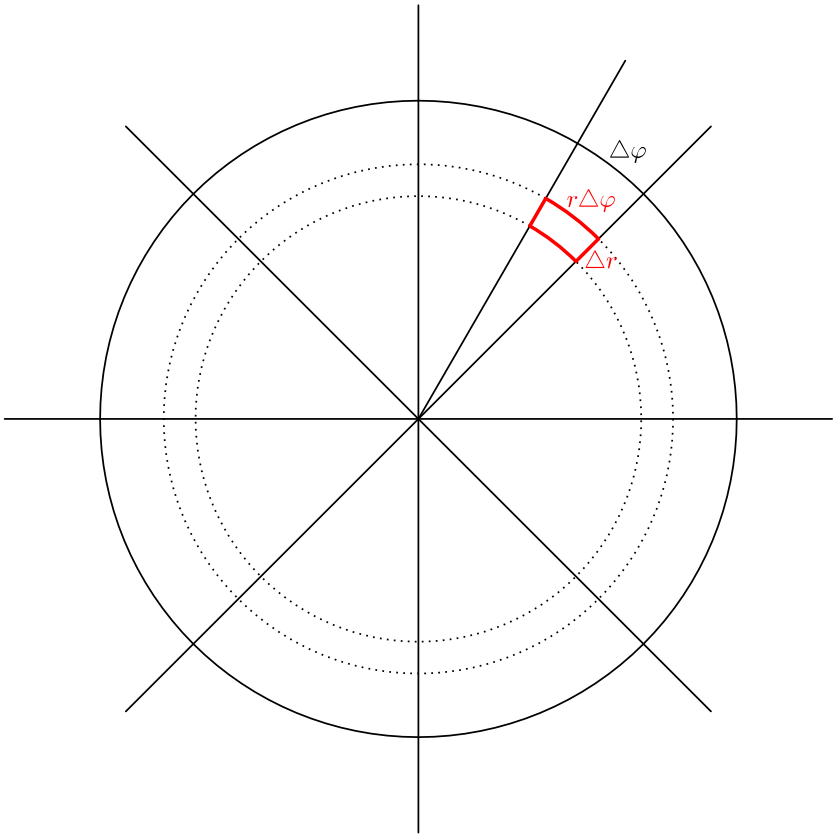
$$\mathrm{D}_{(r,\varphi)^t}\Phi = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}$$

*gegeben, womit die Jacobi-Determinante gleich*

$$\det\left(\mathrm{D}_{(r,\varphi)^t}\Phi\right) = r(\cos^2\varphi + \sin^2\varphi) = r$$

*ist.*

*Leider ist* (12.11) *noch keine Anwendung von Satz 12.56, da wir nicht angenommen haben, dass* $f$ *kompakten Träger hat. Wir werden dieses Manko im nächsten Abschnitt reparieren (siehe auch Übung 12.59). Der Faktor* $r$ *als Jacobi-Determinante hat auch eine klare geometrische Bedeutung, siehe Figur 12.4.*




Figur 12.4: Ein Quader mit Seitenlängen $\triangle r$ und $\triangle\varphi$ entspricht einem fast rechteckigen Ausschnitts eines Kreisrings mit „Seitenlängen“ $\triangle r$ und etwa $r\triangle\varphi$, welche man oft auf Grund der erwünschten Grösse dieser Differenzen auch informell mit $\mathrm{d}r$ und $r\,\mathrm{d}\varphi$ bezeichnet.

**Applet 12.58** (Geometrie der Jacobi-Determinante in Substitution)**.** *In diesen beiden Applets (Polarkoordinaten und ein weiterer Diffeomorphismus) können Sie sich vom Verhalten der Jacobi-Determinante in Beispiel 12.57 und Beispiel 11.9 überzeugen.*

**Übung 12.59.** *Zeigen Sie mit Hilfe von Satz 12.56 die Substitutionsformel* (12.11) *mit Polarkoordinaten für Riemann-integrierbare Funktionen* $f$ *mit kompaktem Träger in* $\mathbb{R}^2$ *(!).*

**Beispiel 12.60** (Substitution für Kugelkoordinaten)**.** *Wir betrachten die offene Kugel*

$$B_R(0) = \{x \in \mathbb{R}^3 \mid \|x\| < R\}$$

*mit Radius $R > 0$ und eine Riemann-integrierbare Funktion $f : B_r(0) \to \mathbb{R}$. Dann gilt unter ähnlichen Argumenten wie in Beispiel 12.57*

$$\int f \,\mathrm{dvol} = \int_0^r \int_0^\pi \int_0^{2\pi} f(r\sin\theta\cos\phi, r\sin\theta\sin\varphi, r\cos\theta) r^2 \sin\theta \,\mathrm{d}\varphi \,\mathrm{d}\theta \,\mathrm{d}r,$$

*siehe auch Abschnitt 11.1.4. Hierbei hat $r^2 \sin\theta \,\mathrm{d}\varphi \,\mathrm{d}\theta \,\mathrm{d}r$ wiederum die geometrische Interpretation einer Menge, die „fast quaderförmig“ ist, siehe wiederum Figur 12.4. Des Weiteren haben wir formal wieder das Problem, dass obige Formel nicht direkt aus Satz 12.56 folgt; wir wollen dieses Problem wiederum im nächsten Abschnitt im Allgemeinen lösen.*

### 12.6.3 Beweis der Substitutionsregel mit kompakten Träger

Bevor wir mit dem detailierten Beweis der Substitutionsregel in Satz 12.56 beginnen, wollen wir die Beweisidee erklären. Seien $X, Y \subseteq \mathbb{R}^n$ offen und sei $\Phi : X \to Y$ ein Diffeomorphismus. Wir wollen annehmen, dass $X = Q$ ein Quader ist. Nun unterteilen wir den offenen Quader $X = Q$ in kleinere Teilquader entsprechend einer „feinen“ Zerlegung $\mathfrak{Z}_X$. Auf Grund der Differenzierbarkeit von $\Phi$ gilt

$$\Phi(x) \approx \Phi(x_{\boldsymbol{\alpha}}) + \mathrm{D}_{x_{\boldsymbol{\alpha}}}\Phi(x - x_{\boldsymbol{\alpha}}),$$

wobei $x, x_{\boldsymbol{\alpha}} \in Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X$ sind und $x_{\boldsymbol{\alpha}}$ der Mittelpunkt von $Q_{\boldsymbol{\alpha}}$ ist. (Das Zeichen $\approx$ bedeutet hier vorerst einfach sehr nahe.) Dies zusammen mit der linearen Substitution in Proposition 12.49 legt den Schluss

$$\begin{aligned} \mathrm{vol}(\Phi(Q_{\boldsymbol{\alpha}})) &\approx \mathrm{vol}(\Phi(x_{\boldsymbol{\alpha}}) + \mathrm{D}_{x_{\boldsymbol{\alpha}}}\Phi(Q_{\boldsymbol{\alpha}} - x_{\boldsymbol{\alpha}})) \\ &= \mathrm{vol}(\mathrm{D}_{x_{\boldsymbol{\alpha}}}\Phi(Q_{\boldsymbol{\alpha}})) \\ &= |\det(\mathrm{D}_{x_{\boldsymbol{\alpha}}}\Phi)| \,\mathrm{vol}(Q_{\boldsymbol{\alpha}}) \end{aligned} \tag{12.12}$$

nahe. Damit sollte dann auch für eine beschränkte stetige Funktion $f : Y \to \mathbb{R}$

$$\begin{aligned} \int_Y f(y) \,\mathrm{dvol}(y) &= \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X} \int_{\Phi(Q_{\boldsymbol{\alpha}})} f(y) \,\mathrm{dvol}(y) \\ &\approx \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X} f(\Phi(x_{\boldsymbol{\alpha}})) \,\mathrm{vol}(\Phi(Q_{\boldsymbol{\alpha}})) \\ &\underset{(12.12)}{\approx} \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X} f(\Phi(x_{\boldsymbol{\alpha}})) |\det(\mathrm{D}_{x_{\boldsymbol{\alpha}}}\Phi)| \,\mathrm{vol}(Q_{\boldsymbol{\alpha}}) \\ &\approx \int_Q (f \circ \Phi) |\det \mathrm{D}\Phi| \,\mathrm{dvol}. \end{aligned}$$

gelten. Die Symbole $\approx$ sollten hier wiederum verdeutlichen, dass dies (noch) keinen Beweis darstellt. Um daraus einen Beweis zu machen, müssen wir zuerst (12.12) besser verstehen.

**Lemma 12.61** (Lokale Verzerrung von einem Würfel)**.** *Sei $\ell > 0$, sei $x_0 \in \mathbb{R}^n$ und sei*

$$Q_0 = x_0 + [-\ell, \ell]^n$$

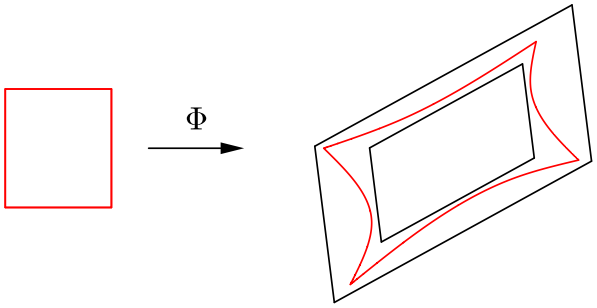
*der achsenparallele ($n$-dimensionale) Würfel mit Seitenlängen gleich $2\ell$ und Mittelpunkt $x_0$. Sei $X \supseteq Q_0$ eine offene Obermenge und sei $\Phi : X \to Y \subseteq \mathbb{R}^n$ ein $C^1$-Diffeomorphismus. Wir setzen $y_0 = \Phi(x_0)$, $L = \mathrm{D}_{x_0}\Phi$ und*

$$\sigma = \max_{x \in Q_0} \| \mathrm{D}_x\Phi - L\|_{\mathrm{op}}. \tag{12.13}$$

*Falls $s = \sigma \|L^{-1}\|_{\mathrm{op}}\sqrt{n} < 1$ ist, dann gilt*

$$y_0 + (1-s)L(Q_0 - x_0) \subseteq \Phi(Q_0) \subseteq y_0 + (1+s)L(Q_0 - x_0). \tag{12.14}$$

Wir bemerken, dass uns (12.14) dabei helfen wird, eine formal korrekte Version von (12.12) zu beweisen. In der Tat können wir $s$ verkleinern, indem wir kleinere Würfel betrachten, wodurch $\sigma$ in (12.13) auf Grund der stetigen Differenzierbarkeit von $\Phi$ kleiner gemacht werden kann. Des Weiteren bemerken wir, dass Lemma 12.61 zum Satz der inversen Funktion (Satz 11.5) verwandt ist. Dementsprechend werden wir wieder die Methoden aus Kapitel 11 für den Beweis des Lemmas verwenden.




*Beweis.* Wir nehmen der Einfachheit halber zuerst an, dass $x_0 = 0$, $y_0 = \Phi(x_0) = 0$ und $L = I_n$ ist. Dann gilt für alle $x_1, x_2 \in Q_0$ wegen

$$\Phi(x_2) - \Phi(x_1) = \int_0^1 (\mathrm{D}_{x_1 + t(x_2 - x_1)}\Phi)(x_2 - x_1)\,\mathrm{d}t$$

und der Dreiecksungleichung für vektorwertige Integrale in (5.16) die Abschätzung

$$\begin{aligned} \|(\Phi(x_2) - x_2) - (\Phi(x_1) - x_1)\| &= \Big\| \int_0^1 \big(\mathrm{D}_{x_1 + t(x_2 - x_1)}\Phi - I_n\big)(x_2 - x_1)\,\mathrm{d}t \Big\| \\ &\le \sigma \|x_2 - x_1\|. \end{aligned} \tag{12.15}$$

Für $x = x_2 \in Q_0$ und $x_1 = 0$ folgt (wegen der Annahme $\Phi(0) = 0$) daraus

$$\|\Phi(x) - x\|_\infty \leq \|\Phi(x) - x\| \leq \sigma\|x\| \leq \sigma\sqrt{n}\ell \tag{12.16}$$

und somit liegen die Koordinaten des Punktes $\Phi(x)$ jeweils in $[-(1+\sigma\sqrt{n})\ell, (1+\sigma\sqrt{n})\ell]$. Also gilt

$$\Phi(Q_0) \subseteq (1 + \sigma\sqrt{n})Q_0,$$

was auf Grund unserer Annahmen genau der zweiten Inklusion im Lemma entspricht.

Für die erste Inklusion wenden wir zu $y \in (1 - \sigma\sqrt{n})Q_0$ den Banachschen Fixpunktsatz (Satz 9.54) auf die Abbildung

$$\begin{aligned} T_y :& Q_0 \to Q_0, \\ & x \mapsto y - (\Phi(x) - x) \end{aligned}$$

an. Für $x \in Q_0$ gilt wegen (12.16)

$$\begin{aligned} \|T_y(x)\|_\infty &\leq \|y\|_\infty + \|\Phi(x) - x\|_\infty \\ &\leq (1 - \sigma\sqrt{n})\ell + \sigma\sqrt{n}\ell = \ell. \end{aligned}$$

Also bildet $T_y$ in der Tat jeden Punkt in $Q_0$ wieder nach $Q_0$ ab. Des Weiteren gilt für $x_1, x_2 \in Q_0$ nach Gleichung (12.15)

$$\begin{aligned} \|T_y(x_1) - T_y(x_2)\| &= \|(\Phi(x_1) - x_1) - (\Phi(x_2) - x_2)\| \\ &\leq \sigma\|x_1 - x_2\|. \end{aligned}$$

Nach Voraussetzung ist aber $\sigma < 1$ und daher können wir den Banachschen Fixpunktsatz (Satz 9.54) anwenden. Es folgt, dass es ein $x \in Q_0$ mit $T_y(x) = y - (\Phi(x) - x) = x$ oder äquivalenterweise $\Phi(x) = y$ gibt. Dies zeigt $(1 - \sigma\sqrt{n})Q_0 \subseteq \Phi(Q_0)$.

Falls nun $x_0 \neq 0$, $y_0 \neq 0$ oder $L = \mathrm{D}_{x_0}\Phi \neq I_n$ ist, dann wollen wir Obiges auf die Abbildung

$$\begin{aligned} \tilde{\Phi} :& Q_0 - x_0 = [-\ell, \ell]^n \to \mathbb{R}^n, \\ & x \mapsto L^{-1}\big(\Phi(x + x_0) - y_0\big) \end{aligned}$$

anwenden. Wir berechnen zuerst

$$\begin{aligned} \tilde{\sigma} &= \max_{x\in[-\ell,\ell]^n} \|\,\mathrm{D}_x\tilde{\Phi} - I_n\|_{\mathrm{op}} \\ &= \max_{x\in[-\ell,\ell]^n} \|L^{-1}(\,\mathrm{D}_{x+x_0}\Phi - L)\|_{\mathrm{op}} \\ &\leq \|L^{-1}\|_{\mathrm{op}}\sigma \end{aligned}$$

und erhalten, dass $\tilde{\Phi}$ die gleichen Eigenschaften wie $\Phi$ im Lemma hat, aber zusätzlich noch $\tilde{\Phi}(0) = 0$, $D_0\tilde{\Phi} = I$ und $\tilde{\sigma}\sqrt{n} < 1$ erfüllt. Aus obigem Fall erhalten wir nun

$$\begin{aligned}(1 - \tilde{\sigma}\sqrt{n}) \cdot [-\ell, \ell]^n &\subseteq \tilde{\Phi}([-\ell, \ell]^n) \\ &\subseteq (1 + \tilde{\sigma}\sqrt{n}) \cdot [-\ell, \ell]^n.\end{aligned}$$

Setzt man nun die Definition von $\tilde{\Phi}$ und die Abschätzung von $\tilde{\sigma}$ ein und wendet $L$ auf alle drei Mengen an, so ergibt sich

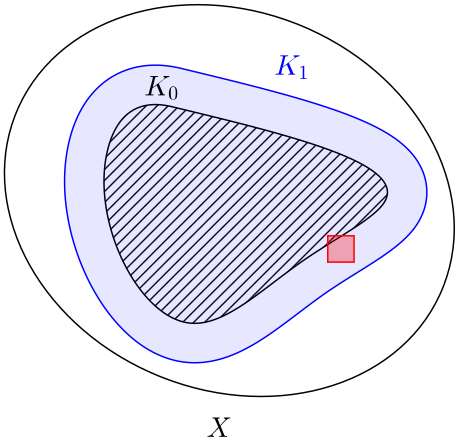
$$\begin{aligned}(1 - s)L(Q_0 - x_0) &\subseteq \Phi(Q_0) - y_0 \\ &\subseteq (1 + s)L(Q_0 - x_0),\end{aligned}$$

woraus (12.14) folgt. □

VO: 27.04.18

Um das letzte Lemma auf ausreichend viele Würfel anwenden zu können benötigen wir folgendes Lemma.

**Lemma 12.62** (Verdickung einer kompakten Menge)**.** *Sei $X \subseteq \mathbb{R}^n$ eine offene Teilmenge und sei $K_0 \subseteq X$ eine kompakte Teilmenge. Dann gibt es ein $\delta_0 > 0$, so dass alle abgeschlossenen Quader $Q$ mit maximaler Kantenlänge kleiner als $\delta_0$ und mit $Q \cap K_0 \neq \emptyset$ in der kompakten Menge $K_1 = K_0 + \overline{B^\infty_{\delta_0}(0)}$ enthalten sind, welche ihrerseits in $X$ liegt.*




Figur 12.5: Die Menge $K_1$ entspricht einer leicht aufgeblähten Version von $K_0$ und ist so definiert, dass genügend kleine Quader, die $K_0$ nicht-trivial schneiden, in $K_1$ liegen wobei aber $K_1$ noch immer eine Teilmenge von $X$ ist.

*Beweis.* Nach Übung 9.41 ist die Abbildung

$$x \in \mathbb{R}^n \mapsto \mathrm{d}_\infty(x, X^c) = \inf_{y \in X^c} \|x - y\|_\infty$$

stetig. Nach Satz 9.66(5) gibt es daher ein $x_0 \in K_0$ mit

$$\mathrm{d}_\infty(x_0, X^c) = \min\left\{\mathrm{d}_\infty(x, X^c) \mid x \in K_0\right\}.$$

Da $x_0 \in K_0$ und $X$ offen ist, gilt $\delta_0 = \frac{1}{2}\mathrm{d}_\infty(x_0, X^c) > 0$. Die Menge $K_1 = K_0 + \overline{B^\infty_{\delta_0}(0)}$ ist als Bild der kompakten Menge $K_0 \times \overline{B^\infty_{\delta_0}(0)}$ unter der stetigen Abbildung $(x, v) \in \mathbb{R}^n \times \mathbb{R}^n \mapsto x + v \in \mathbb{R}^n$ wieder kompakt (nach Satz 9.72) und auf Grund der Definition von $\delta_0$ wieder in $X$ enthalten. Damit ist jeder Quader $Q = [a_1, b_1] \times \ldots \times [a_n, b_n]$ mit maximaler Kantenlänge kleiner $\delta_0$ (das heisst, $\max_{k=1,\ldots,n} |b_k - a_k| \leq \delta_0$) und $Q \cap K_0 \neq \emptyset$ in $K_1$ und somit in $X$ enthalten. □

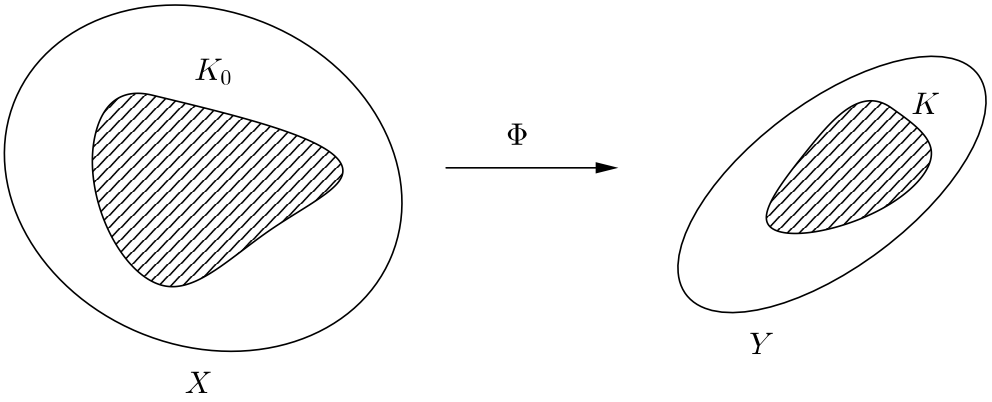
**Lemma 12.63** (Gleichmässige Schranke für Verzerrung von kleinen Würfeln)**.** *Seien $X, Y \subseteq \mathbb{R}^n$ offen, $\Phi : X \to Y$ ein $C^1$-Diffeomorphismus und $K \subseteq Y$ eine kompakten Teilmenge. Dann ist auch $K_0 = \Phi^{-1}(K) \subseteq X$ kompakt, und für jedes $\varepsilon > 0$ existiert ein $\delta > 0$ mit folgender Eigenschaft. Für einen Würfel $Q_0$ mit Kantenlänge kleiner als $\delta$, mit $Q_0 \cap \Phi^{-1}(K) \neq \emptyset$ und mit Mittelpunkt $x_0$ existieren Parallelotope $P^+, P^-$, die*

$$P^- \subseteq \Phi({Q_0}^\circ) \subseteq \Phi(\overline{Q_0}) \subseteq P^+$$

*und ebenso die Abschätzungen*

$$\begin{aligned} \mathrm{vol}(P^-) &> (1 - \varepsilon)|\det D_{x_0}\Phi| \,\mathrm{vol}(Q_0), \\ \mathrm{vol}(P^+) &< (1 + \varepsilon)|\det D_{x_0}\Phi| \,\mathrm{vol}(Q_0) \end{aligned}$$

*erfüllen.*




Figur 12.6: Die Menge $K_0$ ist per Definition das Urbild von $K$.

*Beweis.* Wie bereits im Lemma erwähnt setzen wir $K_0 = \Phi^{-1}(K) \subseteq X$. Da $\Phi^{-1} : Y \to X$ stetig ist und $K$ kompakt ist, ist das Bild $K_0$ ebenso kompakt (siehe Satz 9.72).

Sei nun $\delta_0 > 0$ und $K_1 = K_0 + \overline{B^\infty_{\delta_0}(0)}$ die kompakte Teilmenge von $X$ wie in Lemma 12.62. Nach dem Verhalten von stetigen Funktionen auf kompakten Mengen (Satz 9.66(5) und Proposition 9.77) ist

$$M = \max\left\{\|(\mathrm{D}_x\Phi)^{-1}\|_{\mathrm{op}} \mid x \in K_1\right\} < \infty$$

und $\mathrm{D}\Phi$ auf $K_1$ gleichmässig stetig. Sei nun $\sigma > 0$ beliebig. Wir wählen $\delta \in (0, \delta_0)$ unter Verwendung der gleichmässigen Stetigkeit von $D\Phi$ auf $K_1$, der Operatornorm auf $\mathrm{Mat}_{n,n}(\mathbb{R})$,

und der Norm $\|\cdot\|_\infty$ auf $K_1$. Wenn nun ein Würfel $Q_0$ Kantenlänge kleiner als $\delta$ hat und $Q_0 \cap K_0 \neq \emptyset$, so gilt $Q_0 \subseteq K_1$ (nach Wahl von $\delta_0$) und $\|\mathrm{D}_x\Phi - \mathrm{D}_{x_0}\Phi\|_{\mathrm{op}} < \sigma$ für alle $x, x_0 \in Q_0$ (nach Wahl von $\delta$).

Seien $y_0 = \Phi(x_0)$, $L = \mathrm{D}_{x_0}\Phi$ und $s = s(\sigma) = \sigma M\sqrt{n}$. Aus Lemma 12.61 folgt zumindest für hinreichend kleine $\sigma$ (und damit kleine $s$), dass

$$y_0 + (1-s)L(Q_0 - x_0) \subseteq \Phi(Q_0) \subseteq y_0 + (1+s)L(Q_0 - x_0)$$

gilt. Wir setzen nun $P^+ = y_0 + (1+s)L(\overline{Q_0} - x_0)$ und $P^- = y_0 + (1-2s)L(\overline{Q_0} - x_0)$. Daher gilt $P^- \subseteq \Phi(Q_0^\circ) \subseteq \Phi(\overline{Q_0}) \subseteq P^+$, wobei wir den Faktor 2 in der Definition von $P^-$ eingefügt haben, um jeden Zweifel bezüglich der Inklusion

$$P_- = y_0 + (1-2s)L(\overline{Q_0} - x_0) \subseteq \Phi(Q_0^\circ)$$

zu zerstreuen. Wir haben nach Übung 12.31 und der linearen Substitution in Proposition 12.49

$$\begin{aligned}
\mathrm{vol}(P^-) &= (1-2s)^n \,\mathrm{vol}(L(Q_0)) \\
&= (1-2s)^n |\det(\mathrm{D}_{x_0}\Phi)| \,\mathrm{vol}(Q_0) \\
\mathrm{vol}(P^+) &= (1+s)^n \,\mathrm{vol}(L(Q_0)) \\
&= (1+s)^n |\det(\mathrm{D}_{x_0}\Phi)| \,\mathrm{vol}(Q_0).
\end{aligned}$$

Da $s = \sigma M\sqrt{n}$ für $\sigma \to 0$ gegen Null strebt, können wir für jedes $\varepsilon > 0$ die Zahl $\sigma > 0$ und damit auch $\delta > 0$ so wählen, dass

$$(1-2s)^n > 1 - \varepsilon, \quad (1+s)^n < 1 + \varepsilon$$

ist. Nach obigen Abschätzungen beweist dies die Proposition. □

Wir werden nun die Substitutionsregel in zwei Schritten beweisen, wobei wir zuerst stetige Funktionen mit kompakten Träger untersuchen.

*Beweis von Satz 12.56 für stetige Funktionen.* Sei $f : Y \to \mathbb{R}$ eine stetige Funktion mit kompaktem Träger $K = \mathrm{supp}(f) \subseteq Y$. Wir können ohne Beschränkung der Allgemeinheit annehmen, dass $f$ nicht-negativ ist, denn der allgemeine Fall folgt schliesslich aus der Linearität des Riemann-Integrals und der Zerlegung $f = f^+ - f^-$, wobei $f^+ = \frac{1}{2}(|f| + f)$ und $f^- = \frac{1}{2}(|f| - f)$ nicht-negativ sind. Wir definieren $K_0, K_1 \subseteq X$ und $\delta_0$ wie in Lemma 12.63 und Lemma 12.62. Nach Proposition 9.77 ist $f \circ \Phi$ auf $K_1$ gleichmässig stetig. Wir können somit zu $\varepsilon > 0$ ein $\delta \in (0, \delta_0)$ unter Verwendung der gleichmässigen Stetigkeit von $f \circ \Phi$ so wählen, dass für alle $x_1, x_2 \in K_1$

$$\|x_1 - x_2\|_\infty < \delta \implies |f \circ \Phi(x_1) - f \circ \Phi(x_2)| < \varepsilon$$

gilt. Des Weiteren können wir $\delta$ (wenn nötig) verkleinern, so dass $\delta$ die Aussage zu $\varepsilon > 0$ in Lemma 12.63 für $K = \mathrm{supp}(f)$ erfüllt.

Sei $Q_X$ ein abgeschlossener Würfel, der die Menge $K_0$ enthält. Dann gilt nach Definition

$$\int_X (f \circ \Phi)|\det(\mathrm{D}\Phi)|\,\mathrm{dvol} = \int_{Q_X} \mathbb{1}_{K_0}(f \circ \Phi)|\det(\mathrm{D}\Phi)|\,\mathrm{dvol}.$$

Sei nun $\mathfrak{Z}_X$ eine Zerlegung von $Q_X$ in Würfel mit Kantenlänge kleiner als $\delta$. Wir teilen die der Zerlegung $\mathfrak{Z}_X$ entsprechenden Würfel in zwei Typen ein. Die Würfel $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X$ mit $\overline{Q_{\boldsymbol{\alpha}}} \cap K_0 \neq \emptyset$ erfüllen $\overline{Q_{\boldsymbol{\alpha}}} \subseteq K_1$ und sind im Folgenden „interessant". Für die Würfel $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X$ mit $\overline{Q_{\boldsymbol{\alpha}}} \cap K_0 = \emptyset$ werden wir uns jedoch im Weiteren nicht interessieren, da für diese

$$\big(\mathbb{1}_{K_0} f \circ \Phi\big)|_{\overline{Q_{\boldsymbol{\alpha}}}} = 0$$

gilt. Da Ränder von Würfeln Jordan-Nullmengen sind, erhalten wir

$$\int_X (f \circ \Phi)|\det(\mathrm{D}\Phi)|\,\mathrm{dvol} = \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X}{}' \int_{Q_{\boldsymbol{\alpha}}} (f \circ \Phi)|\det(\mathrm{D}\Phi)|\,\mathrm{dvol}$$

(siehe auch Übung 12.37), wobei $\sum'_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X}$ die Summe über die interessanten Teilwürfel bezeichnet (das heisst, jene $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X$ mit $\overline{Q_{\boldsymbol{\alpha}}} \cap K_0 \neq \emptyset$). Wir wollen nun eine ähnliche Zerlegung von $\int_Y f\,\mathrm{dvol}$ finden, dabei aber auch in $Y$ die Zerlegung $\mathfrak{Z}_X$ von $Q_X$ verwenden.

Dementsprechend verwenden wir die interessanten Teilwürfel $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X$, um $K$ mittels

$$\begin{aligned} K &= \Phi(K_0) \\ &= \Phi\Bigg(\bigcup_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X}{}' K_0 \cap \overline{Q_{\boldsymbol{\alpha}}}\Bigg) \subseteq \bigcup_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X}{}' \Phi(\overline{Q_{\boldsymbol{\alpha}}}) \end{aligned}$$

zu überdecken, wobei $\bigcup'_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X}$ die Vereinigung über die interessanten Teilwürfel andeutet. Nach Wahl von $\delta > 0$ gilt für einen interessanten Teilwürfel $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X$, dass $\overline{Q_{\boldsymbol{\alpha}}} \subseteq K_1 \subseteq X$ und für $x \in \overline{Q_{\boldsymbol{\alpha}}}$ auch

$$f(\Phi(x_{\boldsymbol{\alpha}})) - \varepsilon \leq f(\Phi(x)) \leq f(\Phi(x_{\boldsymbol{\alpha}})) + \varepsilon,$$

wobei $x_{\boldsymbol{\alpha}}$ den Mittelpunkt von $Q_{\boldsymbol{\alpha}}$ bezeichnet. Zur Vereinfachung der Notation setzen wir $y_{\boldsymbol{\alpha}} = \Phi(x_{\boldsymbol{\alpha}})$. Wir verknüpfen obige Ungleichung mit der Aussage über die Verzerrung von kleinen Würfeln in Lemma 12.63. Danach existieren Parallelotope $P^-_{\boldsymbol{\alpha}}, P^+_{\boldsymbol{\alpha}}$ mit

$$P^-_{\boldsymbol{\alpha}} \subseteq \Phi(Q_{\boldsymbol{\alpha}}) \subseteq \Phi(\overline{Q_{\boldsymbol{\alpha}}}) \subseteq P^+_{\boldsymbol{\alpha}}, \tag{12.17}$$

wobei das Volumen der Parallelotope bis auf einen Faktor $1 \pm \varepsilon$ durch $|\det(\mathrm{D}_{x_{\boldsymbol{\alpha}}}\Phi)|\,\mathrm{vol}(Q_{\boldsymbol{\alpha}})$ abgeschätzt wird. Verwenden wir nun noch, dass $f$ nicht-negativ ist, so ergibt sich die Ungleichung

$$\begin{aligned} (f(y_{\boldsymbol{\alpha}}) - \varepsilon)\,\mathbb{1}_{P^-_{\boldsymbol{\alpha}}} \leq f\,\mathbb{1}_{\Phi(Q_{\boldsymbol{\alpha}})} &\leq f\,\mathbb{1}_{\Phi(\overline{Q_{\boldsymbol{\alpha}}})} \\ &\leq (f(y_{\boldsymbol{\alpha}}) + \varepsilon)\,\mathbb{1}_{P^+_{\boldsymbol{\alpha}}} \end{aligned}$$

zwischen Funktionen auf $Y$. Wir summieren über alle interessanten Teilwürfel $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X$ und erhalten

$$\begin{aligned}\sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X}' (f(y_{\boldsymbol{\alpha}}) - \varepsilon)\, \mathbb{1}_{P_{\boldsymbol{\alpha}}^-} \leq f \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X}' \mathbb{1}_{\Phi(Q_{\boldsymbol{\alpha}})} &\leq f \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X}' \mathbb{1}_{\Phi(\overline{Q_{\boldsymbol{\alpha}}})} \\ &\leq \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X}' (f(y_{\boldsymbol{\alpha}}) + \varepsilon)\, \mathbb{1}_{P_{\boldsymbol{\alpha}}^+}.\end{aligned} \tag{12.18}$$

Für $y \in K$ mit $y \in \Phi(Q_{\boldsymbol{\alpha}})$ zu $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X$ ist der Würfel $Q_{\boldsymbol{\alpha}}$ eindeutig bestimmt, womit die beiden Funktionen rechts in (12.18) für solche Punkte $y$ mit $f(y)$ übereinstimmen. Für $y \in K$ mit $y \in \Phi(\partial Q_{\boldsymbol{\alpha}})$ zu $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X$ gibt es möglicherweise mehrere solche Würfel $Q_{\boldsymbol{\alpha}}$ und wir erhalten

$$f(y) \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X}' \mathbb{1}_{\Phi(Q_{\boldsymbol{\alpha}})}(y) = 0 \leq f(y) \leq f(y) \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X}' \mathbb{1}_{\Phi(\overline{Q_{\boldsymbol{\alpha}}})}(y).$$

Da $K \subseteq \bigcup'_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X} \Phi(\overline{Q_{\boldsymbol{\alpha}}})$ gilt, sind dies alle Fälle, die auftreten können und wir können aus (12.18) die Abschätzung

$$\sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X}' (f(y_{\boldsymbol{\alpha}}) - \varepsilon)\, \mathbb{1}_{P_{\boldsymbol{\alpha}}^-} \leq f \leq \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X}' (f(y_{\boldsymbol{\alpha}}) + \varepsilon)\, \mathbb{1}_{P_{\boldsymbol{\alpha}}^+}$$

folgern. Wir integrieren diese Ungleichung nun über $Y$ (oder genauer gesagt über einen Quader der $K$ enthält) und erhalten

$$\sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X}' (f(y_{\boldsymbol{\alpha}}) - \varepsilon) \operatorname{vol}(P_{\boldsymbol{\alpha}}^-) \leq \int_Y f \operatorname{dvol} \leq \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X}' (f(y_{\boldsymbol{\alpha}}) + \varepsilon) \operatorname{vol}(P_{\boldsymbol{\alpha}}^+). \tag{12.19}$$

Mittels der gleichmässigen Abschätzung über die Verzerrung von Würfeln in Lemma 12.63 erhalten wir nun aber für den Ausdruck auf der rechten Seite

$$\sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X}' f(y_{\boldsymbol{\alpha}}) \operatorname{vol}(P_{\boldsymbol{\alpha}}^+) \leq (1+\varepsilon) \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X}' f(y_{\boldsymbol{\alpha}}) |\det(\mathrm{D}_{x_{\boldsymbol{\alpha}}} \Phi)| \operatorname{vol}(Q_{\boldsymbol{\alpha}})$$

sowie

$$\sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X}' \varepsilon \operatorname{vol}(P_{\boldsymbol{\alpha}}^+) \leq \varepsilon(1+\varepsilon) \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X}' |\det(\mathrm{D}_{x_{\boldsymbol{\alpha}}} \Phi)| \operatorname{vol}(Q_{\boldsymbol{\alpha}}) = O(\varepsilon),$$

da die Abbildung $|\det(\mathrm{D}\Phi)|$ auf $K_1$ beschränkt ist. Aus (12.19) erhält man nun

$$\int_Y f \operatorname{dvol} \leq (1+\varepsilon) \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}_X}' f(\Phi(x_{\boldsymbol{\alpha}})) |\det(\mathrm{D}_{x_{\boldsymbol{\alpha}}} \Phi)| \operatorname{vol}(Q_{\boldsymbol{\alpha}}) + O(\varepsilon).$$

Mit den analogen Abschätzungen für die linke Seite in (12.19) zeigt man

$$\int_Y f \,\mathrm{dvol} \geq (1-\varepsilon) \sum_{Q_{\boldsymbol{\alpha}} \sqsubset 3_X}{}' f(\Phi(x_{\boldsymbol{\alpha}}))|\det(\mathrm{D}_{x_{\boldsymbol{\alpha}}}\Phi)| \,\mathrm{vol}(Q_{\boldsymbol{\alpha}}) + O(\varepsilon).$$

Da aber $\mathbb{1}_{K_0}(x_{\boldsymbol{\alpha}})f(\Phi(x_{\boldsymbol{\alpha}}))$ für alle interessanten Quader gleich $f(\Phi(x_{\boldsymbol{\alpha}}))$ und für die restlichen gleich Null ist, so gilt

$$\sum_{Q_{\boldsymbol{\alpha}} \sqsubset 3_X}{}' f(\Phi(x_{\boldsymbol{\alpha}}))|\det(\mathrm{D}_{x_{\boldsymbol{\alpha}}}\Phi)| \,\mathrm{vol}(Q_{\boldsymbol{\alpha}}) = \sum_{Q_{\boldsymbol{\alpha}} \sqsubset 3_X} \mathbb{1}_{K_0}(x_{\boldsymbol{\alpha}})f(\Phi(x_{\boldsymbol{\alpha}}))|\det(\mathrm{D}_{x_{\boldsymbol{\alpha}}}\Phi)| \,\mathrm{vol}(Q_{\boldsymbol{\alpha}})$$

und die rechte Seite stellt eine Riemann-Summe für das Integral $\int_{Q_X} \mathbb{1}_{K_0} f \circ \Phi |\det(\mathrm{D}\Phi)| \,\mathrm{dvol}$ dar, wobei $\mathbb{1}_{K_0} f \circ \Phi |\det(\mathrm{D}\Phi)|$ auf $Q_X$ stetig (siehe Übung 12.64) und insbesondere Riemann-integrierbar ist. Somit folgt die Behauptung für $\varepsilon \to 0$ nach Satz 12.15. □

**Übung 12.64** (Stetige Fortsetzung)**.** *Sei $X \subseteq \mathbb{R}^n$ offen und $f : X \to \mathbb{R}$ stetig mit kompakten Träger $K = \mathrm{supp}(f)$ in $X$. Zeigen Sie, dass $\mathbb{1}_K f$ auf ganz $\mathbb{R}^n$ stetig ist.*

Wir verwenden nun folgende Verallgemeinerung der Charakterisierung von Riemann-Integrierbarkeit mittels stetigen Funktionen aus Proposition 12.13, um die oben bewiesene Aussage auf alle Riemann-integrierbaren Funktionen mit kompaktem Träger zu verallgemeinern.

**Die letzten beiden Schritte wurden in der VO nur skizziert.*

**Lemma 12.65** (Sandwich-Charakterisierung mit stetigen Funktionen*)**.** *Sei $U \subseteq \mathbb{R}^n$ eine offene Teilmenge und sei $f : U \to \mathbb{R}$ eine Funktion mit kompaktem Träger in $U$. Falls die Funktion $f$ Riemann-integrierbar ist, dann gibt es für jedes $\varepsilon > 0$ stetige reellwertige Funktionen $f_-, f_+$ mit kompaktem Träger in $U$ für die $f_- \leq f \leq f_+$ sowie $\int_U (f_+ - f_-) \,\mathrm{dvol} < \varepsilon$ erfüllt ist.*

*Beweis.* Sei $K = \mathrm{supp}(f) \subseteq U$ der Träger der Funktion $f$. Sei $Q$ ein abgeschlossener Quader mit $Q \supseteq K$ und sei $\varepsilon > 0$. Durch Vergrösserung von $Q$ können wir, wenn nötig, erreichen, dass $K \subseteq Q^\circ$. Da $f$ als Riemann-integrierbar angenommen ist, existieren nach Proposition 12.13 stetige Funktionen $f_-, f_+ : Q \to \mathbb{R}$ mit $f_- \leq \mathbb{1}_U f \leq f_+$ und $\int_Q (f_+ - f_-) \,\mathrm{dvol} < \varepsilon$. Wir passen nun $f_-, f_+$ so an, dass ihr Träger in $U \cap Q^\circ$ liegt.

Dazu betrachten wir den minimalen Abstand von $K = \mathrm{supp}(f)$ zum Komplement der offenen Menge $U \cap Q^\circ$ in $\mathbb{R}^n$

$$\delta = \min_{y \in K} \mathrm{d}(y, (U \cap Q^\circ)^c) > 0$$

sowie die stetige Funktion

$$h : x \in \mathbb{R}^n \mapsto \min\big(1, \max(0, \tfrac{2}{\delta}\mathrm{d}(x, (U \cap Q^\circ)^c) - 1)\big).$$

Man bemerke dabei, dass $h(x) = 1$ für alle $x \in K$ und $h(x) = 0$ für alle $x \in \mathbb{R}^n$ mit $\mathrm{d}(x, (U \cap Q^\circ)^c) \leq \frac{\delta}{2}$. (Wieso?) Insbesondere hat $h$ kompakten Träger in $U \cap Q^\circ$ und ist gleich

1 auf dem Träger von $f$. Auch gilt per Definition $h \leq 1$. Also gilt $hf_- \leq \mathbb{1}_U f \leq hf_+$ sowie

$$\int_U (hf_+ - hf_-)\,\mathrm{dvol} = \int_Q h(f_+ - f_-)\,\mathrm{dvol} \leq \int_Q (f_+ - f_-)\,\mathrm{dvol} < \varepsilon.$$

Die stetigen Funktionen $hf_-$ und $hf_+$ haben also die gewünschten Eigenschaften. □

*Beweis von Satz 12.56 für beliebige Riemann-integrierbare Funktionen*.* Sei $f : Y \to \mathbb{R}$ eine Riemann-integrierbare Funktion mit kompaktem Träger $\operatorname{supp}(f) \subseteq Y$ wie im Satz. Sei $\varepsilon > 0$. Dann existieren nach Lemma 12.65 stetige Funktionen $f_-, f_+ : Y \to \mathbb{R}$ mit kompaktem Träger in $Y$ und $f_- \leq f \leq f_+$ sowie $\int_Y (f_+ - f_-)\,\mathrm{dvol} < \varepsilon$.

Aus dem oben bewiesenen Fall folgt nun

$$\int_X \big((f_+ \circ \Phi) - (f_- \circ \Phi)\big)|\det(\mathrm{D}\Phi)|\,\mathrm{dvol} = \int_Y (f_+ - f_-)\,\mathrm{dvol} < \varepsilon.$$

Auch gilt per Definition von $f_-$ und $f_+$, dass

$$(f_- \circ \Phi)\,|\det(\mathrm{D}\Phi)| \leq (f \circ \Phi)\,|\det(\mathrm{D}\Phi)| \leq (f_+ \circ \Phi)|\det(\mathrm{D}\Phi)|$$

erfüllt ist und dass $(f_- \circ \Phi)|\det(\mathrm{D}\Phi)|$ und $(f_+ \circ \Phi)|\det(\mathrm{D}\Phi)|$ kompakten Träger in $X$ haben. Da $\varepsilon$ beliebig war folgt mit Lemma 12.65 also, dass $(f \circ \Phi)|\det(\mathrm{D}\Phi)|$ Riemann-integrierbar ist. Des Weiteren gilt

$$\begin{aligned}\int_X (f \circ \Phi)|\det(\mathrm{D}\Phi)|\,\mathrm{dvol} &\leq \int_X (f_+ \circ \Phi)|\det(\mathrm{D}\Phi)|\,\mathrm{dvol} \\ &= \int_Y f_+\,\mathrm{dvol} < \int_Y f\,\mathrm{dvol} + \varepsilon\end{aligned}$$

und analog

$$\int_X (f \circ \Phi)|\det(\mathrm{D}\Phi)|\,\mathrm{dvol} > \int_Y f\,\mathrm{dvol} - \varepsilon,$$

was den Satz beweist, da $\varepsilon > 0$ beliebig war. □

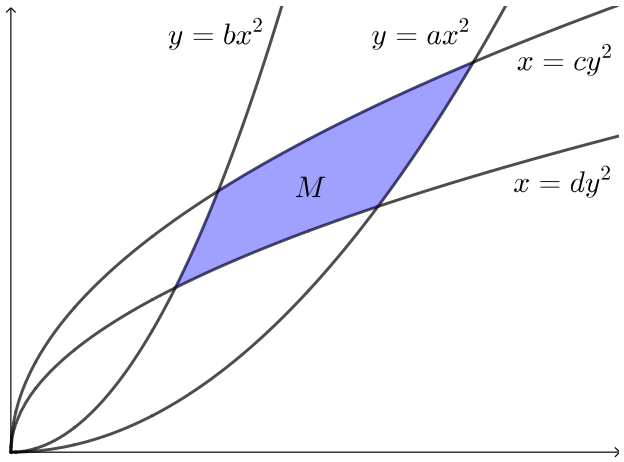
VO: 30.04.18

### 12.6.4 Beispiele

**Beispiel 12.66** (Krummliniges Viereck)**.** *Seien $0 < a < b$ und $0 < c < d$ feste Parameter, die wir verwenden, um das krummlinig begrenzte Viereck $M \subseteq X = \mathbb{R}^2_{>0}$ zu definieren, welches durch die Kurven*

$$\begin{aligned} y &= ax^2 & y &= bx^2 \\ x &= cy^2 & x &= dy^2 \end{aligned}$$

*begrenzt wird; siehe folgendes Bild.*




*Wir wollen die Substitutionsregel verwenden, um den Flächeninhalt* $\mathrm{vol}(M)$ *von* $M$ *zu berechnen. Hierfür bemerken wir, dass* $M$ *auch durch*

$$M = \left\{ (x,y) \in X \mid \frac{1}{b} \leq x^2 y^{-1} \leq \frac{1}{a},\ \frac{1}{d} \leq x^{-1} y^2 \leq \frac{1}{c} \right\}$$

*gegeben ist. Daraus erkennen wir, dass die Variablen* $u = x^2 y^{-1}$ *und* $v = x^{-1} y^2$ *vielleicht nützlich sein könnten. In der Tat definiert*

$$\Psi : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x^2 y^{-1} \\ x^{-1} y^2 \end{pmatrix}$$

*einen Diffeomorphismus* $\Psi : X \to X$*. Die Umkehrabbildung* $\Phi : X \to X$ *ist durch*

$$\Phi : \begin{pmatrix} u \\ v \end{pmatrix} \mapsto \begin{pmatrix} u^{\frac{2}{3}} v^{\frac{1}{3}} \\ u^{\frac{1}{3}} v^{\frac{2}{3}} \end{pmatrix}$$

*gegeben und bestitzt die Jacobi-Determinante*

$$\begin{aligned} \det D_{(u,v)}\Phi &= \det \begin{pmatrix} \frac{2}{3} u^{-\frac{1}{3}} v^{\frac{1}{3}} & \frac{1}{3} u^{\frac{2}{3}} v^{-\frac{2}{3}} \\ \frac{1}{3} u^{-\frac{2}{3}} v^{\frac{2}{3}} & \frac{2}{3} u^{\frac{1}{3}} v^{-\frac{1}{3}} \end{pmatrix} \\ &= \tfrac{4}{9} - \tfrac{1}{9} = \tfrac{1}{3}. \end{aligned}$$

*für alle* $(u,v)^t \in X$*. Dadurch erhalten wir*

$$\begin{aligned} \mathrm{vol}(M) &= \int_X \mathbb{1}_M \, \mathrm{dvol} \\ &= \int_X \mathbb{1}_{[\frac{1}{b},\frac{1}{a}]\times[\frac{1}{d},\frac{1}{c}]}(u,v) \left| \det D_{(u,v)}\Phi \right| \mathrm{dvol}(u,v) \\ &= \tfrac{1}{3}\left(\tfrac{1}{a} - \tfrac{1}{b}\right)\left(\tfrac{1}{c} - \tfrac{1}{d}\right). \end{aligned}$$

**Übung 12.67** (Integralberechnung mittels Substitution)**.**

*(i) Berechnen Sie für vorgegebene Konstanten* $0 < a < b$ *und* $0 < c < d$ *den Flächeninhalt*

$$\operatorname{vol}\left(\left\{(x,y) \in \mathbb{R}^2 \mid a \leq y\mathrm{e}^{-x} \leq b,\ c \leq y\mathrm{e}^{x} \leq d\right\}\right).$$

*(ii) Berechnen Sie das Integral*

$$\int_V x^2 yz \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$$

*über die Teilmenge*

$$V = \left\{(x,y,z) \in \mathbb{R}^3_{>0} \mid x^2 + y^2 \leq 1,\ \tfrac{1}{\sqrt{3}} \leq \tfrac{y}{x} \leq \sqrt{3},\ z \leq 1\right\}.$$

## 12.7 Uneigentliche Mehrfachintegrale

Wir beginnen diesen Abschnitt mit einer Rechnung, die helfen sollte, die danach folgenden Diskussionen zu motivieren.

**Beispiel 12.68** (Gauss'sche Glockenkurve)**.** *Die* ***Gauss'sche Glockenkurve****, die nach Normalisierung des Flächeninhalts unter deren Graphen auch* ***Dichtefunktion der Normalverteilung*** *genannt wird, ist durch*

$$f : \mathbb{R} \to \mathbb{R},\ x \mapsto \mathrm{e}^{-x^2}$$

*definiert. Die Stammfunktion* $x \mapsto F(x) = \int_{-\infty}^{x} \mathrm{e}^{-t^2}\,\mathrm{d}t$ *ist eine, unter anderem in der Statistik, unentbehrliche Funktion, welche sich aber nicht mit den „Standardfunktionen" ausdrücken lässt. Es ist aber möglich, das Integral*

$$I = \int_{-\infty}^{\infty} \mathrm{e}^{-x^2}\,\mathrm{d}x$$

*zu berechnen. Mittels* $\mathrm{e}^{-x^2} \leq \mathrm{e}^{-x}$ *für* $x \geq 1$ *folgt die Konvergenz von* $\int_0^\infty \mathrm{e}^{-x^2}\,\mathrm{d}x$ *(und damit auch die Konvergenz von* $\int_{-\infty}^{\infty} \mathrm{e}^{-x^2}\,\mathrm{d}x$*) unmittelbar aus dem Majorantenkriterium für uneigentliche Integrale. Etwas überraschend verwenden wir dazu Polarkoordinaten im* $\mathbb{R}^2$*, um* $I^2$ *zu berechnen. In der Tat gilt*

$$\begin{aligned}
I^2 &= \Big(\int_{-\infty}^{\infty} \mathrm{e}^{-x^2}\,\mathrm{d}x\Big)\Big(\int_{-\infty}^{\infty} \mathrm{e}^{-y^2}\,\mathrm{d}y\Big) && \text{(12.20a)}\\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{e}^{-x^2-y^2}\,\mathrm{d}y\,\mathrm{d}x && \text{(12.20b)}\\
&= \int_{\mathbb{R}^2} \mathrm{e}^{-(x^2+y^2)}\,\mathrm{dvol}(x,y) && \text{(12.20c)}\\
&= \int_{(0,\infty)\times(0,2\pi)} \mathrm{e}^{-r^2} r\,\mathrm{dvol}(r,\varphi) && \text{(12.20d)}\\
&= \int_0^\infty \mathrm{e}^{-r^2} r \int_0^{2\pi} \mathrm{d}\varphi\,\mathrm{d}r\\
&= \pi \int_0^\infty \mathrm{e}^{-r^2} 2r\,\mathrm{d}r\\
&= \pi \lim_{R\to\infty} \int_0^R \mathrm{e}^{-r^2} \underbrace{2r\,\mathrm{d}r}_{=\mathrm{d}u} = \pi \lim_{R\to\infty} (-\mathrm{e}^{-R^2}+1) = \pi,
\end{aligned}$$

*woraus also* $I = \sqrt{\pi}$ *folgt. Es gibt in dieser Rechnung jedoch ein paar ungeklärte Schritte: Gleichung* (12.20a) *ist bloss die Definition von* $I$ *und Gleichung* (12.20b) *erhalten wir, indem wir die Konstante* $\int_{-\infty}^{\infty} \mathrm{e}^{-y^2}\,\mathrm{d}y$ *in das Integral über* $x$ *hineinziehen und anschliessend den Term* $\mathrm{e}^{-x^2}$ *in das Integral über* $y$ *hineinziehen. In Gleichung* (12.20c) *haben wir den Satz von Fubini*

*auf ganz $\mathbb{R}^2$ angewendet, was wir so begründen wollen:*

$$\begin{aligned}\int_{\mathbb{R}^2} \mathrm{e}^{-(x^2+y^2)}\,\mathrm{dvol}(x,y) &= \lim_{s\to\infty}\int_{[-s,s]^2} \mathrm{e}^{-(x^2+y^2)}\,\mathrm{dvol}(x,y) \\ &= \lim_{s\to\infty}\int_{-s}^{s}\int_{-s}^{s} \mathrm{e}^{-(x^2+y^2)}\,\mathrm{d}y\,\mathrm{d}x = I^2.\end{aligned} \tag{12.21}$$

*In der Zeile* (12.20d) *haben wir die Substitution für Polarkoordinaten auf $\mathbb{R}^2$ angewendet, was wir so begründen wollen:*

$$\begin{aligned}\int_{\mathbb{R}^2} \mathrm{e}^{-(x^2+y^2)}\,\mathrm{dvol}(x,y) &= \lim_{R\to\infty}\int_{B_R(0)} \mathrm{e}^{-(x^2+y^2)}\,\mathrm{dvol}(x,y) \\ &= \lim_{R\to\infty}\int_{(0,R)\times(0,2\pi)} \mathrm{e}^{-r^2} r\,\mathrm{dvol}(r,\varphi) \\ &= \pi \lim_{R\to\infty}\int_0^R \mathrm{e}^{-r^2} 2r\,\mathrm{d}r = \pi.\end{aligned} \tag{12.22}$$

*Leider wirft dies aber eine weitere Frage auf: warum sind die Grenzwerte in den Gleichungen* (12.21) *und* (12.22) *gleich. Unter anderem möchten wir diese Frage in diesem Abschnitt beantworten.*

### 12.7.1 Ausschöpfungen und uneigentliche Integrale

**Definition 12.69** (Ausschöpfungen)**.** Eine **Ausschöpfung** einer Teilmenge $B \subseteq \mathbb{R}^n$ ist eine Folge Jordan-messbarer Teilmengen $(B_m)_m$ mit

$$B_1 \subseteq B_2 \subseteq B_3 \subseteq \dots \quad \text{und} \quad B = \bigcup_{m=1}^{\infty} B_m.$$

In Beispiel 12.68 haben wir bereits zwei Ausschöpfungen von $\mathbb{R}^2$ gesehen, nämlich

$$[-1,1]^2 \subseteq [-2,2]^2 \subseteq [-3,3]^2 \subseteq \dots$$

sowie

$$B_1(0) \subseteq B_2(0) \subseteq B_3(0) \subseteq \dots$$

Wir werden im Folgenden immer Mengen $B \subseteq \mathbb{R}^n$ betrachten, die Ausschöpfungen besitzen.

**Definition 12.70** (Uneigentliches Riemann-Integral)**.** Sei $B \subseteq \mathbb{R}^n$ und sei $f : B \to \mathbb{R}$ eine Funktion. Wir sagen, dass $f$ auf $B$ **uneigentlich Riemann-integrierbar** ist, falls $B$ eine Ausschöpfung $(B_m)_m$ besitzt, so dass für alle $m \in \mathbb{N}$ die Einschränkung $f|_{B_m}$ Riemann-integrierbar ist, der Grenzwert $\lim_{m\to\infty}\int_{B_m} f\,\mathrm{dvol}$ existiert und von der Wahl der Ausschöpfung $(B_m)_m$ mit obiger Eigenschaft unabhängig ist. In diesem Fall schreiben wir

$$\int_B f\,\mathrm{dvol} = \lim_{m\to\infty}\int_{B_m} f\,\mathrm{dvol}$$

für das **uneigentliche Riemann-Integral** von $f$ über $B$.

Diese Definition stellt eine Erweiterung des Riemann-Integrals dar, wie folgende Proposition zeigt. Insbesondere können wir ohne Gefahr auf Widerspruch auch für das uneigentliche Riemann-Integral die Notation $\int_B f \,\mathrm{dvol}$ verwenden.

**Proposition 12.71** (Kompatibilität)**.** *Sei $B$ eine Jordan-messbare Teilmenge, sei $(B_m)_m$ eine Ausschöpfung von $B$ und sei $f : B \to \mathbb{R}$ Riemann-integrierbar. Dann gilt*

$$\begin{aligned} \mathrm{vol}(B) &= \lim_{m\to\infty} \mathrm{vol}(B_m), \\ \int_B f \,\mathrm{dvol} &= \lim_{m\to\infty} \int_{B_m} f \,\mathrm{dvol}. \end{aligned}$$

*Insbesondere ist $f$ über $B$ uneigentlich Riemann-integrierbar.*

*Beweis.* Wegen $B_m \subseteq B_{m+1} \subseteq B$ für alle $m \in \mathbb{N}$ ist die Folge $(\mathrm{vol}(B_m))_m$ monoton wachsend und beschränkt. Also konvergiert sie und es gilt

$$\lim_{m\to\infty} \mathrm{vol}(B_m) \leq \mathrm{vol}(B).$$

Für die andere Ungleichung erinnern wir daran, dass nach Korollar 12.27 die Mengen

$$\partial B_m = \overline{B_m} \setminus B_m^\circ, \quad \partial B = \overline{B} \setminus B^\circ$$

für alle $m \in \mathbb{N}$ Nullmengen sind. Tatsächlich sind sie diese Mengen auch kompakt und somit nach Übung 12.36 Jordan-Nullmengen und also Jordan-messbar und mit Volumen 0. Insbesondere gilt

$$\mathrm{vol}(B_m^\circ) = \mathrm{vol}(B_m), \quad \mathrm{vol}(\overline{B}) = \mathrm{vol}(B)$$

für alle $m \in \mathbb{N}$. Des Weiteren ist

$$N = \bigcup_{m=1}^{\infty} \partial B_m \cup \partial B$$

als abzählbare Vereinigung von Lebesgue-Nullmengen eine Lebesgue-Nullmenge (siehe Lemma 12.17). In anderen Worten existiert für alle $\varepsilon > 0$ eine Überdeckung $N \subseteq \bigcup_{k=1}^{\infty} O_k$ mittels offener Quader $O_1, O_2, \ldots$ mit Gesamtvolumen $\sum_{k=1}^{\infty} \mathrm{vol}(O_k) < \varepsilon$. Dann ist aber

$$\begin{aligned} \overline{B} &= B \cup \partial B \\ &= \bigcup_{m=1}^{\infty} B_m \cup \partial B \\ &\subseteq \bigcup_{m=1}^{\infty} B_m^\circ \cup \bigcup_{k=1}^{\infty} O_k \end{aligned}$$

eine offene Überdeckung der kompakten Menge $\overline{B}$. Nach Satz 9.66(3) kann man von diesen abzählbar vielen offenen Mengen endlich viele auswählen, die immer noch $\overline{B}$ überdecken. Auf Grund der Inklusion $B_m^\circ \subseteq B_{m+1}^\circ$ für alle $m \in \mathbb{N}$ braucht man dazu nur ein $B_m$. Somit gilt für geeignete $m, K \geq 1$

$$\overline{B} \subseteq B_m^\circ \cup \bigcup_{k=1}^{K} O_k.$$

Daraus folgt

$$\operatorname{vol}(B) = \operatorname{vol}(\overline{B}) \leq \operatorname{vol}(B_m^\circ) + \sum_{k=1}^{K} \operatorname{vol}(O_k) < \operatorname{vol}(B_m^\circ) + \varepsilon = \operatorname{vol}(B_m) + \varepsilon.$$

Da $\varepsilon > 0$ beliebig war, folgt der erste Teil der Proposition.

Sei nun $f : B \to \mathbb{R}$ Riemann-integrierbar und sei $M = \sup |f(B)|$. Dann gilt

$$\left| \int_{B_m} f \operatorname{dvol} - \int_B f \operatorname{dvol} \right| \leq \int_{B \setminus B_m} |f| \operatorname{dvol} \leq M \operatorname{vol}(B \setminus B_m) \to 0$$

für $m \to \infty$ wie gewünscht. □

**Übung 12.72.** *Verwenden Sie Proposition 12.71, um das Volumen des Dreiecks*

$$\left\{(x, y) \in [0, 1]^2 \mid x + y \leq 1\right\}$$

*mittels einer Ausschöpfung, die aus einer Folgen von endlichen Vereinigungen von Quadern besteht, zu berechnen.*

**Satz 12.73** (Uneigentliche Riemann-Integrale und nicht-negative Funktionen)**.** *Sei $B \subseteq \mathbb{R}^n$ eine Teilmenge, sei $(B_m)_m$ eine Ausschöpfung von $B$ und sei $f : B \to \mathbb{R}$ eine Funktion so dass $f|_{B_m}$ für jedes $m \in \mathbb{N}$ Riemann-integrierbar ist. Angenommen $f$ ist nicht-negativ und der Grenzwert $I = \lim_{m\to\infty} \int_{B_m} f \operatorname{dvol}$ existiert. Dann existiert der Grenzwert $\lim_{\ell\to\infty} \int_{A_\ell} f \operatorname{dvol}$ für jede weitere Ausschöpfung $(A_\ell)_\ell$ mit der Eigenschaft, dass $f|_{A_\ell}$ für alle $\ell \in \mathbb{N}$ Riemann-integrierbar ist, und ist gleich $I$. Insbesondere ist $f$ auf $B$ uneigentlich Riemann-integrierbar und das uneigentliche Riemann-Integral ist gleich $I$.*

Dies beantwortet die in Beispiel 12.68 gestellte Frage.

*Beweis.* Sei $(A_\ell)_\ell$ eine zweite Ausschöpfung, für die $f|_{A_\ell}$ für alle $\ell \geq 1$ auf $A_\ell$ Riemann-integrierbar ist. Dann gilt auf Grund von $f \geq 0$ und wegen Proposition 12.71, dass

$$\int_{A_\ell} f \operatorname{dvol} = \lim_{m\to\infty} \int_{A_\ell \cap B_m} f \operatorname{dvol} \leq \lim_{m\to\infty} \int_{B_m} f \operatorname{dvol},$$

für alle $\ell \geq 1$, da $A_\ell \cap B_m$ eine Ausschöpfung von $A_\ell$ darstellt. Da die Folge $\int_{A_\ell} f \,\mathrm{dvol}$ monoton wachsend ist und nach obigem auch beschränkt ist, konvergiert sie und es folgt

$$\lim_{\ell\to\infty} \int_{A_\ell} f \,\mathrm{dvol} \leq \lim_{m\to\infty} \int_{B_m} f \,\mathrm{dvol}.$$

Durch Vertauschen der beiden Ausschöpfung folgt dann auch Gleichheit. □

*Bemerkung.* Wir wollen noch bemerken, dass das mehrdimensionale uneigentliche Integral nicht in allen Fällen eine Verallgemeinerung des eindimensionalen uneigentlichen Integrals darstellt. Denn das uneigentliche Integral $\int_0^\infty \frac{\sin(x)}{x}\,\mathrm{d}x$ konvergiert im Sinne der Definitionen von Abschnitt 8.4.1, wobei man allerdings das Leibniz-Kriterium (Proposition 6.25) für die Reihe $\sum_{n=0}^\infty \int_{n\pi}^{(n+1)\pi} \frac{\sin(x)}{x}\,\mathrm{d}x$ anwendet, um die Konvergenz zu zeigen. In der Tat konvergiert diese Reihe aber nicht absolut und mittels dem Riemannschen Umordnungssatz (Satz 6.21) könnten wir jetzt eine Ausschöpfung $B_1 \subseteq B_2 \subseteq \ldots$ von $[0,\infty) \subseteq \mathbb{R}$ finden, für die $\int_{B_m} \frac{\sin(x)}{x}\,\mathrm{dvol}(x)$ gegen einen anderen Wert strebt oder divergiert.

Geht man etwas weiter, so kann man sogar zeigen, dass für ein konvergentes mehrdimensionales uneigentliches Integral die absolute Konvergenz des Integrals impliziert wird, was eben bei dem eindimensionalen uneigentlichen Integral nicht der Fall war.

Der moralische Grund für diesen Unterschied ist folgender Unterschied zwischen $\mathbb{R}$ und $\mathbb{R}^n$ für $n \geq 2$: In $\mathbb{R}$ ist es sehr natürlich über Intervalle zu integrieren, und daher haben wir bedingt konvergente Integrale erlaubt. Doch für teilweise sehr kompliziert aussehende Teilmengen im $\mathbb{R}^n$ gibt es keine kanonischen Ausschöpfungen; deswegen wollen wir, dass der Wert des gesuchten Integrals nicht von der gewählten Ausschöpfung abhängen sollte. Auf Grund von Umordnungsphänomenen erzwingt dies die absolute Konvergenz[1]. In diesem Sinne ist das mehrdimensionale uneigentliche Integral bereits dem Lebesgue-Integral etwas näher, da auch dort bereits in der Definition absolute Integrierbarkeit gefordert wird (siehe die Mass und Integral Vorlesung des vierten Semesters des Mathematikstudiums).

VO: 02.05.18

### 12.7.2 Substitution für uneigentliche Riemann-Integrale

Wir möchten nun Substitution auf nicht-zwingend beschränkten (und insbesondere nicht-zwingend Jordan-messbaren) offenen Teilmengen von $\mathbb{R}^n$ betrachten. Dazu bemerken wir, dass offene Teilmengen tatsächlich ausgeschöpft werden können und wir somit von uneigentlich Riemann-integrierbaren Funktionen auf diesen sprechen können.

**Übung 12.74.** *Zeigen Sie, dass jede offene Teilmenge $U \subseteq \mathbb{R}^n$ eine Ausschöpfung mit Jordan-messbaren Mengen besitzt.*

**Theorem 12.75** (Substitution)**.** *Seien $X, Y \subseteq \mathbb{R}^n$ offene Teilmengen und sei $\Phi : X \to Y$ ein Diffeomorphismus. Weiter sei $f : Y \to \mathbb{R}$ eine Funktion, so dass $f$ und $|f|$ auf $Y$ uneigentlich*

[1] Die Idee hierbei ist den Integrationsbereich $X$ als $X_+ \sqcup X_-$ zu schreiben, so dass $f$ auf $X_+$ nichtnegativ und auf $X_-$ nichtpositiv ist, diese beiden Mengen auszuschöpfen, und dann den Riemannschen Umordnungssatz zu verwenden. Allerdings ist die formal richtige Ausführung dieses Arguments für Riemann-integrierbare Funktionen auf Grund der benötigten Jordan-Messbarkeit der Ausschöpfung eher kompliziert.

*Riemann-integrierbar sind. Dann ist die Funktion* $(f \circ \Phi)|\det(\mathrm{D}\Phi)|$ *uneigentlich Riemann-integrierbar auf* $X$ *und es gilt*

$$\int_Y f \,\mathrm{dvol} = \int_X (f \circ \Phi)|\det(\mathrm{D}\Phi)| \,\mathrm{dvol}.$$

Auf Grund von Proposition 12.71 ist obiges Theorem auch für eigentlich Riemann-integrierbare Funktionen $f$ auf $Y$ gültig, wobei es aber vorkommen kann, dass das neue Integral auf $X$ bloss uneigentlich Riemann-integrierbar ist.

*Beweis.* Wir nehmen zuerst an, dass $f$ nicht-negativ ist, und wählen eine Ausschöpfung $(B_m)_m$ von $Y$, so dass $f|_{B_m}$ für alle $m \in \mathbb{N}$ Riemann-integrierbar ist und der Grenzwert

$$\int_Y f \,\mathrm{dvol} = \lim_{m\to\infty} \int_{B_m} f \,\mathrm{dvol}$$

existiert. Wir wollen die Ausschöpfung $(B_m)_m$ durch eine andere Ausschöpfung ersetzen, so dass wir Satz 12.56 anwenden können.

Um kompakten Träger in $Y$ zu erreichen, ersetzen wir $B_m$ durch $Y_m = B_m \cap K_m$ für alle $m \in \mathbb{N}$. Hierbei ist $K_m \subseteq Y$ für jedes $m \in \mathbb{N}$ Jordan-messbar und kompakt und es gilt $Y = \bigcup_{m=1}^\infty K_m$. Wir konstruieren solche Mengen $K_1, K_2, \ldots$ unten. Damit ist $Y = \bigcup_{m=1}^\infty Y_m$ und $(Y_m)_m$ ist wieder eine Ausschöpfung. Nun berechnet man mit dem Satz über uneigentliche Integrale (Satz 12.73) und der Substitutionsregel mit kompakten Träger (Satz 12.56)

$$\begin{aligned}
\int_Y f \,\mathrm{dvol} &= \lim_{m\to\infty} \int_{Y_m} f \,\mathrm{dvol} \\
&= \lim_{m\to\infty} \int_Y \mathbb{1}_{Y_m} f \,\mathrm{dvol} \\
&= \lim_{m\to\infty} \int_X (\mathbb{1}_{Y_m} f) \circ \Phi |\det(\mathrm{D}\Phi)| \,\mathrm{dvol} \\
&= \lim_{m\to\infty} \int_X \mathbb{1}_{\Phi^{-1}(Y_m)} f \circ \Phi |\det(\mathrm{D}\Phi)| \,\mathrm{dvol} \\
&= \lim_{m\to\infty} \int_{\Phi^{-1}(Y_m)} f \circ \Phi |\det(\mathrm{D}\Phi)| \,\mathrm{dvol}.
\end{aligned}$$

Da $(\Phi^{-1}(Y_m))_m$ eine Ausschöpfung von $X$ ist und der obige Grenzwert existiert, ist die Funktion $f \circ \Phi|\det(\mathrm{D}\Phi)|$ uneigentlich Riemann-integrierbar und es gilt

$$\begin{aligned}
\int_X f \circ \Phi |\det(\mathrm{D}\Phi)| \,\mathrm{dvol} &= \lim_{m\to\infty} \int_{\Phi^{-1}(Y_m)} f \circ \Phi |\det(\mathrm{D}\Phi)| \,\mathrm{dvol} \\
&= \int_Y f \,\mathrm{dvol}.
\end{aligned}$$

Wir konstruieren nun Mengen $K_1, K_2, \ldots$ mit den oben verlangten Eigenschaften und verwenden dabei eine vergleichbare Strategie zum Beweis von Lemma 12.62. Genauer betrachten

wir die stetige Funktion

$$y \mapsto \mathrm{d}_\infty(y, \mathbb{R}^n \setminus Y) = \inf\{\|y - v\|_\infty \mid v \in \mathbb{R}^n \setminus Y\}$$

und zu $m \in \mathbb{N}$ die Menge

$$C_m = \left\{y \in [-m, m]^n \mid \mathrm{d}_\infty(y, \mathbb{R}^n \setminus Y) \geq \tfrac{2}{m}\right\}.$$

Wir wählen nun eine Zerlegung $\mathfrak{Z}$ von $[-m, m]^n$ mit Maschenweite kleiner als $\frac{1}{m}$ und definieren $D_m$ als die Vereinigung aller Quader $\overline{Q_{\boldsymbol{\alpha}}}$ für $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$, die $C_m$ schneiden. Insbesondere ist $D_m$ als eine endliche Vereinigung von kompakten Quadern wiederum kompakt und auf Grund der Definitionen gilt $C_m \subseteq D_m \subseteq Y$. (Wieso?) Auch ist die Vereinigung all dieser $D_m$ gleich $Y$, da die Vereinigung aller $C_m$ gleich $Y$ ist. Nun definieren wir $K_m = \bigcup_{k=1}^m D_k$ und erhalten die gewünschte Ausschöpfung von $Y$.

Falls $f$ auch negative Werte annimmt, so wenden wir obiges auf $f^+ = \frac{1}{2}(f + |f|)$ und $f^- = \frac{1}{2}(|f| - f)$ an und erhalten den Satz. □

**Übung 12.76.** *Sei $A \in \mathrm{Mat}_{n,n}(\mathbb{R})$ symmetrisch und positiv definit. Zeigen Sie*

$$\int_{\mathbb{R}^n} \mathrm{e}^{-\langle Ax, x\rangle} \,\mathrm{dvol}(x) = \frac{\pi^{n/2}}{\sqrt{\det A}}.$$

### 12.7.3 Eine Anwendung: das Basler Problem*

Wir möchten in diesem Teilabschnitt als Anwendung der Theorie dieses Kapitels die Gleichung

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$$

zeigen, der Wert dieser Reihe wird auch als $\zeta(2)$ bezeichnet. Dabei konvergiert die Reihe links absolut, wie in Beispiel 6.14 bewiesen wurde. Die Herangehensweise an dieses Problem (auch **Basler Problem** genannt), die wir hier wählen, ist von Apostol [Apo83]; im Kapitel B des Appendix findet sich eine weitere, die die Theorie der Fourierreihen ausnützt.

Wir möchten auf zwei verschiedene Art und Weisen das Integral

$$\int_0^1 \int_0^1 \frac{1}{1 - xy} \,\mathrm{d}x \,\mathrm{d}y \tag{12.23}$$

berechnen. Dazu werden wir teilweise etwas informell vorgehen und den Rest als Übung überlassen.

Unter Verwendung der geometrischen Reihe beobachten wir zuerst, dass

$$\begin{aligned}\int_0^1\int_0^1 \frac{1}{1-xy}\,\mathrm{d}x\,\mathrm{d}y &= \int_0^1\int_0^1 \sum_{k=0}^{\infty}(xy)^k\,\mathrm{d}x\,\mathrm{d}y = \sum_{k=0}^{\infty}\int_0^1\int_0^1 x^k y^k\,\mathrm{d}x\,\mathrm{d}y \\ &= \sum_{k=0}^{\infty}\int_0^1 \tfrac{1}{k+1}y^k\,\mathrm{d}y = \sum_{k=0}^{\infty}\frac{1}{(k+1)^2} = \zeta(2)\end{aligned} \tag{12.24}$$

**Übung 12.77** (Das Doppelintegral zum Basler Problem)**.** *Wir möchten in dieser Übung zeigen, dass die Funktion $f : (x,y) \mapsto \frac{1}{1-xy}$ auf $[0,1)^2$ uneigentlich Riemann-integrierbar ist und dass die Manipulationen in Gleichung* (12.24) *tatsächlich erlaubt sind. Dabei gilt es zu beachten, dass die Funktion $f$ für $(x,y) \to (1,1)$ gegen Unendlich geht und insbesondere nicht Riemann-integrierbar ist.*

*(i) Sei $s \in (0,1)$ und $B_s = [0,1) \times [0,s)$. Zeigen Sie, dass $f|_{B_s}$ Riemann-integrierbar ist und dass*

$$\int_{B_s} f\,\mathrm{dvol} = \int_0^s\int_0^1 \frac{1}{1-xy}\,\mathrm{d}x\,\mathrm{d}y = \sum_{k=0}^{\infty}\frac{s^{k+1}}{(k+1)^2}$$

*gilt.*

*(ii) Zeigen Sie, dass der Grenzwert $\lim_{s\to 1}\int_{B_s} f\,\mathrm{dvol}$ existiert und gleich $\zeta(2)$ ist.*

*(iii) Schliessen Sie, dass $f$ uneigentlich Riemann-integrierbar ist und dass $\int_0^1\int_0^1 \frac{1}{1-xy}\,\mathrm{d}x\,\mathrm{d}y$ gleich $\zeta(2)$ ist.*

Nun möchten wir das Integral in Gleichung (12.23) mit einer (linearen) Substitution berechnen. Dazu betrachten wir neue Intergrationsvariablen $u, v$ gegeben durch die Rotation

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}.$$

Damit ist der neue Integrationsbereich (in den $u,v$-Koordinaten) also ein um 45 Grad gedrehtes und leicht gestrecktes Quadrat mit Ecken bei $(0,0)^t$, $(1,1)^t$, $(2,0)^t$ und $(1,-1)^t$, wie man leicht durch Einsetzen überprüfen kann. Auch ist per Definition

$$\frac{1}{1-xy} = \frac{1}{1-\frac{1}{4}(u-v)(u+v)} = \frac{4}{4-u^2+v^2}$$

Nach Theorem 12.75 und dem Satz von Fubini (Theorem 12.39) gilt nun

$$\begin{aligned}\int_0^1\int_0^1 \frac{1}{1-xy}\,\mathrm{d}x\,\mathrm{d}y &= \int_0^1\int_{-u}^{u}\frac{4}{4-u^2+v^2}\frac{1}{2}\,\mathrm{d}v\,\mathrm{d}u + \int_1^2\int_{-(2-u)}^{2-u}\frac{4}{4-u^2+v^2}\frac{1}{2}\,\mathrm{d}v\,\mathrm{d}u \\ &= \underbrace{4\int_0^1\int_0^u \frac{1}{4-u^2+v^2}\,\mathrm{d}v\,\mathrm{d}u}_{=I_1} + \underbrace{4\int_1^2\int_0^{2-u}\frac{1}{4-u^2+v^2}\,\mathrm{d}v\,\mathrm{d}u}_{=I_2}\end{aligned}$$

Eine direkte Rechnung (siehe Beispiel 8.18(iii)) zeigt nun

$$I_1 = 4\int_0^1 \frac{1}{\sqrt{4-u^2}} \arctan\left(\frac{u}{\sqrt{4-u^2}}\right) \mathrm{d}u = 4\int_0^{\frac{\pi}{6}} \theta \,\mathrm{d}\theta = 2\left(\tfrac{\pi}{6}\right)^2$$

wobei $u = 2\sin(\theta)$ gesetzt wurde. Ähnlich berechnet man

$$I_2 = 4\int_1^2 \frac{1}{\sqrt{4-u^2}} \arctan\left(\frac{2-u}{\sqrt{4-u^2}}\right) \mathrm{d}u = 8\int_0^{\frac{\pi}{6}} \phi \,\mathrm{d}\phi = 4\left(\tfrac{\pi}{6}\right)^2,$$

wobei $u = 2\cos(2\phi)$ gesetzt wurde.

**Übung 12.78** (Das Integral nach Substitution)**.** *Führen Sie obige Rechnungen vollständig durch, begründen Sie die Substitutionsschritte jeweils auch formal, und argumentieren Sie für das Integral $I_2$ und die Grenze $u = 2$ sorgfältiger.*

Unter dem Strich erhalten wir nun

$$\zeta(2) = \int_0^1 \int_0^1 \frac{1}{1-xy} \,\mathrm{d}x \,\mathrm{d}y = I_1 + I_2 = 2\left(\frac{\pi}{6}\right)^2 + 4\left(\frac{\pi}{6}\right)^2 = \frac{\pi^2}{6}$$

wie behauptet.

**Übung 12.79** (Alternierendes Basler Problem)**.** *Berechnen Sie den Wert der alternierenden Reihe $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2}$.*

VO: 02.04.18

## 12.8 Weitere Lernmaterialien

### 12.8.1 Verwendung des Kapitels

Die mehrdimensionale Integralrechnung ist für viele Anwendungen innerhalb und ausserhalb der Mathematik notwendig, wobei wir für explizite Berechnungen den Satz von Fubini (Theorem 12.39) gemeinsam mit den Methoden zur Berechnung von eindimensionalen Integralen benötigen. Ebenso ist die Substitutionsregel (Satz 12.56 oder Theorem 12.75) manchmal für die Berechnung von Integralen hilfreich, zum Beispiel wenn der Integrationsbereich ursprünglich kompliziert definiert wurde, doch mit Hilfe eines Diffeomorphismus „gerade gebogen werden kann“. Des Weiteren ist die Substitutionsregel für die weitere Theorie notwendig, damit wir im nächsten Kapitel zeigen können, dass ein Oberflächenintegral von der Wahl der Parametrisierung der Fläche (in anderen Worten dem innerhalb der Fläche verwendeten Koordinatensystem) nicht abhängt. In vielen auch für die Praxis relevanten Beispielen wird das Integral einer unbeschränkten Funktion oder über unbeschränkte Teilmengen von $\mathbb{R}^n$ betrachtet, weswegen wir eben den Begriff des uneigentlichen mehrdimensionalen Integrals eingeführt haben.

### 12.8.2 Übungen

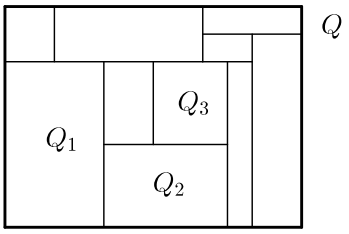
**Übung** (Nullmengen und Vielfachheit). *Sei $N \subseteq \mathbb{R}^n$ eine Teilmenge. Zeigen Sie, dass $N$ genau dann eine Nullmenge ist, wenn eine Folge $(Q_k)_k$ von Quadern existiert, so dass jedes $x \in N$ in abzählbar vielen $Q_k$'s liegt und so dass $\sum_{k=1}^{\infty} \mathrm{vol}(Q_k) < \infty$.*

**Übung** (Nullmengen und Lipschitz-Stetigkeit). *Sei $U \subseteq \mathbb{R}^n$ offen und sei $\Phi : U \to \mathbb{R}^n$ eine Lipschitz-Abbildung. Zeigen Sie, dass das Bild einer Lebesgue-Nullmenge unter $\Phi$ wieder eine Lebesgue-Nullmenge ist.*

**Übung** (Nullmengen und Graphen). *Sei $n \geq 1$, $k \in \{1, \ldots, n-1\}$, $Q \subseteq \mathbb{R}^{n-k}$ ein abgeschlossener Quader und $f$ eine stetige Funktion auf $Q$ mit Werten in $\mathbb{R}^k$. Zeigen Sie, dass $\mathrm{graph}(f) \subseteq \mathbb{R}^n$ eine Nullmenge ist.*

**Übung** (Nullmengen und Teilmannigfaltigkeiten). *Verwenden Sie vorherige Übung, um zu zeigen, dass jede $k$-dimensionale Teilmannigfaltigkeit von $\mathbb{R}^n$ für $k < n$ eine Nullmenge von $\mathbb{R}^n$ ist.*

**Übung** (Ganzzahlige Kantenlängen). *Gegeben sei ein Rechteck $Q \subseteq \mathbb{R}^2$, das eine endliche Überdeckung mit Rechtecken $Q_1, \ldots, Q_n \subseteq Q$ besitzt, die sich jeweils höchstens an den Kanten schneiden.*



*Angenommen jedes der Rechtecke $Q_1, \ldots, Q_n$ hat mindestens eine Kante mit ganzzahliger Länge. Zeigen Sie, dass dann auch $Q$ mindestens eine Kante mit ganzzahliger Länge hat.*

**Übung** (Volumen des Standardsimplex)**.** *Sei $n \in \mathbb{N}$. In dieser Übungen möchten wir das Volumen des $n$-dimensionalen Simplex*

$$\triangle_n = \left\{ x \in \mathbb{R}^n \mid x_1, \ldots, x_n \in \mathbb{R}_{\geq 0} \text{ und } \sum_{k=1}^{n} x_k \leq 1 \right\}$$

*berechnen.*

*(i) Zeigen Sie, dass $\triangle_n$ das gleiche Volumen hat wie die Teilmenge*

$$A_e = \{ y \in \mathbb{R}^n \mid 0 \leq y_1 \leq y_2 \leq \ldots \leq y_n \leq 1 \}.$$

*(ii) Zu $\sigma \in S_n$ definieren wir*

$$A_\sigma = \left\{ y \in \mathbb{R}^n \mid 0 \leq y_{\sigma(1)} \leq y_{\sigma(2)} \leq \ldots \leq y_{\sigma(n)} \leq 1 \right\}.$$

*Verifizieren Sie, dass $\mathrm{vol}(A_\sigma) = \mathrm{vol}(A_e)$ gilt für alle $\sigma \in S_n$.*

*(iii) Begründen Sie, wieso $[0,1]^n = \bigcup_{\sigma \in S_n} A_\sigma$ ist und verwenden Sie dies, um zu zeigen, dass das Volumen von $\triangle_n$ gleich $\frac{1}{n!}$ ist.*

**Übung.** *Berechnen Sie für jedes $n \in \mathbb{N}$ das Volumen $\mathrm{vol}(B_1^{\mathbb{R}^n}(0))$ des Einheitsballs.*

**Übung** (Dreiecksungleichung für uneigentliche Integrale)**.** *Sei $(B_m)_m$ eine Ausschöpfung einer Teilmenge $B$. Seien $f, g : B \to \mathbb{R}$, so dass $|f| \leq g$ gilt und die Funktionen $f|_{B_m}$ und $g|_{B_m}$ für alle $m \in \mathbb{N}$ Riemann-integrierbar sind. Angenommen $g$ ist uneigentlich Riemann-integrierbar. Zeigen Sie, dass dann auch $f$ und $|f|$ auf $B$ uneigentlich Riemann-integrierbar sind und dass*

$$\left| \int_B f \,\mathrm{d}\,\mathrm{vol} \right| \leq \int_B |f| \,\mathrm{d}\,\mathrm{vol} \leq \int_B g \,\mathrm{d}\,\mathrm{vol}$$

*gilt.*

**Übung** (Betafunktion)**.** *Für $x, y > 0$ ist die Betafunktion durch*

$$B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1} \,\mathrm{d}t$$

*definiert. Zeigen Sie, dass*

$$B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$

*gilt.*

### 12.8.3 Lernkarten

Sie können wiederum die Lernkarten oder den Graphen für Ihre Wiederholung der Themen des Kapitels verwenden.

# Kapitel 13

# Mehrdimensionale Integralsätze

Wir wollen in diesem Kapitel Integralsätze für Vektorfelder im $\mathbb{R}^2$ und im $\mathbb{R}^3$ behandeln, welche mehrdimensionale Verallgemeinerungen des Fundamentalsatzes der Integral- und Differentialrechnung (Theorem 8.2) darstellen. Die Einschränkung der Dimension $n$ auf die Fälle $n = 2$ und $n = 3$ ist dabei aus mathematischer Sicht nicht notwendig, macht aber Sinn, da diese Fälle die geometrisch anschaulichsten sind und für die meisten Anwendungen ausreichen. Des Weiteren erfordert der allgemeine Satz von Stokes für Teilmannigfaltigkeiten von $\mathbb{R}^n$ oder sogar für abstrakte Mannigfaltigkeiten zusätzliche Methoden und Begriffe aus der Analysis, der Differentialgeometrie und der Linearen Algebra wie zum Beispiel „glatte Partitionen der Eins“ und „Differentialformen“ (im Wesentlichen „alternierende Multilinearformen“), welche wir hier nur am Rande und nicht in der nötigen Tiefe besprechen können.

Die hier bewiesenen Sätze sind auch weitere Eckpfeiler der Vektoranalysis, die wir eigentlich schon in Abschnitt 10.7 begonnen haben.

## 13.1 Der Divergenzsatz und der Satz von Green in der Ebene

Wir betrachten in diesem Abschnitt ein stetig differenzierbares Vektorfeld,

$$f = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} : U \to \mathbb{R}^2,$$

welches auf einer offenen Menge $U \subseteq \mathbb{R}^2$ definiert ist. Wie schon zuvor, werden wir uns den Vektor $f(p)$ zu $p \in U$ als einen von $p$ ausgehenden Vektor vorstellen. Anders ausgedrückt werden wir die Abbildung $f$ auch als die Abbildung $p \in U \mapsto (p, f(p)) \in \mathrm{T}U$ auffassen (siehe Abschnitt 10.2.2).

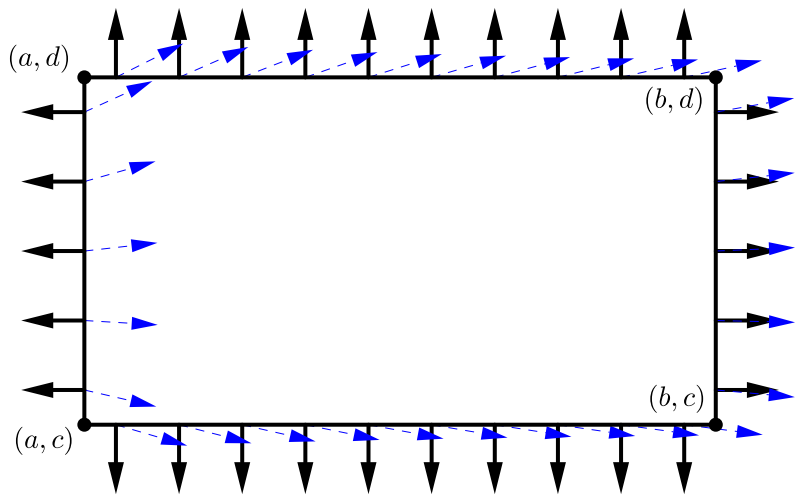
Ist $B \subseteq U$ eine kompakte, Jordan-messbare Teilmenge, deren Rand geeignet durch einen Weg beschrieben wird, so ist das Ziel der Integralsätze dieses Abschnitts, Integrale von $f$ entlang des Randes von $B$ mit einem Integral über $B$ in Verbindung zu bringen. In Anwendungen besteht dabei natürlich meist die Hoffnung, dass sich eines der beiden einfacher berechnen lässt als das andere.

### 13.1.1 Der Divergenzsatz für Rechtecke und Bereiche unter einem Graphen

Wir beginnen mit dem Fall eines achsenparallelen Rechtecks $B = [a,b] \times [c,d] \subseteq U$. In diesem Fall definieren wir das Integral $\int_{\partial B} f \cdot \mathrm{d}\mathbf{n}$ von $f$ bezüglich der (fast überall definierten) Aussennormalen $\mathbf{n}$ auf $\partial B$ durch

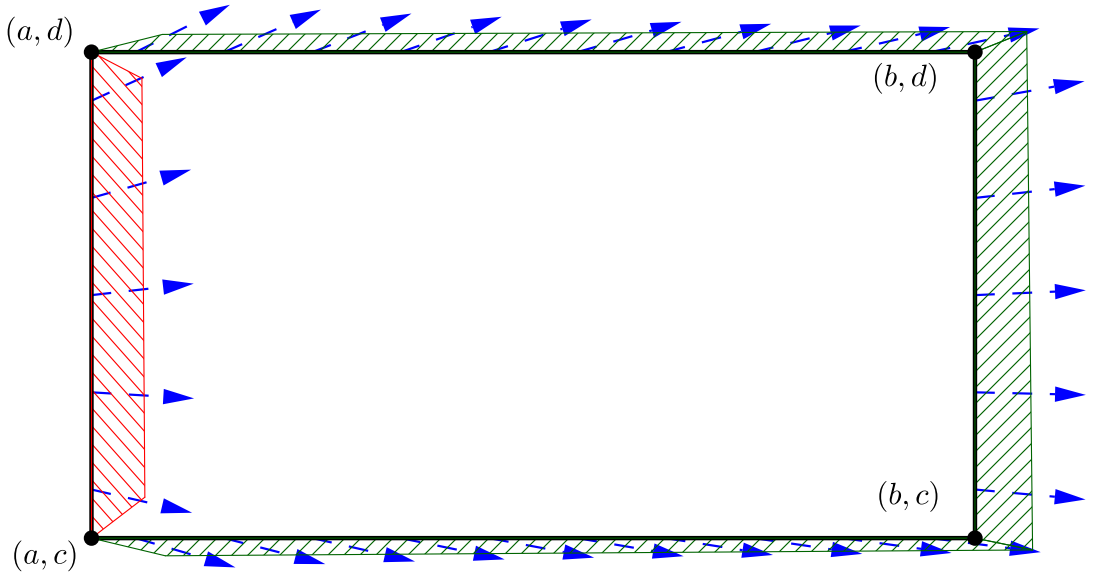
$$\int_{\partial B} f \cdot \mathrm{d}\mathbf{n} = \underbrace{-\int_a^b f_2(x,c)\,\mathrm{d}x}_{\text{unten}} \underbrace{+\int_c^d f_1(b,y)\,\mathrm{d}y}_{\text{rechts}} \underbrace{+\int_a^b f_2(x,d)\,\mathrm{d}x}_{\text{oben}} \underbrace{-\int_c^d f_1(a,y)\,\mathrm{d}y}_{\text{links}}.$$

Das sogenannte **Flussintegral** $\int_{\partial B} f \cdot \mathrm{d}\mathbf{n}$ ist also gegeben als das Integral der Skalarprodukte von $f$ mit der jeweils normalen, nach aussen zeigenden Richtung. Wir illustrieren dies an folgendem Bild.




Figur 13.1: Die dicken schwarzen Pfeile stellen die zum Rand normalen Richtungen dar und die gestrichelten blauen Pfeile stellen $f$ dar. Wir betrachten zum Beispiel die untere Kante des obigen Rechtecks, womit die nach aussen zeigende Normale durch $-(0,1)^t$ gegeben ist. Das entsprechende Integral $\int_a^b \langle f(x,c), -(0,1)^t \rangle \,\mathrm{d}x$ ist gerade $-\int_a^b f_2(x,c)\,\mathrm{d}x$, wie in der Definition des Integrals $\int_{\partial B} f \cdot \mathrm{d}\mathbf{n}$.

Wir wollen uns das Vektorfeld als eine Strömungsgeschwindigkeit eines Mediums (zum Beispiel Luft oder Wasser) vorstellen. Das Flussintegral $\int_{\partial B} f \cdot \mathrm{d}\mathbf{n}$ misst in diesem Fall das Gesamtvolumen des Mediums, welches pro Zeiteinheit aus dem Rechteck strömt (oder bei negativem Integral in den Rechteck strömt), siehe auch folgende Skizze.




Figur 13.2: Wir sehen hier die Strömung aus $B$ (unten, rechts, oben – in Grün) beziehungsweise die Strömung in $B$ (links – in Rot) dargestellt.

Wir möchten diese Quantität nun auf eine andere Art und Weise beschreiben.

**Proposition 13.1** (Divergenzsatz auf Rechtecke)**.** *Sei $U \subseteq \mathbb{R}^2$ eine offene Menge im $\mathbb{R}^2$ und sei $f : (f_1, f_2)^t : U \to \mathbb{R}^2$ ein stetig differenzierbares Vektorfeld. Dann existiert für alle $p \in U$ die sogenannte **Divergenz** oder **Quellenstärke***

$$(\operatorname{div}(f))(p) = \lim_{h \to 0} \tfrac{1}{4h^2} \int_{\partial(p+[-h,h]^2)} f \cdot \mathrm{d}\mathbf{n}$$

*und ist durch*

$$\operatorname{div}(f) = \partial_1 f_1 + \partial_2 f_2$$

*gegeben. Des Weiteren gilt für jedes abgeschlossene Rechteck $B \subseteq U$*

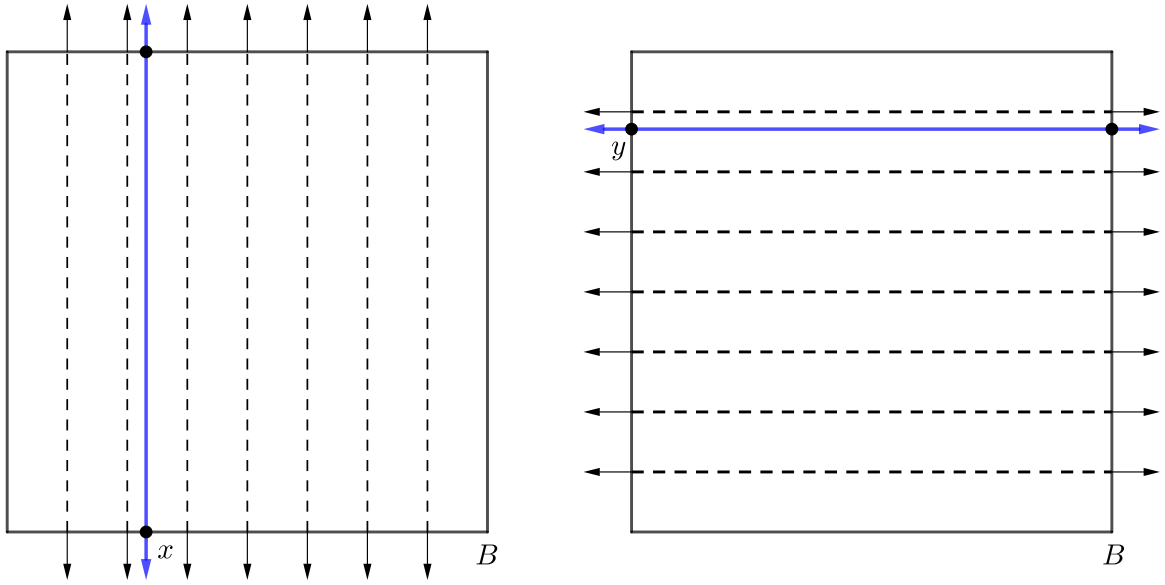
$$\int_B \operatorname{div}(f) \, \mathrm{dvol} = \int_{\partial B} f \cdot \mathrm{d}\mathbf{n}. \tag{13.1}$$

Wir interpretieren die Divergenz $\operatorname{div}(f)$ als eine Messgrösse, die bei jedem Punkt $p \in U$ angibt, wie sehr sich in $p$ die Dichte des Mediums pro Zeiteinheit ändert. Anders formuliert lässt sich damit beurteilen, ob in $p$ eine Quelle (mit $\operatorname{div}(f)(p) > 0$) oder eine Senke (mit $\operatorname{div}(f)(p) < 0$) der Strömung vorhanden ist. Das Vektorfeld $f$ heisst **divergenzfrei**, falls $\operatorname{div}(f) = 0$ gilt. Bei inkompressiblen Flüssigkeiten ist dies eine physikalisch notwendige Bedingung an eine Strömung der Flüssigkeit.

*Beweis.* Sei $B = [a,b] \times [c,d] \subseteq U$ für reelle Zahlen $a < b$ und $c < d$ ein Rechteck. Dann gilt

$$\begin{aligned}\int_{\partial B} f \cdot \mathrm{d}\mathbf{n} &= -\int_a^b f_2(x,c)\,\mathrm{d}x + \int_c^d f_1(b,y)\,\mathrm{d}y + \int_a^b f_2(x,d)\,\mathrm{d}x - \int_c^d f_1(a,y)\,\mathrm{d}y \\ &= \int_a^b (f_2(x,d) - f_2(x,c))\,\mathrm{d}x + \int_c^d (f_1(b,y) - f_1(a,y))\,\mathrm{d}y \\ &= \int_a^b \int_c^d \partial_2 f_2(x,y)\,\mathrm{d}y\,\mathrm{d}x + \int_c^d \int_a^b \partial_1 f_1(x,y)\,\mathrm{d}x\,\mathrm{d}y\end{aligned}$$

nach dem Fundamentalsatz der Integral- und Differentialrechnung (genauer Korollar 8.5).




Figur 13.3: Links die Integration von $\partial_2 f_2$ über Streifen für feste $x$ nach $y$ und rechts die Integration von $\partial_1 f_1$ über Streifen für feste $y$ nach $x$.

Mit dem Satz von Fubini (Theorem 12.39) erhalten wir also

$$\int_{\partial B} f \cdot \mathrm{d}\mathbf{n} = \int_B (\partial_1 f_1 + \partial_2 f_2)\,\mathrm{dvol}.$$

VO: 04.05.18

Wir verwenden dies nun, um die Existenz des Grenzwerts in der Definition der Divergenz und die Formel $\operatorname{div}(f) = \partial_1 f_1 + \partial_2 f_2$ zu beweisen. Sei $p \in U$ und $h > 0$ so dass

$$B_h = p + [-h,h]^2 \subseteq U.$$

Dann gilt (da in obigem $B$ ein beliebiges Rechteck in $U$ war)

$$\int_{\partial B_h} f \cdot \mathrm{d}\mathbf{n} = \int_{B_h} (\partial_1 f_1 + \partial_2 f_2)\,\mathrm{dvol}$$

und damit

$$\begin{aligned}
&\left|\tfrac{1}{4h^2}\int_{\partial B_h} f\cdot \mathrm{d}\mathbf{n} - (\partial_1 f_1+\partial_2 f_2)(p)\right| \\
&\qquad = \left|\tfrac{1}{4h^2}\int_{B_h} (\partial_1 f_1+\partial_2 f_2)\,\mathrm{dvol} - (\partial_1 f_1+\partial_2 f_2)(p)\right| \\
&\qquad = \left|\tfrac{1}{4h^2}\int_{B_h} \Big((\partial_1 f_1+\partial_2 f_2) - (\partial_1 f_1+\partial_2 f_2)(p)\Big)\,\mathrm{dvol}\right| \\
&\qquad \le \tfrac{1}{4h^2}\int_{[-h,h]^2} \Big|(\partial_1 f_1+\partial_2 f_2)(p+v) - (\partial_1 f_1+\partial_2 f_2)(p)\Big|\,\mathrm{dvol}(v) \to 0
\end{aligned}$$

für $h \to 0$ auf Grund der Stetigkeit von $\partial_1 f_1 + \partial_2 f_2$ bei $p$. Also existiert der Grenzwert $\lim_{h\to 0} \frac{1}{4h^2}\int_{\partial B_h} f\cdot \mathrm{d}\mathbf{n}$ und ist gleich $(\partial_1 f_1+\partial_2 f_2)(p)$. Dies beweist die Proposition. □

Der Divergenzsatz gilt auch für allgemeinere Mengen als nur Rechtecke. Die allgemeinste Version, die wir (im zweidimensionalen Falle) behandeln wollen, wird in Theorem 13.14 enthalten sein. Als nächsten Schritt in diese Richtung wollen wir Bereiche unter Graphen betrachten.

**Proposition 13.2** (Divergenzsatz für Bereiche unter Graphen)**.** *Sei $U \subseteq \mathbb{R}^2$ offen und $f$ ein stetig differenzierbares Vektorfeld auf $U$. Seien $a < b$ und $c < d$ reelle Zahlen, so dass $[a,b]\times[c,d] \subseteq U$ ist, und sei $\varphi : [a,b] \to [c,d]$ stetig und stückweise stetig differenzierbar. Für den Bereich*

$$B = \{(x,y) \in U \mid x \in [a,b],\ c \le y \le \varphi(x)\}$$

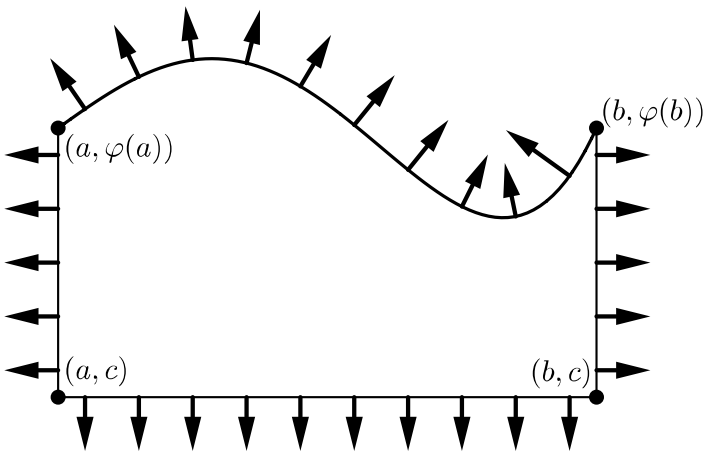
*gilt dann*

$$\int_B \mathrm{div}(f)\,\mathrm{dvol} = \int_{\partial B} f\cdot \mathrm{d}\mathbf{n}.$$

*Hierbei ist per Definition*

$$\begin{aligned}
\int_{\partial B} f\cdot \mathrm{d}\mathbf{n} = &\underbrace{-\int_a^b f_2(x,c)\,\mathrm{d}x}_{\text{unten}} \underbrace{+\int_c^{\varphi(b)} f_1(b,y)\,\mathrm{d}y}_{\text{rechts}} \\
&\underbrace{+\int_a^b \left\langle f(x,\varphi(x)), \begin{pmatrix}-\varphi'(x)\\ 1\end{pmatrix}\right\rangle \mathrm{d}x}_{\text{oben}} \underbrace{-\int_c^{\varphi(a)} f_1(a,y)\,\mathrm{d}y}_{\text{links}}.
\end{aligned}$$

Wie zuvor für Rechtecke, möchten wir obige Definition des Flussintegrals $\int_{\partial B} f\cdot \mathrm{d}\mathbf{n}$ an einem Bild erklären.




Figur 13.4: An einem beliebigen Punkt $(x, \varphi(x))$ im Graphen von $\varphi$ ist $(1, \varphi'(x))^t$ ein nach rechts gerichteter Tangentenvektor. Rotiert man diesen nun um 90 Grad im Gegenuhrzeigersinn, so erhält man den Vektor $(-\varphi'(x), 1)^t$, der also senkrecht auf dem Graphen stehen muss und nach aussen zeigt. Wir bemerken, dass dieser Normalenvektor nicht zwingend Länge 1 besitzt; seine Länge entspricht der Bogenlänge in der Parametrisierung durch die $x$-Koordinate.

*Beweis.* Wir können ohne Beschränkung der Allgemeinheit annehmen, dass $\varphi$ auf ganz $[a, b]$ stetig differenzierbar ist. In der Tat, falls $\varphi$ nur stückweise stetig differenzierbar ist, so zerteilen wir das Intervall $[a, b]$ in Teilintervalle, auf welchen $\varphi$ stetig differenzierbar ist. Kennt man die Proposition für Graphen stetig differenzierbarer Funktionen, so ergibt sich durch Addition der entstehenden Gleichungen die gewünschte Formel.

Wir definieren das Rechteck $Q = [a, b] \times [c, d]$ sowie (unter Verwendung von Satz 9.66(4) und der Stetigkeit der Divergenz nach Proposition 13.1)

$$M = \max\Big\{\|f(x, y)\|, |\operatorname{div}(f)(x, y)|, |\varphi'(x)| \;\Big|\; (x, y) \in [a, b] \times [c, d]\Big\}.$$

Sei $\varepsilon > 0$. Wir wählen zuerst ein $\eta < \varepsilon$ nach der gleichmässigen Stetigkeit von $f_1$ und $f_2$ auf $Q$, so dass für alle $(x, y), (\tilde{x}, \tilde{y}) \in Q$ mit $\|(x, y) - (\tilde{x}, \tilde{y})\|_\infty < \eta$ die Abschätzungen

$$|f_1(x, y) - f_1(\tilde{x}, \tilde{y})| < \varepsilon, \qquad |f_2(x, y) - f_2(\tilde{x}, \tilde{y})| < \varepsilon \tag{13.2}$$

gelten. Des Weiteren wählen wir nach der gleichmässigen Stetigkeit von $\varphi, \varphi'$ ein $\delta < \eta$, so dass für $x, \tilde{x} \in [a, b]$ mit $|x - \tilde{x}| < \delta$ auch

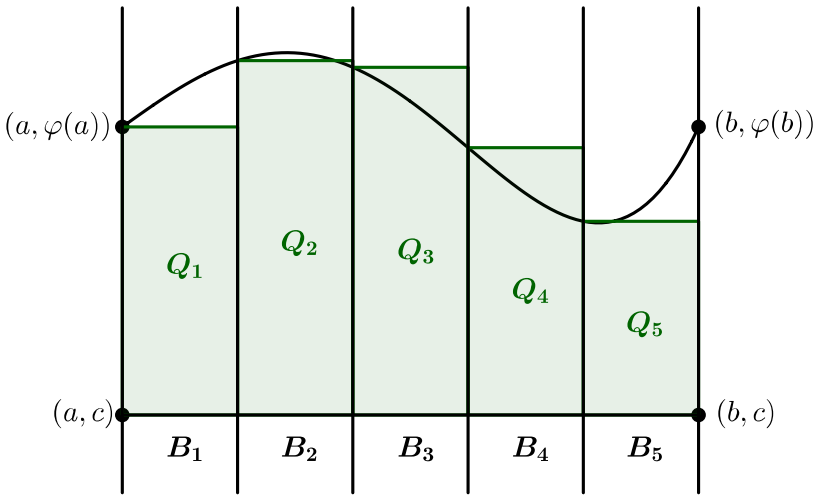
$$|\varphi(x) - \varphi(\tilde{x})| < \eta < \varepsilon, \qquad |\varphi'(x) - \varphi'(\tilde{x})| < \varepsilon$$

gilt.

Sei $\mathfrak{Z} = \{a = x_0 < \ldots < x_J = b\}$ eine Zerlegung von $[a, b]$ mit Maschenweite kleiner als $\delta$. Wir zerlegen $B$ in dünne Streifen $B_1, \ldots, B_J$, wobei

$$B_j = B \cap \big([x_{j-1}, x_j] \times \mathbb{R}\big) = \{(x, y) \in U \mid x \in [x_{j-1}, x_j],\ c \le y \le \varphi(x)\}$$

für $j = 1, \ldots, J$; siehe folgendes Bild.




Auch definieren wir (wie in obigem Bild schon angedeutet)

$$Q_j = [x_{j-1}, x_j] \times [c, \varphi(x_{j-1})]$$

für $j = 1, \ldots, J$. Nach Wahl von $\delta$ und der Zerlegung $\mathfrak{Z}$ erhalten wir die Abschätzung

$$\mathrm{vol}(B_j \triangle Q_j) \leq \varepsilon(x_j - x_{j-1})$$

für $j = 1, \ldots, J$. (Wieso?). Hierbei bezeichnet $B_j \triangle Q_j = (B_j \setminus Q_j) \cup (Q_j \setminus B_j)$ wie in Abschnitt 1.4.1 die symmetrische Differenz. Daher gilt auch

$$\begin{aligned}\left| \int_B \mathrm{div}(f)\,\mathrm{dvol} - \sum_{j=1}^{J} \int_{Q_j} \mathrm{div}(f)\,\mathrm{dvol} \right| &\leq \sum_{j=1}^{J} \left| \int_{B_j} \mathrm{div}(f)\,\mathrm{dvol} - \int_{Q_j} \mathrm{div}(f)\,\mathrm{dvol} \right| \\ &\leq \sum_{j=1}^{J} M\varepsilon(x_j - x_{j-1}) = M(b-a)\varepsilon.\end{aligned}$$

Insbesondere strebt das Flächenintegral über $\bigcup_{j=1}^{J} Q_j$ für $\varepsilon \to 0$ gegen das Flächenintegral über $B$.

Wir wenden nun Proposition 13.1 auf die Rechtecke $Q_1, \ldots, Q_J$ an und erhalten

$$\sum_{j=1}^{J} \int_{Q_j} \mathrm{div}(f)\,\mathrm{dvol} = \sum_{j=1}^{J} \int_{\partial Q_j} f \cdot \mathrm{d}\mathbf{n}.$$

Der Ausdruck rechts besteht vorerst aus $4n$ Integralen (wie sehen diese genau aus?). Wir werden diese Terme nun umorganisieren und dabei feststellen, dass sich einiges weghebt und

zusammenfasst. In der Tat ist

$$\sum_{j=1}^{J}\int_{\partial Q_j} f\cdot \mathrm{d}\mathbf{n} = \underbrace{-\int_a^b f_2(x,c)\,\mathrm{d}x + \int_c^{\varphi(b)} f_1(b,y)\,\mathrm{d}y - \int_c^{\varphi(a)} f_1(a,y)\,\mathrm{d}y}_{\text{unten}\qquad\text{rechts}\qquad\text{links}}$$
$$\underbrace{+\sum_{j=1}^{J}\int_{x_{j-1}}^{x_j} f_2(x,\varphi(x_{j-1})\,\mathrm{d}x - \sum_{j=1}^{J}\int_{\varphi(x_{j-1})}^{\varphi(x_j)} f_1(x_j,y)\,\mathrm{d}y,}_{\text{Treppe}}$$

da sich die Integrale über die „unteren Ränder“ zu einem Integral zusammenfügen und sich das Integral über den „rechten Rand“ von $Q_j$ zum Teil mit dem Integral über den „linken Rand“ von $Q_{j+1}$ weghebt. (Es empfiehlt sich dazu, das obige Bild nochmals zu betrachten.) Die ersten drei Integrale stimmen bereits mit drei Integralen in der Definition von $\int_{\partial B} f\cdot \mathrm{d}\mathbf{n}$ überein.

Es bleibt also die verbleibenden Ausdrücke mit der Beschriftung „Treppe“ mit dem Integral

$$\int_a^b \left\langle f(x,\varphi(x)), \begin{pmatrix}-\varphi'(x)\\ 1\end{pmatrix}\right\rangle \mathrm{d}x = \int_a^b f_2(x,\varphi(x))\,\mathrm{d}x - \int_a^b f_1(x,\varphi(x))\varphi'(x)\,\mathrm{d}x$$

mit der Beschriftung „oben“ in der Definition von $\int_{\partial B} f\cdot \mathrm{d}\mathbf{n}$ zu vergleichen. Wir vergleichen zuerst die Summe für die „horizontalen Teile der Treppe“ mit dem „horizontalen Teil des Integrals oben“ und erhalten

$$\begin{aligned}\Big|\sum_{j=1}^{J}\int_{x_{j-1}}^{x_j} f_2(x,\varphi(x_{j-1}))\,\mathrm{d}x - \int_a^b f_2(x,\varphi(x))\,\mathrm{d}x\Big| \\ \le \sum_{j=1}^{J}\int_{x_{j-1}}^{x_j} \big|f_2(x,\varphi(x_{j-1})) - f_2(x,\varphi(x))\big|\,\mathrm{d}x \le \varepsilon(b-a)\end{aligned}$$

wegen (13.2) und $|\varphi(x_{j-1})-\varphi(x)|<\eta$ für $x\in[x_{j-1},x_j]$ nach Wahl von $\delta$.

Für die zweite Summe zu den „vertikalen Teilen“ der Treppe verwenden wir wiederum (13.2) und erhalten gemeinsam mit dem Mittelwertsatz (Theorem 7.29) ein $\xi_j\in[x_{j-1},x_j]$ mit

$$\begin{aligned}&\Big|\int_{\varphi(x_{j-1})}^{\varphi(x_j)} f_1(x_j,y)\,\mathrm{d}y - f_1(x_j,\varphi(x_j))\big(\varphi(x_j)-\varphi(x_{j-1})\big)\Big| \\ &\qquad = \Big|\int_{\varphi(x_{j-1})}^{\varphi(x_j)} \big(f_1(x_j,y)-f_1(x_j,\varphi(x_j))\big)\,\mathrm{d}y\Big| \\ &\qquad \le \varepsilon\big|\varphi(x_j)-\varphi(x_{j-1})\big| = \varepsilon|\varphi'(\xi_j)(x_j-x_{j-1})| \\ &\qquad \le \varepsilon M(x_j-x_{j-1}).\end{aligned}$$

Ebenso gilt auf Grund des Fundamentalsatzes der Differential und Integralrechnung (in der Form von Korollar 8.5) und (13.2) die Abschätzung

$$\begin{aligned}\left|\int_{x_{j-1}}^{x_j} f_1(x,\varphi(x))\varphi'(x)\,\mathrm{d}x - f_1(x_j,\varphi(x_j))\big(\varphi(x_j)-\varphi(x_{j-1})\big)\right| \\ &= \left|\int_{x_{j-1}}^{x_j} \big(f_1(x,\varphi(x)) - f_1(x_j,\varphi(x_j))\big)\varphi'(x)\,\mathrm{d}x\right| \\ &\leq \varepsilon M(x_j - x_{j-1})\end{aligned}$$

Kombinieren wir diese beiden Abschätzungen und summieren wir über $j \in \{1,\dots,J\}$, so erhalten wir

$$\left|\sum_{j=1}^{J}\int_{\varphi(x_{j-1})}^{\varphi(x_j)} f_1(x_j,y)\,\mathrm{d}y - \int_a^b f_1(x,\varphi(x))\varphi'(x)\,\mathrm{d}x\right| \leq 2\varepsilon M(b-a).$$

Insgesamt haben wir also gezeigt, dass

$$\underbrace{\sum_{j=1}^{J}\int_{x_{j-1}}^{x_j} f_2(x,\varphi(x_{j-1}))\,\mathrm{d}x - \sum_{j=1}^{J}\int_{\varphi(x_{j-1})}^{\varphi(x_j)} f_1(x_j,y)\,\mathrm{d}y}_{\text{Treppe}}$$

für $\varepsilon \to 0$ gegen

$$\underbrace{\int_a^b \left\langle f(x,\varphi(x)), \begin{pmatrix}-\varphi'(x)\\ 1\end{pmatrix}\right\rangle \mathrm{d}x}_{\text{oben}}$$

strebt. Dies impliziert die Proposition in dem Fall einer stetig differenzierbaren Funktion $\varphi$ auf $[a,b]$. □

VO: 07.05.18

Wir möchten noch einen Kommentar im Nachhinein zu obigem Beweis anbringen. Es sieht so aus, als ob die Integrale der Funktion $f_1(x_j,y)$ über die „vertikalen Stücke der Treppe“ stets negatives Vorzeichen hätten; dies ist aber nicht so. Das Vorzeichen hängt davon ab, ob der Funktionsgraph an dieser Stelle „nach oben“ oder „nach unten“ geht oder genauer gesagt, welche der beiden anliegenden Kanten länger ist (wieso?).

**Übung 13.3** (Divergenzdefinition mit Kreisen)**.** *Sei $U \subseteq \mathbb{R}^2$ eine offene Menge, sei $f = (f_1,f_2)^t : U \to \mathbb{R}^2$ ein stetig differenzierbares Vektorfeld und sei $B = \overline{B_r(p)} \subseteq U$ ein abgeschlossener Ball mit $r > 0$. Zeigen Sie, dass*

$$\int_B \operatorname{div}(f)\,\mathrm{dvol} = \int_{\partial B} f\cdot \mathrm{d}\mathbf{n}$$

*gilt, wobei das rechte Integral durch*

$$\int_{\partial B} f\cdot \mathrm{d}\mathbf{n} = \int_0^{2\pi} \left\langle f\left(p + r\begin{pmatrix}\cos(t)\\ \sin(t)\end{pmatrix}\right), r\begin{pmatrix}\cos(t)\\ \sin(t)\end{pmatrix}\right\rangle \mathrm{d}t$$

*gegeben ist. Hier haben wir die Parametrisierung*

$$\gamma : [0, 2\pi] \to \partial B_r(x),\ t \mapsto p + r \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}$$

*des Randes $\partial B$ verwendet. Benutzen Sie dies anschliessend, um zu zeigen, dass die Divergenz von $f$ bei $u \in U$ durch*

$$\operatorname{div}(f)(p) = \lim_{r \to 0} \tfrac{1}{\pi r^2} \int_{\partial B_r(u)} f \cdot \mathrm{d}\mathbf{n}$$

*gegeben ist.*

### 13.1.2 Glatt berandete Bereiche

Um den Divergenzsatz auf allgemeineren Bereichen zu formulieren und zu beweisen, verwenden wir folgenden Begriff (hier für $n = 2$ und im nächsten Abschnitt für $n = 3$).

**Definition 13.4** (Glatt berandet)**.** Eine abgeschlossene Teilmenge $B \subseteq \mathbb{R}^n$ heisst ein **glatt berandeter Bereich**, falls es für jeden Punkt $p \in \partial B$ eine Matrix $P = A_\varepsilon P_\sigma$ für

- eine Diagonalmatrix $A_\varepsilon = \operatorname{diag}((-1)^{\varepsilon_1}, \ldots, (-1)^{\varepsilon_n})$ zu $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)$ mit $\varepsilon_1, \ldots, \varepsilon_n \in \{0, 1\}$ und
- eine Permutationsmatrix $P_\sigma$ zu $\sigma \in S_n$

sowie einen offenen Quader $O = O_1 \times (c, d)$ mit $P(p) \in O$ gibt, so dass der Durchschnitt $P(B) \cap O$ durch eine glatte Funktion $\varphi : O_1 \to \mathbb{R}$ definiert werden kann. Genauer ist

$$P(B) \cap O = \{(x, y) \in O_1 \times (c, d) \mid c < y \leq \varphi(x)\}\,.$$

In Worten ausgedrückt ist ein Bereich $B$ also glatt berandet, wenn $B$ um jeden Randpunkt in geeigneter Lage aussieht wie ein Gebiet unter einem Graphen einer Funktion in $(n-1)$ Koordinaten (der Standard-Koordinaten). Die Matrix $P_\sigma$ erfüllt dann den Zweck, die Koordinaten so zu vertauschen, dass sich $B$ um den Punkt als Bereich unter oder über dem Graphen einer Funktion in den ersten (!) $(n-1)$-Koordinaten auffassen lässt. Weiter spiegelt die Matrix $A_\varepsilon$ entlang der Koordinatenachsen, womit man nur Bereiche unter (!) Graphen betrachten kann. Beispielsweise ist $A_{(0,1)}$ die Spiegelung entlang der zweiten Koordinate, da $A_{(0,1)}(x_1, x_2) = (x_1, -x_2)$ für alle $(x_1, x_2) \in \mathbb{R}^2$ gilt. Die Matrizen wie in obiger Definition sind genau die Elemente von

$$\mathrm{O}_n(\mathbb{Z}) = \left\{P \in \mathrm{Mat}_n(\mathbb{Z}) \mid P^t P = I_n\right\},$$

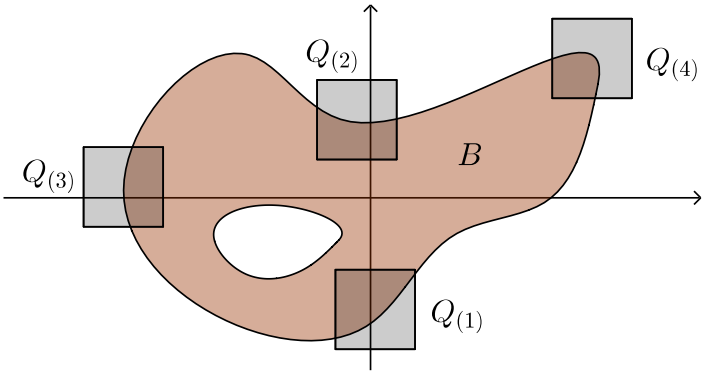
weswegen wir auch schlicht $P \in \mathrm{O}_n(\mathbb{Z})$ schreiben werden. Wir bemerken noch, dass wir auch ohne Beschränkung der Allgemeinheit annehmen könnten, dass die verwendeten Matrizen $P$ zusätzlich $\det P = 1$ erfüllen, also zu

$$\mathrm{SO}_n(\mathbb{Z}) = \left\{P \in \mathrm{Mat}_n(\mathbb{Z}) \mid P^t P = I_n,\ \det P = 1\right\}$$

gehören. In der Tat können wir $P$ und $O$ immer durch $A_{(1,0,\ldots,0)}P$ und $A_{(1,0,\ldots,0)}O$ ersetzen, wobei sich das Vorzeichen von $\det(P)$ ändert aber die Beschreibung der Menge $P(B) \cap O$ sich nicht wesentlich ändert.

Wir bemerken an dieser Stelle noch, dass für die weitere Diskussion dieses Abschnitts die Funktion $\varphi$ zu einem Punkt $p$ in obiger Definition genau so gut nur stückweise stetig differenzierbar sein könnte. Da aber der glatt berandetete Fall bereits alle Beweisideen enthält, werden wir nur diesen Fall behandeln. In den Anwendungen der Sätze werden wir aber nicht auf die Glattheit des Randes bestehen, da in konkreten Fällen meist Proposition 13.2 ausreicht.

Wir möchten die verschiedenen Phänomene, die auftreten können, in folgendem Bild zusammenfassen. Im Folgenden werden wir einen achsenparallelen Würfel im $\mathbb{R}^2$ auch als ein Quadrat bezeichnen.




Figur 13.5: In der durch $Q_{(1)}$ definierte Umgebung lässt sich der Schnitt $B \cap Q_{(1)}$ als die Menge der Punkte in $Q_{(1)}$ oberhalb eines Graphen über der $x$-Achse auffassen. Genauso ist der Schnitt $B \cap Q_{(2)}$ die Menge der Punkte in $Q_{(2)}$ unterhalb eines Graphen über der $x$-Achse. Wir stellen aber auch fest, dass sich $B \cap Q_{(2)}$ (im Gegensatz zu $B \cap Q_{(1)}$) nicht als Gebiet unterhalb eines Graphen über der $y$-Achse auffassen lässt. Letzteres trifft umgekehrt auf den Schnitt $B \cap Q_{(3)}$ zu, welcher nur als Gebiet oberhalb eines (also rechts von dem) Graphen über der $y$-Achse dargestellt werden kann. Auch bemerken wir, dass die gleiche „Grösse“ von Quadraten an anderen Stellen vielleicht nicht passt. Beispielsweise lässt sich $B \cap Q_{(4)}$ weder über der $x$-Achse noch über der $y$-Achse als Menge unterhalb oder oberhalb eines Graphen interpretieren.

Auch stellen wir fest, dass der Rand $\partial B$ eines glatt berandeten Bereiches $B \subseteq \mathbb{R}^n$, wie vielleicht aus obigem Bild anschaulich klar ist, eine $(n-1)$-dimensionale Teilmannigfaltigkeit ist. Die obige Definition liefert aber nicht nur diese Information, sondern erklärt auch, auf welcher „Seite“ des Randes das Innere von $B$ liegt (wieso?).

**Übung 13.5** (Rand eines glatt berandeten Bereiches)**.** *Sei $B \subseteq \mathbb{R}^n$ ein glatt berandeter Bereich. Zeigen Sie, dass der Rand $\partial B$ eine $(n-1)$-dimensionale Teilmannigfaltigkeit bildet.*

Bevor wir uns mit allgemeineren Aussagen beschäftigen, wollen wir in folgendem Lemma eine Kollektion von Beispielen glatt berandeter Mengen einführen.

**Lemma 13.6** (Glatte Ränder und Niveaumengen)**.** *Sei $F : \mathbb{R}^n \to \mathbb{R}$ eine glatte Funktion mit Null als regulären Wert. Dann ist die abgeschlossene Teilmenge*

$$B = \{u \in \mathbb{R}^n \mid F(u) \geq 0\}$$

*glatt berandet und* $\partial B = \{u \in \mathbb{R}^n \mid F(u) = 0\}$.

Beispielsweise ist also der abgeschlossene Einheitsball

$$\overline{B_R(0)} = \left\{u \in \mathbb{R}^n \mid R^2 - (u_1^2 + \ldots + u_n^2) \geq 0\right\}$$

für alle $R > 0$ glatt berandet. Wir merken an, dass Quader nicht glatt berandet sind, was sich anhand von Übung 13.5 und Proposition 11.12 (für $C^1$-Teilmannigfaltigkeiten angepasst) zeigen lässt.

Lemma 13.6 stellt gewissermassen eine Verschärfung des Satz über den konstanten Rang (Theorem 11.16) dar, der im Wesentlichen besagt, dass $\{u \in \mathbb{R}^n \mid F(u) = 0\}$ eine Teilmannigfaltigkeit bildet. Intuitiv ausgedrückt bleibt also zu zeigen, dass die Funktion auf der einen Seite der Niveaumenge positiv und auf der anderen negativ ist.

*Beweis.* Wir imitieren zu grossen Teilen den Beweis von Theorem 11.16. Sei $p \in \partial B$. Nach Annahme hat $\mathrm{D}_p F = \nabla F(p)^t$ vollen Rang und (anders ausgedrückt) es existiert also ein Eintrag, der von Null verschieden ist. Wir können also die Koordinaten permutieren und annehmen, dass $\partial_n F(p) \neq 0$ ist. Wir schreiben $p = (x_0, y_0)$ für $x_0 \in \mathbb{R}^{n-1}$ und $y_0 \in \mathbb{R}$. Genau wie im Beweis von Theorem 11.16 erhalten wir offene Umgebungen $U_0$ von $x_0$ und $V_0$ von $y_0$, so dass der Schnitt von $\{(x, y) \in \mathbb{R}^n \mid F(x, y) = 0\}$ mit $U_0 \times V_0$ durch den Graphen einer glatten Funktion $f : U_0 \to V_0$ gegeben ist (siehe auch Proposition 11.12).

Wir wollen im Folgenden der Einfachheit halber annehmen, dass $\partial_n F(p) < 0$ ist; das Argument ist im anderen Fall analog. Wenn nötig können wir die Umgebungen $U_0, V_0$ verkleinern, so dass $\partial_n F(v) < 0$ für alle $v \in U_0 \times V_0$ ist und so dass $U_0 \times V_0$ ein Quader ist, womit insbesondere $V_0 = (c, d) \subseteq \mathbb{R}$ ein Intervall ist. Wir zeigen nun, dass

$$B \cap (U_0 \times V_0) = \{(x, y) \in U_0 \times V_0 \mid F(x, y) \geq 0\} = \{(x, y) \in U_0 \times V_0 \mid y \leq f(x)\}$$

gilt. In der Tat ist zu $x \in U_0$ die Funktion $y \in (c, d) \mapsto F(x, y)$ strikt monoton fallend und verschwindet bei $y = f(x)$. Daher gilt $y \leq f(x)$ genau dann, wenn $F(x, y) \geq 0$. Insbesondere ist $\{(x, y) \in \mathbb{R}^n \mid F(x, y) = 0\} \subseteq \partial B$. Für $(x, y) \in B$ mit $F(x, y) > 0$ gilt auf Grund der Stetigkeit von $F$ sogar $(x, y) \in B^\circ$ und das Lemma folgt aus obigem. □

Wir möchten nun die Frage diskutieren, ob und inwiefern die Quader in Definition 13.4 für alle Punkte gleich gross gewählt werden können (Lemma 13.7).

**Lemma 13.7** (Geeignete Wahl der Seitenlänge)**.** *Sei $B \subseteq \mathbb{R}^n$ ein glatt berandeter, beschränkter Bereich. Dann existiert eine Länge $\eta_0 > 0$, so dass für jeden Randpunkt $p \in \partial B$ die Umgebung $O$ zu $p$ aus Definition 13.4 als Würfel mit Kantenlänge $2\eta_0$ um $P(p)$ gewählt werden kann (wobei $P$ ebenfalls wie in Definition 13.4 gegeben ist).*

*Beweis.* Wir bemerken zuerst, dass zu $p \in \partial B$ ein Würfel der Form

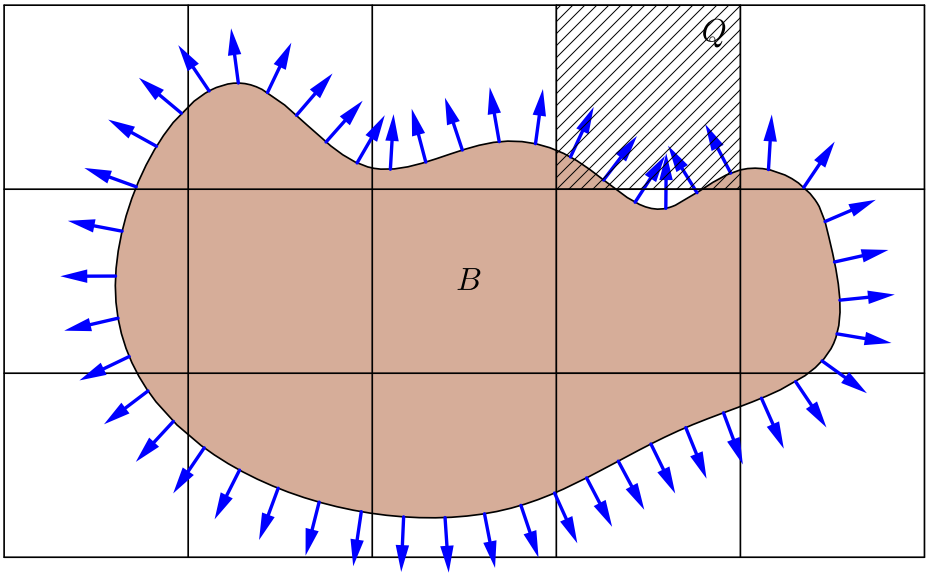
$$O = B_\eta^\infty(P(p)) = \{q \in \mathbb{R}^n \mid \|q - P(p)\|_\infty < \eta\}$$

für ein $P \in \mathrm{O}_n(\mathbb{Z})$ und ein $\eta > 0$ existiert, der alle Eigenschaften aus Definition 13.4 hat. In der Tat gibt es für einen beliebigen Quader $\tilde{O}$ wie in Definition 13.4 stets ein $\eta > 0$ mit $B_\eta^\infty(P(p)) \subseteq \tilde{O}$, da $\tilde{O}$ offen ist und $P(p) \in \tilde{O}$ ist. Die gewünschten Eigenschaften übertragen sich damit von $\tilde{O}$ auf $O = B_\eta^\infty(P(p))$ (wieso?). Wir wollen in diesem Fall sagen, dass $\eta > 0$ für $p$ „geeignet“ ist. Des Weiteren bemerken wir, dass in diesem Fall jedes $\eta' \in (0, \eta)$ ebenso für $p$ geeignet ist (wieso?).

Um die Existenz eines „uniform geeigneten“ $\eta > 0$ zu beweisen, betrachten wir die Überdeckung $\mathcal{O}$ von $\partial B$ mit (relativ) offenen Mengen der Form $\partial B \cap P^{-1}(O)$, wobei $O$ und $P \in \mathrm{O}_n(\mathbb{Z})$ Definition 13.4 für einen Punkt $p \in \partial B$ erfüllen. Nach Proposition 9.67 über die Existenz einer Lebesgue-Zahl und nach der Kompaktheit von $\partial B$ existiert ein $\eta_0 > 0$, so dass für alle $p \in \partial B$ ein $O' \in \mathcal{O}$ mit $B_{\eta_0}^\infty(p) \cap \partial B \subseteq O'$ existiert. Nach dem obigen Argument ist ein solches $\eta_0$ aber für jedes $p \in \partial B$ geeignet. □

### 13.1.3 Glatt berandete Bereiche in der Ebene

Eines der Ziele dieses Abschnitts (Abschnitt 13.1) wird es sein, den Divergenzsatz für glattberandete Bereiche zu formulieren und zu beweisen. Als Ausgangspunkt dazu können Sie folgendes Bild erachten.




Figur 13.6: In Richtung eines Divergenzsatzes für glatt berandete Bereiche.

Gegeben einen glattberandeten Bereich $B$ möchten wir diesen also in Stücke zerlegen, die jeweils in etwa die Form aus Proposition 13.2 haben (nach Drehung und/oder Spiegelung). Wendet man nun Proposition 13.2 auf all diese Stücke an (vorausgesetzt dies ist erlaubt), so erhält man das Integral über den Rand, da sich die Integrale im Inneren entlang der Aussennormalen wegheben. Bei genauerer Betrachtung ergeben sich folgende Probleme:

- Die in obigem Bild bezeichneten Stücke des Bereiches haben nicht alle die Form aus Proposition 13.2; ein solches Beispiel ist $Q \cap B$. Wir werden zur Behebung dieses Problems wie folgt vorgehen:

  - Nach Lemma 13.7 können wir eine Seitenlänge finden, für die jedes Quadrat dieser Seitenlänge (oder kleiner) um einen Punkt auf $\partial B$ als Umgebung wie in Definition 13.4 gewählt werden kann.
  - Anschliessend können wir eine genügend feine Zerlegung konstruieren, so dass jedes Quadrat höchstens mit einem weiteren Quadrat zusammengeklebt werden muss, damit das erhaltene Rechteck den Annahmen von Proposition 13.2 genügt (Proposition 13.8). Beispielsweise kann man $Q$ mit dem Quadrat unter $Q$ zusammenkleben und erhält ein Rechteck $Q'$, für welches $Q' \cap B$ die Anforderungen von Proposition 13.2 erfüllt.

- Das aus der obigen Zerstückelung des Bereiches erhaltene Integral ist a priori von der Zerstückelung abhängig und lässt sich in Beispielen auch schlecht (respektive mühselig) berechnen. Deswegen werden einen allgemeineren Zugang zum Flussintegral über den Rand wählen (Definitionen 13.9, 13.11), welches dann unabhängig ist von den getroffenen Wahlen (Lemma 13.37).

**Proposition 13.8** (Puzzlestein-Domino Überdeckung*)**.** *Sei $B \subseteq \mathbb{R}^2$ ein glatt berandeter Bereich und sei $Q \supseteq B$ ein abgeschlossenes Quadrat. Dann existiert für jede Zerlegungen $\mathfrak{Z}$ von $Q$ in Quadrate mit genügend kleiner Maschenweite eine Äquivalenzrelation $\sim$ auf der Menge $\{Q_{\boldsymbol{\alpha}} \mid Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}\}$, so dass die Äquivalenzklassen aus einem oder zwei benachbarten Quadraten bestehen und so dass für die Vereinigung $P$ über eine Äquivalenzklasse der Durchschnitt $B \cap P$ durch eine stückweise stetig differenzierbare Funktion beschrieben werden kann. Genauer formuliert gibt es für jedes $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$ die folgenden zwei Möglichkeiten:*

- *(Puzzlestein) Entweder $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$ ist zu keinem weiteren Quadrat der Zerlegung $\mathfrak{Z}$ äquivalent, wir nennen in diesem Fall $R = \overline{Q_{\boldsymbol{\alpha}}}$ einen* Puzzlestein.

- *(Domino) Oder $Q_{\boldsymbol{\alpha}}$ ist genau zu einem weiteren Quadrat $Q_{\boldsymbol{\beta}} \sqsubset \mathfrak{Z}$ äquivalent, so dass der Abschluss $R = \overline{Q_{\boldsymbol{\alpha}} \cup Q_{\boldsymbol{\beta}}}$ der Vereinigung ein Rechteck ist, welches wir als* Domino *bezeichnen.*

*In beiden Fällen hat $\partial B \cap \partial R$ höchstens zwei Elemente und der Durchschnitt $B \cap R$ kann, möglicherweise nach Vertauschung der Koordinaten, durch den Graphen einer stückweise stetig differenzierbaren Funktion $\varphi$ wie in Proposition 13.2 beschrieben werden.*

Wie zu Beginn dieses Abschnitts schon angedeutet wurde, haben wir in obiger Proposition den Fall des Dominosteins eingeführt, da eine stetig differenzierbare Funktion oder auch eine glatte Funktion $\varphi : [a, b] \to \mathbb{R}$ einen Wert $d \in \mathbb{R}$ unendlich oft annehmen kann, ohne aber konstant zu sein. Dies führt zu Komplikationen, da der Durchschnitt der Menge unter dem Graph von $\varphi$ mit dem Rechteck $[a, b] \times [c, d]$ möglicherweise nicht mehr durch eine stückweise stetig differenzierbare Funktion beschrieben werden kann. Man beachte, dass diese Problematik im Bild 13.6 nicht dargestellt ist. Allerdings kann durch Vereinigung mit einem zweiten benachbarten Quadrat dieses Problem ebenfalls verhindert werden.

Der Beweis dieser Proposition ist im Wesentlichen eine Verfeinerung des Beweises von Lemma 13.7.

*Beweis.* Wir sagen, dass $\eta > 0$ für $p \in \partial B$ „besser geeignet“ ist, falls $\eta > 0$ geeignet ist (siehe den Beweis von Lemma 13.7) und die stetig differenzierbare Funktion $\varphi$ in der Beschreibung von $P(B) \cap B_\eta^\infty(P(p))$ eine der beiden folgenden Eigenschaften erfüllt:

- (Monoton) $\frac{1}{2} \leq |\varphi'(x)|$ für alle $x$.
- (Flach) $|\varphi'(x)| \leq \frac{2}{3}$ für alle $x$.

Wir bemerken, dass im ersten Fall $\varphi'$ auf Grund der Stetigkeit von $\varphi'$ das Vorzeichen nicht wechseln kann, womit in diesem Fall $\varphi$ in der Tat monoton ist.

Wir zeigen nun, dass jeder Punkt $p \in \partial B$ ein besser geeignetes $\eta > 0$ besitzt. Sei $\eta_1 > 0$ geeignet für $p$ und $(x_0, y_0) = P(p)$.

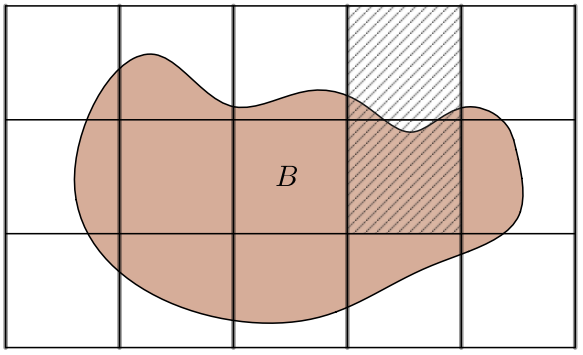
- Falls $|\varphi'(x_0)| > \frac{1}{2}$ ist, so können wir auf Grund der Stetigkeit von $\varphi'$ ein $\eta \in (0, \eta_1)$ wählen, so dass $|\varphi'(x)| \geq \frac{1}{2}$ für alle $x \in (x_0 - \eta, x_0 + \eta)$ ist und wir erhalten den ersten Fall (Monoton) für $P(B) \cap B_\eta^\infty(P(p))$.
- Falls $|\varphi'(x_0)| \leq \frac{1}{2}$ gilt, so können wir analog für ein hinreichend kleines $\eta \in (0, \eta_1)$ den zweiten Fall (Flach) für $P(B) \cap B_\eta^\infty(P(p))$ erzielen.

Wie im Beweis von Lemma 13.7 können wir mit dem Begriff „besser geeignet“ eine Überdeckung von $\partial B$ und eine Lebesgue-Zahl $\eta_0 > 0$ dieser Überdeckung finden, welche dann für alle $p \in \partial B$ besser geeignet ist.

Nun setzen wir $\delta_0 = \frac{1}{4}\eta_0$ und betrachten eine Zerlegung $\mathfrak{Z}$ von $Q$ in Quadrate mit Maschenweite kleiner als $\delta_0$. Wir definieren eine Äquivalenzrelation auf $\{Q_{\boldsymbol{\alpha}} \mid Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}\}$ durch

$$Q_{\boldsymbol{\alpha}} \sim Q_{\boldsymbol{\beta}}$$

für $Q_{\boldsymbol{\alpha}}, Q_{\boldsymbol{\beta}} \sqsubset \mathfrak{Z}$, falls $Q_{\boldsymbol{\alpha}} = Q_{\boldsymbol{\beta}}$ oder falls $\overline{Q_{\boldsymbol{\alpha}}} \cap \overline{Q_{\boldsymbol{\beta}}} \cap \partial B$ mehr als einen Punkt enthält. Wir bemerken zuerst, dass es für ein gegebenes $Q_{\boldsymbol{\alpha}}$ höchstens 4 weitere $Q_{\boldsymbol{\beta}} \sqsubset \mathfrak{Z}$ mit $|\overline{Q_{\boldsymbol{\alpha}}} \cap \overline{Q_{\boldsymbol{\beta}}}| > 1$ gibt, nämlich die vier Quadrate über, unter, links und rechts von $Q_{\boldsymbol{\alpha}}$ (falls diese in $Q$ enthalten sind). Auch stellen wir fest, dass a priori nicht klar ist, wieso die oben definierte Relation transitiv sein sollte.




Figur 13.7: Illustration der Definition der Äquivalenzrelation auf der Menge der $\mathfrak{Z}$ entsprechenden Quadrate. In diesem Bild sind die beiden schraffierten Quadrate zueinandener äquivalent, da der Rand von $B$ die Kante zwischen diesen Quadraten zweimal schneidet. Alle anderen Quadrate sind nur zu sich selbst äquivalent.

Wenn $p \in \overline{Q_{\boldsymbol{\alpha}}} \cap \partial B$ und das Quadrat $Q_p = B^\infty_{\eta_0}(p) \supseteq Q_{\boldsymbol{\alpha}}$ „monoton“ ist, so gibt es auf $\partial Q_{\boldsymbol{\alpha}} \cap \partial B$ nur zwei Punkte auf unterschiedlichen Kanten von $\overline{Q_{\boldsymbol{\alpha}}}$, womit keines der vier benachbarten Quadrate zu $Q_{\boldsymbol{\alpha}}$ äquivalent ist. Auf Grund der Monotonie von $\varphi$ können wir eine glatte Funktion mit Hilfe von $\varphi$ und möglicherweise einer konstanten Funktion finden, so dass $\overline{Q_{\boldsymbol{\alpha}}} \cap B$ durch diese beschrieben wird (aber möglicherweise ein kleineres Intervall als Definitionsbereich besitzt).

Wenn aber $p \in \overline{Q_{\boldsymbol{\alpha}}} \cap \partial B$ und das Quadrat $Q_p = B^\infty_{\eta_0}(p) \supseteq Q_{\boldsymbol{\alpha}}$ „flach“ ist, so kann es mehr als einen Punkt auf $\overline{Q_{\boldsymbol{\alpha}}} \cap \overline{Q_{\boldsymbol{\beta}}} \cap \partial B$ für ein benachbartes $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$ geben. Doch auf Grund der Beschränkung $|\varphi'(x)| \leq \frac{2}{3}$ im Fall „flach“ kann dies nur für einen der vier benachbarten Quadrate eintreten und der Graph der Funktion $\varphi$ schneidet den Rand des Dominos $R = \overline{Q_{\boldsymbol{\alpha}}} \cup \overline{Q_{\boldsymbol{\beta}}}$ nur an den Endpunkten.

Wenn $\overline{Q_{\boldsymbol{\alpha}}} \cap \partial B$ leer ist, so ist das Quadrat entweder komplett innerhalb oder komplett ausserhalb von $B$. Somit definiert $\sim$ also eine Äquivalenzrelation mit den gewünschten Eigenschaften und die Proposition folgt. $\square$

VO: 09.05.18

**Definition 13.9** (Positiv orientierte Parametrisierung)**.** Eine **Parametrisierung des Randes** einer abgeschlossenen Teilmenge $B \subseteq \mathbb{R}^2$ ist eine Kollektion von Wegen $\gamma_k : I_k \to \partial B$ auf abgeschlossenen Intervallen $I_k = [a_k, b_k]$ mit $a_k < b_k$ für $k = 1, \ldots, K$ mit folgenden Eigenschaften:

(i) (Überdeckend) Es gilt[1] $\partial B = \bigcup_{k=1}^K \gamma_k\big([a_k, b_k]\big)$.

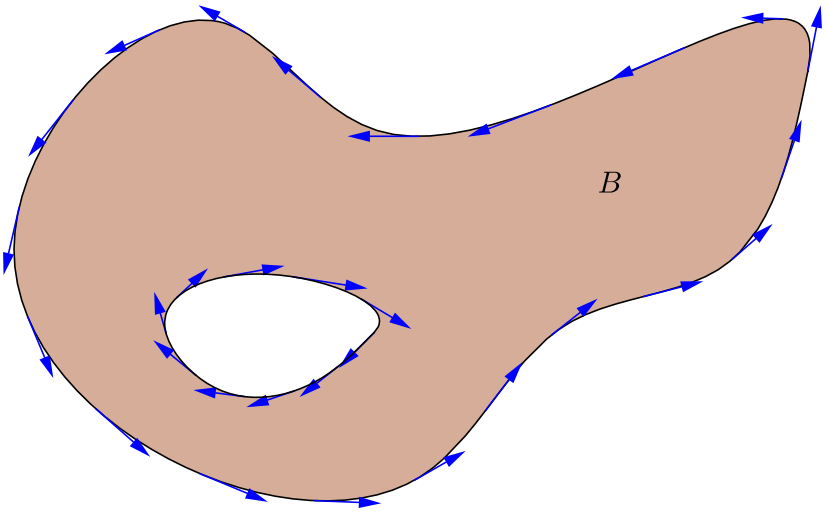
(ii) (Keine Selbstüberschneidungen abgesehen von den Endpunkten) Für alle Indizes $j, k \in \{1, \ldots, K\}$, alle $s \in I_j$ und $t \in I_k^\circ$ mit $(s, j) \neq (t, k)$ gilt $\gamma_j(s) \neq \gamma_k(t)$.

(iii) (Aufeinanderfolgend) Für jedes $k \in \{1, \ldots, K\}$ existiert genau ein $\ell \in \{1, \ldots, K\}$ mit $\gamma_k(b_k) = \gamma_\ell(a_\ell)$.

(iv) (Regularität) Die Wege $\gamma_1, \ldots, \gamma_K$ sind regulär (siehe Abschnitt 8.3.2).

Die Parametrisierung $\gamma_1, \ldots, \gamma_K$ heisst **positiv orientiert**, wenn für jedes $k \in \{1, \ldots, K\}$ und jedes $t \in I_k^\circ$ ein $\varepsilon > 0$ existiert, so dass $\gamma(t) + sR\gamma'(t) \in B^\circ$ für alle $s \in (0, \varepsilon)$ und $\gamma(t) + sR\gamma'(t) \in \mathbb{R}^2 \setminus B$ für alle $s \in (-\varepsilon, 0)$. Dabei ist $R = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right) \in \mathrm{SO}(2, \mathbb{R})$ die Rotationsmatrix zum Winkel 90 Grad **in mathematisch positiver Richtung** (also im Gegenuhrzeigersinn). Intuitiv ausgedrückt ist eine Parametrisierung von $B$ also positiv orientiert, falls die Menge $B$ jeweils links von den Wegen der Parametrisierung liegt.

[1] Auf Grund der folgenden Eigenschaften gilt sogar $\partial B = \bigsqcup_{k=1}^K \gamma_k\big([a_k, b_k)\big)$.




Wir bemerken an dieser Stelle, dass man für einen regulären Weg $\gamma = \gamma_k : I_k \to \partial B$ einer positiv orientierten Parametrisierung $\gamma_1, \ldots, \gamma_K$ von $B$ wie in obiger Definition und einen Zeitpunkt $t_0 \in I_k^\circ$ ein offenes Intervall $J \subseteq I_k^\circ$ um $t_0$ und einen Diffeomorphismus $\psi : J \to \psi(J) = J'$ finden kann, so dass der reparametrisierte Weg $\tilde{\gamma} = \gamma \circ \psi^{-1} : J' \to \partial B$ durch $\tilde{\gamma}(t) = (t, \tilde{\gamma}_2(t))$ oder durch $\tilde{\gamma}(t) = (\tilde{\gamma}_1(t), t)$ für alle $t \in J'$ gegeben ist. Dies folgt im Wesentlichen aus einem Argument wie im Beweis von Proposition 11.12. Man kann das Bild von $\gamma$ also lokal als Graphen auffassen, wobei das Vorzeichen von $\psi'(t_0)$ entscheidet, in welcher Richtung man diesen durchläuft. Ist $\psi'(t_0) > 0$, so liegt $B$ über und sonst unter dem Graphen.

**Wichtige Übung 13.10** (Existenz einer positiv orientierten Parametrisierung)**.** *Zeigen Sie, dass Bereiche wie in Proposition 13.2 und glatt berandete Bereiche eine positiv orientierte Parametrisierung besitzen und geben Sie diese im ersten Fall explizit an.*

**Definition 13.11** (Integral über Ränder)**.** Sei $B \subseteq \mathbb{R}^2$ eine Teilmenge, dessen Rand eine positiv orientierte Parametrisierung $\gamma_1 : I_1 = [a_1, b_1] \to \partial B, \ldots, \gamma_K : I_K = [a_K, b_K] \to \partial B$ besitzt. Weiter sei $f : U \to \mathbb{R}^2$ ein stetig differenzierbares Vektorfeld definiert auf einer offenen Menge $U \supseteq B$. Dann ist das **Wegintegral** von $f$ entlang $\partial B$ durch

$$\int_{\partial B} f \cdot \mathrm{d}\mathbf{s} = \sum_{k=1}^{K} \int_{a_k}^{b_k} \langle f(\gamma_k(t)), \dot{\gamma_k}(t) \rangle \, \mathrm{d}t$$

definiert. Sei $R = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right) \in \mathrm{SO}(2, \mathbb{R})$. Das **Flussintegral** von $f$ durch den Rand $\partial B$ ist dann definiert als

$$\int_{\partial B} f \cdot \mathrm{d}\mathbf{n} = \int_{\partial B} (Rf) \cdot \mathrm{d}\mathbf{s} = \sum_{k=1}^{K} \int_{a_k}^{b_k} \left\langle f(\gamma_k(t)), R^{-1}\dot{\gamma_k}(t) \right\rangle \, \mathrm{d}t.$$

**Wichtige Übung 13.12** (Kompatibilität)**.** *Zeigen Sie, dass unsere Definition von $\int_{\partial B} f \cdot \mathrm{d}\mathbf{n}$ für Rechtecke $B$ am Anfang dieses Abschnitts und für Bereiche $B$ unter Graphen in Proposition 13.2 Spezialfälle obiger Definition darstellen.*

**Lemma 13.13** (Unabhängigkeit von der Parametrisierung)**.** *Sei $B \subseteq \mathbb{R}^2$ eine Teilmenge, die eine positiv orientierte Parametrisierung besitzt, und sei $f : U \to \mathbb{R}^2$ ein stetig differenzierbares Vektorfeld definiert auf einer offenen Menge $U \supseteq B$. Dann hängen sowohl das Wegintegral $\int_{\partial B} f \cdot \mathrm{d}\mathbf{s}$ als auch das Flussintegral $\int_{\partial B} f \cdot \mathrm{d}\mathbf{n}$ nicht von der Wahl der positiv orientierten Parametrisierung des Randes $\partial B$ ab.*

*Beweis.* Per Definition des Flussintegrals reicht es die Aussage für das Wegintegral $\int_{\partial B} f \cdot \mathrm{d}\mathbf{s}$ zu überprüfen. Seien

$$\gamma_1 : I_1 = [a_1, b_1] \to \partial B, \ \dots \ , \gamma_K : I_K = [a_K, b_K] \to \partial B$$
$$\tilde{\gamma}_1 : \tilde{I}_1 = [\tilde{a}_1, \tilde{b}_1] \to \partial B, \ \dots \ , \tilde{\gamma}_{\tilde{K}} : \tilde{I}_{\tilde{K}} = [\tilde{a}_{\tilde{K}}, \tilde{b}_{\tilde{K}}] \to \partial B$$

zwei positiv orientierte Parametrisierungen von $\partial B$. Durch Zerlegen der Wege $\gamma_1, \dots, \gamma_K$ und $\tilde{\gamma}_1, \dots, \tilde{\gamma}_{\tilde{K}}$ und durch Neunummerierung kann man annehmen, dass $K = \tilde{K}$ gilt und für alle $k \in \{1, \dots, K\}$ die Spuren

$$\gamma_k(I_k) = \tilde{\gamma}_k(\tilde{I}_k)$$

gleich und sowohl $\gamma_k$ als auch $\tilde{\gamma}_k$ injektiv sind. Letzteres impliziert auch, dass die Umkehrabbildung $\gamma_k^{-1} : \gamma_k(I_k) \to I_k$ stetig ist. In der Tat ist für jede abgeschlossene Teilmenge $A \subseteq I_k$, das Bild $\gamma_k(A)$ nach Satz 9.72 sogar kompakt und inbesondere abgeschlossen. Nach Proposition 9.37(v) ist damit $\gamma_k^{-1} : \gamma_k(I_k) \to I_k$ stetig.

Sei $k \in \{1, \dots, K\}$. Wir behaupten, dass die Abbildung $\psi = \tilde{\gamma}_k^{-1} \circ \gamma_k : [a_k, b_k] \to [\tilde{a}_k, \tilde{b}_k]$ glatt und monoton wachsend ist. Da Glattheit eine lokale Eigenschaft ist, reicht es für jedes $t_0 \in [a_k, b_k]$ in einer Umgebung von $t_0$ Glattheit zu zeigen.

Sei $\tilde{t}_0 = \tilde{\gamma}_k^{-1}(\gamma_k(t_0))$. Nun lässt sich zum Beispiel mittels

$$\tilde{\Phi} : (\tilde{t}, u) \in (\tilde{t}_0 - \delta, \tilde{t}_0 + \delta) \times (\delta, \delta) \to \tilde{\gamma}_k(\tilde{t}) + R\tilde{\gamma}_k'(\tilde{t})u$$

für ein geeignetes $\delta > 0$ eine stetig differenzierbare Abbildung mit invertierbarer Ableitung bei $(\tilde{t}_0, 0)$ definieren [2], welche $\tilde{\gamma}_k$ erweitert. Nach Satz 11.5 hat diese Abbildung auch eine stetig differenzierbare Umkehrabbildung. Die Existenz der Ableitung von $\psi = \tilde{\Phi}^{-1} \circ \gamma_k$ ergibt sich nun aus der Kettenregel, und aus der Symmetrie zwischen $\gamma_k$ und $\tilde{\gamma}_k$ auch, dass $\psi'(t_0) \neq 0$.

Auf Grund der positiven Orientierung beider Parametrisierungen gilt des Weiteren, dass die beiden Vektoren $\gamma_k'(t)$ und $\tilde{\gamma}_k'(\psi(t))$ bei $t_0$ in dieselbe Richtung zeigen. Dies impliziert $\psi'(t_0) > 0$. Lemma 10.44 zeigt nun, dass

$$\int_{\gamma_k} f \cdot \mathrm{d}\mathbf{s} = \int_{\tilde{\gamma}_k} f \cdot \mathrm{d}\mathbf{s}$$

gilt, was den Beweis abschliesst. □

[2] Falls $\tilde{t}_0$ ein Randpunkt von $\tilde{\gamma}_k$ ist, so kann man zuerst $\tilde{\gamma}_k$ linear über den Randpunkt fortsetzen so dass die Fortsetzung noch immer stetig differenzierbar ist.

### 13.1.4 Divergenzsatz für glatt berandete Bereiche in der Ebene

**Theorem 13.14** (Divergenzsatz in der Ebene)**.** *Sei $B \subseteq \mathbb{R}^2$ ein glatt berandeter, kompakter Bereich und sei $f : U \to \mathbb{R}^2$ ein stetig differenzierbares Vektorfeld definiert auf einer offenen Menge $U \supseteq B$. Dann gilt*

$$\int_B \operatorname{div}(f)\,\mathrm{dvol} = \int_{\partial B} f \cdot \mathrm{d}\mathbf{n}.$$

*Beweis.* Sei $Q \supseteq B$ ein Quadrat und sei $\mathfrak{Z}$ eine genügend feine Zerlegung von $Q$ in Quadrate, so dass die Aussage in Proposition 13.8 gilt (siehe das Bild 13.6). Wir verwenden nun den Divergenzsatz für Bereiche unter Graphen (Proposition 13.2, möglicherweise in rotierter Form) für jedes Puzzlestein $R = \overline{Q_{\boldsymbol{\alpha}}}$ und jedes Domino $R = \overline{Q_{\boldsymbol{\alpha}} \cup Q_{\boldsymbol{\beta}}}$ für (im Sinne von Proposition 13.8) äquivalente $Q_{\boldsymbol{\alpha}}, Q_{\boldsymbol{\beta}} \sqsubset \mathfrak{Z}$. In beiden Fällen erhalten wir

$$\int_{B\cap R} \operatorname{div}(f)\,\mathrm{dvol} = \int_{\partial(B\cap R)} f \cdot \mathrm{d}\mathbf{n}.$$

Die Definition des Flussintegrals über einen Rand ist so gewählt, dass bei Zusammenfügen der Teilbereiche $B \cap R$ für die Puzzlesteine $R$ oder Dominos $R$ die durch die Zerlegung künstlich erzeugten Flussintegrale zweimal auftreten, aber auf Grund der jeweils nach Aussen gewählten Orientierung der Normalenvektoren unterschiedliche Vorzeichen haben. Addieren wir alle Teile auf, so ergibt sich das Theorem, da $B$ bis auf eine Nullmenge durch die Mengen $B\cap R$ überdeckt wird und die verbleibenden Flussintegrale über $\partial B \cap R$ für die einzelnen Puzzlesteine und Dominos gemeinsam das Flussintegral über $\partial B$ ergeben. □

Theorem 13.14 gibt eine Gleichung über zwei Integrale an. Wie bereits zu Beginn des Kapitels erwähnt, kann diese Gleichung in beiden Richtungen nützlich sein, zum Beispiel um Rechnungen möglichst übersichtlich zu halten. Falls die Divergenz des Vektorfelds sehr einfach ist, dann ist es vielleicht von Vorteil das Flächenintegral zu berechnen.

**Beispiel 13.15** (Planimeter)**.** *Wir betrachten einen glatt berandeten, kompakter Bereich $B$ in $\mathbb{R}^2$ und das glatte Vektorfeld*

$$f : (x, y) \in \mathbb{R}^2 \mapsto \begin{pmatrix} x \\ y \end{pmatrix}.$$

*Dann gilt* $\operatorname{div}(f) = 2$ *und der Divergenzsatz ergibt*

$$2\,\mathrm{vol}(B) = \int_B \operatorname{div}(f)\,\mathrm{dvol} = \int_{\partial B} f \cdot \mathrm{d}\mathbf{n}.$$

*Ist also beispielsweise $\partial B$ durch eine geeignete Kurve $\gamma : [a, b] \to \partial B$ parametrisiert, so gilt*

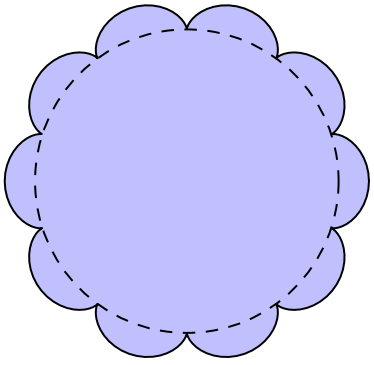
$$\mathrm{vol}(B) = \tfrac{1}{2} \int_a^b \left\langle \gamma(t), R^{-1}\dot{\gamma}(t) \right\rangle \mathrm{d}t,$$

*wobei $R \in \mathrm{SO}(2,\mathbb{R})$ wieder die Rotationsmatrix um 90 Grad im Gegenuhrzeigersinn bezeichnet. Wir erhalten also eine Methode, um die Fläche eines glatt berandeten, beschränkten Bereiches im $\mathbb{R}^2$ mit Hilfe eines Integrals einer Funktion $[a,b] \to \mathbb{R}$ zu berechnen (was ist die Fläche einer Kreisscheibe?). Diese (oder zumindest eine ähnliche) wird auch in der Praxis angewendet und nennt sich ein Planimeter – siehe auch diesen Link.*

**Beispiel 13.16** (Sonnenblume)**.** *Wir wollen in diesem Beispiel den Flächeninhalt einer „Sonnenblume", die von einem Kreis mit innerem Radius $r > 0$ und $n \geq 1$ „Blütenblättern" eingeschlossen wird, berechnen. Hier entstehen die Blütenblätter, indem wir einen Kreis mit Radius $\frac{r}{n}$ am inneren Kreis mit Radius $r$ abrollen und dabei einen Punkt auf dem sich bewegenden Kreis verfolgen. Formal betrachten wir also das Gebiet $S$, welches durch den positiv orientierten geschlossenen Weg*

$$\gamma : t \in [0,2\pi] \mapsto \left(r + \frac{r}{n}\right)\begin{pmatrix}\cos(t)\\ \sin(t)\end{pmatrix} + \frac{r}{n}\begin{pmatrix}\cos((n+1)t)\\ \sin((n+1)t)\end{pmatrix}$$

*eingeschlossen wird.*



*Auf Grund der Diskussion in Beispiel 13.15 können wir den Flächeninhalt $A$ von $S$ mittels des Planimeters berechnen. Mit Hilfe von*

$$R^{-1}\dot{\gamma}(t) = \left(r + \frac{r}{n}\right)\begin{pmatrix}\cos(t)\\ \sin(t)\end{pmatrix} + \frac{r(n+1)}{n}\begin{pmatrix}\cos((n+1)t)\\ \sin((n+1)t)\end{pmatrix}$$

*ergibt sich daraus*

$$\begin{aligned} A &= \frac{1}{2}\int_0^{2\pi} \left\langle \gamma(t), R^{-1}\dot{\gamma}(t)\right\rangle \,\mathrm{d}t \\ &= \frac{1}{2}\int_0^{2\pi} \left( \left(r + \frac{r}{n}\right)^2 \left\| \begin{pmatrix}\cos(t)\\ \sin(t)\end{pmatrix}\right\|^2 + \frac{r^2(n+1)}{n^2}\left\| \begin{pmatrix}\cos((n+1)t)\\ \sin((n+1)t)\end{pmatrix}\right\|^2 \right) \mathrm{d}t \\ &= \pi\left(r + \frac{r}{n}\right)^2 + \pi\frac{r^2(n+1)}{n^2}, \end{aligned}$$

*da die gemischten Terme verschwindendes Integral über $[0,2\pi]$ besitzen, zum Beispiel ist*

$$\int_0^{2\pi} \cos(t)\sin((n+1)t)\,\mathrm{d}t = 0.$$

*Üblicherweise wird der Rand $\partial S$ eine* Epizykloide *genannt und in dem Spezialfall mit nur einem Blütenblatt auch eine* Kardioide.

*Wir bemerken noch, dass wir – streng formal gesehen – den Divergenzsatz in der Form von Theorem 13.14 nicht auf die Sonnenblume anwenden dürfen, da diese nicht glatt berandet ist. Man kann jedoch die Spitzen zwischen den Blütenblättern sehr leicht abrunden ohne Flächen- oder Wegintegral wesentlich zu ändern, und daher stimmt die Aussage trotzdem. Wir werden dies und ähnliche Fälle nicht formalisieren.*

### 13.1.5 Rotation und der Satz von Green

**Definition 13.17** (Rotation eines Vektorfelds)**.** Sei $f : U \to \mathbb{R}^2$ ein stetig differenzierbares Vektorfeld auf einer offenen Menge $U \subseteq \mathbb{R}^2$. Die **Wirbelstärke** oder **Rotation** von $f$ ist durch

$$\operatorname{rot}(f)(u) = \lim_{r\to 0} \frac{1}{\pi r^2} \int_{\partial B_r(u)} f \cdot \mathrm{d}\mathbf{s} \tag{13.3}$$

für $u \in U$ definiert.

**Theorem 13.18** (Satz von Green)**.** *Sei $f : U \to \mathbb{R}^2$ ein stetig differenzierbares Vektorfeld auf einer offenen Menge $U \subseteq \mathbb{R}^2$. Die Wirbelstärke* $\operatorname{rot}(f)$ *existiert auf ganz $U$ und erfüllt*

$$\operatorname{rot}(f) = \partial_1 f_2 - \partial_2 f_1.$$

*Weiter gilt für jeden glatt berandeten, kompakten Bereich $B \subseteq U$*

$$\int_B \operatorname{rot}(f)\,\mathrm{dvol} = \int_{\partial B} f \cdot \mathrm{d}\mathbf{s}.$$

*Beweis.* Sei $R = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right) \in \mathrm{SO}(2,\mathbb{R})$ die Rotationsmatrix zum Winkel 90 Grad im Gegenuhrzeigersinn. Wir definieren ein Vektorfeld $g : U \to \mathbb{R}^2$ durch

$$g = R^{-1} f = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} f_2 \\ -f_1 \end{pmatrix}$$

wobei $f_1, f_2$ die Komponenten von $f$ darstellen. Dann gilt

$$\operatorname{div}(g) = \partial_1 g_1 + \partial_2 g_2 = \partial_1 f_2 - \partial_2 f_1$$

und nach Definition 13.11

$$\int_{\partial B} g \cdot \mathrm{d}\mathbf{n} = \int_{\partial B} (Rg) \cdot \mathrm{d}\mathbf{s} = \int_{\partial B} f \cdot \mathrm{d}s.$$

Nach dem Divergenzsatz (Theorem 13.14) angewendet auf $g$ folgt also

$$\int_B (\partial_1 f_2 - \partial_2 f_1)\,\mathrm{dvol} = \int_{\partial B} f \cdot \mathrm{d}s$$

für alle glatt berandeten Bereiche $B \subseteq U$. Wenden wir dies auf $B = \overline{B_r(p)} \subseteq U$ für $p \in U$ und ein hinreichend kleines $r > 0$ an und verwenden die Stetigkeit von $(\partial_2 f_1 - \partial_1 f_2)$ auf $U$, so erhalten wir wie im Beweis des Divergenzsatzes auf Rechtecken (Proposition 13.1) die Existenz des Grenzwerts in (13.3) und die Gleichung $\operatorname{rot}(f) = \partial_1 f_2 - \partial_2 f_1$. □

Wir sagen, dass ein stetig differenzierbares Vektorfeld auf einer offenen Menge **rotationsfrei** ist, falls $\operatorname{rot}(f) = 0$ gilt. In diesem Extremfall vereinfacht sich Theorem 13.18 zu einer Gleichung zwischen Wegintegralen. Man kann dies manchmal ausnützen, um anstatt einem komplizierten Wegintegral ein deutlich einfacheres Wegintegral zu berechnen (wenn beide zusammen eine Parametrisierung des Randes eines glatt berandeten Bereichs darstellen).

**Applet 13.19** (Divergenz und Rotation)**.** *Dieses Applet illustriert die Begriffe „rotationsfrei“ und „divergenzfrei“ sowie die Sätze dieses Abschnitts.*

### 13.1.6 Eine Anwendung: der Jordansche Kurvensatz für glatte Kurven*

Als Anwendung des Divergenzsatzes in der Ebene (Theorem 13.14) möchten wir in diesem Kapitel den Jordanschen Kurvensatz diskutieren. Die Frage, die dieser Satz beantwortet, ist im Wesentlichen die folgende. Angenommen man hat eine „komplizierte“ geschlossene Kurve in der Ebene gezeichnet. Teilt diese Kurve dann die Ebene in ein „Inneres“ und ein „Äusseres“ auf? Und wie kann man entscheiden, ob ein Punkt im Inneren oder im Äusseren liegt? Für letzteres kann man sich bereits folgenden naiven Algorithmus vorstellen: man beginnt bei dem gegebenen Punkt und zeichnet einen Weg durch das von der Kurve gegebene Labyrinth, welcher die Kurve nie schneidet. Kann man damit „nach Unendlich gehen“, so liegt man im Äusseren. Wenn nicht, so liegt man im Inneren. Dies ist zwar anschaulich klar, doch ist – je nach Struktur des ursprünglichen Weges – überraschend schwierig zu beweisen.

Bevor wir den allgemeinen Satz für glatte Wege diskutieren, möchten wir das wesentliche Beweishilfsmittel in einem Spezialfall diskutieren.

**Übung 13.20** (Umlaufzahl für Kreise)**.** *Wir betrachten zu $r > 0$ die Kurve*

$$\gamma_r : t \in [0, 2\pi] \mapsto r \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix},$$

*welche den Rand des Balles $B_r(0)$ parametrisiert. Für $u \notin \partial B_r(0)$ bezeichnet*

$$I_{\gamma_r}(u) = \tfrac{1}{2\pi} \int_0^{2\pi} \frac{\langle R(\gamma_r(t) - u), \dot{\gamma}_r(t) \rangle}{\|\gamma_r(t) - u\|^2} \, \mathrm{d}t$$

*die sogenannte* Umlaufzahl *von $\gamma_r$ um $u$, wobei $R \in \mathrm{SO}(2, \mathbb{R})$ die Rotationsmatrix um 90 Grad im Gegenuhrzeigersinn ist.*

*(i) Realisieren Sie $I_{\gamma_r}(u)$ für $u \notin \partial B_r(0)$ als Flussintegral eines Vektorfeldes über $\partial B_r(0)$.*

*(ii) Zeigen Sie, dass für alle $u \notin \partial B_r(0)$ die Umlaufzahl durch*

$$I_{\gamma_r}(u) = \begin{cases} 1 & \textit{falls } u \in B_r(0) \\ 0 & \textit{falls } u \notin B_r(0) \end{cases}$$

*gegeben ist.*

Für Kreise ist die Umlaufzahl also in der Lage zu entscheiden, ob ein Punkt innerhalb oder ausserhalb des Kreises liegt. Hat man nun eine allgemeinere Kurve gegeben, so ist es formal gesehen überhaupt nicht klar, ob so ein Inneres und ein Äusseres überhaupt existiert. Man kann nun aber umgekehrt die Umlaufzahl verwenden, um das Innere und das Äussere zu definieren.

**Satz 13.21** (Jordanscher Kurvensatz für glatte Funktionen)**.** *Sei $\gamma : [a, b] \to \mathbb{R}^2$ ein glatter, regulärer, einfacher, geschlossener Weg. Dann kann man das Komplement der Spur $\gamma([a, b])$ des Weges schreiben als*

$$\mathbb{R}^2 \setminus \gamma([a, b]) = \operatorname{Inn}(\gamma) \sqcup \operatorname{Auss}(\gamma),$$

*wobei das Innere $\operatorname{Inn}(\gamma)$ eine offene, beschränkte, zusammenhängende Teilmenge und das Äussere $\operatorname{Auss}(\gamma)$ eine offene, unbeschränkte, zusammenhängende Teilmenge ist. Des Weiteren gilt $\partial \operatorname{Inn}(\gamma) = \partial \operatorname{Auss}(\gamma) = \gamma([a, b])$.*

Wir verlangen von $\gamma$ also unter anderem, dass bis auf die Gleichung $\gamma(a) = \gamma(b)$ keine weiteren Selbstüberschneidungen stattfinden und dass $\dot\gamma(a) = \dot\gamma(b)$ ist (womit man $\gamma$ zu einer „periodischen“ stetig differenzierbaren Funktion auf $\mathbb{R}$ fortsetzen könnte).

Wir bemerken auch noch, dass der Jordansche Kurvensatz in einer viel grösseren Allgemeinheit gilt; die Annahme, dass $\gamma$ glatt und regulär, ist tatsächlich nicht notwendig. Unter diesen Annahmen sind wir allerdings in der Lage, den Beweis zu vereinfachen und, wie schon erwähnt, den Divergenzsatz anzuwenden.

Für den Beweis betrachten wir zu $u \in \mathbb{R}^2 \setminus \gamma([a, b])$ die **Umlaufzahl** von $\gamma$ um $u$

$$I_\gamma(u) = \tfrac{1}{2\pi} \int_a^b \frac{\langle R(\gamma(t) - u), \dot\gamma(t)\rangle}{\|\gamma(t) - u\|^2}\, \mathrm{d}t$$

genauer. Als erstes behaupten wir nun, dass $I_\gamma$ lokal konstant ist oder anders ausgedrückt, dass um jeden Punkt im Definitionsbereich eine Umgebung existiert, auf der $I_\gamma$ konstant ist. Insbesondere ist $I_\gamma$ auf allen zusammenhängenden Teilmengen konstant (wieso?).

**Übung 13.22** (Lokale Konstanz der Umlaufzahl)**.** *Zeigen Sie, dass $I_\gamma$ lokal konstant ist.*

Die obige Übung zeigt insbesondere, dass $I_\gamma$ ausserhalb eines genügend grossen Balles identisch verschwindet. In der Tat geht $I_\gamma(u)$ gegen Null für $\|u\| \to \infty$, da in der Definition von $I_\gamma(u)$ die Norm von $u$ im Nenner zur Potenz 2 und im Zähler zur Potenz 1 vorkommt (wieso?). Ist $B_M(0)$ ein Ball, der $\gamma([a, b])$ enthält, so muss $I_\gamma$ auf $\mathbb{R}^2 \setminus B_M(0)$ konstant sein (wegen obiger Übung) und gleichzeitig für wachsende Radien gegen Null gehen, was die Behauptung zeigt.

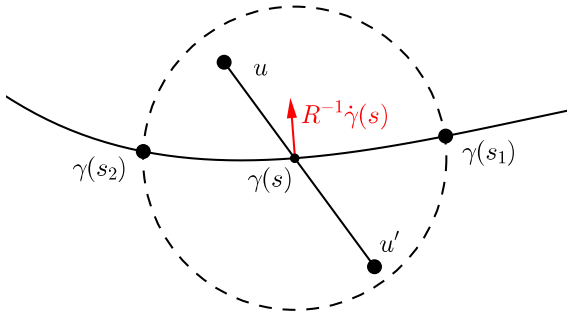
Die wesentliche Idee ist nun, von Aussen nach Innen zu gehen und dabei zu untersuchen, was passiert, wenn man jeweils die Spur $\gamma([a,b])$ von $\gamma$ überquert. Mit Hilfe des Divergenzsatzes kann man zeigen, dass die Umlaufzahl bei Sprüngen über die Spur sich um $\pm 1$ ändert, wobei das Vorzeichen davon abhängt, in welcher Richtung der Normalenvektor von $\gamma$ zeigt.

**Übung 13.23** (Überqueren der Spur)**.** *Seien $u, u' \in \mathbb{R}^2 \setminus \gamma([a,b])$, so dass das Geradensegment von $u$ nach $u'$ die Spur von $\gamma$ in genau einem Punkt $\gamma(s)$ schneidet. Zeigen Sie, dass*

$$I_\gamma(u') - I_\gamma(u) = \begin{cases} 1 & \textit{falls } \left\langle u - u', R^{-1}\dot\gamma(s)\right\rangle > 0 \\ -1 & \textit{falls } \left\langle u - u', R^{-1}\dot\gamma(s)\right\rangle < 0 \end{cases}$$

*gilt.*

*Hinweis: Da die Spur von $\gamma$ in genau einem Punkt geschnitten wird, lässt sich die Aussage auf folgendes Bild reduzieren:*




*Ersetzen Sie nun das Wegstück von $\gamma$ zwischen den (genügend nahe an $s$ gewählten) Zeiten $s_1, s_2$ durch das eine oder das andere Kreissegment wie im Bild.*

Geht man nun von weit aussen nach innen, so kann man mit obiger Übung zeigen, dass $I_\gamma$ auch Werte verschieden von Null annehmen muss.

**Übung 13.24.** *Zeigen Sie, dass es mindestens einen Punkt $u_0 \in \mathbb{R}^2 \setminus \gamma([a,b])$ mit $I_\gamma(u_0) = 1$ oder $I_\gamma(u_0) = -1$ gibt.*

Im Folgenden wollen wir der Einfachheit halber annehmen, dass es ein $u_0$ mit $I_\gamma(u_0) = 1$ gibt. Wir definieren nun die Kandidaten für das Innere und das Äussere von $\gamma$ durch

$$\begin{aligned} \mathrm{Inn}(\gamma) &= \left\{ u \in \mathbb{R}^2 \setminus \gamma([a,b]) \mid I_\gamma(u) = 1 \right\}, \\ \mathrm{Auss}(\gamma) &= \left\{ u \in \mathbb{R}^2 \setminus \gamma([a,b]) \mid I_\gamma(u) = 0 \right\}. \end{aligned}$$

**Übung 13.25.** *Zeigen Sie, dass die Mengen* $\mathrm{Inn}(\gamma)$ *und* $\mathrm{Auss}(\gamma)$ *offen und wegzusammenhängend sind.*

**Übung 13.26.** *Beweisen Sie den glatten Jordanschen Kurvensatz (Theorem 13.21).*

## 13.2 Oberflächenintegrale

In Analogie zum Divergenzsatz in der Ebene (Theorem 13.14) und dem Satz von Green (Theorem 13.18) möchten wir zwei Integralsätze im $\mathbb{R}^3$ in Abschnitten 13.3 und 13.5 formulieren. Dazu benötigen wir (man vergleiche mit Definition 13.11) den Begriff des Oberflächenintegrals.

### 13.2.1 Flächen und deren Karten

Eine **Fläche** $S \subseteq \mathbb{R}^3$ ist eine zweidimensionale Teilmannigfaltigkeit. Wir werden an zwei Arten von Flächen interessiert sein:

(1) $S$ ist der Rand eines kompakten, glatt berandeten Bereiches im $\mathbb{R}^3$. Beispielsweise könnte $S$ also die Sphäre $\mathbb{S}^2$ sein, welche den Rand des Einheitsballes im $\mathbb{R}^3$ darstellt.

(2) Der Abschluss von $S$ ist kompakt und ist eine „glatt berandete Teilmenge“ einer weiteren Fläche $M \supseteq \overline{S}$. Ein solches $S$ könnte beispielsweise die obere Hemisphäre von $\mathbb{S}^2$ sein. In diesem Fall nennt man $\overline{S}$ oft auch eine **Fläche mit Rand**.

VO: 11.05.18

Um die Fläche $S$ lokal mit Teilmengen von $\mathbb{R}^2$ zu beschreiben, werden wir Karten verwenden, an welche wir kurz erinnern werden (siehe auch die Diskussion nach Definition 11.10). Sei $p \in S$, sei $U_p \subseteq \mathbb{R}^3$ von $p$ eine offene Umgebung und sei $\varphi_p : U_p \to V_p = \varphi_p(U_p)$ ein Diffeomorphismus auf eine weitere offene Teilmenge $V_p \subseteq \mathbb{R}^3$ existiert, so dass

$$\varphi_p(U_p \cap S) = \{y \in V_p \mid y_3 = 0\} = V_p \cap (\mathbb{R}^2 \times \{0\}).$$

Dann ist eine **Karte** von $S$ um $p$ durch die offene Menge

$$\big\{(s,t) \in \mathbb{R}^2 \mid (s,t,0) \in V_p\big\}$$

gegeben. Die Abbildung $\Phi = \varphi_p^{-1}$ oder auch deren Einschränkung

$$(s,t) \in \big\{(s,t) \in \mathbb{R}^2 \mid (s,t,0) \in V_p\big\} \mapsto \varphi_p^{-1}(s,t,0) \in S$$

auf die Karte wird dann **Kartenabbildung** genannt. Zur Vereinfachung der Notation werden wir die Koordinatenebene $\mathbb{R}^2 \times \{0\}$ oft mit $\mathbb{R}^2$ identifizieren. Insbesondere schreiben wir zum Beispiel $V_p \cap \mathbb{R}^2$ für obige Karte und identifizieren $(s,t)$ mit $(s,t,0)$. Die Kartenabbildungen werden wir meist mit $\Phi$ (und einem Index) bezeichnen. In der Literatur wird oft auch das Tupel $(V_p \cap \mathbb{R}^2, \varphi_p^{-1}|_{V_p \cap \mathbb{R}^2})$ als Karte bezeichnet.

Auf Grund der Kompaktheit von $\overline{S}$ in beiden Fällen (1) und (2) (und der Enthaltung $\overline{S} \subseteq M$ in (2)) können wir $S$ durch endlich viele offene Mengen $U_1, \ldots, U_L \subseteq \mathbb{R}^3$ überdecken, so dass es endlich viele offene Mengen $V_1, \ldots, V_L$ und Kartenabbildungen $\Phi_\ell : V_\ell \to U_\ell$ gibt mit

$$\Phi_\ell(V_\ell \cap \mathbb{R}^2) = U_\ell \cap S$$

für $\ell = 1, \dots, L$. Die Liste von Kartenabbildungen $\Phi_1, \dots, \Phi_L$ werden wir auch eine **Parametrisierung** von $S$ nennen.

Gegeben zwei Karten $V_{\ell_1} \cap \mathbb{R}^2$ und $V_{\ell_2} \cap \mathbb{R}^2$ kann man auch den **Kartenwechsel**

$$\Phi_{\ell_2}^{-1} \circ \Phi_{\ell_1}|_{\Phi_{\ell_1}^{-1}(U_{\ell_2}) \cap \mathbb{R}^2} : V_{\ell_1} \cap \Phi_{\ell_1}^{-1}(U_{\ell_2}) \cap \mathbb{R}^2 \to V_{\ell_2} \cap \Phi_{\ell_2}^{-1}(U_{\ell_1}) \cap \mathbb{R}^2$$

betrachten. Dieser gibt zu einem Punkt $(s,t)$ in der Karte $V_{\ell_1} \cap \mathbb{R}^2$ jeweils an, welcher Punkt dem Punkt $p = \Phi_{\ell_1}(s,t)$ auf der Fläche in der Karte $V_{\ell_2} \cap \mathbb{R}^2$ zugeordnet wird (falls $p$ überhaupt im Bild der Karte $V_{\ell_2} \cap \mathbb{R}^2$ liegt).

Wir wollen des Weiteren für jedes $\ell \in \{1, \dots, L\}$ eine (im $\mathbb{R}^2$) Jordan-messbare Teilmenge $K_\ell \subseteq V_\ell \cap \mathbb{R}^2$ wählen, so dass

$$S = \bigsqcup_{\ell=1}^{L} \Phi_\ell(K_\ell). \tag{13.4}$$

Obige **Kartenpartition** von $S$ in die Bilder von $K_1, \dots, K_L$ erlaubt uns eine vollständige und eindeutige Parametrisierung der Fläche: Für jeden Punkt $u \in S$ gibt es eine Seitennummer $\ell$ im „Atlas“ und eine eindeutige Koordinate $(s,t) \in K_\ell$ mit $\Phi_\ell(s,t) = u$. Es hilft vielleicht sich die Mengen $\Phi_\ell(K_\ell)$ als ein Land in der Fläche $S$ vorzustellen, wobei $K_\ell \subseteq V_\ell \cap \mathbb{R}^2$ der Darstellung des Landes in der Karte $V_\ell \cap \mathbb{R}^2$ entspricht. Bei der Integration über $S$ sollte also jedes Land einen Beitrag zum Gesamtintegral leisten.

In Anwendungen oder auch in theoretischen Überlegungen kann es teilweise nützlich sein, die Annahme in (13.4) etwas abzuschwächen. Falls für je zwei verschiedene $\ell_1, \ell_2 \in \{1, \dots, L\}$ der Schnitt $\Phi_{\ell_1}(K_{\ell_1}) \cap \Phi_{\ell_2}(K_{\ell_2})$ die Vereinigung einer endlichen Menge und einer eindimensionalen Teilmannigfaltigkeit ist, so spielt dies für das folgende keine Rolle. Selbiges trifft auch zu, wenn $S \setminus \bigcup_{\ell=1}^{L} \Phi_\ell(K_\ell)$ die Vereinigung einer endlichen Menge mit einer eindimensionalen Teilmannigfaltigkeit ist.

### 13.2.2 Skalares Oberflächenintegral

Für eine Fläche $S$ mit einer Parametrisierung wie oben und eine stetige, beschränkte Funktion $f : S \to \mathbb{R}$ definieren wir das **skalare Oberflächenintegral** durch

$$\int_S f \,\mathrm{d}A = \sum_{\ell=1}^{L} \int_{\Phi_\ell(K_\ell)} f \,\mathrm{d}A = \sum_{\ell=1}^{L} \int_{K_\ell} (f \circ \Phi) \|\partial_s \Phi_\ell \times \partial_t \Phi_\ell\| \,\mathrm{d}s \,\mathrm{d}t$$

Hierbei stellt $\mathrm{d}A$ das „Flächenelement“ dar und $\times$ bezeichnet das **Kreuzprodukt** (siehe Übung 13.27). Für eine der Kartenabbildungen $\Phi_\ell$ interpretieren wir $\mathrm{d}A$ als $\|\partial_s \Phi_\ell \times \partial_t \Phi_\ell\| \,\mathrm{d}s \,\mathrm{d}t$. Des Weiteren schreiben wir $\mathrm{d}s \,\mathrm{d}t$ anstatt $\mathrm{dvol}(s,t)$ für die Integration über die „Kartenebene“ $\mathbb{R}^2$, um Verwirrungen mit dem Volumen im dreidimensionalen zu vermeiden.

Dies ist formal gesehen bloss eine Definition und Notation für diese. Trotzdem ist es nützlich, sich die geometrische Bedeutung der auftretenden Ausdrücke klar zu machen, welche im

Spezialfall der konstanten Funktion $f = 1$ zu einer weiteren Definition führen: Der **Flächeninhalt** von $S$ ist

$$\int_S \mathrm{d}A = \sum_{\ell=1}^{L} \int_{K_\ell} \|\partial_s \Phi_\ell \times \partial_t \Phi_\ell\| \,\mathrm{d}s\,\mathrm{d}t.$$

**Übung 13.27** (Das Kreuzprodukt)**.** *Wir erinnern daran, dass das Kreuzprodukt $u \times v$ zweier Vektoren*

$$u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \; v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \in \mathbb{R}^3$$

*durch*

$$u \times v = \begin{pmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{pmatrix} \in \mathbb{R}^3$$

*definiert wird.*

*(i) Zeigen Sie, dass der Vektor $u \times v$ auf $u$ und auf $v$ normal steht.*

*(ii) Zeigen Sie, dass die Länge $\|u \times v\|$ gleich der Wurzel der Gramschen Determinante*

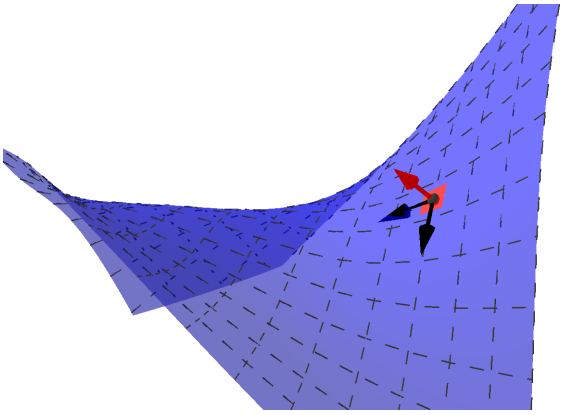
$$\det \begin{pmatrix} \langle u, u \rangle & \langle u, v \rangle \\ \langle v, u \rangle & \langle v, v \rangle \end{pmatrix}$$

*ist.*

Insbesondere zeigen Übung 13.27 und Korollar 12.54, dass die Länge $\|u \times v\|$ des Kreuzprodukts zweier Vektoren $u, v \in \mathbb{R}^3$ als der Flächeninhalt des von $u, v$ aufgespannten Parallelogramms interpretiert werden kann. Betrachten wir dies im Zusammenhang mit einer Parametrisierung $(s,t) \in V \cap \mathbb{R}^2 \mapsto \Phi(s,t) \in S$, so erhalten wir eine Interpretation von

$$\|\partial_s \Phi_\ell \,\mathrm{d}s \times \partial_t \Phi_\ell \,\mathrm{d}t\| = \|\partial_s \Phi_\ell \times \partial_t \Phi_\ell\| \,\mathrm{d}s\,\mathrm{d}t$$

als der Flächeninhalt eines „sehr kleinen“ oder sogar „infinitesimalen“ Teilstücks der Fläche $S$, siehe auch folgendes Bild.



**Beispiel 13.28.** *Wir definieren die Fläche $S$ als den Graphen von*

$$(x, y) \in B_1(0) \mapsto z = x^2 + y^2.$$

*Damit ist der Flächeninhalt von $S$ definiert als das Integral*

$$\begin{aligned} A &= \int_{B_1(0)} \left\| \begin{pmatrix} 1 \\ 0 \\ 2x \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ 2y \end{pmatrix} \right\| \mathrm{d}x\,\mathrm{d}y \\ &= \int_{B_1(0)} \left\| \begin{pmatrix} -2x \\ -2y \\ 1 \end{pmatrix} \right\| \mathrm{d}x\,\mathrm{d}y \\ &= \int_{B_1(0)} \sqrt{1 + 4(x^2 + y^2)}\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_0^{2\pi} \int_0^1 \sqrt{1 + 4r^2}\, r\,\mathrm{d}r\,\mathrm{d}\varphi, \end{aligned}$$

*welches sich mit der eindimensionalen Substitution $u = 1 + 4r^2$ schnell berechnen lässt.*

**Übung 13.29.** *Zeigen Sie, dass das skalare Oberflächenintegral über eine kompakte Fläche nicht von der Wahl der Parametrisierung abhängt.*

**Übung 13.30.** *Berechnen Sie den Flächeninhalt der Sphäre $\mathbb{S}^2$.*

### 13.2.3 Flussintegrale entlang Oberflächen

Im Kontext der Integralsätze im dreidimensionalen Raum wird der Begriff des Flussintegrals entlang Oberflächen für uns wichtiger sein als obiger Begriff des skalaren Oberflächenintegrals. Fundamental für das Flussintegral wird die Annahme sein, dass die gegebene Fläche $S \subseteq \mathbb{R}^3$ eine Orientierung besitzt. Intuitiv ausgedrückt existiert für die Fläche $S$ eine Orientierung, falls man bei jedem Punkt in $S$ einen (nicht-verschwindenden) Normalenvektor wählen kann, so dass dieser stetig vom Punkt abhängt. Falls $S = \partial B$ der Rand eines glatt berandeten Bereiches ist, so wollen wir eine Orientierung von $S$ so wählen, dass der Normalenvektor bei einem Punkt nach aussen zeigt.

Sei also $S \subseteq \mathbb{R}^3$ eine Fläche, sei $\Phi_1 : V_1 \to U_1, \ldots, \Phi_L : V_L \to U_L$ eine Parametrisierung von $S$ wie im letzten Abschnitt. Wir sagen, dass die Parametrisierung $\Phi_1, \ldots, \Phi_L$ **orientiert** ist (oder eine **Orientierung** von $S$ definiert), falls für alle $\ell_1, \ell_2 \in \{1, \ldots, L\}$ der zugehörige Kartenwechsel orientierungserhaltend ist, das heisst,

$$\det\Big(\mathrm{D}\big(\Phi_{\ell_1}^{-1} \circ \Phi_{\ell_2}|_{\Phi_{\ell_2}^{-1}(U_{\ell_1})\cap\mathbb{R}^2}\big)\Big) > 0.$$

Des Weiteren nennen wir $S$ **orientierbar**, falls $S$ eine orientierte Parametrisierung besitzt.

Wir bemerken, dass in dieser Definition die Determinante einer $2 \times 2$-Matrix genommen wird: Denn wir betrachten die Einschränkung von $\Phi_{\ell_2}$ auf die zwei-dimensional offene Teilmenge $\Phi_{\ell_2}^{-1}(U_{\ell_1}) \cap \mathbb{R}^2 \subseteq \mathbb{R}^2$ und die Abbildung $\Phi_{\ell_1}^{-1} \circ \Phi_{\ell_2}$ nimmt dort (wegen $\Phi_\ell(V_\ell \cap \mathbb{R}^2) = U_\ell \cap S$ für $\ell \in \{\ell_1, \ell_2\}$) Werte in $\mathbb{R}^2 \subseteq \mathbb{R}^3$ an. Wir vergessen also die letzte (verschwindende) Komponente und betrachten $\psi = \Phi_{\ell_1}^{-1} \circ \Phi_{\ell_2}|_{\Phi_{\ell_2}^{-1}(U_{\ell_1})\cap\mathbb{R}^2}$ als einen Diffeomorphismus zwischen offenen Teilmengen im $\mathbb{R}^2$.

Anschaulich bedeutet die Positivität der Jacobi-Determinante, dass wir die Fläche zweimal von der gleichen Seite her betrachten (und nicht etwa einmal von oben und einmal von unten). Formal gesehen, wird dies eine wichtige Rolle im Beweis von Lemma 13.37 spielen. Dort werden wir das Kreuzprodukt von partiellen Ableitungen von zwei verschiedenen Kartenabbildungen vergleichen, wobei Positivität bewirken wird, dass die zwei Normalenvektoren in die gleiche Richtung zeigen (also positive Vielfache von einander sind).

Alle Flächen, die wir bisher kennengelernt haben, wie zum Beispiel die Sphäre oder der Torus, sind orientierbar.

**Übung 13.31.** *Zeigen Sie, dass die Sphäre $\mathbb{S}^2 \subseteq \mathbb{R}^3$ orientierbar ist.*

**Wichtige Übung 13.32** (Eine Charakterisierung von Orientierbarkeit)**.** *Sei $S \subseteq \mathbb{R}^3$ eine Fläche. Ein stetiges **normiertes Normalenfeld** ist eine stetige Abbildung $n : S \to \mathbb{R}^3$ mit $n(p) \perp \mathrm{T}_pS$ und $\|n(p)\| = 1$ für alle $p \in S$. Zeigen Sie, dass $S$ genau dann orientierbar ist, falls ein stetiges Normalenfeld existiert.*

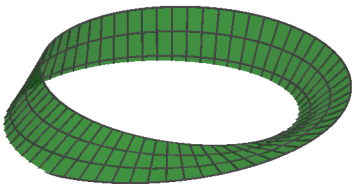
Obige Übung liefert eine nützliche Charakterisierung von Orientierbarkeit. Beispielsweise ist daraus sofort klar, dass die Sphäre orientierbar ist, da ein stetiges Normalenfeld durch

$$n : \mathbb{S}^2 \to \mathbb{R}^3,\ p \mapsto p$$

gegeben ist.

Es gibt auch Flächen, die nicht orientierbar sind.

**Beispiel 13.33** (Das Möbius-Band)**.** *Das Möbius-Band $M \subseteq \mathbb{R}^3$ erhält man, indem man einen rechteckigen Streifen Papier nach einer Drehung an einer der kurzen Enden zusammenklebt.*



*Die so entstehende Fläche hat keine Orientierung. Wenn eine Ameise auf einer Seite den Streifen entlang geht, kommt diese nach einer Umdrehung zuerst am gleichen Ort auf dem Papier, aber auf der anderen Seite vorbei. Erst nach zwei Umdrehungen gelangt sie wieder an den gleichen Ort und die gleiche Seite. Wenn man nun den Fluss einer Strömung durch das Blatt berechnen will, so stellt man fest, dass diese Frage gar keinen Sinn ergibt. An jeder Stelle auf dem Blatt, wo auch immer die Ameise ist, kann man die Ameise verwenden, um die Oberseite oder den Normalenvektor zur Oberfläche an dieser Stelle zu definieren. Wandert die Ameise weiter, so kennt man die Orientierung des Normalenvektors nebenan und beginnt die einzelnen Anteile des Flusses durch die Oberfläche aufzuaddieren. Da aber die Ameise nach einer Umdrehung auf der anderen Seite des Blattes wiedererscheint, sehen wir, dass wir den Fluss durch den ursprünglichen Teil des Blattes von sich selbst subtrahieren sollten.*

*Wir könnten dieses Problem natürlich beseitigen, indem wir das Blatt der kurzen Seite entlang aufschneiden. Dann haben wir aber die Frage geändert, weil wir das Möbiusband durch das ursprüngliche (vielleicht verdrehte, aber nicht zusammengeklebte) Blatt ersetzt haben. Dieses ist orientierbar und das Flussintegral (siehe unten) macht Sinn, hängt aber im Allgemeinen vom gewählten Ort des Durchschneidens des Möbiusbands ab.*

**Applet 13.34** (Nicht-Orientierbarkeit des Möbiusbandes).

Die Situation ist, wenn $S$ der Rand eines glatt berandeten Bereiches ist, allerdings deutlich besser.

**Lemma 13.35** (Existenz einer Aussennormalen). *Sei $B \subseteq \mathbb{R}^3$ ein kompakter glatt berandeter Bereich. Dann existiert eine orientierte Parametrisierung $\Phi_1, \ldots, \Phi_L$ des Randes $S = \partial B$, so dass $\partial_s \Phi_\ell \times \partial_t \Phi_\ell$ eine Aussennormale ist. Das heisst, für alle $\ell \in \{1, \ldots, L\}$ und $(s,t) \in V_\ell$ gilt*

$$\begin{aligned} \Phi_\ell(s,t) + \varepsilon(\partial_s \Phi_\ell \times \partial_t \Phi_\ell)(s,t) &\notin B \\ \Phi_\ell(s,t) - \varepsilon(\partial_s \Phi_\ell \times \partial_t \Phi_\ell)(s,t) &\in B^\circ, \end{aligned}$$

*wenn nur $\varepsilon > 0$ genügend klein ist.*

*Beweis*.* Sei $Q = [-a,a]^3$ ein Würfel, der $B$ enthält, sei $\eta_0 > 0$ geeignet wie in Lemma 13.7 und sei $\mathfrak{Z}$ eine Zerlegung von $Q$ mit Maschenweite kleiner als $\frac{\eta_0}{2}$. Sei $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$ ein $\mathfrak{Z}$ entsprechender Würfel, der $\partial B$ schneidet, und sei $p_{\boldsymbol{\alpha}} \in Q_{\boldsymbol{\alpha}} \cap \partial B$. Wir nehmen zuerst an, dass

$$B^\infty_{\eta_0}(p_{\boldsymbol{\alpha}}) \cap B = \left\{(x,y,z)^t \in B^\infty_{\eta_0}(p_{\boldsymbol{\alpha}}) \mid z \leq \varphi(x,y)\right\}$$

für eine glatte Funktion $\varphi$. In diesem Fall definieren wir

$$U_{\boldsymbol{\alpha}} = \left\{(x,y,z)^t \in B^\infty_{\eta_0}(p_{\boldsymbol{\alpha}}) \mid (x,y,\varphi(x,y))^t \in B^\infty_{\eta_0}(p_{\boldsymbol{\alpha}})\right\}$$

und die Abbildung

$$\Psi_{\boldsymbol{\alpha}} : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in U_{\boldsymbol{\alpha}} \mapsto \begin{pmatrix} x \\ y \\ z - \varphi(x,y) \end{pmatrix}.$$

Man beachte dabei, dass das Bild von $U_{\boldsymbol{\alpha}} \cap B$ unter $\Psi_{\boldsymbol{\alpha}}$ gerade aus jenen Punkten im Bild $V_{\boldsymbol{\alpha}} = \Psi_{\boldsymbol{\alpha}}(U_{\boldsymbol{\alpha}})$ besteht, die nicht-positive $z$-Koordinate haben. Sei weiter

$$\Phi_{\boldsymbol{\alpha}} = \Psi_{\boldsymbol{\alpha}}^{-1} : V_{\boldsymbol{\alpha}} \to U_{\boldsymbol{\alpha}}, \quad \begin{pmatrix} s \\ t \\ w \end{pmatrix} \mapsto \begin{pmatrix} s \\ t \\ w + \varphi(s,t) \end{pmatrix}$$

die Inverse. Daher gilt

$$\partial_s \Phi_{\boldsymbol{\alpha}} = \begin{pmatrix} 1 \\ 0 \\ \partial_s \varphi \end{pmatrix}, \ \partial_t \Phi_{\boldsymbol{\alpha}} = \begin{pmatrix} 0 \\ 1 \\ \partial_t \varphi \end{pmatrix}, \ \partial_s \Phi_{\boldsymbol{\alpha}} \times \partial_t \Phi_{\boldsymbol{\alpha}} = \begin{pmatrix} -\partial_s \varphi \\ -\partial_t \varphi \\ 1 \end{pmatrix}.$$

Da letzteres mit dem Gradienten der Funktion $f : (x,y,z)^t \mapsto z - \varphi(x,y)$ übereinstimmt, deren Niveaumenge zum Wert 0 lokal den Rand von $B$ beschreibt, und der Gradient somit senkrecht auf der Oberfläche stehen muss, ist $\partial_s \Phi_{\boldsymbol{\alpha}}(s,t) \times \partial_t \Phi_{\boldsymbol{\alpha}}(s,t)$ ein Normalenvektor bei $\Phi_{\boldsymbol{\alpha}}(s,t)$. Des Weiteren gilt

$$f\Big(\Phi_{\boldsymbol{\alpha}}(s,t) + \varepsilon \partial_s \Phi_{\boldsymbol{\alpha}}(s,t) \times \partial_t \Phi_{\boldsymbol{\alpha}}(s,t)\Big) > 0$$

für alle hinreichend kleinen $\varepsilon > 0$ und damit zeigt $\partial_s \Phi_{\boldsymbol{\alpha}}(s,t,0) \times \partial_t \Phi_{\boldsymbol{\alpha}}(s,t)$ nach aussen.

Falls eine Koordinatenvertauschung notwendig ist, um $B^\infty_{\eta_0}(p_{\boldsymbol{\alpha}}) \cap B$ mittels einer stetig differenzierbaren Funktion $\varphi$ zu beschreiben, oder auch falls $B^\infty_{\eta_0}(p_{\boldsymbol{\alpha}}) \cap B$ als Teilmenge oberhalb statt unterhalb der Graphen von $\varphi$ gegeben ist, so können wir obiges „Flachdrücken“ mit einer Koordinatenvertauschung respektive einem Koordinatenvorzeichenwechel verknüpfen. Damit können wir erreichen, dass die Ableitung des Diffeomorphismus $\Phi_{\boldsymbol{\alpha}} : V_{\boldsymbol{\alpha}} \to U_{\boldsymbol{\alpha}}$ stets (wie im ersten Fall, den wir betrachtet haben) Determinante 1 hat, die Punkte $(s,t,w)^t \in V_{\boldsymbol{\alpha}}$ mit $w \leq 0$ nach $B$ abgebildet werden und die Punkte $(s,t,w)^t \in V_{\boldsymbol{\alpha}}$ mit $w > 0$ nicht nach $B$ abgebildet werden.

Seien nun zwei Kartenabbildungen $\Phi_{\boldsymbol{\alpha}} : V_{\boldsymbol{\alpha}} \to U_{\boldsymbol{\alpha}}$ und $\Phi_{\boldsymbol{\beta}} : V_{\boldsymbol{\beta}} \to U_{\boldsymbol{\beta}}$, deren Bilder sich in der offenen Menge $U_{\boldsymbol{\alpha}} \cap U_{\boldsymbol{\beta}} \neq \emptyset$ schneiden. Dann hat die Ableitung des Kartenwechsel $\Phi_{\boldsymbol{\beta}}^{-1} \circ \Phi_{\boldsymbol{\alpha}}$ Determinante 1 wegen obigem und ist von der Form $\left(\begin{smallmatrix} A & * \\ 0 & b \end{smallmatrix}\right)$ für $A \in \mathrm{GL}_2(\mathbb{R})$ und $b \in \mathbb{R}^\times$. Auch nach obigem muss aber $b > 0$ und somit $\det(A) = \frac{1}{b} > 0$ gelten. Dies ergibt aber somit genau eine orientierte Parametrisierung, da die offenen Mengen $U_{\boldsymbol{\alpha}}$ für $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$ zusammen $\partial B$

überdecken. □

**Definition 13.36** (Flussintegral durch Flächen)**.** Sei $S \subseteq \mathbb{R}^3$ eine orientierbare Fläche, sei $\Phi_1 : V_1 \to U_1, \ldots, \Phi_L : V_L \to U_L$ eine orientierte Parametrisierung von $S$ und sei $K_1 \subseteq V_1 \cap \mathbb{R}^2, \ldots, K_L \subseteq V_L \cap \mathbb{R}^2$ eine Kartenpartition für diese Parametrisierung. Sei $U \supseteq S$ eine offene Menge und sei $f$ ein stetiges Vektorfeld auf $U$. Dann definieren wir das **Flussintegral** von $f$ über $S$ durch

$$\int_S f \cdot \mathrm{d}\mathbf{n} = \sum_{\ell=1}^{L} \int_{K_\ell} \langle f \circ \Phi_\ell, \partial_s \Phi_\ell \times \partial_t \Phi_\ell \rangle \, \mathrm{d}s \, \mathrm{d}t.$$

**Lemma 13.37** (Wohldefiniertheit des Flussintegrals)**.** *Das Flussintegral über eine orientierbare, zusammenhängende Fläche hängt, abgesehen vom Vorzeichen, nicht von der gewählten orientierten Parametrisierung ab. Für den Rand $S = \partial B$ eines kompakten glatt berandeten Bereiches und einem Atlas wie in Lemma 13.35 gilt sogar, dass das Flussintegral (inklusive dem Vorzeichen) wohldefiniert ist.*

*Beweis.* Gegeben seien zwei orientierte Parametrisierungen $\Phi_\ell : V_\ell \to U_\ell$ für $\ell \in \{1, \ldots, L\}$ und $\Psi_m : \tilde{V}_m \to \tilde{U}_m$ für $m \in \{1, \ldots, M\}$ sowie Karten $K_\ell \subseteq V_\ell \cap \mathbb{R}^2$ für alle $\ell \in \{1, \ldots, L\}$ und $\tilde{K}_m \subseteq \tilde{V}_m \cap \mathbb{R}^2$ für alle $m \in \{1, \ldots, M\}$.

Wir nehmen nun ohne Beschränkung der Allgemeinheit an, dass

$$\Phi_1(V_1 \cap \mathbb{R}^2) \cap \Psi_1(\tilde{V}_1 \cap \mathbb{R}^2) = U_1 \cap \tilde{U}_1 \cap S \neq \emptyset$$

gilt, und wählen ein $p_0 \in U_1 \cap \tilde{U}_1 \cap S$. Falls die Einschränkung der Abbildung

$$\Phi_1^{-1} \circ \Psi_1 : \Psi_1^{-1}(U_1 \cap \tilde{U}_1) \to \Phi_1^{-1}(U_1 \cap \tilde{U}_1)$$

auf den Schnitt $\mathbb{R}^2 \cap \Psi_1^{-1}(U_1 \cap \tilde{U}_1)$ mit der Koordinatenebene in $\Psi_1^{-1}(p_0)$ negative Jacobi-Determinante hat, verknüpfen wir $\Psi_1, \ldots, \Psi_M$ mit einer Koordinatenvertauschung in $\mathbb{R}^2$ und können dadurch erreichen, dass die Jacobi-Determinante positiv ist.

Wir sagen, dass die beiden Parametrisierungen in $p \in S$ „kompatibel“ sind, falls es ein $\ell$ und ein $m$ gibt, so dass $p$ in $U_\ell \cap \tilde{U}_m$ liegt und die Jacobi-Determinante der Einschränkungen von $\Phi_\ell^{-1} \circ \Psi_m$ auf $\mathbb{R}^2 \cap \Psi_m^{-1}(U_\ell \cap \tilde{U}_m)$ positive Jacobi-Determinante besitzt. Da beide Parametrisierungen nach Annahme orientiert sind, hängt diese Eigenschaft weder von der Wahl von $\ell$ noch von der Wahl von $m$ ab. Wir definieren

$$O = \{p \in S \mid \text{die Parametrisierungen sind in } p \text{ kompatibel}\}.$$

Nach obigem können wir also annehmen, dass $p_0 \in O$ liegt. Da die Ableitungen der Kartenabbildungen stetig sind, folgt aus der Definition, dass $O$ eine offene Teilmenge von $S$ ist. Aus demselben Argument folgt aber auch, dass $S \setminus O$ offen ist (wieso?). Der angenommene Zusammenhang von $S$ impliziert nun, dass $O = S$ sein muss, da $O$ nicht leer ist.

Wir vereinfachen nun die Notation und nehmen an, dass $\Phi : V \to U$ und $\tilde{\Phi} : \tilde{V} \to \tilde{U} = U$ Diffeomorphismen sind, welche lokal $S$ parametrisieren und für die die Jacobi-Determinante

der Abbildung

$$\psi = \Phi^{-1} \circ \tilde{\Phi}|_{V \cap \mathbb{R}^2} : V \cap \mathbb{R}^2 \to \tilde{V} \cap \mathbb{R}^2$$

positiv ist. Weiter seien $K \subseteq V \cap \mathbb{R}^2$ und $\tilde{K} \subseteq \tilde{V} \cap \mathbb{R}^2$ Karten zu $S$ mit $\Phi(K) = \tilde{\Phi}(\tilde{K})$. Für ein stetiges Vektorfeld gilt dann auf Grund der Substitutionsregel für $\psi$ (Satz 12.56), dass

$$\begin{aligned}\int_K \langle f \circ \Phi, \partial_s \Phi \times \partial_t \Phi \rangle \, \mathrm{d}s \, \mathrm{d}t &= \int_{\tilde{K}} \langle f \circ \Phi, \partial_s \Phi \times \partial_t \Phi \rangle \circ \psi |\det(\mathrm{D}\psi)| \, \mathrm{d}\tilde{s} \, \mathrm{d}\tilde{t} \\ &= \int_{\tilde{K}} \Big\langle f \circ \tilde{\Phi}, \underbrace{(\partial_s \Phi \times \partial_t \Phi) \circ \psi \det(\mathrm{D}\psi)}_{(*)} \Big\rangle \, \mathrm{d}\tilde{s} \, \mathrm{d}\tilde{t},\end{aligned}$$

wobei wir $(\tilde{s}, \tilde{t})$ für die Koordinaten in $\tilde{V} \cap \mathbb{R}^2$ geschrieben haben. Wir behaupten nun, dass $(*)$ gleich $\partial_{\tilde{s}} \tilde{\Phi} \times \partial_{\tilde{t}} \tilde{\Phi}$ ist. In der Tat gilt

$$\Big(\partial_{\tilde{s}} \tilde{\Phi} \, , \, \partial_{\tilde{t}} \tilde{\Phi}\Big) = \Big(\partial_s \Phi \circ \psi \, , \, \partial_t \Phi \circ \psi\Big) \, \mathrm{D}\psi$$

auf Grund der mehrdimensionalen Kettenregel (Satz 10.13 – wieso?). Wir vereinfachen nun die Notation weiter und betrachten $\tilde{u}, \tilde{v}, u, v \in \mathbb{R}^3$ und $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(\mathbb{R})$ mit

$$(\tilde{u}, \tilde{v}) = (u, v)A.$$

Wir haben also $\tilde{u} = au + cv$ und $\tilde{v} = bu + dv$, womit nach der Bilinearität und Antisymmetrie des Kreuzprodukts gilt

$$\begin{aligned}\tilde{u} \times \tilde{v} &= (au + cv) \times (bu + dv) \\ &= ab(u \times u) + ad(u \times v) + bc(v \times u) + cd(v \times v) \\ &= ad(u \times v) + bc(v \times u) \\ &= (ad - bc)u \times v = \det(A) u \times v\end{aligned}$$

Es folgt daher, dass $(*)$ tatsächlich gleich $\partial_{\tilde{s}} \tilde{\Phi} \times \partial_{\tilde{t}} \tilde{\Phi}$ ist, und somit gilt

$$\int_K \langle f \circ \Phi, \partial_s \Phi \times \partial_t \Phi \rangle \, \mathrm{d}s \, \mathrm{d}t = \int_{\tilde{K}} \Big\langle f \circ \tilde{\Phi}, \partial_{\tilde{s}} \tilde{\Phi} \times \partial_{\tilde{t}} \tilde{\Phi} \Big\rangle \, \mathrm{d}\tilde{s} \, \mathrm{d}\tilde{t}$$

Wir können dies nun für alle Mengen der Form $K = \Phi_\ell^{-1}(\Phi_\ell(K_\ell) \cap \tilde{\Phi}_m(K_m))$ und $\tilde{K} = \tilde{\Phi}_\ell^{-1}(\Phi_\ell(K_\ell) \cap \tilde{\Phi}_m(K_m))$ anwenden, um die erste Aussage vom Lemma zu beweisen.

Die letzte Aussage des Lemmas folgt analog wobei man verwendet, dass zwei Parametrisierungen einer offenen Umgebung eines Punktes $p \in S = \partial B$, die beide einen Aussennormalenvektor von $B$ bei $p$ als Kreuzprodukt der partiellen Ableitungen besitzen, kompatibel sein müssen. □

**Beispiel 13.38.** *Wir betrachten nochmals die Fläche $S$ definiert als den Graphen von*

$$(x, y) \in B_1(0) \mapsto z = x^2 + y^2.$$

*Sei des Weiteren das Vektorfeld $f : \mathbb{R}^3 \to \mathbb{R}^3$ definiert durch*

$$f\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} z \\ x^2 \\ y \end{pmatrix}.$$

*Damit ist das Flussintegral durch $S$ definiert als das Integral*

$$\begin{aligned} A &= \int_{B_1(0)} \left\langle f\left(\begin{pmatrix} x \\ y \\ x^2+y^2 \end{pmatrix}\right), \begin{pmatrix} -2x \\ -2y \\ 1 \end{pmatrix} \right\rangle \mathrm{d}x\,\mathrm{d}y \\ &= \int_{B_1(0)} \big((x^2+y^2)(-2x) + x^2(-2y) + y\big)\,\mathrm{d}x\,\mathrm{d}y, \end{aligned}$$

*welches auf Grund der vorhandenen Symmetrien gleich Null ist.*

VO: 14.05.18

## 13.3 Der Divergenzsatz im dreidimensionalen Raum

**Definition 13.39** (Divergenz)**.** Sei $U \subseteq \mathbb{R}^3$ eine offene Menge und sei $f : U \to \mathbb{R}^3$ ein stetig differenzierbares Vektorfeld. Die **Divergenz** von $f$ bei $p \in U$ ist definiert als

$$\operatorname{div}(f)(p) = \lim_{h \to 0} \tfrac{1}{8h^3} \int_{\partial B_h^\infty(p)} f \cdot \mathrm{d}\mathbf{n},$$

falls der Grenzwert existiert.

Man beachte, dass dies die analoge Definition ist wie im zweidimensionalen (siehe Proposition 13.1), da das Volumen von $B_h^\infty(p)$ gerade $8h^3$ ist. Des Weiteren sind die Aussennormalen des Quaders $B_h^\infty(p)$ sowie das Integral $\int_{\partial B_h^\infty(p)} f \cdot \mathrm{d}\mathbf{n}$ in Analogie zum zweidimensionalen Fall definiert.

### 13.3.1 Der Divergenzsatz für Bereiche unter Graphen

**Proposition 13.40** (Erste Version des Divergenzsatzes)**.** *Sei $f : U \to \mathbb{R}^3$ ein stetig differenzierbares Vektorfeld auf einer offenen Menge $U \subseteq \mathbb{R}^3$. Dann existiert die Divergenz* $\operatorname{div}(f)$ *auf ganz $U$ und ist durch*

$$\operatorname{div}(f) = \partial_1 f_1 + \partial_2 f_2 + \partial_3 f_3$$

*gegeben. Sei weiter $Q \subseteq \mathbb{R}^2$ ein abgeschlossenes Rechteck und sei $\varphi : Q \to [z_0, \infty)$ eine stetig differenzierbare Funktion, so dass der abgeschlossene Bereich unter dem Graphen von $\varphi$*

$$B = \left\{(x, y, z)^t \in \mathbb{R}^3 \mid (x, y) \in Q,\ z_0 \leq z \leq \varphi(x, y)\right\}$$

*in $U$ liegt. Dann gilt*

$$\int_B \operatorname{div}(f)\,\mathrm{dvol} = \int_{\partial B} f \cdot \mathrm{d}\mathbf{n}.$$

Dabei ist $\int_{\partial B} f \cdot \mathrm{d}\mathbf{n}$ das Flussintegral bezüglich der Aussennormalen und wird wie in Proposition 13.2 als Summe der Flussintegrale über die Seitenflächen von $B$ definiert (siehe auch den Beweis unten für eine explizite Darstellung).

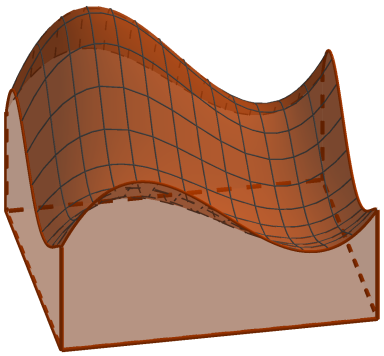
*Beweis.* Wir beweisen zuerst für ein abgeschlossenes Rechteck $Q = [a, b] \times [c, d]$ und einen Bereich $B$ definiert durch eine Funktion $\varphi : Q \to [z_0, \infty)$ die Gleichheit

$$\int_{\partial B} f \cdot \mathrm{d}\mathbf{n} = \int_B (\partial_1 f_1 + \partial_2 f_2 + \partial_3 f_3)\,\mathrm{dvol}.$$

Der Rand $\partial B$ von $B$ hat 6 Seiten: „links“ ($x = a$), „rechts“ ($x = b$), „vorne“ ($y = c$), „hinten“ ($y = d$), „unten“ ($z = z_0$) und „oben“. Der obere Teil des Randes ist durch $(x, y) \in Q$ parametrisiert, da er durch $\{(x, y, \varphi(x, y)) \mid (x, y) \in Q\}$ gegeben ist. Der Normalenvektor an

dem durch $(x, y) \in Q$ gegebenen Punkt ist somit

$$\begin{pmatrix} 1 \\ 0 \\ \partial_x \varphi(x,y) \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ \partial_y \varphi(x,y) \end{pmatrix} = \begin{pmatrix} -\partial_x \varphi(x,y) \\ -\partial_y \varphi(x,y) \\ 1 \end{pmatrix}$$



Wir betrachten drei Spezialfälle für das Vektorfeld $f$, aus welchen wir dann den allgemeinen Fall erhalten.

Falls das Vektorfeld $f$ von der Form $f = \left(\begin{smallmatrix} 0 \\ 0 \\ f_3 \end{smallmatrix}\right)$ ist, so gilt

$$\begin{aligned} \int_{\partial B} f \cdot \mathrm{d}\mathbf{n} &= \int_Q \big(f_3(x,y,\varphi(x,y)) - f_3(x,y,z_0)\big)\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_Q \int_{z_0}^{\varphi(x,y)} \partial_3 f_3(x,y,z)\,\mathrm{d}z\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_B \partial_3 f_3(x,y,z)\,\mathrm{dvol}(x,y,z) \end{aligned}$$

auf Grund des Fundamentalsatzes der Differential- und Integralrechnung (Korollar 8.5) und des Satzes von Fubini (in der Form von Korollar 12.43).

Falls das Vektorfeld $f$ von der Form $f = \left(\begin{smallmatrix} f_1 \\ 0 \\ f_3 \end{smallmatrix}\right)$, so verschwinden die Integrale „vorne“ und „hinten“ und jedes der verbleibenden Integrale enthält die Integration über $y \in [c, d]$. Verwenden wir für die verbleibenden vier Integral die entsprechenden Aussennormalen ergibt sich damit

$$\begin{aligned} \int_{\partial B} f \cdot \mathrm{d}\mathbf{n} = \int_c^d \left( \int_a^b \left\langle \begin{pmatrix} f_1(x,y,\varphi(x,y)) \\ 0 \\ f_3(x,y,\varphi(x,y)) \end{pmatrix}, \begin{pmatrix} -\partial_x \varphi(x,y) \\ -\partial_y \varphi(x,y) \\ 1 \end{pmatrix} \right\rangle - \left\langle \begin{pmatrix} f_1(x,y,z_0) \\ 0 \\ f_3(x,y,z_0) \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle \mathrm{d}x \right) \mathrm{d}y \\ + \int_c^d \left( \int_{z_0}^{\varphi(b,y)} \left\langle \begin{pmatrix} f_1(b,y,z) \\ 0 \\ f_3(b,y,z) \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle \mathrm{d}z - \int_{z_0}^{\varphi(a,y)} \left\langle \begin{pmatrix} f_1(a,y,z) \\ 0 \\ f_3(a,y,z) \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle \mathrm{d}z \right) \mathrm{d}y \end{aligned}$$

nach Definition des Flussintegrals über $\partial B$ und dem Satz von Fubini (Theorem 12.39). Wir löschen nun aus den inneren Produkten die zweite Koordinate, da diese verschwindenden Beitrag zur Summe haben, und interpretieren die inneren eindimensionalen Integrale als das

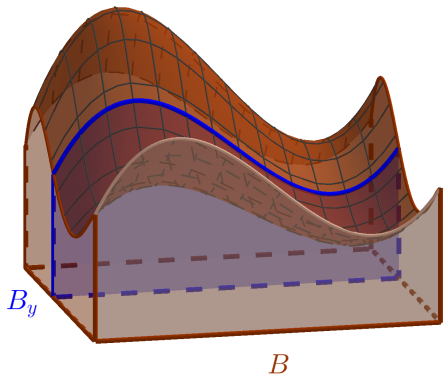
Flussintegral über $\partial B_y$, wobei

$$B_y = B \cap (\mathbb{R} \times \{y\} \times \mathbb{R}) \widehat{=} \{(x,z) \mid a \leq x \leq b \text{ und } z_0 \leq z \leq \varphi(x,y)\}\,.$$

Also erhalten wir

$$\begin{aligned}\int_{\partial B} f \cdot \mathrm{d}\mathbf{n} &= \int_c^d \int_{\partial B_y} \begin{pmatrix} f_1(x,y,z) \\ f_3(x,y,z) \end{pmatrix} \cdot \mathrm{d}\mathbf{n}(x,z)\,\mathrm{d}y \\ &= \int_c^d \int_{B_y} (\partial_1 f_1 + \partial_3 f_3)\,\mathrm{dvol}(x,z)\,\mathrm{d}y\end{aligned}$$

auf Grund des zweidimensionalen Divergenzsatz (Proposition 13.2) für Bereiche unter Graphen.




Figur 13.8: Tomografie in den Ebenen $\{(x', y', z') : y' = y\}$.

Nochmalige Anwendung des Satzes von Fubini ergibt nun

$$\int_{\partial B} f \cdot \mathrm{d}\mathbf{n} = \int_B (\partial_1 f_1 + \partial_3 f_3)\,\mathrm{dvol}$$

wie gewünscht.

Falls das Vektorfeld $f$ von der Form $f = \begin{pmatrix} 0 \\ f_2 \\ f_3 \end{pmatrix}$ ist, so erhält man analog

$$\int_{\partial B} f \cdot \mathrm{d}\mathbf{n} = \int_B (\partial_2 f_2 + \partial_3 f_3)\,\mathrm{dvol}$$

mittels Tomografie in den Ebenen $\{(x', y', z') : x' = x\}$.

Ist das Vektorfeld $f$ nun beliebig, so können $f$ schreiben als

$$f = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} = \begin{pmatrix} f_1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ f_2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ f_3 \end{pmatrix}$$

und es ergibt sich mit Linearität

$$\int_{\partial B} f \cdot \mathrm{d}\mathbf{n} = \int_B (\partial_1 f_1 + \partial_2 f_2 + \partial_3 f_3)\,\mathrm{dvol}$$

wie behauptet.

Für $B = B_h^\infty(p)$ können wir dies gemeinsam mit der Stetigkeit von $\partial_1 f_1 + \partial_2 f_2 + \partial_3 f_3$ bei $p$ verwenden, um

$$\tfrac{1}{8h^3}\int_{\partial B_h^\infty(p)} f \cdot \mathrm{d}\mathbf{n} = \tfrac{1}{8h^3}\int_{B_h^\infty(p)} (\partial_1 f_1 + \partial_2 f_2 + \partial_3 f_3)\,\mathrm{dvol} \to (\partial_1 f_1 + \partial_2 f_2 + \partial_3 f_3)(p)$$

für $h \to 0$ zu folgern. □

### 13.3.2 Der Divergenzsatz auf glatt berandeten Bereichen

**Theorem 13.41** (Divergenzsatz im dreidimensionalen Raum – Satz von Gauss)**.** *Sei $B$ ein kompakter, glatt berandeter Bereich und $f$ ein stetig differenzierbares Vektorfeld definiert auf einer offenen Obermenge von $B$. Dann gilt*

$$\int_B \mathrm{div}(f)\,\mathrm{dvol} = \int_{\partial B} f \cdot \mathrm{d}\mathbf{n}.$$

Dabei stellt $\int_{\partial B} f \cdot \mathrm{d}\mathbf{n}$ das Flussintegral bezüglich der Aussennormalen wie in Definition 13.36 dar.

*Beweis*.* Sei $\eta_0 > 0$ geeignet wie in Lemma 13.7, sei $Q$ ein Würfel, der $B$ enthält, und sei $\mathfrak{Z}$ eine Zerlegung von $Q$ in Würfel mit Maschenweite kleiner als $\frac{\eta_0}{3}$. Wir wollen wiederum den Satz auf den Fall eines Bereiches unter einem Graphen zurückführen. Allerdings ist es technisch einfacher, nicht den Bereich $B$ wie im Beweis von Theorem 13.14 in Puzzleblöcke zu zerlegen, sondern das Vektorfeld $f$ in „lokale Teile“ $f_{\boldsymbol{\alpha}}$ für alle $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$ zu zerlegen. Hierbei müssen wir die Vektorfelder $f_{\boldsymbol{\alpha}}$ aber so wählen, dass

$$f = \sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} f_{\boldsymbol{\alpha}}$$

gilt und dass diese Funktionen ebenfalls stetig differenzierbar sind. Wenn der Satz dann bereits für jedes $f_{\boldsymbol{\alpha}}$ gilt, dann impliziert Linearität den Satz wiederum auch für $f$.

Wir definieren die Funktionen $f_{\boldsymbol{\alpha}}$ mittels einer sogenannten „glatten Partition der Eins“. Hierfür benötigen wir eine glatte Funktion $\psi_1 : \mathbb{R} \to [0,1]$ mit $\psi_1(t) = 1$ für alle $t \in [-1,1]$ und $\psi_1(t) = 0$ für alle $t$ mit $|t| \geq \frac{3}{2}$ (siehe Beispiel 7.23 und Übung 7.24). Mit dieser können wir dann eine glatte Funktion $\psi_3$ auf $\mathbb{R}^3$ durch

$$\psi_3 : (x,y,z) \in \mathbb{R}^3 \mapsto \psi_1(x)\psi_1(y)\psi_1(z) \in \mathbb{R}$$

definieren, welche auf $B_1^\infty(0)$ gleich 1 ist und ausserhalb von $B_{\frac{3}{2}}^\infty(0)$ verschwindet. Wir verschieben und strecken diese Funktion, so dass wir für jedes $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$ eine Funktion $\psi_{3,\boldsymbol{\alpha}} : \mathbb{R}^3 \to \mathbb{R}$

finden, die auf $Q_{\boldsymbol{\alpha}}$ gleich 1 ist und Träger innerhalb eines grösseren offenen Quaders $U_{\boldsymbol{\alpha}}$ hat, wobei $B \cap U_{\boldsymbol{\alpha}}$ durch eine stetig differenzierbare Funktion beschrieben werden kann (im Sinne von Definition 13.4). Wir definieren nun

$$\psi_{\boldsymbol{\alpha}} = \left( \sum_{Q_{\boldsymbol{\beta}} \sqsubset \mathfrak{Z}} \psi_{3,\boldsymbol{\beta}} \right)^{-1} \psi_{3,\boldsymbol{\alpha}}$$

und erhalten glatte Funktionen auf einer offenen Obermenge $U$ von $Q$, welche Werte in $[0,1]$ annehmen und

$$\sum_{Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}} \psi_{\boldsymbol{\alpha}} = 1$$

auf ganz $U$ erfüllen. Des Weiteren definieren wir

$$f_{\boldsymbol{\alpha}} = \psi_{\boldsymbol{\alpha}} f$$

für alle $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$.

Falls $U_{\boldsymbol{\alpha}} \subseteq B^\circ$ für ein $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$, so gilt

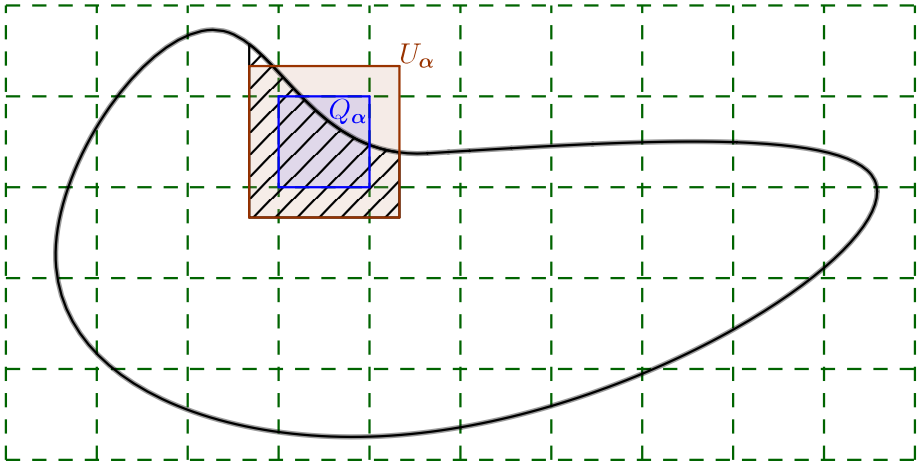
$$\int_B \operatorname{div}(f_{\boldsymbol{\alpha}}) \operatorname{dvol} = \int_{U_{\boldsymbol{\alpha}}} \operatorname{div}(f_{\boldsymbol{\alpha}}) \operatorname{dvol} = \int_{\partial U_{\boldsymbol{\alpha}}} f_{\boldsymbol{\alpha}} \cdot \mathrm{d}\mathbf{n} = 0 = \int_{\partial B} f_{\boldsymbol{\alpha}} \cdot \mathrm{d}\mathbf{n},$$

da $f_{\boldsymbol{\alpha}}$ Träger innerhalb von $U_{\boldsymbol{\alpha}}$ hat.

Falls hingegen $U_{\boldsymbol{\alpha}} \cap \partial B \neq \emptyset$ für ein $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$ ist, so erhalten wir

$$\int_B \operatorname{div}(f_{\boldsymbol{\alpha}}) \operatorname{dvol} = \int_{B \cap U_{\boldsymbol{\alpha}}} \operatorname{div}(f_{\boldsymbol{\alpha}}) \operatorname{dvol} = \int_{B_{\boldsymbol{\alpha}}} \operatorname{div}(f_{\boldsymbol{\alpha}}) \operatorname{dvol},$$

wobei wir $B_{\boldsymbol{\alpha}}$ (nach Vertauschung der Koordinaten wenn nötig) als den Bereich unter dem Graphen einer stetig differenzierbaren Funktion so definieren, dass $B_{\boldsymbol{\alpha}} \cap U_{\boldsymbol{\alpha}} = B \cap U_{\boldsymbol{\alpha}}$ gilt.



Auf Grund von Proposition 13.40 gilt dann

$$\int_{B_{\boldsymbol{\alpha}}} \operatorname{div}(f_{\boldsymbol{\alpha}})\,\mathrm{dvol} = \int_{\partial B_{\boldsymbol{\alpha}}} f_{\boldsymbol{\alpha}} \cdot \mathrm{d}\mathbf{n} = \int_{\partial B \cap U_{\boldsymbol{\alpha}}} f_{\boldsymbol{\alpha}} \cdot \mathrm{d}\mathbf{n} = \int_{\partial B} f_{\boldsymbol{\alpha}} \cdot \mathrm{d}\mathbf{n},$$

da $\partial B_{\boldsymbol{\alpha}} \cap U_{\boldsymbol{\alpha}} = \partial B \cap U_{\boldsymbol{\alpha}}$ und $f_{\boldsymbol{\alpha}}$ Träger innerhalb von $U_{\boldsymbol{\alpha}}$ hat und somit die Integrale über die Ränder von $\partial B_{\boldsymbol{\alpha}}$ und von $\partial B$ ausserhalb von $U_{\boldsymbol{\alpha}}$ verschwinden. Summieren wir nun diese Formeln über alle $Q_{\boldsymbol{\alpha}} \sqsubset \mathfrak{Z}$, die $B$ schneiden, so folgt das Theorem. $\square$

*Bemerkung.* Der Divergenzsatz funktioniert ebenso für $\mathbb{R}^n$ und $n \geq 4$. Doch erfordert dieser eine Definition des Flussintegrals über den Rand eines $n$-dimensionalen glatt berandeten Bereichs. Hierzu ist der Begriff der alternierenden Tensoren und der Differentialformen äusserst nützlich, worauf wir später noch einmal kurz zu sprechen kommen werden.

VO: 16.05.18

## 13.4 Der Laplace-Operator*

Wir wollen hier einen weiteren wichtigen Differentialoperator einführen, welcher aber im Unterschied zu allen bisher betrachteten Operatoren wie $\nabla, \operatorname{div}, \operatorname{rot}$ für eine reellwertige zweimal differenzierbare Funktion auf einer offenen Menge $U \subseteq \mathbb{R}^n$ eine weitere reellwertige Funktion auf $U$ definiert.

**Definition 13.42** (Laplace-Operator)**.** Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ zweimal stetig differenzierbar. Dann definieren wir für $p \in U$

$$\Delta f(p) = \lim_{r \searrow 0} \tfrac{c_n}{r^{n+2}} \int_{B_r(p)} \big(f(p+h) - f(p)\big)\, \mathrm{d}h$$

falls der Grenzwert existiert. Es ist üblich die Konstante $c_n$ mittels $c_n = \frac{2(n+2)}{\operatorname{vol}(B_1(0))}$ zu definieren.

In Worten ausgedrückt vergleicht also der Laplace-Operator den Funktionswert in einem Punkt mit den Funktionswerten in kleinen Umgebungen um den Punkt.

**Proposition 13.43** (Laplace-Operator als Differentialoperator)**.** *Sei $U \subseteq \mathbb{R}^n$ offen und sei $F : U \to \mathbb{R}$ zweimal stetig differenzierbar. Dann existiert $\Delta F$ auf ganz $U$ und erfüllt*

$$\Delta F = \sum_{j=1}^{n} \partial_j^2 F.$$

*Beweis.* Zur Vereinfachung der Notation betrachten wir ohne Beschränkung der Allgemeinheit den Punkt $p = 0 \in U$. Wir verwenden die Taylor-Approximation zweiter Ordnung in Korollar 10.25 und erhalten für $r > 0$

$$\int_{B_r(0)} \big(F(h) - F(0)\big)\, \mathrm{dvol}(h) = \int_{B_r(0)} \big(\langle \nabla F(0), h\rangle + \tfrac{1}{2} h^t H(0) h + o(r^2)\big)\, \mathrm{dvol}(h),$$

wobei $H(0)$ die Hesse-Matrix der zweiten Ableitungen von $F$ in 0 darstellt. Auf Grund der Symmetrie von $B_r(0)$ gilt aber

$$\int_{B_r(0)} h_j \, \mathrm{dvol}(h) = 0,$$

womit auch

$$\int_{B_r(0)} \langle \nabla F(0), h\rangle \, \mathrm{dvol}(h) = 0$$

ist. Dies gilt ebenso für die gemischten quadratischen Terme, das heisst $\int_{B_r(0)} h_j h_k \, \mathrm{dvol}(h) = 0$ für $j \neq k$. Daher ist

$$\int_{B_r(0)} \big(F(h) - F(0)\big)\, \mathrm{d}h \, \mathrm{dvol}(h) = \tfrac{1}{2} \sum_{j=1}^{n} \partial_j^2 F(0) \int_{B_r(0)} h_j^2 \, \mathrm{dvol}(h) + \operatorname{vol}(B_r(0)) o(r^2).$$

Des Weiteren gilt

$$\begin{aligned}\int_{B_r(0)} h_j^2 \,\mathrm{dvol}(h) &= \tfrac{1}{n}\sum_{k=1}^{n}\int_{B_r(0)} h_k^2 \,\mathrm{dvol}(h) = \tfrac{1}{n}\int_{B_r(0)} \|h\|^2 \,\mathrm{dvol}(h)\\ &= \tfrac{1}{n}\int_{B_1(0)} r^2\|u\|^2 r^n \,\mathrm{dvol}(u) = Cr^{n+2}\end{aligned}$$

auf Grund der Substitution $h = ru$ für $u \in B_1(0)$, wobei $C = C(n) = \frac{1}{n}\int_{B_1(0)} \|u\|^2 \,\mathrm{dvol}(u)$. Daher existiert der Grenzwert in Definition von $\Delta F(0)$ und es gilt

$$\lim_{r\searrow 0} \tfrac{1}{r^{n+2}} \int_{B_r(0)} \big(F(h) - F(0)\big)\,\mathrm{d}h = \tfrac{C}{2}\sum_{j=1}^{n} \partial_j^2 F(0).$$

Um den Wert der (von $n$ abhängigen) Konstante $C$ zu berechnen, beginnen wir mit

$$C = \frac{1}{n}\int_{B_1(0)} \|u\|^2 \,\mathrm{dvol}(u) = \frac{1}{n}\sum_{k=1}^{K}\int_{B_k} \|u\|^2 \,\mathrm{dvol}(u)$$

für $B_k = B_{\frac{k}{K}}(0) \setminus B_{\frac{k-1}{K}}(0)$. Nun schätzen wir für $u \in B_k$ die Zahl $\|u\|^2$ durch $\left(\frac{k}{K}\right)^2$ ab und erhalten mit dem Mittelwertsatz für $g(t) = t^n$ und dem Satz über Riemann-Summen (Satz 5.79)

$$\begin{aligned}C &\le \frac{1}{n}\sum_{k=1}^{K}\int_{B_k}\left(\frac{k}{K}\right)^2 \mathrm{dvol}(h) = \frac{1}{n}\sum_{k=1}^{K}\left(\frac{k}{K}\right)^2 \underbrace{\left(\left(\frac{k}{K}\right)^n - \left(\frac{k-1}{K}\right)^n\right)}_{=g\left(\frac{k}{K}\right)-g\left(\frac{k-1}{K}\right)} \mathrm{vol}(B_1(0))\\ &\le \frac{1}{n}\sum_{k=1}^{K}\left(\frac{k}{K}\right)^2 n\left(\frac{k}{K}\right)^{n-1}\frac{1}{K}\,\mathrm{vol}(B_1(0))\\ &= \mathrm{vol}(B_1(0))\sum_{k=1}^{K}\left(\frac{k}{K}\right)^{n+1}\frac{1}{K} \to \mathrm{vol}(B_1(0))\int_0^1 t^{n+1}\,\mathrm{d}t = \frac{\mathrm{vol}(B_1(0))}{n+2}\end{aligned}$$

für $K \to \infty$. Auf dieselbe Art und Weise ergibt sich aber auch die umgekehrte Ungleichung und damit die Proposition. □

### 13.4.1 Schwingungen der Trommel

Mit der Formel $\Delta F = \sum_{j=1}^{n} \partial_j^2 F$ ergibt sich unmittelbar der Zusammenhang $\Delta F = \mathrm{div}(\nabla F)$ des Laplace-Operators zum Gradienten und zum Divergenzoperator. Die Definition mittels dem asymptotischen Verhalten von $F$ bei Mittelung über kleine Bälle hat aber auch Vorteile. Erstens ist es recht schnell klar, dass sich $\Delta$ bei der Rotation des Koordinatensystems nicht ändert. Zweitens erklärt diese Definition, warum der Laplace-Operator so oft in (partiellen) Differentialgleichungen vorkommt. Nehmen wir zum Beispiel an, dass die Funktion $F : \overline{B_1(0)} \to \mathbb{R}$ am Rand verschwindet und die vertikale Auslenkung einer elastischen Membran über $B_1(0) \subseteq \mathbb{R}^2$ beschreibt – dies wird auch eine Trommel genannt. Dann

bestimmt das Mittel

$$\int_{B_r(p)} (F(p+h) - F(p))\,\mathrm{dvol}(h),$$

ob der Punkt $(p, F(p))$ auf der ausgelenkten Membran im Schnitt über oder unter anderen Punkten auf der Membran in der Nähe liegt. Dies bestimmt die Richtung (nach oben oder nach unten) und die Stärke der Krafteinwirkung auf den Punkt $(p, F(p))$. Doch welchen Radius sollte man hier verwenden, vor allem wenn der Punkt näher an den Rand der Trommel rutscht? Um eine vernünftige Theorie zu entwickeln, betrachtet man statt dem Mittel über eine Nachbarschaft mit vorgegebenem Radius $r$ den Laplace-Operator. Die Wellengleichung für die Beschreibung der Auslenkung $F(p,t)$ der Trommel über $p \in \overline{B_1(0)}$ zum Zeitpunkt $t \in \mathbb{R}$ hat damit nach geeigneter Wahl der physikalischen Einheiten die Form

$$\partial_t^2 F(p,t) = \Delta_p F(p,t),$$

wobei $\Delta_p F(p,t) = (\partial_1^2 F + \partial_2^2 F)(p,t)$ den Laplace-Operator in den Raumkoordinaten bezeichnet.

Können wir mit unseren Methoden mehrere mögliche Bewegungen der Trommel beschreiben? Die positive Antwort ist in folgender Übung enthalten.

**Übung 13.44** (Laplace-Eigenfunktionen auf der Kreisscheibe)**.**

*(a) Zeigen Sie, dass der Laplace-Operator in Polarkoordinaten die Form $\Delta = \partial_r^2 + \frac{1}{r}\partial_r + \frac{1}{r^2}\partial_\theta^2$ annimmt.*

*(b) Angenommen $g$ ist eine glatte Funktion auf $B_1(0)$, welche am Rand von $B_1(0)$ verschwindet und welche der Gleichung $\Delta g = \lambda g$ für ein $\lambda < 0$ genügt. Zeigen Sie, dass die Funktion $F$ definiert durch $F(p,t) = g(p)\sin(|\lambda|^{\frac{1}{2}}t)$ eine Lösung der Wellengleichung darstellt.*

*(c) Angenommen $g$ ist von der Form $g(r,\varphi) = J_n(r)\sin(n\varphi)$ für ein $n \in \mathbb{N}_0$ und löst $\Delta g = \lambda g$ für ein $\lambda < 0$. Welche Differentialgleichung muss $J_n$ damit erfüllen?*

*(d) Verknüpfen Sie die obigen Beobachtungen mit unserer Diskussion der Bessel-Differentialgleichung in Abschnitt 10.5.1.*

### 13.4.2 Harmonische Funktionen

Falls wir nun die Vorgabe am Rand ändern, also die Membran am Rand in eine bestimmte Form zwingen, so stellt sich die Frage, welche Form die Membran im Innern annehmen wird. Zuerst kann es natürlich zu Schwingungen der Membran kommen, doch existieren in der realen Welt Reibungskräfte, womit die Membran nach einiger Zeit zu Ruhe kommt. In dieser Ruhelage sollte also an keinem Ort eine Krafteinwirkung vorhanden sein, das heisst, die Funktion $F(p,t) = g(p)$ sollte der **Dirichlet-Gleichung**

$$\Delta_p F = \Delta g = 0$$

genügen. In diesem Fall nennen wir die Funktion $g$ auch **harmonisch**. Aus dieser Diskussion leitet sich die Vermutung ab, dass es für jede stetige, vorgegebene Funktion auf dem Rand $\partial B_1(0)$ eine eindeutig bestimmte, stetige harmonische Funktion gibt, deren Einschränkung am Rand die vorgegebene Funktion ist. Dies wird das **Dirichlet-Problem** genannt. Mittels einer Seifenlauge und einem geschlossenen Stück Draht lässt sich dieses Problem sehr schnell für alle möglichen Vorgaben am Rand lösen, doch die mathematische Lösung des Problems geht weit über diese Vorlesung hinaus. Wir begnügen uns mit folgender Anwendung der Divergenzsätze.

**Korollar 13.45** (Mittelwerteigenschaft harmonischer Funktionen)**.** *Sei $U \subseteq \mathbb{R}^n$ für $n = 2$ oder $n = 3$ offen und $g : U \to \mathbb{R}$ harmonisch (das heisst, $\Delta g = 0$). Dann gilt für jeden abgeschlossenen Ball $\overline{B_r(p)} \subseteq U$ die Mittelwerteigenschaft*

$$\frac{1}{\operatorname{vol}(\partial B_r(p))} \int_{\partial B_r(p)} g(h) \operatorname{dvol}(h) = g(p).$$

*Beweis für $n = 2$.* Zur Vereinfachung der Notation nehmen wir wieder an, dass $p = 0$ ist. Wir definieren die Hilfsfunktion

$$x \mapsto \nu(x) = \log(\|x\|)$$

Diese ist überall glatt ausser beim Ursprung und es gilt

$$\nabla\nu(x, y) = \frac{1}{x^2 + y^2} \begin{pmatrix} x \\ y \end{pmatrix}, \qquad \Delta\nu(x, y) = 0$$

für alle $(x, y)^t \in \mathbb{R}^2 \setminus \{0\}$. Des Weiteren betrachten wir das Vektorfeld

$$f = g\nabla\nu - \nu\nabla g,$$

für welches gilt

$$\operatorname{div}(f) = \big(\langle\nabla g, \nabla\nu\rangle + g\Delta\nu\big) - \big(\langle\nabla g, \nabla\nu\rangle + (\Delta g)\nu\big) = 0$$

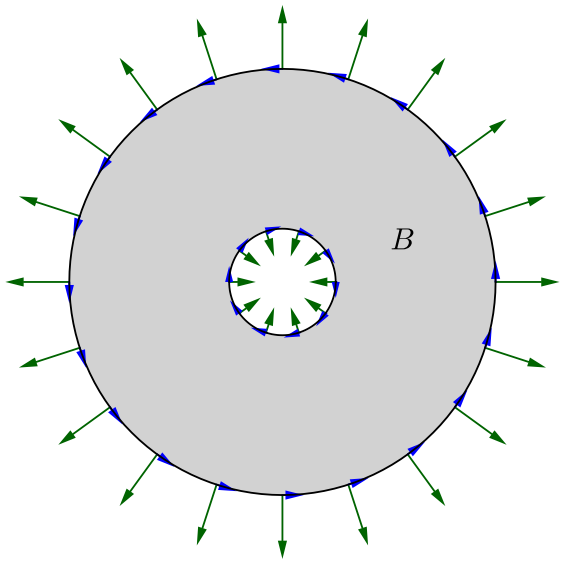
auf $\mathbb{R}^2 \setminus \{0\}$, da $g$ harmonisch ist. Wir verwenden nun den Divergenzsatz in der Ebene (Theorem 13.14) für den Kreisring $B = \overline{B_r(0)} \setminus B_\varepsilon(0)$ für ein $\varepsilon > 0$ mit $\varepsilon < r$. Damit erhalten wir für den Weg

$$\gamma_r : t \in [0, 2\pi] \mapsto r(\cos(t), \sin(t))^t$$

mit Ableitung $\gamma'(t) = r(-\sin(t), \cos(t))^t$ und Aussennormale $\mathbf{n}(t) = \gamma_r(t)$ die Gleichung

$$\int_B \operatorname{div}(f) \operatorname{dvol} = 0 = \int_0^{2\pi} \langle f(\gamma_r(t)), \gamma_r(t)\rangle \, \mathrm{d}t - \int_0^{2\pi} \langle f(\gamma_\varepsilon(t)), \gamma_\varepsilon(t)\rangle \, \mathrm{d}t,$$

wobei $\gamma_\varepsilon$ wie $\gamma_r$ definiert ist und auf Grund der Orientierung negativ in obige Gleichung eingeht; siehe auch folgendes Bild.




Wir betrachten zuerst das erste Integral, indem wir die Definition von $f$ einsetzen und erhalten

$$\begin{aligned}\int_0^{2\pi} \langle f(\gamma_r(t)), \gamma_r(t)\rangle \,\mathrm{d}t &= \int_0^{2\pi} g(\gamma_r(t))\underbrace{\langle \nabla\nu(\gamma_r(t)), \gamma_r(t)\rangle}_{=1} \,\mathrm{d}t - \int_0^{2\pi} \log(r)\, \langle \nabla g(\gamma_r(t)), \gamma_r(t)\rangle \,\mathrm{d}t \\ &= \int_0^{2\pi} g(\gamma_r(t))\,\mathrm{d}t - \log(r) \int_{B_r(0)} \underbrace{\operatorname{div}(\nabla g)}_{=\Delta g=0} \,\mathrm{dvol} \\ &= \int_0^{2\pi} g(\gamma_r(t))\,\mathrm{d}t\end{aligned}$$

nach dem Divergenzsatz für das Vektorfeld $\nabla g$. Da dies analog auch für den Radius $\varepsilon$ gilt, erhalten wir mit obigem

$$\tfrac{1}{2\pi}\int_0^{2\pi} g(\gamma_r(t))\,\mathrm{d}t = \tfrac{1}{2\pi}\int_0^{2\pi} g(\gamma_\varepsilon(t))\,\mathrm{d}t \to g(0)$$

für $\varepsilon \to 0$ und damit das Korollar. □

**Übung 13.46** (Mittelwerteigenschaft im dreidimensionalen)**.** *Beweisen Sie Korollar 13.45 für $n = 3$ unter Verwendung der Hilfsfunktion $\nu(p) = \|p\|^{-n}p$ und dem Divergenzsatz im $\mathbb{R}^3$ (Theorem 13.41).*

Satz 13.45 gilt analog für $n \geq 4$, der Beweis verwendet wieder die Hilfsfunktion wie in obiger Übung und den $n$-dimensionalen Divergenzsatz.

## 13.5 Der Satz von Stokes im dreidimensionalen Raum

Der verbleibende Integralsatz, den wir in dieser Vorlesung genau besprechen wollen, ist der Satz von Stokes, welcher eine Übertragung des Satzes von Green (Theorem 13.18) in den dreidimensionalen Raum darstellt.

### 13.5.1 Die Wirbelstärke in einem infinitesimalen Parallelogramm

Wir wollen diese Diskussion mit einer geometrisch motivierten Definition der Wirbelstärke beginnen. Hierfür bezeichnen wir das Parallelogramm

$$p + \{su + tv \mid s, t \in [0,1]\}$$

mit Ausgangspunkt $p \in \mathbb{R}^3$ und aufspannenden Vektoren $u, v \in \mathbb{R}^3$ als $P(p, u, v)$. Weiter parametrisieren wir den Rand des Parallelogramms durch den Weg

$$\gamma : t \in [0,4] \mapsto \begin{cases} p + tu & \text{falls } t \in [0,1] \\ p + u + (t-1)v & \text{falls } t \in [1,2] \\ p + (3-t)u + v & \text{falls } t \in [2,3] \\ p + (4-t)v & \text{falls } t \in [3,4] \end{cases}$$

und schreiben für ein gegebenes Vektorfeld $f$ auf einer offenen Menge $U$ im $\mathbb{R}^3$ mit $P(p, u, v) \subseteq U$ auch

$$\int_{\partial P(p,u,v)} f \cdot \mathrm{d}\mathbf{s} = \int_0^4 \langle f(\gamma(t)), \dot{\gamma}(t) \rangle \, \mathrm{d}t$$

für das Wegintegral von $f$ über den Rand $\partial P(p, u, v)$. Wir bemerken allerdings, dass hier der Begriff Rand als Rand der Menge in einem geeigneten zweidimensionalen affinen Unterraum aufgefasst werden muss, da ansonsten die Bezeichnung nicht mit unserer Definition des Randes übereinstimmt (wieso?).

**Definition 13.47** (Wirbelstärke in Parallelogrammen)**.** Sei $f$ ein stetig differenzierbares Vektorfeld auf einer offenen Teilmenge $U \subseteq \mathbb{R}^3$. Für $p \in U$ und $u, v \in \mathbb{R}^3$ definieren wir die **Wirbelstärke** von $f$ bei $p$ **im infinitesimalen Parallelogramm** zu $u, v \in \mathbb{R}^3$ durch

$$\operatorname{rot}(f, p, u, v) = \lim_{h \to 0} \tfrac{1}{h^2} \int_{\partial P(p,hu,hv)} f \cdot \mathrm{d}\mathbf{s},$$

falls der Grenzwert existiert.

Wir bemerken an dieser Stelle, dass man sich die Normalisierung $h^2$ in obiger Definition nicht als die Fläche von $P(p, hu, hv)$ vorstellen sollte, sondern viel eher als die Rate, mit der die Fläche für $h \to 0$ kleiner wird.

**Proposition 13.48** (Alternierend und Multilinear)**.** *Sei $f$ ein stetig differenzierbares Vektorfeld auf einer offenen Teilmenge $U \subseteq \mathbb{R}^3$ und sei $p \in U$. Dann gelten folgende Aussagen:*

- *Für alle* $p \in U$ *und* $u, v \in \mathbb{R}^3$ *existiert die Wirbelstärke* $\operatorname{rot}(f, p, u, v)$.
- *Die Wirbelstärke* $\operatorname{rot}(f, p, u, v)$ *hängt bei festem* $p \in U$ *und* $v \in \mathbb{R}^3$ *linear von* $u$ *ab.*
- *Die Wirbelstärke* $\operatorname{rot}(f, p, u, v)$ *hängt bei festem* $p \in U$ *und* $u \in \mathbb{R}^3$ *linear von* $v$ *ab.*
- *Es gilt* $\operatorname{rot}(f, p, u, v) = -\operatorname{rot}(f, p, v, u)$ *für alle* $u, v \in \mathbb{R}^3$ *und* $p \in U$.

Wir benötigen obiges Resultat bloss als Motivation für einige weitere Diskussionen; deswegen verweisen wir für den Beweis auf folgende Übung oder Übung 13.53.

**Übung 13.49.** *Beweisen Sie Proposition 13.48.*

### 13.5.2 Glatt berandete orientierbare Flächen im dreidimensionalen Raum

Wir erinnern daran, dass wir hier nur Flächen im $\mathbb{R}^3$ betrachten wollen, die mit einer endlichen Listen kompatibler Kartenabbildungen

$$\Phi_1 : V_1 \to U_1, \ldots, \Phi_L : V_L \to U_L$$

parametrisiert werden können (siehe Abschnitt 13.2.1, wo die Notation erklärt wurde). Obwohl dies nicht notwendig ist, wollen wir hier zur Vereinfachung der Diskussion und vor allem der Notation davon ausgehen, dass wir nur einen Diffeomorphismus

$$\Phi : V \to U$$

benötigen, um die gegebene Fläche $S$ zu parametrisieren. Genauer wollen wir annehmen, dass es einen kompakten, glatt berandeten Bereich $K \subseteq V \cap \mathbb{R}^2$ mit $S = \Phi(K)$ gibt. In diesem Fall wollen wir $S$ **glatt berandet** nennen. Man beachte, dass $S$ nicht im Sinne unserer Definition eine Teilmannigfaltigkeit bildet (sondern nur $\Phi(K^\circ)$), weswegen wir $S$ manchmal auch eine **Fläche mit Rand** nennen wollen. Der **Rand der Fläche** $S$ ist dann definiert als $\partial S = \Phi(\partial K)$ (obwohl dies nicht mit der Definition des Randes einer Menge übereinstimmt) und bildet eine kompakte eindimensionale Teilmannigfaltigkeit (wieso?).

Das Wegintegral eines Vektorfelds auf $U$ über den Rand von $S$ wird mit Hilfe von Wegen der Form $\Phi \circ \gamma_1, \ldots, \Phi \circ \gamma_L$ definiert, wobei $\gamma_1, \ldots, \gamma_L$ eine positiv orientierte Parametrisierung des Randes $\partial K$ ist.

### 13.5.3 Der Satz von Stokes

Mit diesen Definitionen können wir nun den Satz von Stokes formulieren.

**Theorem 13.50** (Satz von Stokes)**.** *Sei* $f$ *ein stetig differenzierbares Vektorfeld auf einer offenen Teilmenge* $U \subseteq \mathbb{R}^3$. *Dann existiert die Wirbelstärke* $\operatorname{rot}(f, p, u, v)$ *von* $f$ *zu jedem Punkt* $p \in U$ *und zu allen* $u, v \in \mathbb{R}^3$ *und ist durch*

$$\operatorname{rot}(f, p, u, v) = \langle \operatorname{rot}(f)(p), u \times v \rangle$$

*gegeben, wobei das Vektorfeld*

$$\operatorname{rot}(f) = \begin{pmatrix} \partial_2 f_3 - \partial_3 f_2 \\ \partial_3 f_1 - \partial_1 f_3 \\ \partial_1 f_2 - \partial_2 f_1 \end{pmatrix}$$

*auf $U$ die* ***Rotation*** *von $f$ genannt wird. Ist $S \subseteq U$ eine glatt berandete, orientierbare Fläche (mit Rand), so gilt des Weiteren*

$$\int_S \operatorname{rot}(f) \cdot \mathrm{d}\mathbf{n} = \int_{\partial S} f \cdot \mathrm{d}\mathbf{s}.$$

Um sich die Formel für die Rotation $\operatorname{rot}(f)$ merken zu können, ist vielleicht folgende (sehr) informelle Rechnung nützlich. Schreiben wir $\nabla = \left(\begin{smallmatrix} \partial_1 \\ \partial_2 \\ \partial_3 \end{smallmatrix}\right)$, so erhalten wir

$$\operatorname{rot}(f) = \begin{pmatrix} \partial_2 f_3 - \partial_3 f_2 \\ \partial_3 f_1 - \partial_1 f_3 \\ \partial_1 f_2 - \partial_2 f_1 \end{pmatrix} = \begin{pmatrix} \partial_1 \\ \partial_2 \\ \partial_3 \end{pmatrix} \times \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} = \nabla \times f.$$

Auch möchten wir anmerken, dass streng genommen der Begriff „glatt berandet“ für allgemeine Flächen nicht eingeführt wurde, sondern wie schon erwähnt nur für Flächen der Form $S = \Phi(K)$. Die allgemeine Definition genauso wie der Beweis des Satzes lassen sich im Wesentlichen durch Zusammenstückeln von Flächen mit dieser Eigenschaft erhalten, weswegen wir uns (auch im Beweis) auf diesen Spezialfall einschränken möchten. Genau wie der Beweis des Satz von Gauss (Theorem 13.41) verwendet der Beweis im allgemeinen Fall dann eine Zerlegung der Eins, um den Beweis auf folgenden Beweis zurückzuführen.

*Beweis.* Sei $S = \Phi(K)$ wie in Abschnitt 13.5.2. Wir möchten via der Abbildung $\Phi$ das Vektorfeld $f$ auf $V \cap \mathbb{R}^2$ zurückziehen und definieren also ein Vektorfeld $\tilde{f}$ auf der Karte $V \cap \mathbb{R}^2$ durch

$$\tilde{f}(s,t) = \begin{pmatrix} \langle f \circ \Phi(s,t), \partial_s \Phi(s,t) \rangle \\ \langle f \circ \Phi(s,t), \partial_t \Phi(s,t) \rangle \end{pmatrix}$$

für alle $(s,t) \in V \cap \mathbb{R}^2$. Um den Satz von Green (Theorem 13.18) geeignet anzuwenden, müssen wir die Rotation von $\tilde{f}$ in $V \cap \mathbb{R}^2$ mittels der Rotation von $f$ ausdrücken und behaupten deswegen, dass

$$\operatorname{rot}(\tilde{f}) = \langle \operatorname{rot}(f) \circ \Phi, \partial_s \Phi \times \partial_t \Phi \rangle \tag{13.5}$$

gilt. Wir berechnen dazu

$$\begin{aligned}
\operatorname{rot}(\tilde{f}) &= \partial_s \tilde{f}_2 - \partial_t \tilde{f}_1 \\
&= \partial_s \langle f \circ \Phi, \partial_t \Phi \rangle - \partial_t \langle f \circ \Phi, \partial_s \Phi \rangle \\
&= \langle (\mathrm{D} f) \circ \Phi\, \partial_s \Phi, \partial_t \Phi \rangle + \langle f \circ \Phi, \partial_s \partial_t \Phi \rangle - \langle (\mathrm{D} f) \circ \Phi\, \partial_t \Phi, \partial_s \Phi \rangle - \langle f \circ \Phi, \partial_s \partial_t \Phi \rangle \\
&= \langle (\mathrm{D} f) \circ \Phi\, \partial_s \Phi, \partial_t \Phi \rangle - \langle (\mathrm{D} f) \circ \Phi\, \partial_t \Phi, \partial_s \Phi \rangle .
\end{aligned}$$

Des Weiteren können wir die Rotation von $f$ auch in der Form

$$\operatorname{rot}(f) = \sum_{\sigma \in A_3} \big(\partial_{\sigma(2)} f_{\sigma(3)} - \partial_{\sigma(3)} f_{\sigma(2)}\big) e_{\sigma(1)}$$

schreiben, wobei $A_3 \subseteq S_3$ die Gruppe der zyklikschen Permutationen der Menge $\{1, 2, 3\}$ bezeichnet (wieso?). Damit gilt nun

$$\begin{aligned}
\operatorname{rot}(\tilde{f}) &= \sum_{j,k=1}^{3} \underbrace{(\partial_j f_k) \circ \Phi}_{=a_{jk}} \big(\underbrace{\partial_s \Phi_j}_{=u_j} \underbrace{\partial_t \Phi_k}_{=v_k} - \underbrace{\partial_t \Phi_j}_{=v_j} \underbrace{\partial_s \Phi_k}_{=u_k}\big) \\
&= \sum_{j,k=1}^{3} a_{jk} \underbrace{(u_j v_k - v_j u_k)}_{=0 \text{ falls } j=k} \\
&= \sum_{\sigma \in A_3} \Big( a_{\sigma(2)\sigma(3)} \big(u_{\sigma(2)} v_{\sigma(3)} - u_{\sigma(3)} v_{\sigma(2)}\big) + a_{\sigma(3)\sigma(2)} \big(u_{\sigma(3)} v_{\sigma(2)} - u_{\sigma(2)} v_{\sigma(3)}\big) \Big) \\
&= \sum_{\sigma \in A_3} \big(a_{\sigma(2)\sigma(3)} - a_{\sigma(3)\sigma(2)}\big) \big(u_{\sigma(2)} v_{\sigma(3)} - u_{\sigma(3)} v_{\sigma(2)}\big) \\
&= \langle \operatorname{rot}(f) \circ \Phi, \partial_s \Phi \times \partial_t \Phi \rangle .
\end{aligned}$$

Aus dem Satz von Green (Theorem 13.18), der Kettenregel für die Ableitung von $\Phi \circ \gamma_\ell$, und aus unseren Definitionen folgt nun

$$\begin{aligned}
\int_S \operatorname{rot}(f) \cdot \mathrm{d}\mathbf{n} &= \int_K \langle \operatorname{rot}(f) \circ \Phi, \partial_s \Phi \times \partial_t \Phi \rangle \,\mathrm{dvol} = \int_K \operatorname{rot}(\tilde{f}) \,\mathrm{dvol} \\
&= \int_{\partial K} \tilde{f} \cdot \mathrm{d}\mathbf{s} = \sum_{\ell=1}^{L} \int_{I_\ell} \Big\langle \tilde{f} \circ \gamma_\ell, \dot{\gamma}_\ell \Big\rangle \,\mathrm{d}t \\
&= \sum_{\ell=1}^{L} \int_{I_\ell} \Big\langle \Big( \begin{smallmatrix} \langle f \circ \Phi \circ \gamma_\ell, (\partial_s \Phi) \circ \gamma_\ell \rangle \\ \langle f \circ \Phi \circ \gamma_\ell, (\partial_t \Phi) \circ \gamma_\ell \rangle \end{smallmatrix} \Big), \dot{\gamma}_\ell \Big\rangle \,\mathrm{d}t \\
&= \sum_{\ell=1}^{L} \int_{I_\ell} \sum_{j=1}^{3} (f_j \circ \Phi \circ \gamma_\ell) \Big( \big((\partial_s \Phi_j) \circ \gamma_\ell\big) \dot{\gamma}_{\ell,1} + \big((\partial_t \Phi_j) \circ \gamma_\ell\big) \dot{\gamma}_{\ell,2} \Big) \,\mathrm{d}t \\
&= \sum_{\ell=1}^{L} \int_{I_\ell} \big\langle f, \tfrac{d}{dt} (\Phi \circ \gamma_\ell) \big\rangle \,\mathrm{d}t = \int_{\partial S} f \cdot \mathrm{d}\mathbf{s},
\end{aligned}$$

wobei $\int_K (\cdot) \,\mathrm{dvol} = \int_K (\cdot) \,\mathrm{d}s \,\mathrm{d}t$ das zwei-dimensionale Integral über $K \subseteq \mathbb{R}^2$ bezeichnet.

Wenden wir dies nun auf ein Parallelogramm $P(p, hu, hv)$ für $p \in U$, $u, v \in \mathbb{R}^3$ und genügend kleine $h > 0$ an (wieso dürfen wir das?), so erhalten wir

$$\begin{aligned}
\operatorname{rot}(f, p, u, v) &= \lim_{h\to 0} \tfrac{1}{h^2} \int_{\partial P(p,hu,hv)} f \cdot \mathrm{d}\mathbf{n} \\
&= \lim_{h\to 0} \tfrac{1}{h^2} \int_0^1 \int_0^1 \langle \operatorname{rot}(f)(p + shu + thv), (hu) \times (hv) \rangle \, \mathrm{d}s \, \mathrm{d}t \\
&= \langle \operatorname{rot}(f)(p), u \times v \rangle \, .
\end{aligned}$$

auf Grund der Stetigkeit von $\operatorname{rot}(f)$. □

VO: 18.05.18

**Beispiel 13.51.** *Für $t \in [0,1]$ sei*

$$\begin{pmatrix} \gamma_1(t) \\ \gamma_2(t) \end{pmatrix} = \begin{cases} t(1,0)^t & \text{falls } t \in [0,1] \\ (1,0)^t + (t-1)(0,1)^t & \text{falls } t \in [1,2] \\ (1,1)^t - (t-2)(1,0)^t & \text{falls } t \in [2,3] \\ (4-t)(0,1)^t & \text{falls } t \in [3,4] \end{cases},$$

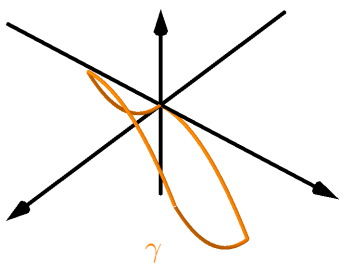
*womit $t \in [0,4] \mapsto (\gamma_1(t), \gamma_2(t))^t$ eine Parametrisierung des Randes des Einheitsquadrats $[0,1]^2$ darstellt. Wir betrachten nun den geschlossenen Weg*

$$\gamma : t \in [0,4] \mapsto \begin{pmatrix} \gamma_1(t) \\ \gamma_2(t) \\ \gamma_1(t)^2 - \gamma_2(t)^2 \end{pmatrix}.$$

*und das Vektorfeld*

$$f : (x,y,z) \in \mathbb{R}^3 \mapsto \begin{pmatrix} yz + \cos(x) \\ xz + \sin(y) \\ 2xy \end{pmatrix};$$

*siehe folgendes Bild.*




*Möchte man nun im Prinzip das Wegintegral $\int_\gamma f \cdot \mathrm{d}\mathbf{s}$ direkt aus der Definition berechnen, wozu man dieses also als Summe von vier Wegintegralen schreibt. Letztere können und sind in diesem Fall aber relativ kompliziert zu berechnen. Mit etwas Glück ist die Rotation des Vektorfelds $f$*

*jedoch von einer einfacheren Form als $f$, womit man den Satz von Stokes (Theorem 13.50) anwenden kann. In der Tat gilt*

$$\operatorname{rot} f(x,y,z) = \begin{pmatrix} x \\ -y \\ 0 \end{pmatrix}$$

*für alle $(x,y,z) \in \mathbb{R}^3$. Der Weg $\gamma$ parametrisiert gerade den Rand der Fläche*

$$S = \left\{(x,y,z) \in \mathbb{R}^3 \mid x,y \in [0,1]^2,\ z = x^2 - y^2\right\},$$

*welche wiederum durch die Abbildung $\Phi : (s,t) \in \mathbb{R}^2 \mapsto (s,t,s^2-t^2) \in \mathbb{R}^3$ und den Bereich $K = [0,1]^2$ parametrisiert ist. Man berechnet*

$$\partial_s \Phi = \begin{pmatrix} 1 \\ 0 \\ 2s \end{pmatrix}, \quad \partial_t \Phi = \begin{pmatrix} 0 \\ 1 \\ -2t \end{pmatrix}, \quad \partial_s \Phi \times \partial_t \Phi = \begin{pmatrix} -2s \\ 2t \\ 1 \end{pmatrix}.$$

*Nach dem Satz von Stokes (Theorem 13.50) gilt somit*

$$\begin{aligned} \int_\gamma f \cdot \mathrm{d}\mathbf{s} = \int_S \operatorname{rot}(f) \cdot \mathrm{d}\mathbf{n} &= \int_0^1 \int_0^1 \left\langle \begin{pmatrix} s \\ -t \\ 0 \end{pmatrix}, \begin{pmatrix} -2s \\ 2t \\ 1 \end{pmatrix} \right\rangle \mathrm{d}s\,\mathrm{d}t \\ &= \int_0^1 \int_0^1 (-2s^2 - 2t^2)\,\mathrm{d}s\,\mathrm{d}t = -\tfrac{4}{3}. \end{aligned}$$

**Übung 13.52.** *Berechnen Sie mittels dem Satz von Stokes das Wegintegral $\int_\gamma f\,\mathrm{d}\mathbf{s}$, wobei $f(x,y,z) = \begin{pmatrix} 2xz\cos(x^2) \\ z \\ \sin(x^2) \end{pmatrix}$ und $\gamma : t \in [0,1] \to \begin{pmatrix} t \\ t^2 \\ (t)^2 - (t^2)^2 \end{pmatrix}$. Als Hinweis bemerken wir noch, dass $\gamma$ Werte in der Fläche $z = x^2 - y^2$ annimmt.*

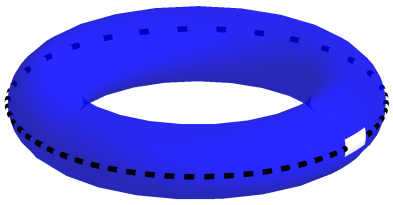
**Übung 13.53.** *Verwenden Sie den Satz von Stokes (Theorem 13.50), um Proposition 13.48 zu beweisen.*

**Übung 13.54** (Vektorfelder mit Potential sind rotationsfrei)**.** *Sei $F$ eine zweimal stetig differenzierbare Funktion auf einer offenen Menge $U \subseteq \mathbb{R}^3$. Zeigen Sie auf folgende zwei Art und Weisen, dass $\operatorname{rot}(\nabla F) = 0$ gilt.*

*(i) Berechnen Sie dies direkt mit den nötigen partiellen Ableitungen.*

*(ii) Zeigen Sie zuerst, dass $\operatorname{rot}(\nabla F, p, u, v) = 0$ gilt für alle $p \in U$ und $u, v \in \mathbb{R}^3$.*

Für obigen Beweis des Satzes von Stokes (Theorem 13.50) war die Annahme hilfreich, dass die Fläche $S$ durch eine Karte gegeben ist. Es gibt aber auch interessante Beispiele, die eben nicht von dieser Form sind. Ein einfaches solches Beispiel wäre zum Beispiel der Torus mit einem Loch, siehe folgendes Bild:



In vielen konkreten Fällen wie auch diesem ist es ein leichtes, die Fläche so zu zerschneiden, dass man Theorem 13.50 trotzdem anwenden kann; den Torus mit einem Loch kann man entlang der oben eingezeichneten, gestrichelten Linie aufschneiden und erhält zwei Flächen, für die der Satz von Stokes (wie er oben formuliert ist) anwenden kann. Die Wegintegrale entlang der gestrichelten Linien kompensieren sich aber gerade, womit der Satz von Stokes in der Formulierung von Theorem 13.50 direkt verwendet werden kann.

## 13.6 Zwei Werkzeuge der weiteren Analysis*

Im Beweis des Divergenzsatzes im $\mathbb{R}^3$ (Theorem 13.41) haben wir das gegebene Vektorfeld unter Verwendung einer „Partition der Eins“ in Teile zerteilt, die einfacher zu untersuchen waren. Hierbei war es aber wichtig, die stetige Differenzierbarkeit des Vektorfelds nicht zu zerstören, weswegen wir nicht einfach charakteristische Funktionen, sondern eben eine „glatte Partition der Eins“ verwendet haben. Dies ist ein wichtiges technisches Hilfsmittel für die weitere Analysis, weswegen wir dies nochmals kurz allgemeiner besprechen möchten.

Hierfür aber auch aus weiteren Gründen wollen wir auch kurz die Faltung zwischen zwei Funktionen besprechen und zeigen, dass sich Stetigkeit und Glattheit einer der beiden Funktionen unmittelbar auf das Faltungsprodukt überträgt.

### 13.6.1 Glättung durch Faltung

**Proposition 13.55** (Glattheit der Faltung)**.** *Sei $F : \mathbb{R}^n \to \mathbb{R}$ eine stetige Funktion. Sei $\psi : \mathbb{R}^n \to \mathbb{R}$ eine glatte (oder auch nur $C^k$ für $k \in \mathbb{N}_0$) Funktion mit kompaktem Träger. Dann definiert*

$$\psi * F(p) = \int_{\mathbb{R}^n} \psi(p-v)F(v)\,\mathrm{d}\,\mathrm{vol}(v)$$

*für $p \in U$ eine glatte (oder $C^k$) Funktion mit Träger in* $\operatorname{supp}(\psi) + \operatorname{supp}(F)$.

*Beweis.* Da $\psi$ kompakten Träger besitzt gibt es ein $R$ mit $\operatorname{supp}\psi \subseteq B_R^\infty(0)$. Sei $p_0 \in \mathbb{R}^n$. Für alle $p \in B_1^\infty(p_0)$ gilt dann

$$\psi * F(p) = \int_Q \psi(p-v)F(v)\,\mathrm{d}\,\mathrm{vol}(v)$$

für $Q = \overline{B_{R+1}^\infty(p_0)}$, da jedes $v \in \mathbb{R}^n$ mit $\psi(p-v) \neq 0$ auch

$$v = p - (p-v) \in B_1^\infty(p_0) + B_R^\infty(0) \subseteq Q$$

erfüllt. Stetigkeit und Glattheit von $\psi * h$ folgt nun aus der Verallgemeinerung des Satzes über Parameterintegrale (Satz 10.36) für mehrdimensionale Integrale – siehe Übung 13.56.

Für die letzte Behauptung nehmen wir an, dass $p \notin \operatorname{supp}(\psi) + \operatorname{supp}(F)$ ist. Für $v \in \mathbb{R}^n$ gilt dann $v \notin \operatorname{supp}(F)$ oder $p - v \notin \operatorname{supp}(\psi)$, womit auf jeden Fall $\psi(p-v)F(v) = 0$ ist. Dies impliziert $\psi * F(p) = 0$. □

**Übung 13.56.** *Verallgemeinern Sie Satz 10.36 auf den oben benötigten Fall.*

*Hinweis: Kombinieren Sie Satz 10.36 mit dem Satz von Fubini (Theorem 12.39).*

**Proposition 13.57** (Annäherung durch Faltung)**.** *Sei $F : \mathbb{R}^n \to \mathbb{R}$, $K \subseteq \mathbb{R}^n$ und $\varepsilon > 0$ und $\delta > 0$ so dass für $p \in K$ und $q \in \mathbb{R}^n$ mit $\|q - p\| < \delta$ die Abschätzung $|F(q) - F(p)| < \varepsilon$ gilt. Falls $\psi : \mathbb{R}^n \to [0, \infty)$ Träger* $\operatorname{supp}\psi \subseteq B_\delta(0)$ *hat und $\int_{B_\delta(0)} \psi\,\mathrm{d}\,\mathrm{vol} = 1$ gilt, dann gilt*

$$\|\psi * F - F\|_{K,\infty} = \max_{p \in K} \|\psi * F(p) - F(p)\| \leq \varepsilon.$$

*Des Weiteren existieren für jedes $\delta > 0$ derartige glatte Funktionen $\psi$.*

*Beweis.* Sei $p \in K$. Dann erhalten wir mittels der Substitution $w = p - v$ (siehe Satz 12.56), dass

$$\begin{aligned}\big|\psi * F(p) - F(p)\big| &= \Big|\int_{\mathbb{R}^n} \psi(p-v)F(v)\,\mathrm{d}\,\mathrm{vol}(v) - F(p)\Big| \\ &= \Big|\int_{\mathbb{R}^n} \psi(w)F(p-w)\,\mathrm{d}\,\mathrm{vol}(v) - \int_{\mathbb{R}^n} \psi(w)F(p)\,\mathrm{d}\,\mathrm{vol}(w)\Big| \\ &\leq \int_{B_\delta(0)} \psi(w)\underbrace{|F(p-w) - F(p)|}_{\leq \varepsilon}\,\mathrm{d}\,\mathrm{vol}(w) \leq \varepsilon\end{aligned}$$

gilt.

Sei nun $\Psi$ die glatte Funktion mit $\Psi(t) = 0$ für alle $t \in (-\infty, 0]$ und $\Psi(t) > 0$ für alle $t > 0$ in Beispiel 7.23. Für ein $\delta > 0$ definieren wir $\psi$ als ein skalares Vielfaches der Funktion $v \in \mathbb{R}^n \mapsto \Psi\left(\frac{1}{2}\delta^2 - \|v\|^2\right)$ so dass $\int_{\mathbb{R}^n} \psi\,\mathrm{d}\,\mathrm{vol} = 1$. □

**Korollar 13.58** (Glatte Funktionen sind dicht)**.** *Für jede stetige Funktion $F : \mathbb{R}^n \to \mathbb{R}$ mit kompakten Träger und jedes $\varepsilon > 0$ existiert eine glatte Funktion $F_{\text{glatt}} : \mathbb{R}^n \to \mathbb{R}$ mit kompakten Träger und $\|F - F_{\text{glatt}}\|_\infty < \varepsilon$.*

**Übung 13.59.** *Beweisen Sie Korollar 13.58.*

*Hinweis: Zeigen Sie zuerst, dass eine stetige Funktion mit kompakten Träger gleichmässig stetig ist.*

### 13.6.2 Glatte Partitionen der Eins

**Satz 13.60** (Glatte Partition der Eins)**.** *Sei $K \subseteq \mathbb{R}^n$ eine kompakte Teilmenge und $O_1, \ldots, O_L$ eine offene Überdeckung von $K$. Dann existieren glatte Funktionen*

$$\psi_0, \psi_1, \ldots, \psi_L : \mathbb{R}^n \to [0, 1],$$

*so dass*

- $\operatorname{supp}(\psi_0) \cap K = \emptyset$.
- $\operatorname{supp}(\psi_\ell) \subseteq O_\ell$ *für* $\ell \in \{1, \ldots, L\}$
- $\sum_{\ell=0}^{L} \psi_\ell = 1$.

Eine solche Kollektion von Funktionen $\psi_0, \psi_1, \ldots, \psi_L$ nennt sich eine **glatte Partition der Eins** für die offene Überdeckung $O_1, \ldots, O_L$ von $K$.

Als ersten Schritt beweisen wir die schwächere, stetigen Version des Satzes.

*Beweis der Existenz einer „stetigen“ Partition der Eins.* Wir wollen zuerst eine Liste an stetigen Funktionen $h_0, h_1, \ldots, h_L$ finden, die abgesehen von der Glattheit alle Eigenschaften im Satz erfüllen. Falls $K$ die leere Menge ist, so können wir einfach $h_0 = 1$ und $h_1 = \cdots = h_L = 0$ setzen. Sei nun $K$ nichtleer und $R > 0$ so dass $K \subseteq B_R(0)$. Falls nötig können wir für

$\ell = 1, \ldots, L$ die Mengen $O_\ell$ durch $O_\ell \cap B_R(0)$ erstzen, ohne die Annahmen oder Aussagen des Satzes zu beeinflussen, und können daher annehmen, dass die Mengen $O_1, \ldots, O_L$ beschränkt sind.

Wir verwenden wieder die stetigen Abstandsfunktionen

$$\begin{aligned} p \in \mathbb{R}^n &\mapsto \mathrm{d}(p, K) = \inf \{\mathrm{d}(p, q) \mid q \in K\} \\ p \in \mathbb{R}^n &\mapsto \mathrm{d}(p, O_\ell^c) = \inf \{\mathrm{d}(p, q) \mid q \in \mathbb{R}^n \setminus O_\ell\} \end{aligned}$$

für $\ell = 1, \ldots, L$ sowie die stetige Funktion

$$f : p \in \mathbb{R}^n \mapsto \max \{\mathrm{d}(p, O_1^c), \ldots, \mathrm{d}(p, O_L^c)\}.$$

Da $K$ kompakt und $f$ stetig ist, existiert ein $p_0 \in K$, bei dem $f$ ein Minimum annimmt. Da $O_1, \ldots, O_L$ eine offene Überdeckung von $K$ ist, gibt es ein $\ell$ mit $p \in O_\ell$ und daher ist $\eta = f(p_0) > 0$. Die Zahl $\eta$ wird wieder eine Lebesgue-Zahl der Überdeckung $O_1, \ldots, O_L$ genannt und hat die Eigenschaft, das für jeden Punkt $p \in K$ ein $\ell \in \{1, \ldots, L\}$ existiert mit $B_\eta(p) \subseteq O_\ell$.

Wir betrachten nun die stetigen Funktionen

$$\begin{aligned} \tilde{h}_0(p) &= \big(\mathrm{d}(p, K) - \tfrac{\eta}{2}\big)^+, \\ \tilde{h}_\ell(p) &= \big(\mathrm{d}(p, O_\ell^c) - \tfrac{\eta}{2}\big)^+ \end{aligned}$$

für $p \in \mathbb{R}^n$ und $\ell = 1, \ldots, L$, welche bereits die ersten beiden Eigenschaften im Satz erfüllen. Wir dividieren diese nun durch die positive stetige Funktion $H = \sum_{\ell=0}^L \tilde{h}_\ell$ und erhalten eine stetige Partition der Eins. Des Weiteren gilt sogar $\operatorname{supp}(h_0) + B_{\frac{\eta}{2}}(0) \subseteq K^c$ und $\big(\operatorname{supp}(h_\ell) + B_{\frac{\eta}{2}}(0)\big) \subseteq O_\ell$ für $\ell \in \{1, \ldots, L\}$. □

*Beweis der Existenz der glatten Partition der Eins.* Sei $h_0, \ldots, h_L$ die stetige Partition der Eins wie im ersten Teil des Beweises. Sei $\psi : \mathbb{R}^n \to [0, \infty)$ eine glatte Funktion mit kompakten Träger. Dann sind die Funktionen

$$\psi_\ell = \psi * h_\ell$$

für $\ell = 0, \ldots, L$ nach Proposition 13.55 glatt und nehmen nach Definition Werte in $[0, \infty)$ an.

Auf Grund der Bilinearität der Faltung gilt des Weiteren

$$\psi_0 + \ldots + \psi_L = \psi * (h_0 + \cdots + h_L) = \psi * \mathbb{1} = \Big(\int_{\mathbb{R}^n} \psi \,\mathrm{d}\,\mathrm{vol}\Big)\mathbb{1}.$$

Sei $\psi$ wie in Proposition 13.57 mit Träger $\operatorname{supp} \psi \subseteq \overline{B_{\eta/2}(0)}$. Wir errinnern daran, dass

$$\begin{aligned} \operatorname{supp}(h_0) + B_{\frac{\eta}{2}}(0) &\subseteq K^c, \\ \operatorname{supp}(h_\ell) + B_{\frac{\eta}{2}}(0) &\subseteq O_\ell \end{aligned}$$

für $\ell = 1, \ldots, L$ gilt. Die Aussage in Proposition 13.55 über den Träger von $\psi_\ell = \psi * h_\ell$ beweist nun den Satz. $\square$

Die glatte Partition der Eins kann zum Beispiel dafür verwendet werden um den Beweis vom Satz von Stokes auf orientierte glatt berandetet Flächen zu erweitern, die mittels mehr als nur einem Diffeomorphismus parametrisiert werden.

## 13.7 Konservative Vektorfelder*

Sei $f$ ein stetig differenzierbares Vektorfeld $f$ auf einem Gebiet $U \subseteq \mathbb{R}^n$. In diesem Abschnitt möchten wir zeigen, dass die Integrabilitätsbedingungen

$$\partial_k f_j = \partial_j f_k$$

für alle $j, k \in \{1, \ldots, n\}$ für viel allgemeinere Gebiete als nur sternförmige (siehe Satz 10.52) bereits implizieren, dass das Vektorfeld konservativ ist. Hierzu benötigen wir zuerst einige weitere Begriffe.

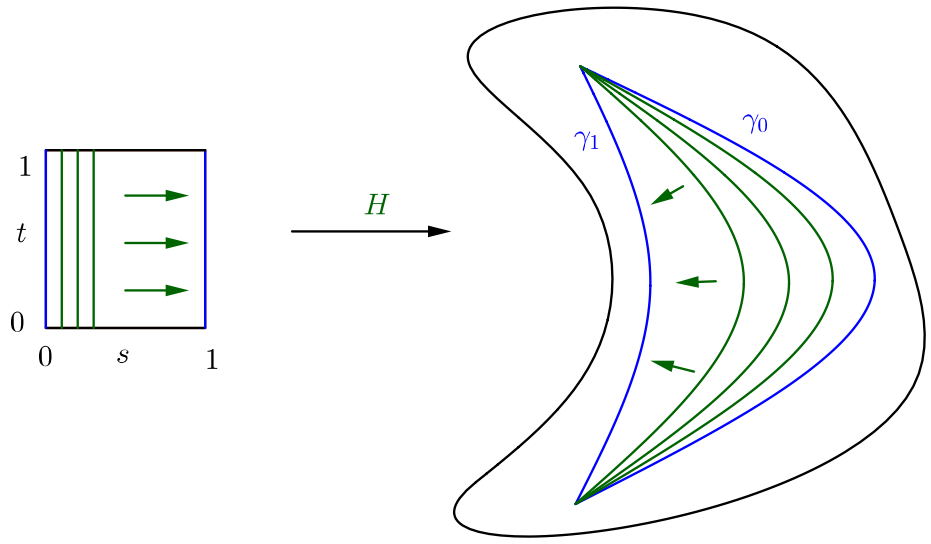
### 13.7.1 Homotopie

**Definition 13.61** (Homotopie)**.** Sei $U \subseteq \mathbb{R}^n$ ein Gebiet (also eine zusammenhängende, offene Teilmenge) und seien $\gamma_0, \gamma_1 : [0,1] \to U$ zwei Wege in $U$ mit gleichem Anfangspunkt $\gamma_0(0) = \gamma_1(0)$ und gleichem Endpunkt $\gamma_0(1) = \gamma_1(1)$. Dann heissen $\gamma_0$ und $\gamma_1$ **homotop** (in $U$), falls es eine stetige Abbildung $H : [0,1]^2 \to U$ gibt, die wir eine **Homotopie** nennen und die folgende Eigenschaften erfüllt:

- $H(0,t) = \gamma_0(t)$ für alle $t \in [0,1]$.
- $H(1,t) = \gamma_1(t)$ für alle $t \in [0,1]$.
- $H(s,0) = \gamma_0(0) = \gamma_1(0)$ für alle $s \in [0,1]$.
- $H(s,1) = \gamma_0(1) = \gamma_1(1)$ für alle $s \in [0,1]$.

Falls $\gamma_0$ und $\gamma_1$ glatte Wege sind und die Homotopie glatt gewählt werden kann, so nennen wir die Wege **auf glatte Weise homotop**.

In Worten ausgedrückt sind zwei glatte Wege $\gamma_0, \gamma_1$ homotop, falls es eine Kollektion von glatten Wegen $\gamma_s = H(s, \cdot)$ für $s \in [0,1]$ gibt, die „stetig vom Parameter $s$ abhängt“, so dass die Wege $\gamma_s$ auch die gleichen Anfangs- und Endpunkte wie $\gamma_0, \gamma_1$ haben und so dass die Kollektion zwischen den Wegen $\gamma_0$ und $\gamma_1$ interpoliert. Siehe auch folgendes Bild:



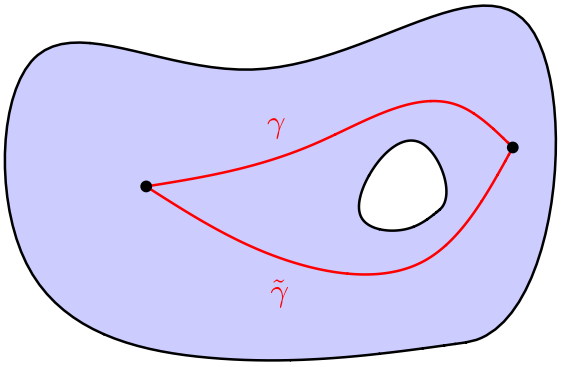
### 13.7.2 Integrabilität auf einfach zusammenhängenden Gebieten

Sternförmige Gebiete haben die Eigenschaft, dass je zwei Wege homotop sind.

**Übung 13.62** (Sternförmige Gebiete und Homotopien)**.** *Sei $U$ ein sternförmiges Gebiet und seien $\gamma_0, \gamma_1 : [0,1] \to U$ zwei Wege in $U$ mit $\gamma_0(0) = \gamma_1(0)$ und $\gamma_0(1) = \gamma_1(1)$. Finden Sie eine Homotopie zwischen $\gamma_0$ und $\gamma_1$. Zeigen Sie auch, dass Sie eine glatte Homotopie finden können, falls $\gamma_0$ und $\gamma_1$ glatt sind.*

Sobald die Menge aber „Löcher hat“, findet man vielleicht Wege, die nicht homotop sind.

**Definition 13.63** (Einfach zusammenhängend)**.** Ein Gebiet $U \subseteq \mathbb{R}^n$ heisst **einfach zusammenhängend**, falls je zwei Wege mit gleichem Anfangs- und Endpunkt homotop sind.




Figur 13.9: Illustration eines nicht einfach zusammenhängenden Gebiets. Die beiden eingezeichneten Wege $\gamma$ und $\tilde{\gamma}$ sind nicht homotop. Wer im Wald mit Hund an der Leine spazieren geht, kennt diesen Unterschied sehr gut.

**Satz 13.64** (Integrabilitätsbedingungen)**.** *Sei $U \subseteq \mathbb{R}^n$ ein einfach zusammenhängendes Gebiet. Unter dieser Annahme ist ein stetig differenzierbares Vektorfeld $f : U \to \mathbb{R}^n$ genau dann konservativ, wenn $f$ den Integrabilitätsbedingungen*

$$\partial_k f_j = \partial_j f_k \tag{13.6}$$

*für alle $j, k \in \{1, \ldots, n\}$ genügt.*

Wir erinnern daran, dass die Notwendigkeit der Integrabilitätsbedingungen bereits aus Satz 10.49 bekannt ist (und für beliebige Gebiete gilt). Das Ziel dieses Abschnittes ist zu beweisen, dass diese auf einem einfach zusammenhängenden Gebiet $U$ auch hinreichend sind. Also angenommen $f : U \to \mathbb{R}^n$ ist stetig differenzierbar und erfüllt die Integrabilitätsbedingungen in (13.6). Wir fixieren $p_0 \in U$ und wollen einen Kandidaten $F : U \to \mathbb{R}$ für ein Potential von $f$ durch

$$F(p) = \int_{\gamma_p} f \cdot \mathrm{d}\mathbf{s} \tag{13.7}$$

für $p \in U$ definieren, wobei $\gamma_p : [a,b] \to U$ ein glatter Weg von $p_0$ nach $p$ ist. Die Wahl des Intervalles und die konkrete Parametrisierung beeinflussen das Integral auf Grund von

Lemma 10.44 nicht, und wir werden daher oft $[a, b] = [0, 1]$ annehmen. Die wichtigste Frage hier ist, warum $F(p)$ auch nicht von der Wahl des Weges $\gamma_p$ abhängt. Wir werden die Annahme des einfachen Zusammenhangs, die Integrabilitätsbedingungen (13.6) und den Satz von Green (Theorem 13.18) verwenden, um diese Frage (positiv) zu beantworten.

### 13.7.3 Invarianz unter Homotopien

**Lemma 13.65** (Invarianz, glatter Fall)**.** *Sei $U \subseteq \mathbb{R}^n$ ein Gebiet und $f : U \to \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld, welches die Integrabilitätsbedingungen erfüllt. Seien weiter $\gamma_0, \gamma_1$ zwei glatte Wege von $p_0 \in U$ nach $p \in U$, die auf glatte Weise in $U$ homotop sind. Dann gilt*

$$\int_{\gamma_0} f \cdot \mathrm{d}\mathbf{s} = \int_{\gamma_1} f \cdot \mathrm{d}\mathbf{s}.$$

Wir werden etwas weiter unten (in Proposition 13.66) sehen, dass die Aussage des obigen Lemmas auch gilt, wenn die obige Homotopie nicht zwingend glatt ist. Da der Beweis von Lemma 13.65 sich auf die Methoden dieses Kapitels (hier den Satz von Green) konzentriert, wollen wir trotzdem zuerst nur die etwas schwächere Aussage in Lemma 13.65 beweisen.

*Beweis.* Nach Annahme existiert eine glatte Homotopie $H$ zwischen $\gamma_0$ und $\gamma_1$. Wir werden den Satz von Green (Theorem 13.18) auf das Quadrat $[0, 1]^2$ und das Vektorfeld

$$\tilde{f} : (s, t) \in [0, 1]^2 \mapsto \begin{pmatrix} \langle f(H(s, t)), \partial_s H(s, t)\rangle \\ \langle f(H(s, t)), \partial_t H(s, t)\rangle \end{pmatrix}$$

wie bereits im Beweis des Satzes von Stokes anwenden.

Für dies bemerken wir zuerst, dass auf Grund der Integrabilitätsbedingungen (13.6) und der Kettenregel (Satz 10.13) die Gleichung

$$\begin{aligned} \langle \partial_v f, u\rangle &= \sum_{j=1}^n \partial_v f_j u_j = \sum_{j,k=1}^n v_k \partial_k f_j u_j \\ &= \sum_{j,k=1}^n v_k \partial_j f_k u_j = \langle \partial_u f, v\rangle \end{aligned}$$

für $u, v \in \mathbb{R}^n$ gilt. Für das Vektorfeld $\tilde{f}$ auf $[0, 1]^2$ und $(s, t) \in [0, 1]^2$ ist daher

$$\begin{aligned} \partial_s \tilde{f}_2(s, t) &= \left\langle \partial_{\partial_s H(s,t)} f(H(s, t)), \partial_t H(s, t)\right\rangle + \langle f(H(s, t)), \partial_s \partial_t H(s, t)\rangle \\ &= \left\langle \partial_{\partial_t H(s,t)} f(H(s, t)), \partial_s H(s, t)\right\rangle + \langle f(H(s, t)), \partial_t \partial_s H(s, t)\rangle \\ &= \partial_t \tilde{f}_1(s, t). \end{aligned}$$

Also ist $\tilde{f}$ rotationsfrei. Des Weiteren gilt

$$\partial_s H(s, 0) = \partial_s H(s, 1) = 0$$

sowie

$$H(0,t) = \gamma_0(t), \quad H(1,t) = \gamma_1(t)$$

für alle $s, t \in [0,1]$ nach Definition der Homotopie. Wir erhalten also

$$\begin{aligned} 0 = \int_{[0,1]^2} \operatorname{rot}(\tilde{f}) \operatorname{dvol} &= 0 + \int_0^1 \tilde{f}_2(1,t)\,\mathrm{d}t - 0 - \int_0^1 \tilde{f}_2(0,t)\,\mathrm{d}t \\ &= \int_0^1 \langle f(H(1,t)), \partial_t H(1,t)\rangle\,\mathrm{d}t - \int_0^1 \langle f(H(0,t)), \partial_t H(0,t)\rangle\,\mathrm{d}t \\ &= \int_0^1 \langle f(\gamma_1(t)), \dot{\gamma}_1(t)\rangle\,\mathrm{d}t - \int_0^1 \langle f(\gamma_0(t)), \dot{\gamma}_0(t)\rangle\,\mathrm{d}t \end{aligned}$$

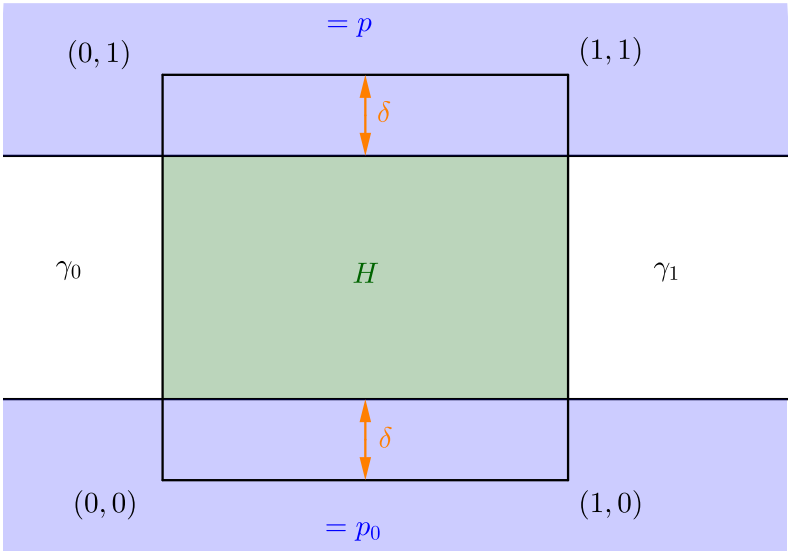
und das Lemma folgt. □

**Proposition 13.66** (Invarianz, stetiger Fall)**.** *Sei $U \subseteq \mathbb{R}^n$ ein Gebiet und $f : U \to \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld, welches die Integrabilitätsbedingungen erfüllt. Des Weiteren seien $\gamma_0, \gamma_1$ zwei glatte Wege von $p_0 \in U$ nach $p \in U$, die in $U$ (nicht zwingend auf glatte Weise) homotop sind. Dann gilt*

$$\int_{\gamma_0} f \cdot \mathrm{d}\mathbf{s} = \int_{\gamma_1} f \cdot \mathrm{d}\mathbf{s}.$$

*Beweisskizze.* Seien $\gamma_0, \gamma_1$ zwei glatte Wege, die in $U$ homotop sind. Wir werden zeigen, dass es auch eine glatte Homotopie zwischen $\gamma_0$ und $\gamma_1$ gibt, womit die Proposition aus Lemma 13.65 folgt.

Sei $H : [0,1]^2 \to U$ eine stetige Homotopie zwischen $\gamma_0$ und $\gamma_1$. Wir verwenden die Homotopie, um eine stetige Funktion $\tilde{H} : \mathbb{R}^2 \to U$ zu definieren, wobei wir dies mittels Fallunterscheidung und folgendem Bild machen.



Nun definieren wir

$$\tilde{H}_{\text{glatt}} = \psi * \tilde{H},$$

wobei wir eine Funktion $\psi$ wie in Proposition 13.57 für ein geeignetes $\delta > 0$ gewählt haben und die Faltung komponentenweise durch ein Integral über $\mathbb{R}^2$ definiert ist. Wir bemerken, dass auf Grund der Definition von $\tilde{H}$ in obiger Fallunterscheidung

$$\tilde{H}_{\text{glatt}}(s,0) = p_0, \quad \tilde{H}_{\text{glatt}}(s,1) = p$$

für alle $s \in [0,1]$. Nun verwenden wir drei glatte Funktionen $g_0, g, g_1 : [0,1] \to [0,1]$ mit

$$\operatorname{supp}(g_0) \subseteq [0,\delta), \qquad \operatorname{supp}(g_1) \subseteq (1-\delta,\delta],$$

sowie $g_0(0) = g_1(1) = 1$ und $g_0 + g + g_1 = 1$ (welche wir zum Beispiel aus Satz 13.60 für $K = [0,1]$, $O_0 = (-\delta,\delta)$, $O = (0,1)$, $O_1 = (1-\delta, 1+\delta)$ erhalten), um

$$H_{\text{glatt}}(s,t) = g_0(s)\gamma_0(t) + g(s)\tilde{H}_{\text{glatt}}(s,t) + g_1(s)\gamma_1(t)$$

für $(s,t) \in [0,1]^2$ zu definieren. Dies ist eine glatte Homotopie zwischen $\gamma_0$ und $\gamma_1$. Für genügend kleine $\delta > 0$ sind $\tilde{H}_{\text{glatt}}$ und $H_{\text{glatt}}$ uniform nahe an $H$ und nehmen deswegen ebenso Werte in $U$ an. □

### 13.7.4 Beweis der Konservativität

*Beweis von Satz 13.64.* Wie bereits erwähnt sind die Integrabilitätsbedingungen sogar auf allgemeinen Gebieten notwendig (siehe Satz 10.49). Sei also $U$ ein einfach zusammenhängendes Gebiet und $f$ ein stetig differenzierbares Vektorfeld, welches die Integrabilitätsbedingungen erfüllt. Proposition 13.66 beweist, dass $F(p)$ wie in (13.7) definiert in der Tat unabhängig ist von der Wahl des glatten Weges von $p_0$ nach $p$ und somit wohldefiniert ist. Wir bemerken ebenso, dass ein stückweise glatter Weg so reparametrisiert werden kann, dass dieser glatt ist (siehe auch die ähnliche Übung in Abschnitt 10.8.2).

Es bleibt nun zu zeigen, dass $F$ eine glatte Funktion ist, welche ein Potential von $f$ darstellt. Wir können für ein festes $j \in \{1,\dots,n\}$, $\varepsilon > 0$ klein genug und $p \in U$ einen glatten Weg $\gamma : [0,2] \to U$ mit $\gamma(0) = p_0$ so wählen, dass er für $t \in (1-\varepsilon, 1+\varepsilon)$ die Form $\gamma_p(t) = p + (t-1)e_j$ hat. Wir können also zum einen den Weg $\gamma_p = \gamma|_{[0,1]}$ von $p_0$ nach $p$ betrachten und zum anderen für $p + he_j$ für $|h| < \varepsilon$ den Weg $\gamma_{p+he_j}$ wählen. Also gilt

$$\partial_j F(p) = \lim_{h\to 0} \frac{F(p+he_j) - F(p)}{h} = \lim_{h\to 0} \tfrac{1}{h} \int_0^h \langle f(p+te_j), e_j\rangle \,\mathrm{d}t = f_j(p).$$

Da $j \in \{1,\dots,n\}$ und $p \in U$ beliebig waren, beweist dies den Satz auf Grund von Satz 10.49. □

## 13.8 Integrale über Teilmannigfaltigkeiten und das alternierende Tensorprodukt*

Wie können wir Volumen von $k$-dimensionalen Teilmannigfaltigkeiten definieren und berechnen? Für $n \geq 4$ und $k \geq 2$ ist unsere räumliche Anschauung meist nicht mehr ausreichend, um ohne Zweifel die richtige Definition einfach zu sehen. Deswegen benötigen wir einen Formalismus, der zweifelsfrei in allen Fällen zur richtigen Definition führt. Wir werden sehen, dass die Gramsche Determinante für diesen Formalismus ein geeignetes Hilfsmittel darstellt.

Wie sieht der Divergenzsatz im $\mathbb{R}^n$ für $n \geq 4$ aus? Ziemlich genauso wie für $n = 2$ und $n = 3$, ausser dass wir ein Flussintegral über den $(n-1)$-dimensionalen Rand verwenden müssen. Um dieses gerichtete Integral über den Rand definieren zu können, benötigen wir eine Verallgemeinerung des Kreuzprodukts auf $\mathbb{R}^n$, was einen Spezialfall einer alternierenden Multilinearform darstellt.

Wir erhielten den Satz von Stokes (Theorem 13.50), indem wir den zweidimensionalen Integralsatz in der Form des Satzes von Green (Theorem 13.18) in den $\mathbb{R}^3$ geschoben haben. Um dies für $k$-dimensionale Teilmannigfaltigkeiten im $\mathbb{R}^n$ bewerkstelligen zu können, benötigt man den Begriff einer Differentialform, welche wiederum ohne alternierende Multilinearformen nicht definiert werden kann.

### 13.8.1 Das skalare Integral über eine Teilmannigfaltigkeit

Sei $M \subseteq \mathbb{R}^n$ eine $k$-dimensionale Teilmannigfaltigkeit, $\Phi : V \to U$ ein Diffeomorphismus offener Teilmengen von $\mathbb{R}^n$ mit $\Phi(V \cap \mathbb{R}^k) = U \cap M$, wobei wir $\mathbb{R}^k$ mit $\mathbb{R}^k \times \{0\}^{n-k}$ identifiziert haben. Weiter sei $K \subseteq V \cap \mathbb{R}^k$ Jordan-messbar. Für eine stetige Funktion $f : M \to \mathbb{R}$ definieren wir

$$\int_{\Phi(K)} f \,\mathrm{d}\,\mathrm{vol}_M = \int_K f \circ \Phi \sqrt{\mathrm{gram}_k(\,\mathrm{D}\Phi)}\,\mathrm{d}\,\mathrm{vol}_k,$$

wobei

$$\mathrm{gram}_k(\,\mathrm{D}\Phi) = \det\big(\,\langle \partial_i \Phi, \partial_j \Phi\rangle\,\big)_{i,j\in\{1,\ldots,k\}}.$$

die Gramsche Determinante der partiellen Ableitungen darstellt. Das skalare Integral über die Teilmannigfaltigkeit $M$ wird dann mittels mehreren solchen Integralen definiert, wozu man $M$ als die disjunkte Vereinigung von Bildmengen des Typs $\Phi(K)$ schreibt. Man beachte an dieser Stelle, dass eine Begründung für die obige Definition des Integrals $\int_{\Phi(K)} f \,\mathrm{d}\,\mathrm{vol}_M$ in der mehrdimensionalen Substitutionsregel (Satz 12.56) und Korollar 12.54 gefunden werden kann.

**Übung 13.67** (Wohldefiniertheit)**.** *Nehmen Sie an, dass $\tilde{\Phi} : \tilde{V} \to \tilde{U}$ ein weiterer Diffeomorphismus ist und dass es eine Jordan-messbare Teilmenge $\tilde{K} \subseteq \tilde{V} \cap \mathbb{R}^k$ mit $\tilde{\Phi}(\tilde{K}) = \Phi(K)$ gibt. Zeigen Sie, dass*

$$\int_K f \circ \Phi \sqrt{\mathrm{gram}_k(\,\mathrm{D}\Phi)}\,\mathrm{d}\,\mathrm{vol}_M = \int_{\tilde{K}} f \circ \tilde{\Phi} \sqrt{\mathrm{gram}_k(\,\mathrm{D}\tilde{\Phi})}\,\mathrm{d}\,\mathrm{vol}_k$$

*gilt. Schliessen Sie daraus, dass die Integrale* $\int_{\Phi(K)} f \,\mathrm{d}\,\mathrm{vol}_M$ *und* $\int_M f \,\mathrm{d}\,\mathrm{vol}_M$ *nicht von den gewählten Parametrisierungen von* $M$ *abhängen.*

**Übung 13.68** (Volumen der 3-Sphäre)**.** *Sei* $M = \mathbb{S}^3 \subseteq \mathbb{R}^4$*. Verwenden Sie eine Verallgemeinerung von Kugelkoordinaten, um das drei-dimensionale Volumen dieser Mannigfaltigkeit zu berechnen.*

### 13.8.2 Das alternierende Tensorprodukt

Alle unsere mehrdimensionale Integralsätze (Theorem 13.14, Theorem 13.18, Theorem 13.41 und Theorem 13.50) benötigen vorzeichenbehaftete oder gerichtete Integrale. Dies trifft sogar für den Fundamentalsatz der Integral- und Differentialrechnung (Theorem 8.2 und seine Korollare) zu. Zum Beispiel ist die Vereinbarung

$$\int_{x_0}^{x_0+h} f(t)\,\mathrm{d}t = -\int_{x_0+h}^{x_0} f(t)\,\mathrm{d}t$$

für $x_0, x_0 + h \in I = [a, b]$ und eine stetige Funktion $f : I \to \mathbb{R}$ notwendig, damit die (links- und rechtseitige) Ableitung von $x \mapsto \int_{x_0}^{x} f(t)\,\mathrm{d}t$ bei $x = x_0$ durch $f(x_0)$ gegeben ist.

Des Weiteren haben wir in Proposition 13.48 die linearen Eigenschaften gewisser Grenzübergänge gesehen, die im direkten Zusammenhang mit dem Satz von Stokes (Theorem 13.50) stehen. Es ist daher zu erwarten, dass höherdimensionale Verallgemeinerungen unserer Integralsätze ebenso multilineare Ausdrücke enthalten. Wir wollen deswegen das alternierende Tensorprodukt $\bigwedge^k \mathbb{R}^n$ von $\mathbb{R}^n$ besprechen und den Zusammenhang zur Gramschen Determinante von $k$ Vektoren im $\mathbb{R}^n$ erklären, wobei wir uns hier nicht um eine basisunabhängige Definition des $k$-ten alternierenden Produkts $\bigwedge^k V$ eines Vektorraumes $V$ bemühen werden.

Wir definieren für $0 \leq k \leq n$ das **alternierende Tensorprodukt** $\bigwedge^k \mathbb{R}^n$ durch

$$\bigwedge^k \mathbb{R}^n = \mathbb{R}^{\mathcal{P}(n,k)} \cong \mathbb{R}^{\binom{n}{k}},$$

wobei

$$\mathcal{P}(n, k) = \{A \subseteq \{1, \ldots, n\} \mid |A| = k\}$$

die Menge der Teilmengen von $\{1, \ldots, n\}$ mit $k$ Elementen bezeichnet. Wir schreiben den Standardbasisvektor zu dem Index $A = \{j_1 < j_2 < \ldots < j_k\} \in \mathcal{P}(n, k)$ als

$$e_{j_1} \wedge \ldots \wedge e_{j_k} \in \bigwedge^k \mathbb{R}^n.$$

Für $A = \{j_1 < j_2 < \ldots, < j_k\} \in \mathcal{P}(n, k)$ und eine Permutation $\sigma \in S_k$ der Menge $\{1, \ldots, k\}$ definieren wir

$$e_{j_{\sigma(1)}} \wedge \ldots \wedge e_{j_{\sigma(k)}} = \mathrm{sgn}(\sigma) e_{j_1} \wedge \ldots \wedge e_{j_k}.$$

Falls die Teilmenge $\{j_1, \ldots, j_k\} \subseteq \{1, \ldots, n\}$ weniger als $k$ Elemente hat (weil $j_a = j_b$ für gewisse $a \neq b$ in $\{1, \ldots, k\}$), so setzen wir $e_{j_1} \wedge \ldots \wedge e_{j_k} = 0$. Damit ist der Ausdruck $e_{j_1} \wedge \ldots \wedge e_{j_k}$ für beliebige $j_1, \ldots, j_k \in \{1, \ldots, n\}$ definiert. Des Weiteren definieren wir für beliebige Vektoren $v_1, \ldots, v_k \in \mathbb{R}^n$

$$v_1 \wedge \ldots \wedge v_k = \sum_{j_1=1}^{n} \cdots \sum_{j_k=1}^{n} v_{1,j_1} \cdots v_{k,j_k} e_{j_1} \wedge \ldots \wedge e_{j_k},$$

wobei wir für $i \in \{1, \ldots, n\}$ den Vektor $v_i$ als $v_i = (v_{i,1}, \ldots, v_{i,n})^t$ schreiben. Wir erhalten somit die Abbildung

$$(v_1, \ldots, v_k) \in (\mathbb{R}^n)^k \mapsto v_1 \wedge \ldots \wedge v_k \in \bigwedge^k \mathbb{R}^n. \tag{13.8}$$

**Wichtige Übung 13.69** (Multilinear und alternierend)**.** *Zeigen Sie, dass die Abbildung in* (13.8) *multilinear und alternierend ist.*

**Proposition 13.70** (Universelle Eigenschaft des alternierenden Tensorprodukts)**.** *Das alternierende Tensorprodukt $\bigwedge^k \mathbb{R}^n$ besitzt folgende **universelle Eigenschaft**. Falls*

$$A : (\mathbb{R}^n)^k \to W$$

*eine multilineare, alternierende Abbildung mit Werten in einem $\mathbb{R}$-Vektorraum $W$ ist, so gibt es eine eindeutig bestimmte lineare Abbildung $L : \bigwedge^k \mathbb{R}^n \to W$ mit*

$$A(v_1, \ldots, v_k) = L(v_1 \wedge \ldots \wedge v_k) \tag{13.9}$$

*für alle $v_1, \ldots, v_k \in \mathbb{R}^n$.*

*Beweis.* Sei $A : (\mathbb{R}^n)^k \to W$ eine multilineare, alternierende Abbildung wie in der Proposition. Falls $L : \bigwedge^k \mathbb{R}^n \to W$ eine lineare Abbildung mit der Eigenschaft in Gleichung (13.9) ist, so gilt für alle $j_1 < \ldots < j_k$ in $\{1, \ldots, n\}$ die Gleichung

$$L(e_{j_1} \wedge \ldots \wedge e_{j_k}) = A(e_{j_1}, \ldots, e_{j_k}). \tag{13.10}$$

Da aber die Vektoren der Form $e_{j_1} \wedge \ldots \wedge e_{j_k}$ mit $j_1 < \ldots < j_k$ eine Basis von $\bigwedge^k \mathbb{R}^n$ bilden, folgt, dass $L$ durch die Eigenschaft in Gleichung (13.9) eindeutig bestimmt ist.

Für die Existenz der linearen Abbildung $L$ wie oben betrachten wir (13.10) als Definition einer linearen Abbildung $L$, wobei wir wieder verwenden, dass die Vektoren $e_{j_1} \wedge \ldots \wedge e_{j_k}$ mit $j_1 < \ldots < j_k$ in $\{1, \ldots, n\}$ eine Basis von $\bigwedge^k \mathbb{R}^n$ darstellen. Für $v_1, \ldots, v_k \in \mathbb{R}^n$ mit $v_i = (v_{i,1}, \ldots, v_{i,n})^t$ für alle $i \in \{1, \ldots, n\}$ gilt

$$A(v_1, \ldots, v_k) = A\Big(\sum_{j_1=1}^{n} v_{1,j_1} e_{j_1}, \ldots, \sum_{j_k=1}^{n} v_{k,j_k} e_{j_k}\Big) = \sum_{j_1=1}^{n} \cdots \sum_{j_k=1}^{n} v_{1,j_1} \cdots v_{k,j_k} A(e_{j_1}, \ldots, e_{j_k}),$$

da $A$ multilinear ist. Da $A$ aber auch alternierend ist, verschwindet $A(e_{j_1}, \ldots, e_{j_k})$ sobald ein Basisvektor zweimal in der Liste $e_{j_1}, \ldots, e_{j_k}$ vorkommt. Des Weiteren gilt

$$A(e_{1,j_{\sigma(1)}}, \ldots, e_{k,j_{\sigma(k)}}) = \operatorname{sgn}(\sigma) A(e_{j_1}, \ldots, e_{j_k}) \tag{13.11}$$

für $j_1 < \ldots < j_k$ und eine beliebige Transposition oder auch eine beliebige Permutation $\sigma \in S_k$. Dies zeigt, dass (13.10) sogar für beliebige $j_1, \ldots, j_k \in \{1, \ldots, n\}$ gilt. Wir schliessen

$$\begin{aligned} A(v_1, \ldots, v_k) &= \sum_{j_1=1}^{n} \cdots \sum_{j_k=1}^{n} v_{1,j_1} \cdots v_{k,j_k} L(e_{j_1} \wedge \ldots \wedge e_{j_k}) \\ &= L\Bigg( \sum_{j_1=1}^{n} \cdots \sum_{j_k=1}^{n} v_{1,j_1} \cdots v_{k,j_k} e_{j_1} \wedge \ldots \wedge e_{j_k} \Bigg) = L(v_1 \wedge \ldots \wedge v_k) \end{aligned}$$

wie erwünscht. □

Die geometrische Bedeutung des alternierenden Tensorprodukts wird klarer, wenn wir $\bigwedge^k \mathbb{R}^n$ mit dem inneren Produkt ausstatten, in dem die Standardbasisvektoren $e_{j_1} \wedge \ldots \wedge e_{j_k}$ für $j_1 < \ldots < j_k$ in $\{1, \ldots, n\}$ eine Orthonormalbasis bilden. Wir bemerken, dass Übung 13.27 zur Länge des Kreuzproduktes einen Spezialfall folgender Proposition bildet.

**Proposition 13.71** (Volumen und alternierende Tensoren)**.** *Für $v_1, \ldots, v_k \in \mathbb{R}^n$ gilt*

$$\|v_1 \wedge \ldots \wedge v_k\|_2 = \sqrt{\operatorname{gram}(v_1, \ldots, v_k)} = \sqrt{\det(\langle v_i, v_j \rangle)_{i,j=1,\ldots,k}}.$$

*In Worten ausgedrückt können wir die Länge des alternierenden Tensors $v_1 \wedge \ldots \wedge v_k$ als das Volumen des von $v_1, \ldots, v_k$ aufgespannten Parallelotops auffassen.*

Prinzipiell ist nach Quadrieren obige Aussage eine Identität zwischen polynomiellen Ausdrücken und kann für gegebenes $k$ direkt durch eine (lange) Rechnung verifiziert werden. Eine ungeordnete Rechnung dieser Form empfiehlt sich nicht wirklich; stattdessen möchten wir die universelle Eigenschaft des alternierenden Tensorprodukts in Proposition 13.70 verwenden, um dem Beweis Struktur zu verleihen.

*Beweis.* Wir betrachten für $v_1, \ldots, v_k, w_1, \ldots, w_k \in \mathbb{R}^n$ den Ausdruck $\det(\langle v_i, w_j \rangle)_{i,j=1,\ldots,k}$ und bemerken, dass dieser multilinear ist. Des Weiteren ist für feste $w_1, \ldots, w_k \in \mathbb{R}^n$ die Abbildung

$$A_{w_1,\ldots,w_k} : (v_1, \ldots, v_k) \in (\mathbb{R}^n)^k \mapsto \det(\langle v_i, w_j \rangle)_{i,j=1,\ldots,k} \in \mathbb{R}$$

alternierend und multilinear. Nach Proposition 13.70 gibt es daher eine eindeutig bestimmte lineare Abbildung $L_{w_1,\ldots,w_k} : \bigwedge^k \mathbb{R}^n \to \mathbb{R}$ mit

$$L_{w_1,\ldots,w_k}(v_1 \wedge \ldots \wedge v_k) = \det(\langle v_i, w_j \rangle)_{i,j=1,\ldots,k}$$

für alle $v_1, \ldots, v_k \in \mathbb{R}^n$. Des Weiteren gelten für $w_1, \ldots, w_k \in \mathbb{R}^n$ und $s_1, \ldots, s_k \in \mathbb{R}$ die Gleichungen

$$\begin{aligned}
L_{w_{\sigma(1)},\ldots,w_{\sigma(k)}}(v_1 \wedge \ldots \wedge v_k) &= \det(\langle v_i, w_{\sigma(j)} \rangle)_{i,j=1,\ldots,k} \\
&= \operatorname{sgn}(\sigma) \det(\langle v_i, w_j \rangle)_{i,j=1,\ldots,k} \\
&= \operatorname{sgn}(\sigma) L_{w_1,\ldots,w_k}(v_1 \wedge \ldots \wedge v_k) \\
L_{s_1 w_1,\ldots,s_k w_k}(v_1 \wedge \ldots \wedge v_k) &= \det(\langle v_i, s_j w_j \rangle)_{i,j=1,\ldots,k} \\
&= s_1 \ldots s_k \det(\langle v_i, w_j \rangle)_{i,j=1,\ldots,k} \\
&= s_1 \ldots s_k L_{w_1,\ldots,w_k}(v_1 \wedge \ldots \wedge v_k)
\end{aligned}$$

auf Grund der Eigenschaften der Determinante für alle $v_1, \ldots, v_k \in \mathbb{R}^n$, woraus mit der Eindeutigkeit der linearen Abbildung $L_{w_1,\ldots,w_k}$ die Gleichungen

$$\begin{aligned}
L_{w_{\sigma(1)},\ldots,w_{\sigma(k)}} &= \operatorname{sgn}(\sigma) L_{w_1,\ldots,w_k} \\
L_{s_1 w_1,\ldots,s_k w_k} &= s_1 \ldots s_k L_{w_1,\ldots,w_k}
\end{aligned}$$

folgen. Die Gleichungen

$$\begin{aligned}
L_{w_1+w,w_2,\ldots,w_k} &= L_{w_1,w_2,\ldots,w_k} + L_{w,w_2\ldots,w_k} \\
L_{w_1,w_2+w,w_3,\ldots,w_k} &= L_{w_1,w_2,w_3,\ldots,w_k} + L_{w_1,w,w_3\ldots,w_k} \\
&\vdots \\
L_{w_1,\ldots,w_k+w} &= L_{w_1,\ldots,w_k} + L_{w_1,\ldots,w}
\end{aligned}$$

für ein weiteres $w \in \mathbb{R}^n$ folgen analog. Daher ist

$$(w_1, \ldots, w_k) \in (\mathbb{R}^n)^k \mapsto L_{w_1,\ldots,w_k}$$

eine multilineare, alternierende Abbildung mit Werten im Dualraum von $\bigwedge^k \mathbb{R}^n$. Dies induziert mittels der universellen Eigenschaft in Proposition 13.70 eine lineare Abbildung von $\bigwedge^k \mathbb{R}^n$ in den Dualraum von $\bigwedge^k \mathbb{R}^n$. Insgesamt erhält man damit eine bilineare Abbildung

$$B : \bigwedge^k \mathbb{R}^n \times \bigwedge^k \mathbb{R}^n \to \mathbb{R}$$

mit der Eigenschaft

$$B(v_1 \wedge \ldots \wedge v_k, w_1 \wedge \ldots \wedge w_k) = L_{w_1,\ldots,w_k}(v_1 \wedge \ldots \wedge v_k) = \det(\langle v_i, w_j \rangle)_{i,j=1,\ldots,k}$$

für alle $v_1, \ldots, v_k, w_1, \ldots, w_k \in \mathbb{R}^n$. Verwenden wir nun die Standardbasisvektoren $e_{j_1} \wedge \ldots \wedge e_{j_k}$ für $j_1 < \ldots < j_k$ in $\{1, \ldots, n\}$, so erkennen wir, dass $B$ das innere Produkt ist, bezüglich dem die Standardbasis eine Orthonormalbasis darstellt. Dies beweist die Proposition. $\square$

Aus der Diskussion in Abschnitt 13.2.2 und Proposition 13.71 ergibt sich, dass wir

$$v_1 \wedge \ldots \wedge v_k$$

als gerichtetes Volumenelement, das dem von $v_1, \ldots, v_k \in \mathbb{R}^n$ aufgespannten Parallelotop entspricht, auffassen können.

### 13.8.3 Differentialformen

Der allgemeine Integralsatz von Stokes ist ein Satz, der ein $(k-1)$-dimensionales Integral über den Rand $\partial M$ einer orientierbaren $k$-dimensionalen Teilmannigfaltigkeit einer „$(k-1)$-Differentialform“ $\omega$ mit dem $k$-dimensionalen Integral über $M$ der „abgeleiteten $k$-Differentialform“ $\mathrm{d}\omega$ in Verbindung bringt. Genauer gilt

$$\int_B \mathrm{d}\omega = \int_{\partial B} \omega$$

schreiben, wobei hinter dieser Gleichung natürlich einige Arbeit steckt. Wir verweisen für diese auf die bestehende Literatur, wie zum Beispiel auf Zorich's Buch [Zor07].

Dennoch wollen wir den Begriff der $k$-Differentialform oder kurz $k$-Form mit den uns bekannten Begriffen in Verbindung bringen. Sei also $U \subseteq \mathbb{R}^n$ offen (die Definitionen machen analog auch für Teilmannigfaltigkeiten von $\mathbb{R}^n$ Sinn).

Eine 1-**Form** ist eine glatte Abbildung, die jedem $p \in U$ ein Element $\omega_p$ im Dualraum von $\mathrm{T}_pU$, also eine Linearform auf $\mathrm{T}_pU$ zuweist. Das wichtigste Beispiel ist der Gradient einer stetig differenzierbaren Funktion $\varphi : U \to \mathbb{R}$ oder genauer die Abbildungen

$$\omega_p : (p, v) \in \mathrm{T}_pU \mapsto \tfrac{d}{dt}(t \mapsto \varphi(p + tv)) = \langle \nabla\varphi(p), v \rangle$$

für $p \in U$. Ein weiteres Beispiel ist durch Kraftfelder gegeben, welche Richtung und Stärke der Krafteinwirkung bei jedem $p \in U$ beschreiben. Das Integral über eine eindimensionale Teilmannigfaltigkeit ist in diesem Fall ein Wegintegral und der Satz von Stokes ist in diesem Fall bloss der Hauptsatz der Integral- und Differentialrechnung (Theorem 8.2) in den $\mathbb{R}^n$ geschoben (also Satz 10.49). Eine 1-Form im $\mathbb{R}^n$ wird meist als

$$g_1 \,\mathrm{d}x_1 + \ldots + g_n \,\mathrm{d}x_n$$

geschrieben, wobei $g_1, \ldots, g_n$ Funktionen auf $U$ sind. Eine solche 1-Form wirkt dann an einem Punkt $p \in U$ durch $v \in \mathrm{T}_pU \mapsto \sum_{k=1}^n g_k(p)v_i$.

Eine 2-**Form** ist eine glatte Abbildung, die jedem $p \in U$ ein Element $\omega_p$ im Dualraum von $\bigwedge^2 \mathrm{T}_pU$ zuordnet, wobei wegen der universellen Eigenschaft in Proposition 13.70 $\omega_p$ mit einer bilinearen, alternierenden Abbildung $\mathrm{T}_pU \times \mathrm{T}_pU \to \mathbb{R}$ identifiziert werden kann. Ein wichtiges Beispiel einer solchen Abbildung haben wir in Proposition 13.48 gesehen.

**Übung 13.72.** *Übertragen Sie Proposition 13.48 auf „infinitesimale“ Parallelogramme im $\mathbb{R}^n$ für $n \geq 4$.*

Wir sind 2-Formen sowohl im Satz von Stokes (Theorem 13.50) als auch im Satz von Gauss (Theorem 13.41) begegnet. Auf Grund des Zusammenhangs zwischen Integration und alternierenden Tensoren wird auch das $n$-dimensionale Riemann-Integral einer Riemann-integrierbaren Funktion $F$ auf einer Jordan-messbaren Menge oft in der Form

$$\int_B F \, \mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}x_n$$

geschrieben.

Wie bereits erwähnt verweisen wir für jegliche Details für diese Themen auf [Zor07].

## 13.9 Weitere Lernmaterialien

### 13.9.1 Verwendung des Kapitels

Die Integralsätze des $\mathbb{R}^2$ und $\mathbb{R}^3$ sind für Anwendungen innerhalb und ausserhalb der Mathematik von grosser Bedeutung und stellen gemeinsam mit dem Begriff der Konservativität von Vektorfeldern die Standardtools für die Vektoranalysis dar.

Eine häufige Anwendungen der Sätze besteht darin, dass nach Anwendung des Satzes das Integral mitunter viel einfacher zu berechnen ist als das ursprüngliche Integral. Zum Beispiel kann ein Flussintegral durch eine komplizierte Fläche für ein divergenzfreies Vektorfeld vielleicht einfacher berechnet werden, wenn man eine einfachere Fläche findet, die gemeinsam mit der ursprünglichen Fläche genau den Rand eines (stückweise) glatt berandeten Bereichs darstellt.

### 13.9.2 Übungen

**Übung.** *Seien die glatte Funktion $f(x,y,z) = (yz, x^2, 1)$ und die Fläche*

$$S = \{(x,y,z)^t \in \mathbb{R}^3 \mid x^2 + y^2 = 1,\ 0 < z < 1\}$$

*gegeben. Berechnen Sie das Flussintegral (nach aussen) $\int_S \operatorname{rot} f \cdot \mathrm{d}\mathbf{n}$ mit und ohne dem Satz von Stokes.*

**Übung.** *Bestimmen Sie Parameter $\alpha, \beta, \gamma \in \mathbb{R}$, sodass das Vektorfeld*

$$f(x,y,z) = (x + 2y + \alpha z, \beta x - 3y - z, 4x + \gamma y + 2z)$$

*auf $\mathbb{R}^3$ rotationsfrei wird. Bestimmen Sie nach Wahl dieser Parameter ein Potential von $f$.*

**Übung.** *Sei $0 < d < 1$ und sei $S$ jener Teil von $\mathbb{S}^2 \subseteq \mathbb{R}^3$, welcher $z > -d$ erfüllt. Sei $f(x,y,z) = (-y, x, 0)^T$ und $v = \operatorname{rot} f$. Berechnen Sie den Fluss $\int_S v \cdot \mathrm{d}\mathbf{n}$ nach aussen wie folgt:*

*(i) direkt,*

*(ii) mit dem Satz von Stokes,*

*(iii) mit dem Satz von Gauss.*

**Übung.** *Betrachten Sie auf $\mathbb{R}^3 \setminus (\{(0,0)\} \times \mathbb{R})$ das Vektorfeld*

$$f(x,y,z) = \left( \frac{2(xz+y)}{x^2+y^2}, \frac{2(yz-x)}{x^2+y^2}, \log\left(x^2+y^2\right) \right)$$

*und die Kurven*

$$\gamma_1 : [0, 2\pi] \to \mathbb{R}^3, \varphi \mapsto (\cos\varphi, \sin\varphi, 0)$$
$$\gamma_2 : [0, 2\pi] \to \mathbb{R}^3, \varphi \mapsto (2\cos\varphi, 2\sin\varphi, 2)\ .$$

*(i) Berechnen Sie die Differenz* $\int_{\gamma_1} f \cdot \mathrm{d}\mathbf{s} - \int_{\gamma_2} f \cdot \mathrm{d}\mathbf{s}$ *mit Hilfe des Satzes von Stokes.*

*(ii) Ist* $f$ *konservativ?*

**Übung.** *Sei* $f : \mathbb{R}^2 \to \mathbb{R}^2$ *gegeben durch* $f(x,y) = e^{\frac{1}{2}(y^2-x^2)}(\cos(xy), \sin(xy))$. *Des Weiteren sei für jede Wahl von Parametern* $a \in \mathbb{R}, b > 0$ *die Kurve* $\gamma_{a,b} : [-b,b] \to \mathbb{R}^2$, $\gamma_{a,b}(t) = (t,a)^T$ *gegeben. Zeigen Sie, dass*

$$\int_{\gamma_{a,b}} f \cdot \mathrm{d}\mathbf{s} = \int_{\gamma_{0,b}} f \cdot \mathrm{d}\mathbf{s} + o(1) \text{ für } b \to \infty.$$

**Übung.** *Sei* $\Omega$ *das nicht-einfach zusammenhängende Gebiet* $\mathbb{R}^2 \setminus \overline{B_1(0)}$ *und sei* $f$ *ein rotationsfreies, stetig differenzierbares Vektorfeld auf* $\Omega$. *Zeigen sie*

$$\int_{\partial B_2(0)} f \cdot \mathrm{d}\mathbf{s} = 0 \iff \exists\, \varphi\ C^2 \text{ mit } f = \nabla\varphi$$

*in mehreren Schritten.*

*(i) Die Implikation* $\Leftarrow$ *gilt.*

*(ii) Die Mengen* $\Omega_\pm = \left\{\mathbf{x} \in \Omega \middle| \pm y < \frac{1}{2}\right\}$ *sind einfach zusammenhängend.*

*(iii) Es existieren* $C^2$*-Funktionen* $\varphi_\pm$ *auf* $\Omega_\pm$ *mit* $\nabla\varphi_\pm = f_{|\Omega_\pm}$. *Insbesondere ist* $\nabla(\varphi_+ - \varphi_-) = 0$ *auf* $\Omega_+ \cap \Omega_-$.

*(iv) Die Implikation* $\Rightarrow$ *gilt: Konstruieren Sie* $\varphi$ *mit Hilfe von* $\varphi_\pm$.

**Übung.** *Sei* $f : \mathbb{R}^3 \to \mathbb{R}^3$ *ein stetig differenzierbares Vektorfeld und sei* $\varphi : \mathbb{R}^3 \to \mathbb{R}$ *eine* $C^2$*-Funktion. Wir erinnern daran, dass der Laplace-Operator auf* $\mathbb{R}^3$ *durch* $\Delta = \sum_{j=1}^3 \frac{\partial^2}{\partial x_j^2}$ *gegeben ist und bezeichnen mit* $\nabla\varphi$ *den Gradienten von* $\varphi$. *Beweisen Sie die folgenden Identitäten.*

*(i)* $\operatorname{rot}(\varphi f) = (\nabla\varphi) \times f + \varphi \operatorname{rot} f$,

*(ii)* $\operatorname{div}(\nabla\varphi) = \Delta\varphi$,

*(iii)* $\operatorname{rot}(\nabla\varphi) = 0$.

**Übung** (Existenz eines Potentials)**.** *Sei* $U \subseteq \mathbb{R}^2$ *ein sternförmiges Gebiet und sei* $f : U \to \mathbb{R}^2$ *ein stetig differenzierbares Vektorfeld. Zeigen Sie als Anwendung des Satzes von Green (Theorem 13.18), dass* $f$ *genau dann konservativ ist, wenn* $f$ *rotationsfrei ist.*

**Übung.** *Sei* $f$ *ein zweimal stetig differenzierbares Vektorfeld auf einer offenen Menge* $U$. *Zeigen Sie auf folgende zwei Art und Weisen, dass* $\operatorname{div}(\operatorname{rot}(f)) = 0$ *gilt.*

*(i) Berechnen Sie dies direkt mit den nötigen partiellen Ableitungen.*

*(ii) Verwenden Sie, den Satz von Stokes und den Satz von Gauss, um dies zu beweisen.*

**Übung.** *Sei* $A \in \mathrm{GL}_3(\mathbb{R})$ *und* $f : \mathbb{R}^3 \to \mathbb{R}^3$ *ein stetig differenzierbares Vektorfeld. Wir definieren das Vektorfeld* $f_A := A \circ f \circ A^{-1}$.

*(i) Beweisen Sie* $\operatorname{div}(f_A)(p) = \operatorname{div}(f)(A^{-1}p)$ *für alle* $p \in \mathbb{R}^3$.

*(ii) Beweisen Sie* $\operatorname{rot}(f_A)(p) = \operatorname{rot}(f)(A^{-1}p)$ *für alle* $p \in \mathbb{R}^3$, *falls* $A \in \mathrm{SO}(3, \mathbb{R})$.

**Übung** (Ein Energieerhaltungsgesetz für Wellen)**.** *In dieser Übung möchten wir zeigen, dass einer Lösung* $F(p,t)$ *der Wellengleichung (siehe Abschnitt 13.4.1) eine natürliche Energie* $E(t)$ *zugewiesen werden kann, welche dann in der Zeit konstant ist.*

*Sei* $B \subseteq \mathbb{R}^3$ *ein glatt berandeter Bereich.*

*(i) (Greensche Formel) Seien* $F, G : U \to \mathbb{R}$ *zweimal stetig differenzierbare Funktion definiert auf einer offenen Obermenge* $U$ *von* $B$. *Zeigen Sie die Formel*

$$\int_B \langle \nabla F, \nabla G \rangle \, \mathrm{dvol} = \int_{\partial B} G \, (\nabla F) \cdot \mathrm{d}\mathbf{n} - \int_B G \, (\Delta F) \, \mathrm{dvol}.$$

*Wir betrachten nun eine zweimal differenzierbare Funktion* $F : (p,t) \in U \times \mathbb{R} \mapsto F(p,t) \in \mathbb{R}$, *womit also* $u$ *eine Funktion in drei „Ortsvariablen“* $p = (x,y,z)$ *und einer „Zeitvariable“* $t$ *ist. Weiter schreiben wir* $\Delta_p F = \partial_x^2 F + \partial_y^2 F + \partial_z^2 F$. *Angenommen* $F$ *genügt der Wellengleichung*

$$\begin{aligned} \partial_t^2 F(p,t) &= \Delta_p F(p,t) \text{ für alle } (p,t) \in B \times \mathbb{R} \\ F(p,t) &= 0 \text{ für alle } (p,t) \in \partial B \times \mathbb{R} \end{aligned}$$

*(ii) (Energieerhaltungsgesetz) Zeigen Sie, dass die Energiefunktion* $E : \mathbb{R} \to \mathbb{R}_{>0}$ *gegeben durch*

$$E(t) = \int_B \Big( (\partial_t F(p,t))^2 + (\partial_x F(p,t))^2 + (\partial_y F(p,t))^2 + (\partial_z F(p,t))^2 \Big) \, \mathrm{dvol}(p)$$

*für* $t \in \mathbb{R}$ *konstant ist.*

**Übung.** *Sei* $A \in \mathrm{GL}_n(\mathbb{R})$ *und* $F : \mathbb{R}^n \to \mathbb{R}$ *eine zweimal stetig differenzierbare Funktion. Wir definieren die Funktion* $F_A := F \circ A^{-1}$.

*(i) Beweisen Sie* $(\Delta F_A)(p) = (\Delta F)(A^{-1}p)$ *für alle* $p \in \mathbb{R}^n$ *und* $A \in \mathrm{O}(n, \mathbb{R})$.

*(ii) Bestimmen Sie die maximale Untergruppe*

$$\begin{aligned} H = \{ A \in \mathrm{GL}_n(\mathbb{R}) \mid (\Delta F_A)(p) = (\Delta F)(A^{-1}p) \text{ für alle } p \in \mathbb{R}^n \\ \text{und } C^2\text{-Abbildungen } F : \mathbb{R}^n \to \mathbb{R} \}. \end{aligned}$$

### 13.9.3 Lernkarten

Sie können wiederum die Lernkarten oder den Graphen für Ihre Wiederholung der Themen des Kapitels verwenden.

# Kapitel 14

# Gewöhnliche Differentialgleichungssysteme

Wir werden nun das Thema der Differentialgleichungen, welches sowohl für vielfältige Anwendungen in den Naturwissenschaften und der Technik als auch für viele weitere mathematische Fragestellungen wichtig ist, ausführlicher besprechen. Dabei werden wir auch die Begriffe und Fragestellungen von Differentialgleichungen wie in Abschnitt 7.5 zu gekoppelten Differentialgleichungssystemen erweitern. Des Weiteren wollen wir die Existenz und Eindeutigkeit für entsprechende Anfangswertprobleme unter sehr schwachen und natürlichen Annahmen beweisen. Hierfür werden wir mittels einer geeigneten Umformulierung des Anfangswertproblems nochmals den Banachschen Fixpunktsatz anwenden.

## 14.1 Mehrdeutigkeit der Lösung

Wir wollen hier anhand eines einfachen Beispiels zeigen, dass die Lösung eines Anfangswertproblems wie in Abschnitt 7.5 nicht immer eindeutig bestimmt ist.

**Beispiel 14.1** (Ein Anfangswertproblem mit überabzählbar vielen Lösungen)**.** *Wir betrachten das Anfangswertproblem*

$$\begin{aligned} y' &= 3y^{\frac{2}{3}} \\ y(0) &= -1, \end{aligned}$$

*wobei wir die Definition* $a^{\frac{2}{3}} = (a^2)^{\frac{1}{3}} = |a|^{\frac{2}{3}}$ *für alle* $a \in \mathbb{R}$ *verwenden. Wir verwenden nochmals die Leibniz-Notation und die Methode der Trennung der Variablen (siehe auch Abschnitt 14.1 für eine allgemeinere Diskussion), um einen Kandidaten für eine Lösung zu finden. Es ergibt*

*sich daraus*

$$\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}x} &= 3y^{\frac{2}{3}} \\
\frac{\mathrm{d}y}{y^{\frac{2}{3}}} &= 3\,\mathrm{d}x \\
\int \frac{\mathrm{d}y}{y^{\frac{2}{3}}} &= 3\int \mathrm{d}x \\
\frac{1}{-\frac{2}{3}+1} y^{-\frac{2}{3}+1} &= 3x + C' \\
y^{\frac{1}{3}} &= x + C \\
y &= (x+C)^3,
\end{aligned}$$

*wobei* $C'$ *ebenso wie* $C = \frac{1}{3}C'$ *eine unbekannte Konstante darstellt. Für* $x = 0$ *soll* $y(0) = (0+C)^3 = -1 = C^3$ *sein, weswegen* $C = -1$ *ist und wir die Funktion*

$$x \mapsto y_1(x) = (x-1)^3$$

*erhalten. Obwohl die Schritte der obigen Rechnung nicht unserem Standard für Argumentationen genügen, kann (und soll) man nun schnell überprüfen, dass* $y_1$ *tatsächlich eine Lösung des Anfangswertproblems darstellt.*

*Wir möchten jetzt aber für jedes* $s \geq 1$ *eine Lösung* $y_s$ *des Anfangswertproblems definieren, so dass* $y_s$ *für* $s = 1$*, wie die Notation suggeriert,* $y_1$ *ist, und so dass die* $y_s$ *jeweils verschieden sind. Somit besitzt dieses Anfangswertproblem überabzählbar viele Lösungen. Zu* $s \in (1, \infty]$ *setzen wir*

$$x \mapsto y_s(x) = \begin{cases} (x-1)^3 & \textit{falls } x \leq 1 \\ 0 & \textit{falls } 1 < x < s \\ (x-s)^3 & \textit{falls } x \geq s \end{cases}$$

*und zeigen, dass* $y_s$ *das Anfangswertproblem löst. Es gilt*

$$\begin{aligned}
y_s'(x) &= \begin{cases} 3(x-1)^2 & \textit{falls } x < 1 \\ 0 & \textit{falls } 1 < x < s \\ 3(x-s)^2 & \textit{falls } x > s \end{cases} \\
&= 3y_s^{\frac{2}{3}}(x)
\end{aligned}$$

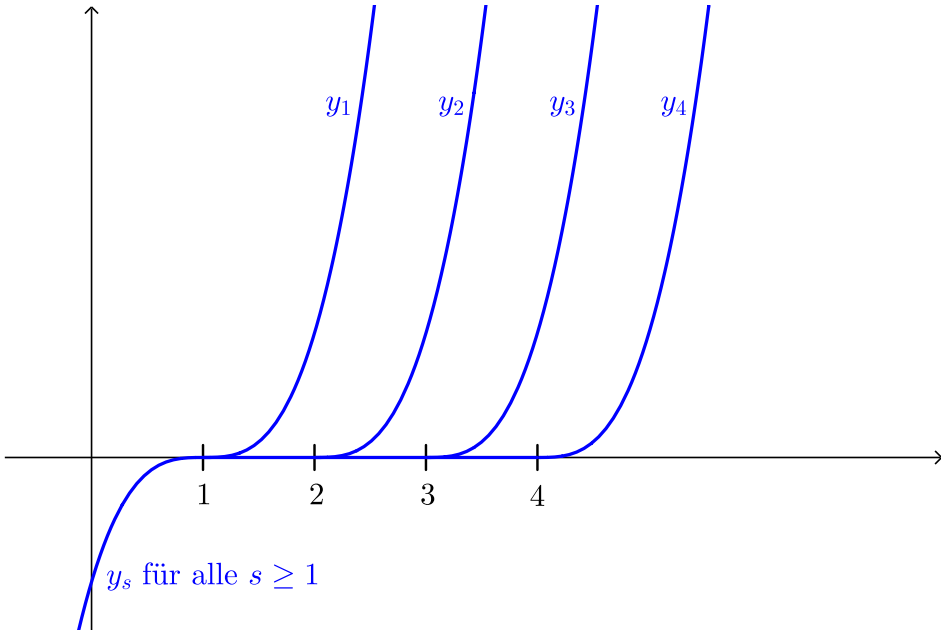
*für alle* $x \in \mathbb{R} \setminus \{1, s\}$*. Bei* $x = 1$ *gilt für die links- respektive rechtseitige Ableitung*

$$\begin{aligned}
(y_s)_-'(1) &= 3(1-1)^2 = 0 = 3y_s(1)^{\frac{2}{3}}, \\
(y_s)_+'(1) &= 0 = 3y_s(1)^{\frac{2}{3}} = (y_s)_-'(1)
\end{aligned}$$

*und bei $x = s$ gilt*

$$(y_s)'_-(s) = 0 = 3y_s(1)^{\frac{2}{3}},$$
$$(y_s)'_+(s) = 3(s-s)^2 = 0 = 3y_s(1)^{\frac{2}{3}} = (y_s)'_-(s).$$

*Somit ist $y_s$ stetig differenzierbar und löst das Anfangswertproblem. In der Tat ist der Anfangswert von $y_s$ für alle $s > 1$ durch $y_s(0) = (0-1)^3 = -1$ gegeben.*




Figur 14.1: Die Differentialgleichung $y' = 3y^{\frac{2}{3}}$ hat „Weichen im Richtungsfeld“ entlang der ganzen $x$-Achse.

Wie wir später sehen werden, ist die Struktur der Funktion $(x, y) \mapsto y^{\frac{2}{3}}$ für die Mehrdeutigkeit in obigem Beispiel verantwortlich. In der Tat werden wir in Abschnitt 14.4 zeigen, dass Lipschitz-Stetigkeit der Funktion $(x, y) \mapsto f(x, y)$ die Eindeutigkeit der Lösung des Anfangwertproblems

$$y' = f(x, y), \quad y(x_0) = y_0$$

impliziert. Die Differentialgleichung in Beispiel 14.1 erfüllt diese Bedingung allerdings in der Nähe der $x$-Achse nicht und dies ist auch genau der Ort wo die Mehrdeutigkeit der Lösungen entsteht.

Wir wollen noch bemerken, dass es je nach Anwendung vielleicht genügt eine Lösung gefunden zu haben. Es könnte aber auch sein, dass in der Anwendung die Mehrdeutigkeit eine spezielle Bedeutung hat. Das Theorem von Picard-Lindelöf (Theorem 14.23) über Existenz und Eindeutigkeit der Lösung (und dessen Annahmen) kann in diesem Sinne auch als Hilfsmittel zur Auffindung der möglichen „Weichen eines Richtungsfeldes“ verstanden werden.

### 14.1.1 Separierbare Differentialgleichungen: ein Leibniz Kochrezept

Die in Beispiel 14.1 angewandte (und bereits in Abschnitt 7.5.3 für lineare Differentialgleichungen erster Ordnung angedeutete) Methode zur Berechnung von Lösungen gewisser Differentialgleichungen kann allgemeiner wie folgt formuliert werden.

Wie wollen annehmen, dass die Differentialgleichung erster Ordnung die Form

$$y' = f(x)g(y) \tag{14.1}$$

annimmt. In diesem Fall nennen wir die Differentialgleichung separierbar und können in der Leibniz-Notation wie folgt die Variablen trennen.

$$\begin{aligned} \frac{\mathrm{d}y}{\mathrm{d}x} &= f(x)g(y) \\ \frac{\mathrm{d}y}{g(y)} &= f(x)\,\mathrm{d}x \\ \int \frac{\mathrm{d}y}{g(y)} &= \int f(x)\,\mathrm{d}x \end{aligned}$$

Dabei hoffen wir nun, dass die Integrale auf den beiden Seiten der Gleichung konkret berechnet werden können. In diesem Fall ergibt sich ein Gleichungssystem für $y$, welches man anschliessend versucht nach $y$ aufzulösen (siehe Abschnitt 11.1.1). Wir wollen diese Methode weiter nicht formal begründen, was aber abgesehen von der Möglichkeit, dass $g(y)$ wie in Beispiel 14.1 verschwinden könnte, mittels der Kettenregel möglich ist.

**Übung 14.2** (Methode der Trennung der Variablen)**.** *Angenommen obige Integrale können berechnet werden und es ergibt sich die Gleichung $G(y) = F(x) + C$, wobei $G$ eine Stammfunktion von $\frac{1}{g}$ und $F$ eine Stammfunktion von $f$ ist. Angenommen diese Gleichung lässt sich nach Wahl von $C$ und innerhalb einer gewissen offenen Teilmenge $U \subseteq \mathbb{R}^2$ eindeutig nach $y = y(x)$ für $x$ in einem Intervall $I \subseteq \mathbb{R}$ und eine differenzierbare Funktion $y : I \to \mathbb{R}$ auflösen. Nehmen Sie schlussendlich noch an, dass $g(y(x)) \neq 0$ gilt für alle $x \in I$. Zeigen Sie, dass dann $y = y(x)$ die Differentialgleichung* (14.1) *löst.*

**Übung 14.3** (Trennung der Variablen)**.** *Finden Sie eine Lösung des Anfangswertproblems*

$$\sqrt{1 - x^2}\, y' - y^2 = 1, \quad y(0) = 0.$$

## 14.2 Lineare Differentialgleichungen mit konstanten Koeffizienten

Wir wollen hier das Verfahren aus Abschnitt 7.5.4 (für den harmonischen Oszillator und ähnliche linearen Differentialgleichungen zweiter Ordnung) nochmals allgemeiner diskutieren. Insbesondere empfiehlt es sich Abschnitte 7.5.3–7.5.4 und das Beispiel der gedämpften Schwingung (mit oder ohne Reibung oder Krafeinwirkung) in Erinnerung zu rufen.

Sei $d \in \mathbb{N}$. Eine **lineare**, **gewöhnliche Differentialgleichung** $d$**-ter Ordnung mit konstanten Koeffizienten** ist eine Differentialgleichung der Form

$$y^{(d)} + a_{d-1}y^{(d-1)} + \ldots + a_0 y^{(0)} = g(x), \tag{14.2}$$

wobei die Koeffizienten $a_{d-1}, \ldots, a_0$ reelle Zahlen und $g : I \to \mathbb{R}$ eine **Störfunktion** $g$ auf einem Intervall $I \subseteq \mathbb{R}$ ist. Falls $g \neq 0$ ist, so nennt man die Differentialgleichung auch **inhomogen**. Falls die Störfunktion verschwindet, nimmt die Gleichung die Gestalt

$$y^{(d)} + a_{d-1}y^{(d-1)} + \ldots + a_0 y^{(0)} = 0 \tag{14.3}$$

an und wird **homogen** genannt.

Das zugehörige **Anfangswertproblem** hat zusätzlich die Bedingungen

$$y(x_0) = w_0,\ y'(x_0) = w_1, \ldots,\ y^{(d-1)}(x_0) = w_{d-1} \tag{14.4}$$

für ein $x_0 \in \mathbb{R}$ (beziehungsweise $x_0 \in I$ falls die Störfunktion $g$ nur auf $I \subseteq \mathbb{R}$ definiert ist) und **Anfangswerte** $w_0, w_1, \ldots, w_{d-1} \in \mathbb{R}$. Wie schon für lineare gewöhnliche Differentialgleichungen erster Ordnung in Abschnitt 7.5.3 lassen sich für obige Differentialgleichungen alle Lösungen der inhomogenen Gleichung (14.2) mit Hilfe einer partikulären Lösungen und der Lösungen der homogenen Gleichung (14.3) bestimmen. Wir erinnern hier daran, dass eine partikuläre Lösung der inhomogenen Gleichung schlicht eine beliebige Lösung ist (es reicht also, nur eine solche zu finden). Wie wir schon in Abschnitt 7.5.4 gesehen haben ist es von grossem Vorteil, hier nicht nur reellwertige, sondern auch komplexwertige[1] Lösungen $y$ zu erlauben.

**Wichtige Übung 14.4** (Struktur der Lösungsmengen)**.** *Sei $I \subseteq \mathbb{R}$ ein Intervall und $g : I \to \mathbb{R}$ die Störfunktion. Zeigen Sie, dass die komplexwertigen Lösungen von* (14.3) *auf $I \subseteq \mathbb{R}$ einen Vektorraum $V$ über $\mathbb{C}$ bilden und dass die Lösungsmenge von* (14.2) *die Form $y_{\text{part}} + V$ hat, falls $y_{\text{part}}$ eine partikuläre Lösung von* (14.2) *ist.*

### 14.2.1 Die Lösungen für die homogene Gleichung

Um die Lösungen der homogenen Gleichung in (14.3) zu erraten (siehe auch Beispiel 7.65), setzen wir $y = \mathrm{e}^{\alpha x}$ für ein $\alpha \in \mathbb{C}$ und berechnen $y' = \alpha \mathrm{e}^{\alpha x}$, $y'' = \alpha^2 \mathrm{e}^{\alpha x}$, $\ldots$, $y^{(d)} = \alpha^d \mathrm{e}^{\alpha x}$.

[1] Wir können ebenso die Koeffizienten $a_0, \ldots, a_{d-1}$ und die Anfangswerte $w_0, \ldots, w_{d-1}$ als komplex annehmen, doch bleibt der Fall mit reellen Koeffizienten und Anfangswerten weiterhin von speziellem Interesse.

Damit ist

$$y^{(d)} + a_{d-1}y^{(d-1)} + \ldots + a_0 y^{(0)} = (\alpha^d + a_{d-1}\alpha^{d-1} + \ldots + a_0)\mathrm{e}^{\alpha x}.$$

Daher ist $y = \mathrm{e}^{\alpha x}$ genau dann eine Lösung der Differentialgleichung (14.3), wenn $\alpha \in \mathbb{C}$ eine Nullstelle des sogenannten **charakteristischen Polynoms** der Differentialgleichung (14.3)

$$p(T) = T^d + a_{d-1}T^{d-1} + \ldots + a_0 \in \mathbb{R}[T]$$

ist.

Ist $\alpha$ eine reelle Nullstelle von $p(T)$, dann ist $y = e^{\alpha x}$ eine reellwertige Lösung. Bei reellen Koeffizienten und einer komplexen Nullstelle $\alpha = \beta + \gamma\mathrm{i} \in \mathbb{C}$ mit $\beta, \gamma \in \mathbb{R}$ ist $\overline{\alpha} = \beta - \gamma\mathrm{i}$ ebenso eine Nullstelle von $p(T)$. (Wieso?) Damit sind (nach Übung 14.4) alle Linearkombinationen von

$$\mathrm{e}^{\alpha x} = \mathrm{e}^{\beta x}\big(\cos(\gamma x) + \mathrm{i}\sin(\gamma x)\big), \quad \mathrm{e}^{\overline{\alpha} x} = \mathrm{e}^{\beta x}\big(\cos(\gamma x) - \mathrm{i}\sin(\gamma x)\big)$$

oder auch von

$$\frac{1}{2}(\mathrm{e}^{\alpha x} + e^{\overline{\alpha} x}) = \mathrm{e}^{\beta x}\cos(\gamma x), \quad \frac{1}{2\mathrm{i}}(\mathrm{e}^{\alpha x} - \mathrm{e}^{\overline{\alpha} x}) = \mathrm{e}^{\beta x}\sin(\gamma x)$$

Lösungen von (14.3). Man beachte, dass $\mathrm{e}^{\alpha x}$ und $\mathrm{e}^{\overline{\alpha} x}$ über $\mathbb{C}$ den gleichen Unterraum von Lösungen zu (14.3) aufspannen wie $\mathrm{e}^{\beta x}\cos(\gamma x)$ und $\mathrm{e}^{\beta x}\sin(\gamma x)$. (Wieso?)

Wir werden in der folgenden allgemeinen Diskussion diesen Wechsel von den Exponentialfunktionen zu Produkten von Exponentialfunktionen und trigonometrischen Funktionen nicht mehr durchführen, wollen dies aber meist in konkreten Beispielen machen um am Ende die allgemeine Lösung mittels reellwertige Funktionen zu beschreiben. Hat $p(T)$ genau $d$ verschiedene Nullstellen über $\mathbb{C}$, dann ergibt obige Diskussion $d$ linear unabhängige Lösungen auf ganz $\mathbb{R}$.

**Übung 14.5** (Lineare Unabhängigkeit)**.** *Seien $\alpha_1, \ldots, \alpha_d \in \mathbb{C}$ genau $d$ (das heisst, paarweise verschiedene) komplexe Zahlen. Zeigen Sie, dass die komplexwertigen Funktionen*

$$f_k : x \in \mathbb{R} \mapsto \mathrm{e}^{\alpha_k x} \in \mathbb{C}$$

*für $k = 1, \ldots, d$ linear unabhängig über $\mathbb{C}$ sind.*

Wir möchten nun den Fall, wo eine Nullstelle des charakteristischen Polynoms der Differentialgleichung (14.3) mehrfach erscheint, allgemeiner thematisieren. Dazu beginnen wir, Funktionen des Typs $x^n\mathrm{e}^{\alpha x}$, wie wir sie im letzten Fall von Beispiel 7.67 gesehen hatten, genauer zu untersuchen.

**Proposition 14.6** (Eine Basis eines Teilraumes)**.** *Die komplexwertigen Funktionen*

$$x \in \mathbb{R} \mapsto x^n\mathrm{e}^{\alpha x}$$

*für $n \in \mathbb{N}_0$ und $\alpha \in \mathbb{C}$ bilden eine linear unabhängige Teilmenge des komplexen Vektorraum $F_{\mathbb{C}}(\mathbb{R})$ der komplexwertigen Funktionen auf $\mathbb{R}$.*

Wir bezeichnen die lineare Hülle dieser Funktionen als

$$\mathrm{PE}_{\mathbb{C}}(\mathbb{R}) = \left\langle x^n \mathrm{e}^{\alpha x} \middle| n \in \mathbb{N}_0, \alpha \in \mathbb{C} \right\rangle,$$

da die Elemente von $\mathrm{PE}_{\mathbb{C}}(\mathbb{R})$ aus **P**olynomen und der **E**xponentialabbildung konstruiert werden. Für $\alpha \in \mathbb{C}$ definieren wir des Weiteren den Teilraum

$$\mathrm{PE}^{\alpha}_{\mathbb{C}}(\mathbb{R}) = \left\langle x^n \mathrm{e}^{\alpha x} \middle| n \in \mathbb{N}_0 \right\rangle = \mathbb{C}[x]\mathrm{e}^{\alpha x} \subseteq \mathrm{PE}_{\mathbb{C}}(\mathbb{R}).$$

Die Ableitung definiert eine lineare Abbildung

$$D : \mathrm{PE}_{\mathbb{C}}(\mathbb{R}) \to \mathrm{PE}_{\mathbb{C}}(\mathbb{R}) \quad f \mapsto D(f) = f',$$

die des Weiteren $D(\mathrm{PE}^{\alpha}_{\mathbb{C}}(\mathbb{R})) \subseteq \mathrm{PE}^{\alpha}_{\mathbb{C}}(\mathbb{R})$ für alle $\alpha \in \mathbb{C}$ erfüllt. In der Tat ist

$$D(\mathrm{e}^{\alpha x}) = \alpha \mathrm{e}^{\alpha x}, \quad D(x^n \mathrm{e}^{\alpha x}) = n x^{n-1} \mathrm{e}^{\alpha x} + \alpha x^n \mathrm{e}^{\alpha x} \tag{14.5}$$

für alle $n \in \mathbb{N}$ und $\alpha \in \mathbb{C}$. Für die Funktion $f : x \mapsto p(x)\mathrm{e}^{\alpha x}$ zu $p(t) \in \mathbb{C}[t]$ bezeichnen wir $\deg(p)$ als den Grad von $f$ und $\alpha$ als den Exponenten von $f$. Nach Obigem ist für $f \in \mathrm{PE}^{\alpha}_{\mathbb{C}}(\mathbb{R})$ mit Exponenten $\alpha \neq 0$ der Grad von $D(f)$ gleich dem Grad von $f$.

*Beweis.* Wir nehmen indirekt an, dass es paarweise verschiedene Parameter $\alpha_1, \ldots, \alpha_\ell \in \mathbb{C}$ und von Null verschiedene Polynome $q_1(T), \ldots, q_\ell(T) \in \mathbb{C}[T]$ gibt, so dass

$$\sum_{k=1}^{\ell} q_k(x)\mathrm{e}^{\alpha_k x} = 0$$

für alle $x \in \mathbb{R}$. Des Weiteren dürfen wir annehmen, dass $\ell \geq 1$ minimal ist mit dieser Eigenschaft (nach Satz 2.22).

Falls $\ell = 1$ ist, multiplizieren wir obige Gleichung mit $\mathrm{e}^{-\alpha_1 x}$ und erhalten die Gleichung $q_1(x) = 0$ für alle $x \in \mathbb{R}$, die $q_1(T) \neq 0$ widerspricht (siehe Abschnitt 3.2).

Sei also $\ell > 1$. Wir multiplizieren die Gleichung $\sum_{k=1}^{\ell} q_k(x)\mathrm{e}^{\alpha_k x} = 0$ mit $\mathrm{e}^{-\alpha_\ell x}$ und erhalten

$$\sum_{k=1}^{\ell-1} q_k(x)\mathrm{e}^{(\alpha_k - \alpha_\ell)x} + q_\ell(x) = 0.$$

für alle $x \in \mathbb{R}$. Wir wenden nun die Ableitung $D$ genau $\deg(q_\ell) + 1$-mal an. Der Term $q_\ell(x)$ verschwindet und wegen Gleichung (14.5) wird $q_k(x)\mathrm{e}^{(\alpha_k - \alpha_\ell)x}$ für $k \in \{1, \ldots, \ell - 1\}$ zu einem Ausdruck der Form $\tilde{q}_k(x)\mathrm{e}^{(\alpha_k - \alpha_\ell)x}$, wobei $\tilde{q}_k(T) \in \mathbb{C}[T]$ den gleichen Grad wie $q_k$ hat (wieso?).

Somit gilt

$$\sum_{k=1}^{\ell-1} \tilde{q}_k(x) \mathrm{e}^{(\alpha_k - \alpha_\ell)x} = 0$$

für alle $x \in \mathbb{R}$, was der minimalen Wahl von $\ell$ widerspricht. □

**Übung 14.7** (Matrixdarstellung des Differentiationsoperators)**.** *Finden Sie eine „Matrixdarstellung“ von $D$ eingeschränkt auf $\mathrm{PE}^{\alpha}_{\mathbb{C}}(\mathbb{R})$ für $\alpha \in \mathbb{C}$ bezüglich der Basis*

$$\mathrm{e}^{\alpha x}, x\mathrm{e}^{\alpha x}, \frac{x^2}{2!}\mathrm{e}^{\alpha x}, \frac{x^3}{3!}\mathrm{e}^{\alpha x}, \ldots.$$

*Um die Anführungszeichen zu vermeiden, können Sie $D$ auch auf den Unterraum*

$$\{e^{\alpha x} q(x) \mid p \in \mathbb{C}[x],\ \deg(q) \leq n\}$$

*von $\mathrm{PE}^{\alpha}_{\mathbb{C}}(\mathbb{R})$ für $n \in \mathbb{N}$ einschränken. Vergleichen Sie ihr Ergebnis mit dem Begriff der Jordan-Normalform aus der Linearen Algebra.*

Für ein Polynom $p(T) = \sum_{k=0}^{d} a_k T^k \in \mathbb{C}[T]$ definieren wir die lineare Abbildung

$$p(D) = \sum_{k=0}^{d} a_k D^k : f \in \mathrm{PE}_{\mathbb{C}}(\mathbb{R}) \mapsto \sum_{k=0}^{d} a_k f^{(k)} \in \mathrm{PE}_{\mathbb{C}}(\mathbb{R}),$$

wobei $D^0 = I$ die Identitätsabbildung auf $\mathrm{PE}_{\mathbb{C}}(\mathbb{R})$ darstellt (wir haben $f^{(0)}$ als $f$ definiert). Die homogene Differentialgleichung (siehe auch (14.3))

$$a_n y^{(d)} + \ldots + a_0 y^{(0)} = \sum_{k=0}^{d} a_k y^{(k)} = 0$$

lässt sich mit dem **Differentialoperator** $p(D)$ auch in der angenehm kurzen Form

$$p(D)y = 0$$

schreiben. Die Abbildung $p \in \mathbb{C}[T] \mapsto p(D)$ ist linear. Des Weiteren gilt für zwei Polynome $p(T) = \sum_{k=0}^{m} a_k T^k$, $q(T) = \sum_{\ell=0}^{n} b_\ell T^\ell$ und der Konvention $a_k = 0 = b_\ell$ für $k > m$ und $\ell > n$, dass

$$p(T)q(T) = \sum_{j=0}^{m+n} \left( \sum_{k=0}^{j} a_k b_{j-k} \right) T^j.$$

Somit gilt die multiplikative Eigenschaft

$$\begin{aligned} p(D)(q(D)f) &= p(D)\left(\sum_{\ell=0}^{n} b_\ell D^\ell f\right) \\ &= \sum_{k=0}^{m} a_k D^k \left(\sum_{\ell=0}^{n} b_\ell D^\ell f\right) = \sum_{k=0}^{m}\sum_{\ell=0}^{n} a_k b_\ell D^{k+\ell} f \\ &= \sum_{j=0}^{m+n}\left(\sum_{k=0}^{j} a_k b_{j-k}\right) D^j f = (pq)(D)(f) \end{aligned}$$

für alle $f \in \mathrm{PE}_{\mathbb{C}}(\mathbb{R})$ oder kurz $p(D)q(D) = (pq)(D)$. Diese Eigenschaft macht die **Auswertungsabbildung** $p \in \mathbb{C}[T] \mapsto p(D)$ zu einem sogenannten **Ringhomomorphismus**.

Mit der oben eingeführten Abstraktion können wir nun Lösungen zu linearen homogenen Differentialgleichungen mit konstanten Koeffizienten auf zufriedenstellende Art und Weise angeben.

**Proposition 14.8** (Lineare homogene Differentialgleichungen mit konstanten Koeffizienten)**.** *Sei $p(T) = \sum_{k=0}^{d} a_k T^k \in \mathbb{C}[T]$ ein Polynom mit $a_d = 1$, das mittels $p(T) = \prod_{k=1}^{\ell}(T-\alpha_k)^{n_k}$ in Linearfaktoren zerfällt, wobei wir annehmen, dass $\alpha_j \neq \alpha_k$ für $j \neq k \in \{1,\dots,\ell\}$. Dann hat die zugehörige homogene Differentialgleichung*

$$p(D)y = 0$$

*die folgenden $d$ linear unabhängigen Lösungen:*

$$\begin{array}{llll} \mathrm{e}^{\alpha_1 x}, & x\mathrm{e}^{\alpha_1 x}, & \dots\ , & x^{n_1-1}\mathrm{e}^{\alpha_1 x} \\ \vdots & & & \\ \mathrm{e}^{\alpha_\ell x}, & x\mathrm{e}^{\alpha_\ell x}, & \dots\ , & x^{n_\ell-1}\mathrm{e}^{\alpha_\ell x} \end{array}$$

Wir bemerken, dass nach dem Fundamentalsatz der Algebra (Theorem 9.81) jedes Polynom über $\mathbb{C}$ in Linearfaktoren zerfällt und somit die Annahme an $p$ in obiger Proposition keine Einschränkung ist. Des Weiteren werden wir später sehen, dass die Lösungsmenge der Gleichung $p(D)y = 0$ genau $d$-dimensional ist. Somit haben wir mit obigem Resultat eine Basis des Lösungsraumes der Differentialgleichung $p(D)y = 0$ erhalten.

*Beweis.* Nach Annahme hat das Polynom $p$ Grad $d = \sum_{k=0}^{\ell} n_k$, womit wir in der Tat $d$ Funktionen angegeben haben, die nach Proposition 14.6 linear unabhängig sind. Für ein festes $k \in \{1,\dots,\ell\}$ gilt weiter

$$\begin{aligned} (D-\alpha_k I)\mathrm{e}^{\alpha_k x} &= \alpha_k \mathrm{e}^{\alpha_k x} - \alpha_k \mathrm{e}^{\alpha_k x} = 0 \\ (D-\alpha_k I)x^m \mathrm{e}^{\alpha_k x} &= m x^{m-1}\mathrm{e}^{\alpha_k x} + \alpha_k x^m \mathrm{e}^{\alpha_k x} - \alpha_k x^m \mathrm{e}^{\alpha_k x} \\ &= m x^{m-1}\mathrm{e}^{\alpha_k x} \end{aligned}$$

für alle $m \in \mathbb{N}$. Für $m \in \{0, \dots, n_k - 1\}$ folgt nun aus dieser Rechnung und Induktion, dass $(D - \alpha_k I)^{n_k} x^m \mathrm{e}^{\alpha_k x} = 0$. Weil die Auswertungsabbildung ein Ringhomomorphismus ist, folgt

$$p(D)x^m \mathrm{e}^{\alpha_k x} = \left( \prod_{\substack{j=1,\dots,n \\ j \neq k}} (D - \alpha_j)^{n_j} \right) (D - \alpha_k I)^{n_k} x^m \mathrm{e}^{\alpha_k x} = 0,$$

was zu beweisen war. □

**Übung 14.9** (Reellwertige Lösungen)**.** *Beschreiben Sie einen reellen Vektorraum* $\mathrm{PESC}_{\mathbb{R}}(\mathbb{R}) \subseteq \mathrm{PE}_{\mathbb{C}}(\mathbb{R})$ *von reellwertigen Funktionen auf* $\mathbb{R}$*, so dass alle reellwertigen Lösungen in* $\mathrm{PE}_{\mathbb{C}}(\mathbb{R})$ *von homogenen linearen Differentialgleichungen mit konstanten Koeffizienten in* $\mathrm{PESC}_{\mathbb{R}}(\mathbb{R})$ *enthalten sind.*

VO: 23.05.18

### 14.2.2 Eine Lösung der inhomogenen Differentialgleichung

Wir wenden uns nun dem inhomogenen Problem zu. Für eine Störfunktion $g \in \mathrm{PE}_{\mathbb{C}}(\mathbb{R})$ gibt es ein einfaches Rezept, wie man eine partikuläre Lösung $y_{\text{part}}$ der inhomogenen Differentialgleichung

$$p(D)y = g,$$

wobei $p(T) \in \mathbb{C}[T]$, finden kann.

- Falls $g(x) = q(x)\mathrm{e}^{\alpha x}$ für ein Polynom $q(T)$ vom Grad $n$ und $\alpha \in \mathbb{C}$ mit $p(\alpha) \neq 0$, dann definiert man $y_{\text{part}} = Q(x)\mathrm{e}^{\alpha x}$, wobei $Q(T)$ ein Polynom vom Grad $n$ mit noch zu bestimmenden Koeffizienten ist. Nun berechnet man $p(D)y_{\text{part}}$, setzt dies gleich $g$ und verwendet diese Gleichung, um die Koeffizienten von $Q$ zu bestimmen.

- Falls $g(x) = q(x)\mathrm{e}^{\alpha x}$ für ein Polynom $q(T)$ vom Grad $n$ und $\alpha \in \mathbb{C}$ mit $p(\alpha) = 0$, dann wiederholt man obiges Verfahren, allerdings mit dem Ansatz $y_{\text{part}} = Q(x)x^{\ell}\mathrm{e}^{\alpha x}$, wobei $\ell$ die Vielfachheit der Nullstelle $\alpha$ von $p(T)$ angibt.

- Ein allgemeines $g \in \mathrm{PE}_{\mathbb{C}}(\mathbb{R})$ lässt sich als Linearkombination von Ausdrücken wie oben darstellen. Auf Grund der Linearität von $p(D)$ kann man also obiges Verfahren für Summanden der Form $q(x)\mathrm{e}^{\alpha x}$ in $g$ anwenden und dann die resultierenden Funktionen addieren.

- Falls $\alpha \in \mathbb{C} \setminus \mathbb{R}$ ist und wir an reellwertigen Lösungen einer Differentialgleichung mit reellen Koeffizienten interessiert sind, dann kann man $\mathrm{e}^{\alpha x}, \mathrm{e}^{\overline{\alpha} x}$ in obigen Diskussionen durch $\mathrm{e}^{\beta x}\cos(\gamma x)$, $\mathrm{e}^{\beta x}\sin(\gamma x)$ für $\beta, \gamma \in \mathbb{R}$ mit $\alpha = \beta + \gamma \mathrm{i}$ ersetzen.

**Übung 14.10** (Dies funktioniert immer)**.** *Zeigen Sie, dass das oben erklärte Verfahren tatsächlich zum Ziel (das heisst, zu einer partikulären Lösung) führt.*

### 14.2.3 Lösen des Anfangswertproblems

Die allgemeine Lösung einer inhomogenen Differentialgleichung ergibt sich nun wiederum als die Summe einer partikulären Lösung und der allgemeinen Lösung der zugehörigen homogenen Differentialgleichung (siehe Übung 14.4). Anschliessend kann man die Anfangsbedingungen eines Anfangswertproblems (siehe (14.4)) verwenden, um die Lösung näher zu bestimmen.

Wir wollen das Verfahren zur Lösung eines Anfangswertproblems für eine inhomogene lineare Differentialgleichung mit konstanten Koeffizienten nochmals an einem Beispiel erklären, welches etwas komplexer ist, damit wir eben einige der verschiedenen Möglichkeiten sehen können. Wir überspringen allerdings das Lösen der sich ergebenden linearen Gleichungssysteme.

**Beispiel 14.11.** *Wir betrachten das Polynom* $p(T) = (T^2+4)(T-1)T = T^4 - T^3 + 4T^2 - 4T$ *und dazu die inhomogene Differentialgleichung*

$$p(D)y = y^{(4)} - y^{(3)} + 4y^{(2)} - 4y^{(1)} = \sin(x) + \mathrm{e}^x + x.$$

- *Für die homogene Differentialgleichung bemerken wir, dass die Nullstellen von* $p(T)$ *durch* $2\mathrm{i}, -2\mathrm{i}, 1, 0$ *(jeweils mit Vielfachheit* $1$*) gegeben sind und somit die allgemeine Lösung der homogenen Differentialgleichung* $p(D)y = 0$ *durch*

  $$y_{\text{hom}} = A_1 \sin(2x) + A_2 \cos(2x) + A_3 \mathrm{e}^x + A_4$$

  *gegeben ist.*

- *Für die Störfunktion* $g_1(x) = \sin(x)$ *wählen wir gemäss Rezept den Ansatz* $y_{p,1} = B\sin(x) + C\cos(x)$*. Nun berechnet man die Ableitungen* $y'_{p,1}, \ldots, y^{(4)}_{p,1}$ *und setzt dies in* $p(D)y_{p,1} = \sin(x)$ *ein. Dann bestimmt man* $B, C$ *mit dem resultierenden Gleichungssystem und erhält* $B = -\frac{1}{6}$, $C = \frac{1}{6}$*. Somit ist* $y_{p,1} = \frac{1}{6}(-\sin(x) + \cos(x))$*.*

- *Für die Störfunktion* $g_2(x) = \mathrm{e}^x$ *wählen wir gemäss Rezept den Ansatz* $y_{p,2} = Exe^x$*. Wie zuvor berechnet man nun die Ableitungen* $y'_{p,2}, \ldots, y^{(4)}_{p,2}$ *und setzt dies in* $p(D)y_{p,2} = \mathrm{e}^x$ *ein. Man erhält* $E = \frac{1}{5}$ *und* $y_{p,2} = \frac{1}{5}xe^x$*.*

- *Für die Störfunktion* $g_3(x) = x$ *wählen wir den Ansatz* $y_{p,3} = Fx^2 + Gx$*. Man berechnet* $y'_{p,3} = 2Fx + G$, $y''_{p,3} = 2F$, $y^{(3)}_{p,3} = y^{(4)}_{p,3} = 0$ *und setzt dies in* $p(D)y_{p,3} = x$ *ein. Dies ergibt*

  $$4(2F) - 4(2Fx + G) = x = 8F - 8Fx - 4G$$

  *und somit* $F = -\frac{1}{8}$, $G = -\frac{1}{4}$*. Also ist* $y_{p,3} = -\frac{1}{8}x^2 - \frac{1}{4}x$*.*

*Für die allgemeine Lösung der inhomogenen Differentialgleichung* $p(D)y = \sin(x) + \mathrm{e}^x + x$ *erhalten wir daraus*

$$\begin{aligned} y_{\text{inhom}} &= y_{\text{hom}} + y_{\text{part}} = y_{\text{hom}} + y_{p,1} + y_{p,2} + y_{p,3} \\ &= A_1 \sin(2x) + A_2 \cos(2x) + A_3 \mathrm{e}^x + A_4 + \tfrac{1}{6}(-\sin(x) + \cos(x)) + \tfrac{1}{5} x\mathrm{e}^x - \tfrac{1}{8}x^2 - \tfrac{1}{4}x \end{aligned}$$

*Wenn jetzt noch Anfangswerte*

$$y(0) = \tfrac{1}{24},\ y'(0) = -\tfrac{1}{20},\ y''(0) = \tfrac{1}{12},\ y^{(3)}(0) = \tfrac{1}{10}$$

*gegeben sind, so ergibt sich nach Lösung eines Gleichungssystems in den Variablen* $A_1, A_2, A_3, A_4$ *die Lösung*

$$y = \tfrac{1}{12}\sin(2x) - \tfrac{1}{40}\cos(2x) - \tfrac{1}{10} + \tfrac{1}{6}(-\sin(x) + \cos(x)) + \tfrac{1}{5}x\mathrm{e}^x - \tfrac{1}{8}x^2 - \tfrac{1}{4}x.$$

**Übung 14.12** (Mehrfache Nullstellen)**.** *Bestimmen Sie die allgemeine Lösung der inhomogenen Differentialgleichung* $p(D)y = x + \mathrm{e}^x + \sin(2x)$ *wobei das Polynom* $p(T) \in \mathbb{R}[T]$ *durch* $p(T) = (T-1)^2(T^2+4)$ *gegeben ist. Gehen Sie dabei in folgenden Schritten vor:*

*(i) Bestimmen Sie zuerst die allgemeine Lösung von* $p(D)y_0 = 0$.

*(ii) Finden Sie ein Polynom* $q_1(T) \in \mathbb{R}[T]$, *so dass* $y_1 = q_1(x)$ *die inhomogene Differentialgleichung* $p(D)y_1 = x$ *löst (und multiplizieren Sie hierfür* $p(T)$ *aus).*

*(iii) Finden Sie als nächstes ein Polynom* $q_2(T) \in \mathbb{R}[T]$, *so dass* $y_2 = q_2(x)\mathrm{e}^x$ *eine Lösung von* $p(D)y_2 = \mathrm{e}^x$ *ist.*

*(iv) Finden Sie schlussendlich zwei Polynome* $q_3(T), q_4(T) \in \mathbb{R}[T]$, *so dass für die Funktion* $y_3 = q_3(x)\sin(2x) + q_4(x)\cos(2x)$ *die Gleichung* $p(D)y_3 = \sin(2x)$ *gilt.*

*(v) Nehmen Sie die Summe von all diesen Funktionen (welche auf Grund der Funktion* $y_0$ *vier unbestimmte Konstanten enthalten wird).*

**Übung 14.13** (Eindeutigkeit)**.** *Sei* $p \in \mathbb{C}[t]$ *ein Polynom von Grad* $d \in \mathbb{N}$ *und seien die Anfangswerte durch* $w_0, \ldots, w_{d-1} \in \mathbb{C}$ *gegeben. Sei des Weiteren* $g \in \mathrm{PE}_{\mathbb{C}}(\mathbb{R})$ *eine Störfunktion. Zeigen Sie, dass das Anfangswertproblem*

$$\begin{aligned} &p(D)y = g \\ &y(0) = w_0,\ y'(0) = w_1,\ \ldots\ ,\ y^{(d-1)}(0) = w_{d-1} \end{aligned}$$

*in* $\mathrm{PE}_{\mathbb{C}}(\mathbb{R})$ *eine eindeutig bestimmte Lösung besitzt.*

**Übung 14.14** (Glattheit)**.** *Sei* $p \in \mathbb{R}[t]$ *ein Polynom von Grad* $d \in \mathbb{N}$ *und sei* $g : \mathbb{R} \to \mathbb{R}$ *glatt. Zeigen Sie, dass jede Lösung* $y$ *der Differentialgleichung* $p(D)y = g$ *eine glatte Funktion ist.*

## 14.3 Differentialgleichungsysteme

Wir wollen hier den Begriff der Differentialgleichung auf vektorwertige Funktionen oder (äquivalenterweise) auf mehrere, verschiedene Funktionen mit Querbeziehungen verallgemeinern. Sei $I \subseteq \mathbb{R}$ ein Intervall, $U \subseteq I \times \mathbb{R}^d$ offen und $f : U \to \mathbb{R}^d$ eine stetige Funktion. Dann heisst die Gleichung

$$\dot{x}(t) = f(t, x(t))$$

für eine unbekannte (gesuchte) Funktion (Kurve) $x$ ein **$d$-dimensionales Differentialgleichungssystem erster Ordnung**. Dabei ist der Definitionsbereich einer Lösung $x$ möglicherweise nur ein Teilintervall von $I$ (insbesondere muss ja $(t, x(t)) \in U$ für alle $t$ im Definitionsbereich von $x$ gelten). Die Funktion $f$ wird oft auch die **rechte Seite** genannt. Des Weiteren nennen wir zu $(t_0, x_0) \in U$ die Gleichungen

$$\dot{x}(t) = f(t, x(t)), \quad x(t_0) = x_0$$

ein **Anfangswertproblem** zum **Anfangswert** $x_0$ bei $t_0 \in I$.

Die Differentialgleichung ist **autonom** (sich selbst überlassen), falls $f$ zeitunabhängig ist, das heisst, falls $f(t, x) = f(t', x)$ für alle $(t, x), (t', x) \in U$. Wie schon zuvor nennen wir die Funktion $f$ ein **Vektorfeld**, selbst wenn dieses zeitabhängig ist. Das Vektorfeld gibt bei $(t, x) \in U$ die Richtung und die Geschwindigkeit des gesuchten Weges $t \mapsto x(t)$ an. Eine praktische Applikation zur Darstellung von Vektorfeldern (und der Lösungen von Anfangswertproblemen dazu) können Sie hier finden.

### 14.3.1 Umwandlung einer Differentialgleichung höherer Ordnung

Es stellt sich die Frage, warum wir oben (und auch im Folgenden) nur Differentialgleichungssysteme erster Ordnung besprechen. Der Grund dafür ist die folgende Konstruktion.

Angenommen

$$y^{(d)}(t) = f\big(t, y(t), \dot{y}(t), \ldots, y^{(d-1)}(t)\big) \tag{14.6}$$

ist eine explizite Differentialgleichung der Ordnung $d$ und $f : U \to \mathbb{R}$ ist eine stetige Funktion auf einer offenen Teilmenge $U \subseteq \mathbb{R}^{d+1}$. Dann können wir diese Differentialgleichung in ein $d$-dimensionales Differentialgleichungssystem erster Ordnung umwandeln, indem wir

$$x(t) = (x_0(t), x_1(t), \ldots, x_{d-1}(t))^t$$

und

$$\begin{cases} \dot{x}_0(t) = x_1(t) \\ \dot{x}_1(t) = x_2(t) \\ \vdots \\ \dot{x}_{d-1}(t) = f(t, x_0(t), \ldots, x_{d-1}(t)) \end{cases} \tag{14.7}$$

setzen. In der Tat sind mit $y(t) = x_0(t)$ die Gleichungen (14.6) und (14.7) äquivalent. Die in (14.7) definierte (stetige) rechte Seite $\tilde{f} : U \to \mathbb{R}^d$ „erbt“ des Weiteren die meisten Eigenschaften von $f$. Ist beispielsweise $f$ differenzierbar, so ist auch $\tilde{f}$ differenzierbar und selbiges gilt unter anderem für Lipschitz-Stetigkeit (siehe die Annahmen von Theorem 14.23 wieter unten).

Selbiges Verfahren, leicht angepasst, liefert zu einer Differentialgleichungen höherer Ordnung in vektorwertigen Funktionen ebenfalls eine äquivalentes Differentialgleichungssystem erster Ordnung. Die gleiche Diskussion trifft auf die Übertragung von Anfangswerten zu.

Obige Äquivalenz erklärt, wieso wir in den Diskussionen weiter unten uns meist auf Differentialgleichungen (oder Differentialgleichungssysteme) erster Ordnung einschränken werden. Wir überlassen es dabei den Leserinnen und Lesern, die Resultate (siehe vor allem Theorem 14.23) auf Differentialgleichungen höherer Ordnung übertragen.

### 14.3.2 Lineare autonome Differentialgleichungssysteme

Wir betrachten hier den Fall eines linearen Differentialgleichungssystem mit konstanten Koeffizienten.

**Proposition 14.15** (Existenz und Eindeutigkeit von Lösungen)**.** *Sei $A \in \mathrm{Mat}_{d,d}(\mathbb{R})$, $x_0 \in \mathbb{R}^d$ und $t_0 \in \mathbb{R}$. Das Anfangswertproblem*

$$\dot{x}(t) = Ax(t), \quad x(t_0) = x_0$$

*hat die eindeutig bestimmte Lösung*

$$t \in \mathbb{R} \mapsto x(t) = \exp(A(t - t_0))x_0.$$

*Hierbei ist die (Matrix-) Exponentialabbildung* $\exp : \mathrm{Mat}_{d,d}(\mathbb{R}) \to \mathrm{GL}_d(\mathbb{R})$ *durch*

$$\exp(B) = \sum_{n=0}^{\infty} \frac{1}{n!} B^n$$

*für $B \in \mathrm{Mat}_{d,d}(\mathbb{R})$ definiert ist.*

Auf Grund der Konstruktion aus Abschnitt 14.3.1 gibt Proposition 14.15 auch eine Lösungsmethode für lineare Differentialgleichungen $d$-ter Ordnung mit konstanten Koeffizienten wie in (14.3) an (siehe auch den Beweis von Korollar 14.18 unten). Hier möchten wir andeuten, wie sich solche Differentialgleichungen als Differentialgleichungssystem erster Ordnung

schreiben lassen. Sei also die Differentialgleichung

$$y^{(d)} + a_{d-1}y^{(d-1)} + \ldots + a_0 y^{(0)} = 0$$

für $a_0, \ldots, a_{d-1} \in \mathbb{R}$ gegeben. Wandelt man dies in Differentialgleichungssystem erster Ordnung um, so ergibt sich (siehe (14.7)) die Gleichung $\dot{x}(t) = Ax(t)$ für

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ -a_0 & -a_1 & \cdots & -a_{d-2} & -a_{d-1} \end{pmatrix}.$$

Dies ist auch die Begleitmatrix des charakteristischen Polynoms der obigen Differentialgleichung – siehe folgende Übung.

**Übung 14.16** (Begleitmatrix eines Polynoms)**.** *Zeigen Sie, dass das charakteristische Polynom der obigen Matrix $A$ genau das Polynom $T^d + a_{d-1}T^{d-1} + \ldots + a_1 T + a_0$ ist.*

Die Aussage von Proposition 14.15 „konkurriert“ gewissermassen mit der Aussage von Proposition 14.8, weswegen man sich fragen könnte, ob die Exponentialabbildung für Matrizen auch für Differentialgleichungen wie Proposition 14.8 entsprechende Lösungen liefert. Dafür möchten wir auf Übung 14.21 verweisen.

*Beweis.* Wir zeigen zuerst, dass die Exponentialabbildung in der Tat durch eine absolut konvergente Reihe definiert wurde. Dabei ist auf Grund der Normäquivalenz auf endlichdimensionalen Vektorräumen die Wahl der Norm uns überlassen. Wir verwenden die Euklidsche Norm $\|\cdot\|$ auf $\mathbb{R}^d$ und die Matrixnorm (Operatornorm – siehe Definition 9.75)

$$\|B\|_{\mathrm{op}} = \max_{v \in \mathbb{R}^d,\ \|v\| \leq 1} \|Bv\|$$

für $B \in \mathrm{Mat}_{d,d}(\mathbb{R})$. Nach Lemma 9.76 ist die Matrixnorm submultiplikativ, das heisst, für $B_1, B_2 \in \mathrm{Mat}_{d,d}(\mathbb{R})$ gilt

$$\|B_1 B_2\|_{\mathrm{op}} \leq \|B_1\|_{\mathrm{op}} \|B_2\|_{\mathrm{op}}.$$

Daraus folgt für $B \in \mathrm{Mat}_{d,d}(\mathbb{R})$

$$\sum_{n=0}^{\infty} \|\tfrac{1}{n!} B^n\|_{\mathrm{op}} \leq \sum_{n=0}^{\infty} \tfrac{1}{n!} \|B\|_{\mathrm{op}}^n = \exp(\|B\|_{\mathrm{op}}) < \infty.$$

In anderen Worten ist die Reihe $\sum_{n=0}^{\infty} \frac{1}{n!} B^n$ also für jedes $B \in \mathrm{Mat}_{d,d}(\mathbb{R})$ (bezüglich $\|\cdot\|_{\mathrm{op}}$ oder äquivalenterweise in jedem Matrixeintrag) absolut konvergent und daher ist die Exponentialabbildung exp auf Matrizen wohldefiniert.

Für $A, B \in \mathrm{Mat}_{d,d}(\mathbb{R})$ mit $AB = BA$ (das heisst, $A$ und $B$ kommutieren) gilt der erweiterte Binomialsatz

$$(A+B)^n = \sum_{\ell=0}^{n} \binom{n}{\ell} A^\ell B^{n-\ell}$$

für alle $n \in \mathbb{N}_0$. (Wieso?) Daraus folgt

$$\begin{aligned}
\exp(A+B) &= \sum_{n=0}^{\infty} \frac{1}{n!}(A+B)^n \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{\ell=0}^{n} \binom{n}{\ell} A^\ell B^{n-\ell} \\
&= \sum_{n=0}^{\infty} \sum_{\ell=0}^{n} \frac{1}{\ell!} A^\ell \frac{1}{(n-\ell)!} B^{n-\ell} \\
&= \left( \sum_{\ell=0}^{\infty} \frac{1}{\ell!} A^\ell \right) \left( \sum_{m=0}^{\infty} \frac{1}{m!} B^m \right) \\
&= \exp(A)\exp(B),
\end{aligned}$$

wobei wir die Verallgemeinerung des Produktsatzes für absolut konvergente Reihen (Satz 6.36 und Korollar 6.37) für Matrizen verwendet haben. Diese Verallgemeinerung folgt aus demselben Beweis oder durch Anwendung des Satzes auf die einzelnen Komponenten

$$\begin{aligned}
(\exp(A))_{ij} &= \sum_{\ell=0}^{\infty} \frac{1}{\ell!} (A^\ell)_{ij} \\
(\exp(B))_{jk} &= \sum_{m=0}^{\infty} \frac{1}{m!} (B^m)_{jp}
\end{aligned}$$

für $i, j, k \in \{1, \ldots, d\}$. Insbesondere gilt

$$\exp(B)\exp(-B) = \exp(B-B) = \exp(0) = I_d$$

und daher ist $\exp(B) \in \mathrm{GL}_d(\mathbb{R})$ invertierbar für alle $B \in \mathrm{Mat}_{d,d}(\mathbb{R})$.

Wir betrachten nun die Abbildung $t \in \mathbb{R} \mapsto \exp(At) \in \mathrm{GL}_d(\mathbb{R}) \subseteq \mathrm{Mat}_{d,d}(\mathbb{R})$ und wollen die Ableitung bestimmen. Es gilt

$$\begin{aligned}
\frac{d}{dt}\big(\exp(At)\big)|_{t=0} &= \lim_{h\to 0} \frac{\exp(Ah) - \exp(0)}{h} \\
&= \lim_{h\to 0} \sum_{n=1}^{\infty} \tfrac{1}{n!} A^n h^{n-1} = A
\end{aligned}$$

und daher allgemeiner

$$\begin{aligned}\frac{d}{dt}\big(\exp(At)\big)|_{t=t_0} &= \lim_{h\to 0}\frac{\exp(A(t_0+h))-\exp(At_0)}{h}\\ &= \lim_{h\to 0}\frac{\exp(Ah)-\exp(0)}{h}\exp(At_0) = A\exp(At_0).\end{aligned}$$

Für die Funktion $t\in\mathbb{R}\mapsto x(t)=\exp(A(t-t_0))x_0$ gilt damit $x(t_0)=I_dx_0=x_0$ und

$$\dot{x}(t) = A\exp(A(t-t_0))x_0 = Ax(t).$$

VO: 25.05.18

Um die Eindeutigkeit zu zeigen, nehmen wir an, dass zu einem Intervall $I\subseteq\mathbb{R}$ mit $t_0\in I$ eine beliebige Lösung $y: I\to\mathbb{R}^d$ des Anfangswertproblems gegeben ist. Wir betrachten die Funktion $t\in I\mapsto\exp(-A(t-t_0))y(t)$ und berechnen die Ableitung mit der Produktregel. Es gilt

$$\begin{aligned}\frac{d}{dt}\big(\exp(-A(t-t_0))y(t)\big) &= -A\exp(-A(t-t_0))y(t)+\exp(-A(t-t_0))\dot{y}(t)\\ &= -A\exp(-A(t-t_0))y(t)+\exp(-A(t-t_0))Ay(t)\\ &= -A\exp(-A(t-t_0))y(t)+A\exp(-A(t-t_0))y(t)\\ &= 0,\end{aligned}$$

da $A$ und $\exp(-A(t-t_0))$ kommutieren. Also ist die Abbildung $t\in I\mapsto\exp(-A(t-t_0))y(t)$ konstant und damit gleich ihrem Wert $y(t_0)=x_0$ bei $t_0$. Dies beweist, dass wie gewünscht $y(t)=\exp(A(t-t_0))x_0$ für alle $t\in I$ gilt. □

**Übung 14.17** (Produktregel für matrixwertige Funktionen)**.** *Seien $\ell,m,n\in\mathbb{N}$. Seien $a<b$ reelle Zahlen und seien $t\in[a,b]\mapsto A(t)\in\mathrm{Mat}_{\ell,m}(\mathbb{R})$ und $t\in[a,b]\mapsto B(t)\in\mathrm{Mat}_{m,n}(\mathbb{R})$ differenzierbare Funktionen. Zeigen Sie, dass auch $t\in[a,b]\mapsto A(t)B(t)$ differenzierbar ist mit Ableitung*

$$\tfrac{d}{dt}\big(A(t)B(t)\big) = \big(\tfrac{d}{dt}A(t)\big)B(t)+A(t)\big(\tfrac{d}{dt}B(t)\big)$$

*für alle $t\in[a,b]$. Wo wurde dies in obigem Beweis verwendet?*

**Korollar 14.18** (Eindeutigkeit)**.** *Sei $d\in\mathbb{N}$. Ein Anfangswertproblem* (14.4) *zu einer linearen homogenen gewöhnlichen Differentialgleichung mit konstanten Koeffizienten wie in* (14.3) *hat eine eindeutig bestimmte Lösung. Falls das Anfangswertproblem zu der inhomogenen Version* (14.2) *eine Lösung mit einem Intervall $I$ als Definitionsbereich besitzt, so ist diese Lösung ebenso eindeutig bestimmt.*

*Beweis.* Wie vor dem Beweis von Proposition 14.15 erwähnt, erhält man mittels Umwandlung (siehe die Diskussion in Abschnitt 14.3.1) aus einer linearen homogenen gewöhnlichen Differentialgleichung mit konstanten Koeffizienten ein äquivalentes Differentialgleichungssystem $z'=Az$, wobei $A$ hier die Begleitmatrix des charakteristischen Polynoms ist. Um konsistent

mit der Notation in (14.3) zu sein, wollen wir hier in beiden Fällen $x$ als unabhängige Variable verwenden wollen. Bei vorgegebenen Anfangswerten für (14.3) wie in (14.4) erhalten wir ebenso einen Anfangswert $z_0$ für $z$ bei $x_0$.

Nun impliziert aber Proposition 14.15 die eindeutige Lösbarkeit des Anfangswertproblems

$$z'(x) = Az(x), \quad z(x_0) = z_0$$

(und gibt auch eine konkrete Lösung mittels der Exponentialabbildung für Matrizen an). Dadurch sehen wir, dass auch das Anfangswertproblem zu (14.3),(14.4) eine eindeutig bestimmte Lösung besitzt.

Falls nun das Anfangswertproblem (14.4) für die inhomogene Gleichung (14.2) zwei Lösungen $y_1, y_2$ hat, so ist $y_1 - y_2$ eine Lösung für die Differentialgleichung (14.3) mit trivialen (d.h. verschwindenden) Anfangswerten. In diesem Fall ist aber auf Grund von Obigem $y = 0$ die einzige Lösung und wir erhalten $y_1 = y_2$. □

Vergleicht man den Beweis des obigen Korollars mit der (auf einem Ansatz beruhenden) Diskussion in Abschnitt 14.2 so ergibt sich, dass die Exponentialabbildung für Matrizen (gemeinsam mit der Jordan-Normalform von Matrizen) die Struktur der in Proposition 14.8 gefunden Lösungen sehr gut erklärt (siehe auch Übung 14.21).

### 14.3.3 Beispiele linearer autonomer Differentialgleichungssysteme

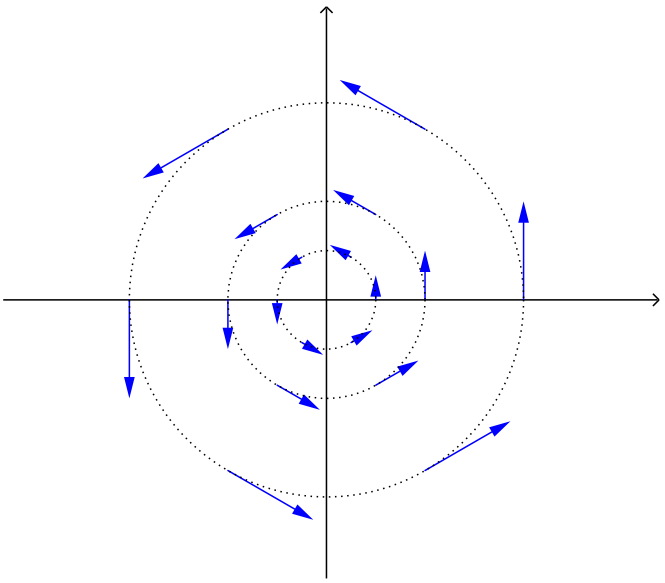
**Beispiel 14.19** (Rotation)**.** *Wir betrachten das gekoppelte Differentialgleichungssystem*

$$\dot{x}_1 = -x_2, \quad \dot{x}_2 = x_1,$$

*welches sich auch kurz als*

$$\dot{x} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} x$$

*schreiben lässt. Auf Grund der Zeitunabhängigkeit der rechten Seite ist dieses autonom. Das Vektorfeld lässt sich wie in folgendem Bild darstellen.*



*Weiter ist die eindeutig bestimmte Lösung $x$ zum Anfangswertproblem mit einem beliebigen Anfangswert $x_0 \in \mathbb{R}^2$ durch*

$$x(t) = \exp\left(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} t\right) x_0 = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix} x_0$$

*gegeben. (Wieso?). Startet man also beispielsweise bei $(1,0)^t$, so bewegt man sich entlang des Einheitskreises mit Geschwindigkeit 1, wobei man stets den oben illustrierten Pfeilen folgt, die den Kreis tangential berühren.*

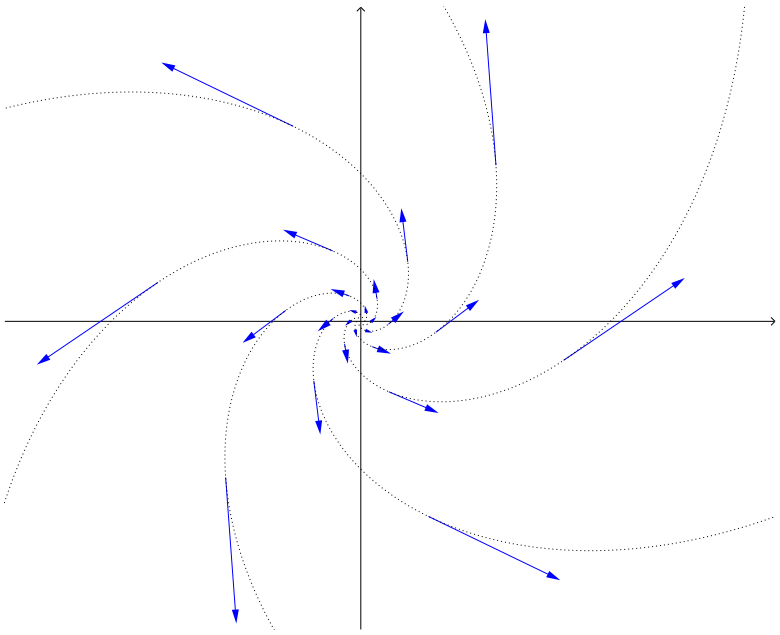
**Beispiel 14.20** (Spirale)**.** *Wir betrachten das gekoppelte, autonome Differentialgleichungssystem*

$$\dot{x}_1 = x_1 - x_2, \quad \dot{x}_2 = x_1 + x_2,$$

*welches sich auch kurz als*

$$\dot{x} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} x$$

*schreiben lässt. Wir stellen das Vektorfeld wieder in einer Grafik dar.*



*Die eindeutig bestimmte Lösung $x$ zum Anfangswertproblem mit einem beliebigen Anfangswert $x_0 \in \mathbb{R}^2$ ist durch*

$$\begin{aligned} x(t) &= \exp\left(\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} t\right) x_0 = \exp\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} t\right) \exp\left(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} t\right) x_0 \\ &= \begin{pmatrix} \mathrm{e}^t \cos(t) & -\mathrm{e}^t \sin(t) \\ \mathrm{e}^t \sin(t) & \mathrm{e}^t \cos(t) \end{pmatrix} x_0 \end{aligned}$$

*gegeben. (Wieso?)*

**Übung 14.21** (Mehr zur Exponentialabbildung)**.** *In dieser Übung möchten wir die Exponentialabbildung und Proposition 14.15 weiter erkundschaften.*

*(i) Berechnen Sie zu $a_1, \ldots, a_d \in \mathbb{R}$ die Diagonalmatrix*

$$\exp \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_d \end{pmatrix}.$$

*Interpretieren Sie die Differentialgleichung in dem Fall*

$$\dot{x} = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix} x$$

*und das zugehörige Vektorfeld.*

*(ii) Berechnen Sie* $\exp\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}$ *und interpretieren Sie wie zuvor die Differentialgleichung*

$$\dot{x} = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix} x$$

*und das zugehörige Vektorfeld.*

*(iii) Sei* $\mathcal{J}$ *die* $d \times d$*-Matrix, die über der Diagonalen Einsen und sonst nur Nullen stehen hat. Zeigen Sie, dass*

$$\exp(t\mathcal{J}) = I_d + t\mathcal{J} + \tfrac{t^2}{2}\mathcal{J}^2 + \ldots + \tfrac{t^{d-1}}{(d-1)!}\mathcal{J}^{d-1}$$

*und kommentieren Sie die Differentialgleichung* $\dot{x} = \mathcal{J}x$.

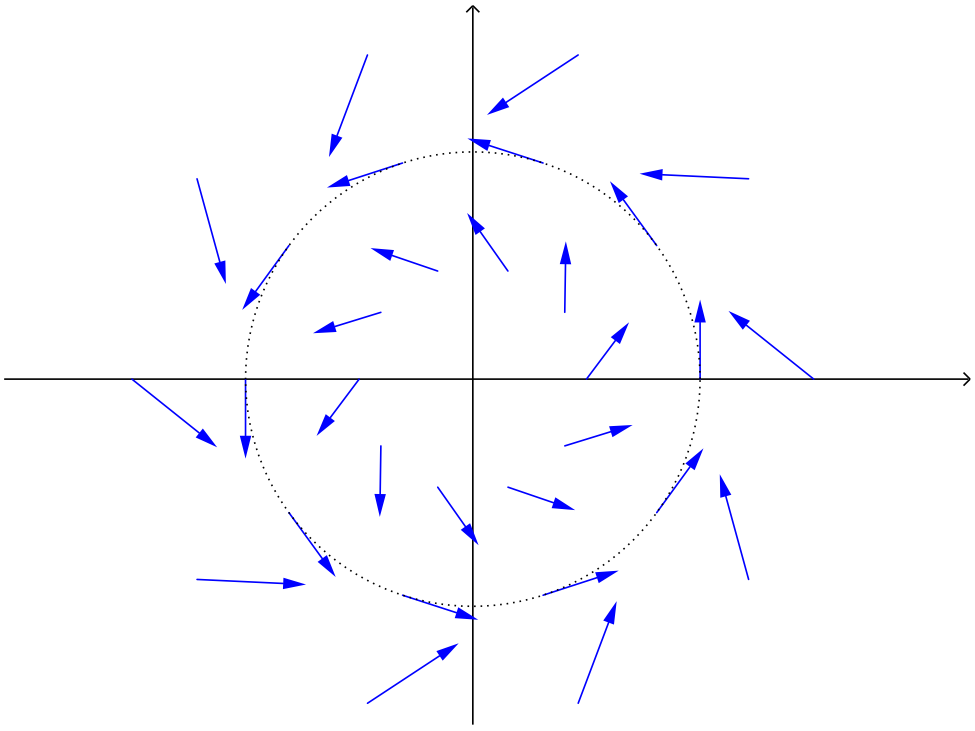
**Übung 14.22** (Allgemeiner Fall)**.** *Sei* $A \in \mathrm{Mat}_{d,d}(\mathbb{R})$. *Verwenden Sie die Jordan-Normalform von Matrizen (über* $\mathbb{C}$*) um* $\exp(At)$ *für alle* $t \in \mathbb{R}$ *zu berechnen. Liefern Sie damit eine explizite Beschreibung der Lösungen in Proposition 14.15 und beweisen Sie Proposition 14.8 erneut.*

### 14.3.4 Ein Beispiel eines nicht-linearen Systems

Wir betrachten das nicht-lineare autonome Differentialgleichungssystem

$$\begin{aligned}\dot{x} &= x - y - (x^2 + y^2)x\\ \dot{y} &= x + y - (x^2 + y^2)y\end{aligned} \tag{14.8}$$

Das involvierte Vektorfeld lässt sich dann wie in folgendem Bild darstellen.



Wir möchten zuerst zwei spezielle Fälle betrachten:

(a) Wir suchen die konstanten Lösungen der Differentialgleichung (14.8). Seien also $x, y$ definiert (und differenzierbar) auf einem Intervall mit

$$\begin{aligned} 0 &= x - y - (x^2 + y^2)x \\ 0 &= x + y - (x^2 + y^2)y \end{aligned} .$$

Insbesondere ist dann

$$\begin{aligned} 0 &= \big(x - y - (x^2 + y^2)x\big) \cdot y - \big(x + y - (x^2 + y^2)y\big) \cdot x \\ &= -x^2 - y^2 \end{aligned}$$

und somit $x = y = 0$.

(b) Wir betrachten nun das Anfangwertproblem zu (14.8) mit einem Anfangswert $(x_0, y_0)^t$ auf dem Einheitskreis. Das obige Bild lässt dann vermuten, dass die Lösung aus Beispiel 14.19 auch hier eine Lösung darstellt;

$$(x(t), y(t))^t = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix} (x_0, y_0)^t.$$

In der Tat bewegt sich diese Kurve auf dem Einheitskreis, womit $x^2 + y^2 = 1$ stets erfüllt ist und somit die zu lösende Gleichung (14.8) zur Gleichung aus Beispiel 14.19 äquivalent ist (welche ja von $(x(t), y(t))^t$ gelöst wird).

Um nun die Differentialgleichung (14.8) allgemeiner behandeln zu können, betrachten wir Polarkoordinaten und hoffen, dass in diesen (14.8) gelöst werden kann. Das heisst, wir schreiben $r(t)^2 = x(t)^2 + y(t)^2$ und

$$x(t) = r(t)\cos(\varphi(t)), \quad y(t) = r(t)\sin(\varphi(t))$$

für eine geeignete Funktion $t \mapsto \varphi(t) \in \mathbb{R}$.

Wir betrachten zuerst den Radius $r$ und finden eine Differentialgleichung für diesen. Man berechnet

$$\begin{aligned} (r^2)^{\cdot} &= (x^2 + y^2)^{\cdot} = 2x\dot{x} + 2y\dot{y} \\ &= 2\big(x^2 - xy - (x^2 + y^2)x^2\big) + 2\big(xy + y^2 - (x^2 + y^2)y^2\big) \\ &= 2(r^2 - r^4) \end{aligned}$$

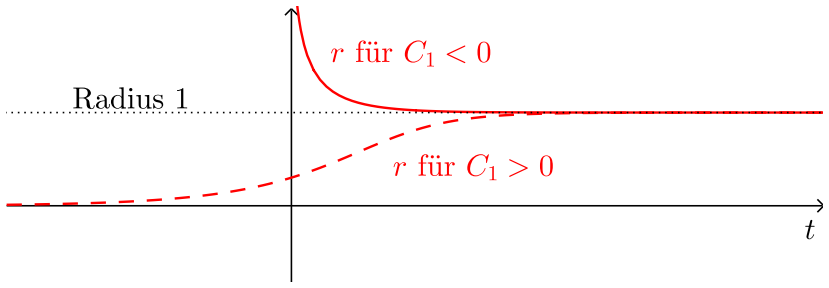
und somit, da $(r^2)^{\cdot} = 2r\dot{r}$, erhält man

$$\dot{r} = r - r^3.$$

Eine Lösung davon lässt sich mit der Trennung der Variablen aus Abschnitt 14.1.1 ermitteln; man erhält

$$r(t) = \frac{\mathrm{e}^t}{\sqrt{C_1 + \mathrm{e}^{2t}}}$$

für eine passend zum Anfangswert gewählte Konstante $C_1$. Der Radius verhält sich also wie in folgendem Bild.




Insbesondere kommt die Lösung dem Einheitskreis also immer näher, wie man es schon aus der obigen Illustration des Vektorfelds erahnen könnte.

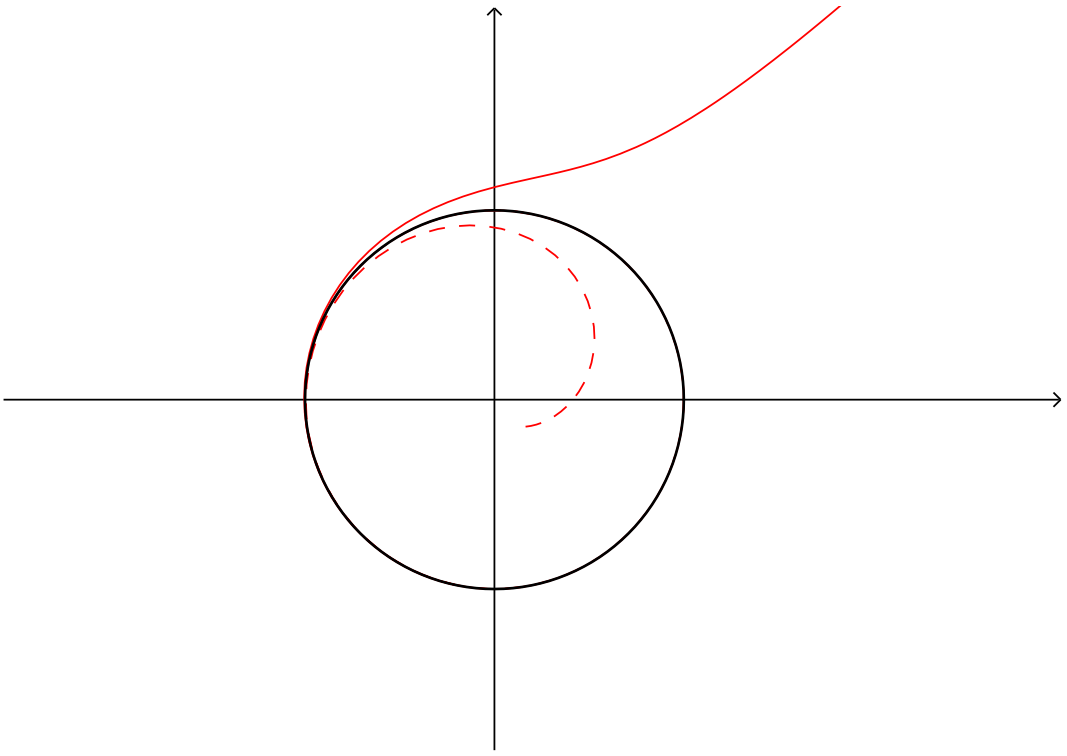
Wir versuchen nun ein ähnliches Vorgehen, um eine Formel für den Winkel $\varphi$ zu finden. Intuitiv gesehen könnte man ein ähnliches Verfahren wie oben für $x^2 + y^2$ auf beispielsweise $\frac{x}{y}$ anwenden, was zwar auf das richtige Resultat führt, aber nur dann zulässig ist, wenn $y(t)$ für kein $t$ verschwindet. Wir wollen dies aber nicht voraussetzen. Stattdessen berechnen wir mit der Kettenregel

$$\begin{aligned}\dot{x} &= \dot{r}\cos(\varphi) - r\sin(\varphi)\dot{\varphi} = x - y - r^2 x\\ \dot{y} &= \dot{r}\sin(\varphi) + r\cos(\varphi)\dot{\varphi} = x + y - r^2 y.\end{aligned}$$

Es gilt jedoch gemeinsam mit $\dot{r} = (1 - r^2)r$ ebenso

$$\begin{aligned}\dot{x} &= (1 - r^2)\underbrace{r\cos(\varphi)}_{=x} - \underbrace{r\sin(\varphi)}_{=y}\dot{\varphi}\\ &= x - y\dot{\varphi} - r^2 x\end{aligned}$$

und damit $-y\dot{\varphi} = -y$. Analog zeigt man $x\dot{\varphi} = x$. Unter dem Strich ist also entweder $x = y = 0$ (was der konstanten Lösung entspricht) oder $\dot{\varphi} = 1$ und somit $\varphi(t) = t + C_2$ für eine passend zum Anfangswert gewählte Konstante $C_2$. Wie man also aus der Darstellung des Vektorfelds hätte vermuten können, bilden die Lösungen Spiralen, die sich dem Einheitskreis annähern.



Wir bemerken noch, dass die Lösung $x(t)$ des Anfangwertproblems mit $\|x_0\| > 1$ (das heisst, $C_1 < 0$) nie auf ganz $\mathbb{R}$, sondern nur immer auf einem Intervall $(a, \infty)$ definiert sind und $\lim_{t \searrow a} \|x(t)\| = +\infty$ erfüllen. Dies hat damit zu tun, dass die rechte Seite der Differentialgleichung (14.8) für $\|(x, y)^t\| \to \infty$ sehr schnell anwächst.

## 14.4 Existenz- und Eindeutigkeitssatz von Picard-Lindelöf

Wie wir bereits in Abschnitt 14.1 gesehen haben, muss die Lösung eines Anfangswertproblems (selbst für eine relativ einfache Differentialgleichung wie $y' = 3y^{\frac{2}{3}}$) nicht immer eindeutig bestimmt sein. Wenn man hingegen zu einem Differentialgleichungssystem

$$\dot{x}(t) = f(t, x(t))$$

eine etwas stärkere Annahme an das Vektorfeld $f$ verlangt, so erhält man sowohl Existenz als auch Eindeutigkeit der Lösung nach folgendem fundamentalen Satz (siehe auch [Lin94]).

**Theorem 14.23** (Picard-Lindelöf). *Es sei $d \geq 1$, $U \subseteq \mathbb{R} \times \mathbb{R}^d$ offen und $f : U \to \mathbb{R}^d$ stetig. Angenommen $f$ ist „lokal Lipschitz-stetig im Ort", das heisst, für alle $(t_0, x_0) \in U$ existiert ein $\varepsilon > 0$ mit $B_\varepsilon(t_0, x_0) \subseteq U$ und $M > 0$, so dass für alle $(t, x_1), (t, x_2) \in B_\varepsilon(t_0, x_0)$ die Abschätzung*

$$\|f(t, x_1) - f(t, x_2)\| \leq M\|x_1 - x_2\|$$

*gilt. Sei weiters $t_0 \in \mathbb{R}$ und ein Anfangswert $x_0 \in \mathbb{R}^d$ mit $(t_0, x_0) \in U$ gegeben.*

Existenz: *Dann existiert ein Zeitintervall $I = I_{\max} = (a, b) \subseteq \mathbb{R}$ und eine differenzierbare Funktion $x(\cdot) : I \to \mathbb{R}^d$ mit*

- *$(t, x(t)) \in U$ für alle $t \in I$,*
- *$\dot{x}(t) = f(t, x(t))$ für alle $t \in I$,*
- *$t_0 \in I$ und $x(t_0) = x_0$.*

Eindeutigkeit: *Für jede weitere Lösung $y(\cdot) : J \to \mathbb{R}^d$ desselben Anfangswertproblems definiert auf einem offenen Intervall $J$ mit $t_0 \in J$ gilt $J \subseteq I$ und $x(\cdot)|_J = y(\cdot)$.*

Maximalität: *Die Grenzwerte $\lim_{t \searrow a} (t, x(t))$ und $\lim_{t \nearrow b} (t, x(t))$ existieren in $U$ nicht.*

Zusammenfassend hat jedes Anfangswertproblem für ein (im Ort lokal) Lipschitz-stetiges Vektorfeld eine eindeutige Lösung. Des Weiteren lässt sich diese entweder beliebig lange fortsetzen ($b = \infty$) oder solange, bis sie „explodiert", das heisst, in endlicher Zeit den Definitionsbereich $U$ der Differentialgleichung verlässt (siehe auch Übung 14.28 für eine schärfere Formulierung dieser Behauptung). Wir werden (wie bereits oben) hier häufig $x(\cdot)$ für eine Funktion auf einem Intervall schreiben (welche möglicherweise unser Anfangswertproblem löst), um etwaige Verwirrungen mit dem Anfangspunkt $x_0 = x(t_0) \in \mathbb{R}^d$ zu vermeiden.

Mit Hilfe der mehrdimensionalen Differentialrechnung (Korollar 10.17) sehen wir, dass stetig differenzierbare Funktionen lokal Lipschitz-stetig. Insbesondere lässt sich das obige Theorem auf eine grosse Klasse von Vektorfeldern $f$ anwenden.

**Übung 14.24** (Lokale Lipschitz-Stetigkeit). *Sei $U \subseteq \mathbb{R} \times \mathbb{R}^d$ offen und $f : U \to \mathbb{R}^d$ stetig, so dass die partiellen Ableitungen $\partial_k f$ für $k \in \{2, \ldots, d+1\}$ auf $U$ existieren und stetig sind. Zeigen Sie, dass $f$ dann auch „lokal im Ort Lipschitz-stetig" ist und somit die Voraussetzungen von Theorem 14.23 erfüllt.*

Für den Beweis des Theorems werden wir folgende lokale Version zuerst beweisen.

**Proposition 14.25** (Lokale Existenz und Eindeutigkeit)**.** *Seien $r_1, r_2 > 0$, $t_0 \in \mathbb{R}$, $x_0 \in \mathbb{R}^d$ und $f : (t_0 - r_1, t_0 + r_1) \times B_{r_2}(x_0) \to \mathbb{R}^d$ eine beschränkte stetige Funktion. Angenommen es gibt eine Lipschitz-Konstante $M > 0$ mit*

$$\|f(t, x_1) - f(t, x_2)\| \leq M\|x_1 - x_2\|$$

*für alle $t \in (t_0 - r_1, t_0 + r_1)$ und $x_1, x_2 \in B_{r_2}(x_0)$. Dann existiert ein $\delta \in (0, r_1)$ und eine eindeutig bestimmte Lösung $x(\cdot) : [t_0 - \delta, t_0 + \delta] \to \mathbb{R}^d$ des Anfangswertproblems*

$$\begin{cases} x(t_0) = x_0, \\ \dot{x}(t) = f(t, x(t)) \end{cases} \tag{14.9}$$

*für $t \in [t_0 - \delta, t_0 + \delta]$.*

*Genauer formuliert, es existiert ein $\delta_0$ (welches von $r_1, r_2$, von der oberen Schranke von $\|f(t, x)\|$ für $(t, x) \in (t_0 - r_1, t_0 + r_1) \times B_{r_2}(x_0)$ und von der Lipschitz-Konstante $M$ abhängt) so dass obige eindeutige Existenz auf allen Intervallen der Form $[t_0 - \delta, t_0 + \delta]$ für $\delta \in (0, \delta_0]$ gilt.*

Wie wir sehen werden, besteht der Beweis der obigen Proposition aus dem Fundamentalsatz der Differential- und Integralrechnung (siehe Abschnitt 8.1) und dem Banachschen Fixpunktsatz (siehe Satz 9.54), welcher ja ebenfalls eine eindeutige Existenz behauptet. Der Beweis des Theorems von Picard-Lindelöf wird die Aussage der obigen Proposition jeweils lokal anwenden und eine Lösung dann geeignet „zusammenstückeln“. Bei dem ersten Lesen der Beweise kann es helfen sich den Fall $d = 1$ vorzustellen, an der Beweisidee geht dabei nichts verloren.

*Beweis.* Nach dem Fundamentalsatz der Differential- und Integralrechnung (Korollar 8.3) (angewandt auf jede Komponente von $x(\cdot)$) ist das Anfangwertproblem in (14.9) zur Integralgleichung

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))\,\mathrm{d}s \tag{14.10}$$

für alle $t \in [t_0 - \delta, t_0 + \delta]$ äquivalent. Wir möchten diese Gleichung nun als Fixpunktgleichung interpretieren und wollen dazu also einen geeigneten vollständigen metrischen Raum $V$ und eine geeignete Lipschitz-Kontraktion $T : V \to V$ definieren.

Der vollständige Raum $V_\delta$: Für $\delta \in (0, r_1)$ definieren wir den Raum

$$V_\delta = \Big\{ y(\cdot) : [t_0 - \delta, t_0 + \delta] \to \overline{B_{r_2/2}(x_0)} \mid y(\cdot) \text{ ist stetig} \Big\} \subseteq C([t_0 - \delta, t_0 + \delta], \mathbb{R}^d),$$

wobei wir mit $\overline{B_{r_2/2}(x_0)} = \big\{ z \in \mathbb{R}^d \mid \|x_0 - z\| \leq \frac{r_2}{2} \big\}$ den abgeschlossenen Ball um $x_0$ von Radius $\frac{r_2}{2}$ bezeichnen. Wir statten $V_\delta$ dabei mit der von der Maximumsnorm $\|\cdot\|_\infty$ induzierten

Metrik d aus, womit also

$$\mathrm{d}(y(\cdot), z(\cdot)) = \|y(\cdot) - z(\cdot)\|_\infty = \max_{t\in[t_0-\delta,t_0+\delta]} \|y(t) - z(t)\|.$$

für $y(\cdot), z(\cdot) \in V_\delta$. Nach Proposition 9.86 ist $C([t_0 - \delta, t_0 + \delta], \mathbb{R}^d)$ vollständig. Wir zeigen nun, dass $V_\delta$ ebenfalls vollständig ist (oder äquivalent dazu eine abgeschlossene Teilmenge ist). Sei $(y_n(\cdot))_n$ eine Cauchy-Folge in $V_\delta$. Da $C([t_0 - \delta, t_0 + \delta], \mathbb{R}^d)$ vollständig ist, existiert der Grenzwert $y(\cdot) = \lim_{n\to\infty} y_n(\cdot) \in C([t_0 - \delta, t_0 + \delta], \mathbb{R}^d)$. Für $t \in [t_0 - \delta, t_0 + \delta]$ konvergiert dann wegen $\|y_n(t) - y(t)\| \leq \|y_n(\cdot) - y(\cdot)\|_\infty$ die Folge $(y_n(t))_n$ gegen $y(t)$, womit

$$\|y(t) - x_0\| = \lim_{n\to\infty} \|y_n(t) - x_0\| \leq \tfrac{r_2}{2}.$$

Daher gilt für $t \in [t_0 - \delta, t_0 + \delta]$ also $y(t) \in \overline{B_{r_2/2}(x_0)}$ und somit $y(\cdot) \in V_\delta$. Da $(y_n)_n$ eine beliebige Cauchy-Folge war, ist $V_\delta$ vollständig ist. VO: 28.05.18

Die Picard-Abbildung $T$ auf $V_\delta$: Zu einer Funktion $y(\cdot) \in V_\delta$ definieren wir eine Funktion $Ty(\cdot)$ durch

$$(Ty)(t) = x_0 + \int_{t_0}^{t} f(s, y(s))\,\mathrm{d}s$$

für $t \in [t_0 - \delta, t_0 + \delta]$. (Die Operation $y(\cdot) \mapsto Ty(\cdot)$ nennt sich auch Picard-Abbildung und bildet eine stetige Funktion auf eine neue stetige Funktion ab.) Da $y(s) \in \overline{B_{r_2/2}(x_0)} \subseteq B_{r_2}(x_0)$ und die Funktionen $y$ und $f$ stetig sind, ist auch $s \in [t_0 - \delta, t_0 + \delta] \mapsto f(s, y(s))$ definiert und stetig. Daher ist $Ty(\cdot)$ nach Korollar 8.3 differenzierbar und insbesondere stetig. Des Weiteren gilt

$$\|Ty(t) - x_0\| = \left\| \int_{t_0}^{t} f(s, y(s))\,\mathrm{d}s \right\| \leq \left| \int_{t_0}^{t} \|f(s, y(s))\|\,\mathrm{d}s \right|$$

für $t \in [t_0 - \delta, t_0 + \delta]$, wobei wir die Dreiecksungleichung für das Riemann-Integral in (5.16) verwendet haben und der Absolutbetrag rechts für den Fall $t_0 - \delta \leq t < t_0$ notwendig ist. Nach Annahme ist $f$ beschränkt, also gibt es eine Konstante $C > 0$, so dass $\|f(s, y(s))\| \leq C$ für alle $s \in [t_0 - \delta, t_0 + \delta]$. Daher gilt

$$\|Ty(t) - x_0\| \leq C|t - t_0| \leq C\delta$$

für alle $t \in [t_0 - \delta, t_0 + \delta]$. Diese Abschätzung hängt wohlgemerkt nicht von $y(\cdot)$ ab. Wir sehen hier bereits eine Einschränkung an $\delta > 0$: Falls $\delta \in \left(0, \frac{r_2}{2C}\right)$, dann gilt $Ty(t) \in \overline{B_{C\delta}(x_0)} \subseteq \overline{B_{r_2/2}(x_0)}$ und damit $Ty(\cdot) \in V_\delta$ wie gewünscht.

Die Kontraktionseigenschaft: Seien nun $y_1, y_2 \in V_\delta$. Dann gilt

$$\begin{aligned}\|Ty_1(t) - Ty_2(t)\| &= \left\| \int_{t_0}^{t} f(s, y_1(s)) - f(s, y_2(s))\,\mathrm{d}s \right\| \leq \left| \int_{t_0}^{t} \|f(s, y_1(s)) - f(s, y_2(s))\|\,\mathrm{d}s \right| \\ &\leq \left| \int_{t_0}^{t} M\|y_1(s) - y_2(s)\|\,\mathrm{d}s \right| \leq M\delta\|y_1(\cdot) - y_2(\cdot)\|_\infty\end{aligned}$$

für alle $t \in [t_0 - \delta, t_0 + \delta]$. Insbesondere gilt also

$$\|Ty_1(\cdot) - Ty_2(\cdot)\|_\infty \leq M\delta\|y_1(\cdot) - y_2(\cdot)\|_\infty.$$

Falls nun $\delta$ klein genug ist, so dass $\lambda = M\delta < 1$ gilt, dann definiert $T : V_\delta \to V_\delta$ eine Kontraktion auf einem vollständigen metrischen Raum.

Konklusion: Nach dem Banachschen Fixpunktsatz (Satz 9.54) gibt es also ein eindeutig bestimmtes $x(\cdot) \in V_\delta$ mit $Tx(\cdot) = x(\cdot)$. Dies ist aber genau Gleichung (14.10), die (wie oben erklärt) auf Grund des Fundamentalsatzes der Integral- und Differentialrechnung (Korollar 8.3) zum Anfangswertproblem in Gleichung (14.9) äquivalent ist.

Für die letzte Behauptung in der Proposition definieren wir zum Beispiel

$$\delta_0 = \min\left(\frac{r_1}{2}, \frac{r_2}{2C}, \frac{1}{2M}\right),$$

dann erfüllt jedes $\delta \in (0, \delta_0]$ alle Bedingungen des obigen Beweises. □

*Beweis von Theorem 14.23.* Wir beweisen zuerst die *Eindeutigkeit*. Seien also $x_1(\cdot) : I_1 \to \mathbb{R}$ und $x_2(\cdot) : I_2 \to \mathbb{R}$ zwei Lösungen des Anfangswertproblems

$$\dot{x}(t) = f(t, x(t)), \quad x(t_0) = x_0 \tag{14.11}$$

auf offenen Intervallen $I_1, I_2 \subseteq \mathbb{R}$ mit $t_0 \in I := I_1 \cap I_2 = (\alpha, \beta)$. Wir behaupten, dass $x_1(t) = x_2(t)$ für alle $t \in I$ gilt. Wir definieren

$$s = \sup\left\{t \in I \mid x_1(\cdot)|_{[t_0,t]} = x_2(\cdot)|_{[t_0,t]}\right\}.$$

Falls $s < \beta$ und damit $s \in I$ ist, dann folgt $x_1(s) = x_2(s)$ nach Stetigkeit von $x_1(\cdot)$ und $x_2(\cdot)$ und nach Definition von $s$. Damit lösen aber sowohl $x_1(\cdot)$ als auch $x_2(\cdot)$ das Anfangwertproblem $\dot{x}(t) = f(t, x(t))$, $x(s) = x_1(s)$. Somit impliziert der lokale Existenz- und Eindeutigkeitssatz (Proposition 14.25), dass $x_1|_{(s-\delta,s+\delta)} = x_2|_{(s-\delta,s+\delta)}$ für ein $\delta > 0$. Da dies aber der Definition von $s$ widerspricht, gilt $s = \beta$ und daher

$$\begin{aligned}x_1(\cdot)|_{[t_0,\beta)} &= x_2(\cdot)|_{[t_0,\beta)} \\ x_1(\cdot)|_{(\alpha,t_0]} &= x_2(\cdot)|_{(\alpha,t_0]}, \\ x_1(\cdot)|_{(\alpha,\beta)} &= x_2(\cdot)|_{(\alpha,\beta)}.\end{aligned}$$

In der Tat folgt die zweite Gleichung völlig analog und gemeinsam ergibt sich die dritte Gleichung.

Wir wollen nun die *Existenz* einer maximalen Lösung wie im Theorem zeigen. Dafür definieren wir

$$\begin{aligned} a &= \inf\{\alpha \in (-\infty, t_0) \mid \text{es existiert eine Lösung auf } (\alpha,\beta) \text{ mit } \beta > t_0\}, \\ b &= \sup\{\beta \in (t_0, \infty) \mid \text{es existiert eine Lösung auf } (\alpha,\beta) \text{ mit } \alpha < t_0\}. \end{aligned}$$

Zu $t \in (a,b)$ setzen wir $x_{\max}(t) = x(t)$, wobei $x$ eine Lösung auf einem Teilintervall $(\alpha,\beta)$ mit $a < \alpha$ und $\beta < b$ ist mit $t \in (\alpha,\beta)$ (welche nach Definition von $a, b$ existiert). Auf Grund der oben bewiesenen Eindeutigkeit ist $x_{\max}(t)$ nicht von der Wahl der Lösung $x(\cdot) : (\alpha,\beta) \to \mathbb{R}^d$ abhängig und somit ist $x_{\max} : (a,b) \to \mathbb{R}^d$ wohldefiniert.

Dies definiert tatsächlich eine Lösung $x_{\max}(\cdot)$ des Anfangswertes: für $t = t_0$ gilt

$$x_{\max}(t_0) = x(t_0) = x_0$$

für jede Lösung $x(\cdot)$, die in der Definition von $x_{\max}(\cdot)$ betrachtet wird. Des Weiteren gibt es nach Definition für $t \in (a,b)$ eine Lösung $x(\cdot)$, die bei $t$ und damit auch auf einer Umgebung von $t$ definiert ist. Verwenden wir diese Lösung in der ganzen Umgebung so sehen wir auch

$$\dot{x}_{\max}(t) = \dot{x}(t) = f(t, x(t)) = f(t, x_{\max}(t)),$$

was zu zeigen war.

Es verbleibt den Beweis des behaupteten Verhaltens der maximalen Lösung $x_{\max}(\cdot)$ in der Nähe von $a$ und $b$ zu erbringen. Falls $b = \infty$ gilt, so gibt es nichts zeigen. Also angenommen $b < \infty$. Wir wollen zeigen, dass der Grenzwert $\lim_{t \nearrow b}(t, x_{\max}(t))$ in $U$ nicht existiert und nehmen deswegen indirekt an, dass

$$\lim_{t \nearrow b}(t, x_{\max}(t)) = (b, x_b) \in U.$$

Dann gibt es nach der Existenzaussage in Proposition 14.25 eine Lösung $y(\cdot)$ des Anfangswertproblems

$$\begin{cases} y(b) = x_b, \\ \dot{y}(t) = f(t, y(t)), \end{cases} \tag{14.12}$$

welches zumindest auf $(b-\delta, b+\delta)$ für ein $\delta > 0$ definiert ist. Wir verwenden diese Lösung, um eine Lösung des ursprünglichen Anfangswertproblems zu definieren. Sei $x_{\text{neu}}(\cdot) : (a, b+\delta) \to \mathbb{R}^d$ gegeben durch

$$x_{\text{neu}}(t) = \begin{cases} x_{\max}(t) & \text{falls } t \in (a,b) \\ y(t) & \text{falls } t \in [b, b+\delta) \end{cases}$$

für $t \in (a, b+\delta)$. Wir behaupten nun, dass $x_{\text{neu}}(\cdot)$ ebenfalls das Anfangswertproblem (14.11) löst, was dann aber einen Widerspruch zur Definition von $b$ darstellt.

Dazu bemerken wir zuerst, dass sowohl $x_{\max}(\cdot)$ als auch $y(\cdot)$ stetig sind und somit $x_{\text{neu}}(\cdot)$ sicherlich an allen Punkten in $(a, b+\delta) \setminus \{b\}$ stetig ist. Weiter gilt per Annahme

$$\lim_{t \nearrow b} x_{\max}(t) = x_b = y(b) = \lim_{t \to b} y(t)$$

und somit ist $x_{\text{neu}}(\cdot) : (a, b+\delta) \to \mathbb{R}^d$ stetig. Da $x_{\max}(\cdot)$ und $y(\cdot)$ auf $(a, b)$ respektive $(b, b+\delta)$ die Differentialgleichung $\dot{x}(t) = f(t, x(t))$ lösen, erfüllt auch $x_{\text{neu}}(\cdot)$ diese bei allen Punkten in $(a, b+\delta)$ ausser möglicherweise bei $b$.

Es bleibt zu zeigen, dass $x_{\text{neu}}(\cdot)$ bei $b$ differenzierbar ist und dort ebenfalls $\dot{x}(t) = f(t, x(t))$ genügt. Die rechtsseitige Ableitung von $x_{\text{neu}}(\cdot)$ bei $b$ erfüllt

$$\big(x_{\text{neu}}\big)^{\cdot}_{+} = \dot{y}(b) = f(b, y(b)) = f(b, x_{\text{neu}}(b))$$

wie gewünscht. Die linksseitige Ableitung erfordert ein kleines Argument, welches auf der Stetigkeit von $t \in [t_0, b] \mapsto f(t, x_{\text{neu}}(t))$ beruht (siehe Übung 14.26) und verläuft dann analog. Wie erwähnt zeigt dieser Widerspruch zur Definition von $b$, dass $x_{\max}(\cdot)$ das behauptete Verhalten für $t \nearrow b$ besitzt. Der Nachweis des Verhaltens für $t \searrow a$ erfolgt analog. □

VO: 30.05.18

**Übung 14.26** (Fortsetzung der Ableitung auf Randpunkte)**.** *Seien $t_0 < b$ reelle Zahlen und $x : [t_0, b] \to \mathbb{R}^d$ auf $[t_0, b)$ differenzierbar, so dass $\dot{x}(t) = f(t)$ für eine stetige Funktion $f : [t_0, b] \to \mathbb{R}^d$ und alle Punkte $t \in [t_0, b)$. Dann existiert die linksseitige Ableitung $\dot{x}_{-}(b) = f(b)$.*

Wir wollen noch kurz erwähnen, wie man obigen Existenz- und Eindeutigkeitssatz (Theorem 14.23) auch auf die Differentialgleichung $y' = |y|^{\frac{2}{3}}$ in Beispiel 14.1 anwenden kann. Da dies unser Beispiel für Mehrdeutigkeit ist, können die Vorraussetzungen des Satzes nicht erfüllt sein. In der Tat ist $f(y) = |y|^{\frac{2}{3}}$ bei $y = 0$ nicht lokal Lipschitz-stetig (wieso?). Wenn man allerdings den Definitionsbereich der Differentialgleichung als $U = \mathbb{R} \times (\mathbb{R} \setminus \{0\})$ definiert, dann sind alle Vorraussetzungen erfüllt (überprüfen Sie dies) und die eindeutige Existenz der Lösungen der zugehörigen Anfangswertprobleme garantiert. Dies ist kein Widerspruch, da die Lösungen in Beispiel 14.1 abseits der "Weiche entlang der horizontalen Achse" eindeutig sind, aber bei der Weiche (das heisst ausserhalb von $U$) eben länger oder weniger lang verbleiben können.

**Applet 14.27** (Diverse Differentialgleichungen und ihre Lösungen)**.** *Illustration der Lösungen einiger Differentialgleichungen mit jeweils unterschiedlichem Verhalten. In den nächsten paar Abschnitten werden wir uns unter anderem mit der Frage beschäftigen, ob sich nur anhand der Differentialgleichung die Lösungen (in Abhängigkeit des Anfangswerts) beschreiben lassen (ohne explizites Lösen der Gleichung).*

**Übung 14.28** (Challenge)**.** *Seien $U$ und $f$ wie im Existenz- und Eindeutigkeitstheorem von Picard-Lindelöf (Theorem 14.23) und $x(\cdot) : (a, b) \to \mathbb{R}^d$ die zugehörige maximale Lösung eines Anfangswertproblems für $(t_0, x_0) \in U$. Zeigen Sie als Verschärfung von Theorem 14.23, dass*

*es für jede kompakte Menge $K \subseteq U$ Zeitpunkte $\alpha, \beta \in (a,b)$ gibt, so dass $(t, x(t)) \notin K$ für alle $t \in (a,b) \setminus (\alpha, \beta)$. In diesem Sinn verlässt die Lösung jede kompakte Teilmenge von $U$ (und sozusagen damit auch $U$).*

### 14.4.1 Mögliches Verhalten in der Nähe eines Fixpunktes*

Wir nehmen hier, nebst den Annahmen in des Existenz- und Eindeutigkeitssatzes von Picard-Lindelöf, an, dass die Differentialgleichung $\dot{x} = f(x)$ autonom und in einer offenen Umgebung $U \subseteq \mathbb{R}^2$ von $0 \in \mathbb{R}^2$ definiert ist. Des Weiteren setzen wir voraus, dass $f(0) = 0$ ist. Das heisst, dass die konstante Funktion $t \in \mathbb{R} \mapsto x(t) = 0$ das Anfangswertproblem $\dot{x} = f(x)$, $x(0) = 0$ löst. Wir wollen das Verhalten von Lösungen zu einem Anfangswert nahe bei der 0 in verschiedenen Situationen untersuchen.

**Übung 14.29** (Schwarzes Loch)**.** *Seien $U$ und $f$ wie oben. Angenommen*

$$f(x) = Ax + o(\|x\|) \quad \text{für } x \to 0$$

*für $A \in \mathrm{Mat}_{2,2}(\mathbb{R})$. Sei $x : I \to \mathbb{R}^2$ eine maximale Lösung.*

*a) Sei $A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ mit $\lambda_1, \lambda_2 < 0$. Zeigen Sie, dass es um den Nullpunkt einen Attraktivitätsbereich gibt, d.h. eine Umgebung von $U$ von Null, so dass folgendes gilt: wenn $x(t_0) \in U$ für ein $t_0 \in I$, dann gilt $\lim_{t\to\infty} x(t) = 0$.*

*b) Sei $A = \begin{pmatrix} \lambda & -1 \\ 1 & \lambda \end{pmatrix}$ mit $\lambda < 0$. Zeigen Sie, dass es um den Nullpunkt einen Attraktivitätsbereich gibt.*

*Hinweis: Betrachten Sie die Hilfsfunktion $H(x) = \|x\|^2$.*

*Bemerkung.* Übung 14.29 lässt ich ebenso auf höhere Dimensionen und auf alle Matrizen mit nur negativen Realteilen der Eigenwerte verallgemeinern.

Sogenannte hyperbolische Fixpunkte (wo negative und positive Eigenwerte auftreten) sind genauso wie die Fixpunkte der obigen Übung „stabil“. Das heisst, das Verhalten in der Nähe der Null ist analog zu der “zugehörigen linearen Differentialgleichung”. Wie die folgende Übung zeigt ist dies bei Eigenwerten mit verschwindendem Realteil nicht der Fall.

**Übung 14.30** (Nach aussen rotierende Scheiben)**.** *Seien $U = B_2(0)$ und $f$ gegeben durch*

$$f(x) = \begin{pmatrix} -x_2 + x_1^3 g(\|x\|^2) \\ x_1 + x_2^3 g(\|x\|^2) \end{pmatrix},$$

*wobei $g : [0,4] \to \mathbb{R}$ Lipschitz-stetig ist. Seien $\alpha < \beta$ in $[0,4)$ mit $g(\alpha) = g(\beta) = 0$ und $g(r) > 0$ für alle $r \in (\alpha, \beta)$. Zeigen Sie, dass jede Lösung des Anfangwertproblems $\dot{x} = f(x)$ mit Anfangswert $x_0 \in U$ und $\|x\|^2 \in (\alpha, \beta)$ auf ganz $\mathbb{R}$ definiert ist und die Lösung im Gegenuhrzeigersinn um den Ursprung rotiert und von Innen die Lösung $y_\beta(t) = \sqrt{\beta} \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}$*

*approximiert. Falls stattdessen $g(r) < 0$ für alle $r \in (\alpha, \beta)$, gilt dies analogerweise von aussen und der Lösung* $y_\alpha(t) = \sqrt{\alpha}\begin{pmatrix}\cos(t)\\ \sin(t)\end{pmatrix}$.

### 14.4.2 Stetige Abhängigkeit*

Wir wollen hier den Existenz- und Eindeutigkeitssatz (Theorem 14.23) in einigen Übungen erweitern.

**Übung 14.31** (Glattheit)**.** *Sei $U \subseteq \mathbb{R}^d$ offen und angenommen $f : U \subseteq \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ ist glatt. Zeigen Sie, dass jede Lösung eines Anfangswertproblems zur rechten Seite $f$ glatt ist.*

Wir möchten nun genauer untersuchen, was geschieht, wenn man an den „Angaben“ in Theorem 14.23 etwas „rüttelt“. Wir beginnen damit, den Anfangswert leicht zu ändern und fragen also, wie stark sich die eindeutige Lösung (welche nach Theorem 14.23 existiert) ändert.

**Proposition 14.32** (Stetige Abhängigkeit der Lösung vom Anfangswert)**.** *Seien $U$ und $f$ wie in Theorem 14.23. Sei $(t_0, x_0) \in U$ ein Anfangswert und sei $x$ eine auf dem kompakten Interval $[a, b] \ni t_0$ definierte Lösung zum Anfangswertproblem $\dot{x} = f(t, x(t))$, $x(t_0) = x_0$. Sei $\varepsilon > 0$ beliebig. Dann existiert ein $\delta > 0$ mit folgender Eigenschauft: Falls ein weiterer Anfangswert $\tilde{x}_0$ mit $(t_0, \tilde{x}_0) \in U$ die Abschätzung $\|x_0 - \tilde{x}_0\| < \delta$ erfüllt, dann ist die zugehörige maximale Lösung $\tilde{x}$ ebenfalls auf dem Intervall $[a, b]$ definiert und erfüllt die Abschätzung $\|x|_{[a,b]} - \tilde{x}|_{[a,b]}\|_\infty < \varepsilon$.*

Tatsächlich lässt sich beweisen, dass die Lösung gewissermassen „differenzierbar“ vom Anfangswert abhängt. Wir skizzieren den Beweis der obigen Proposition in folgenden Übung.

**Übung 14.33.** *Beweisen Sie Propoosition 14.32 in folgenden Schritten.*

*(i) Zeigen Sie dazu, dass es genügt eine lokale Version zu beweisen: Angenommen es existiert für jedes $(t_1, x_1) \in U$ ein $\eta > 0$ so dass die zugehörige Lösung $x$ auf $[t_1 - \eta, t_1 + \eta]$ definiert ist und dort die Stetigkeitsaussage von Proposition 14.32 erfüllt. Nehmen Sie weiters für ein kompaktes Intervall $[a, b]$ an, dass $\eta > 0$ gleichmässig für alle $t_1 \in [a, b]$ und $x_1 = x(t_1)$ gewählt werden kann. Schliessen Sie daraus auf die Aussage in Proposition 14.32.*

*(ii) Sei $(X, \mathrm{d})$ ein vollständiger metrischer Raum, seien $T, \tilde{T} : X \to X$ Lipschitz-Kontraktionen mit gleicher Lipschitz-Konstante $\lambda < 1$ und sei $\delta > 0$. Angenommen $T, \tilde{T}$ erfüllen die Eigenschaft, dass für alle $x \in X$ die Abschätzung $\mathrm{d}(Tx, \tilde{T}x) < \delta$ gilt. Seien $x \in X$ der Fixpunkt von $T$ und sei $\tilde{x} \in X$ der Fixpunkt von $\tilde{T}$. Zeigen Sie, dass eine Konstante $C$ in Abhängigkeit von $\lambda$ existiert, so dass $\mathrm{d}(x, \tilde{x}) \leq C\delta$.*

*(iii) Verwenden Sie (ii), um den Beweis von Proposition 14.25 zu einem Beweis der lokalen Version in (i) zu adaptieren.*

Wir erwähnen noch, dass man genauso wie am Anfangswert man auch an dem Vektorfeld leicht „rütteln“ kann, ohne dass man dabei die Lösung stark verändert. Der Beweis dafür kann ebenfalls durch eine genaue Analyse der Beweise des Banachschen Fixpunktsatzes (Satz 9.54) und des lokalen Existenz- und Eindeutigkeitssatzes (Proposition 14.25) erbracht werden. Wir wollen dies aber hier nicht weiter verfolgen.

### 14.4.3 Ein weiteres Beispiel: der rotierende Tropfen*

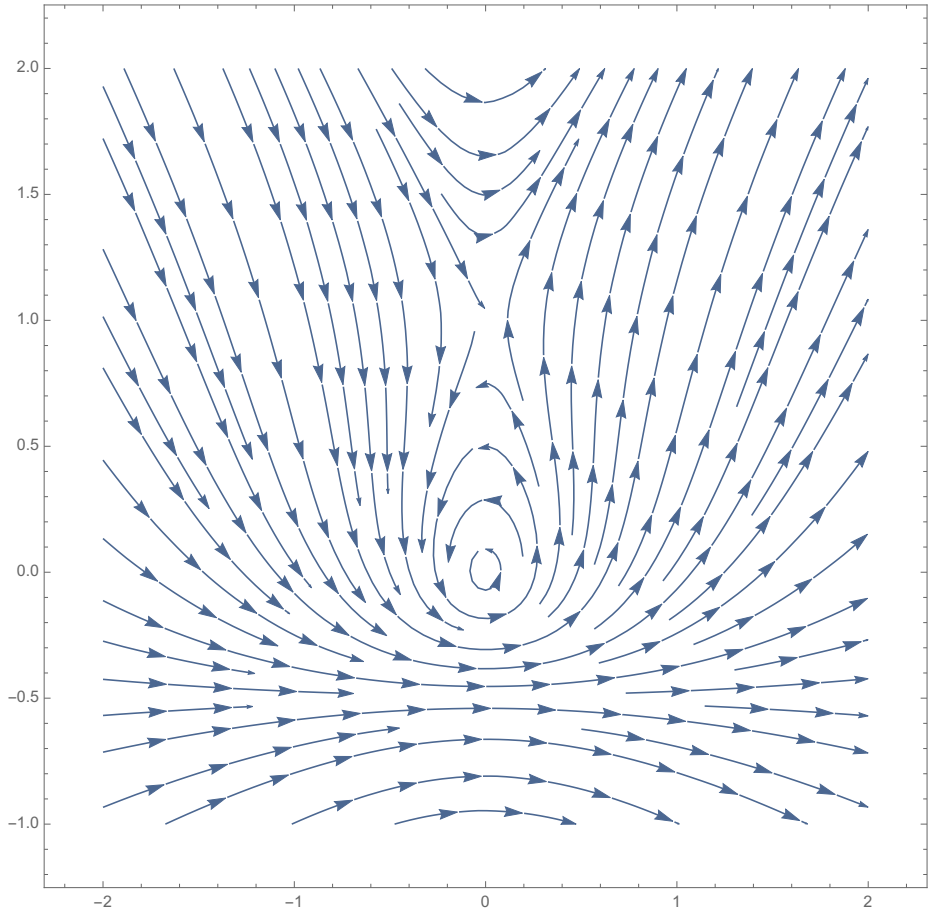
Wir wollen hier das nicht-lineare, autonome Differentialgleichungssystem

$$\begin{aligned} \dot{x} &= -y + x^2 + y^2 \\ \dot{y} &= x + 2xy \end{aligned} \tag{14.13}$$

betrachten. Im Gegensatz zu dem Beispiel in Abschnitt 14.3.4 werden wir dieses System qualitativ untersuchen und die verschiedenen Lösungen in Kategorien unterteilen ohne diese zu berechnen. In der Tat da wir (abgesehen von drei speziellen Lösungen) keine konkreten Formeln angeben werden, beziehen wir unsere Lösungen einzig und allein aus dem Existenz- und Eindeutigkeitssatz von Picard-Lindelöf (Theorem 14.23).

**Übung 14.34.** *Zeigen Sie, dass* (14.13) *der Voraussetzung der lokalen Lipschitz-Stetigkeit in Theorem 14.23 genügt.*

Um die Lösungen der verschiedenen Anfangswertprobleme zu (14.13) zu verstehen, bedienen wir uns unter anderem der Darstellung des Vektorfeld (genauer, der **Strömungslinien**) im folgenden Bild.




Figur 14.2: Visualisierung des Vektorfelds der Differentialgleichung (14.13)

Aus dem Bild entnimmt man mehrere Vermutungen, die wir in verschiedenen Schritten überprüfen wollen.

**Übung 14.35** (Konstante Lösungen und Lösungen auf einer Geraden)**.**

*(i) Zeigen Sie, dass die konstanten Lösungen von* (14.13) *durch*

$$\begin{cases} x(t) = 0 \\ y(t) = 0 \end{cases} \qquad \textit{und} \qquad \begin{cases} x(t) = 0 \\ y(t) = 1 \end{cases}$$

*gegeben sind.*

*(ii) Zeigen Sie, dass es (abgesehen von der konstanten Lösung und einer Zeitverschiebung) nur eine Lösung gibt, die sich entlang einer Geraden bewegt, und dass die entsprechende Gerade durch $y = -\frac{1}{2}$ gegeben ist.*

*(iii) Bestimmen Sie die Lösung zum Anfangswert $x(0) = 0$, $y(0) = -\frac{1}{2}$.*

Das Verhalten der weiteren Lösungen erfordert nebst kleinen Rechnungen auch weitere Argumente.

**Übung 14.36** (Lösungen unterhalb $y_0 = -\frac{1}{2}$)**.**

*(i) Zeigen Sie, dass für jede Lösung $t \mapsto (x(t), y(t))^t$ eines Anfangswertproblems zu* (14.13) *mit Anfangswert $x_0 \in \mathbb{R}$, $y_0 < -\frac{1}{2}$ auch $y(t) < -\frac{1}{2}$ für alle $t$ im Definitionsbereich der Lösung erfüllt ist.*

*(ii) Zeigen Sie, dass für eine solche Lösung $x$ auch $\dot{x}(t) \geq \frac{1}{2}$ für alle $t$ im Definitionsbereich der Lösung erfüllt ist.*

*(iii) Zeigen Sie, dass für jede dieser Lösungen der Definitionsbereich ein beschränktes Intervall bildet.*

Wie bereits das Bild 14.2 zeigt, ist das Verhalten der Lösungen für $y_0 > -\frac{1}{2}$ interessanter (das heisst, schwieriger zu verstehen). Es sieht so aus, als ob es geschlossene (periodische) Lösungen rund um den Ursprung gäbe. Aber wie wir bereits in Übung 14.30 gesehen haben, genügt es dafür nicht, nur den linearen Anteil der rechten Seite der Differentialgleichung rund um den Ursprung zu betrachten. Es könnte also immer noch sein, dass die Lösungen entlang von Spiralen nach 0 streben oder umgekehrt von 0 ausgehend (vielleicht sogar in endlicher Zeit) nach Unendlich streben. Wir wollen dies ausschliessen und auch die Existenz der "Tropfenfunktion" $t \mapsto (x(t), y(t))^t$ mit

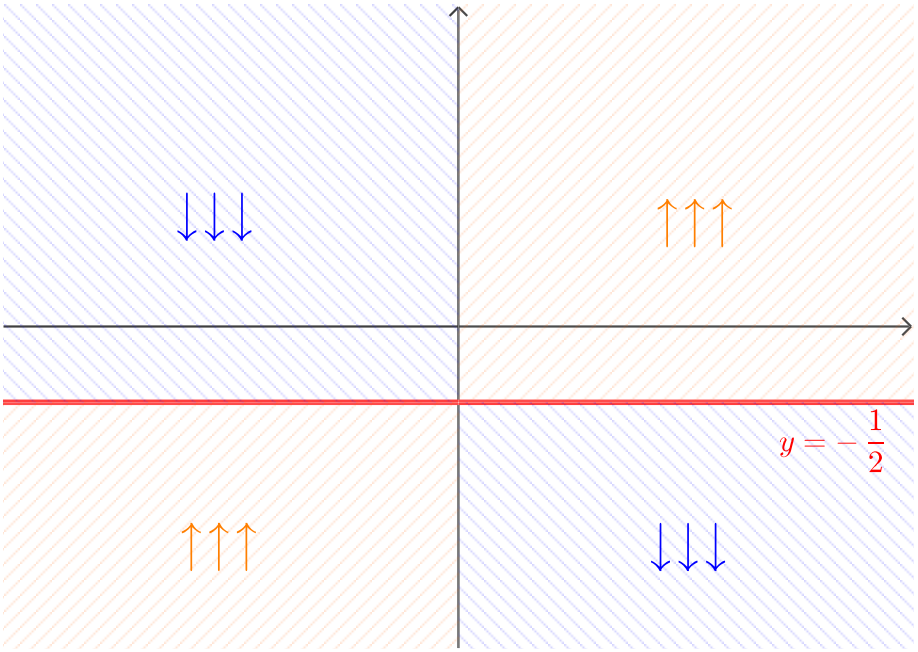
$$\begin{aligned} &\lim_{t \to -\infty} (x(t), y(t))^t = (1, 1)^t \\ &x(0) = 0, \ y(0) < 0 \\ &\lim_{t \to \infty} (x(t), y(t))^t = (1, 1)^t \end{aligned}$$

zeigen.

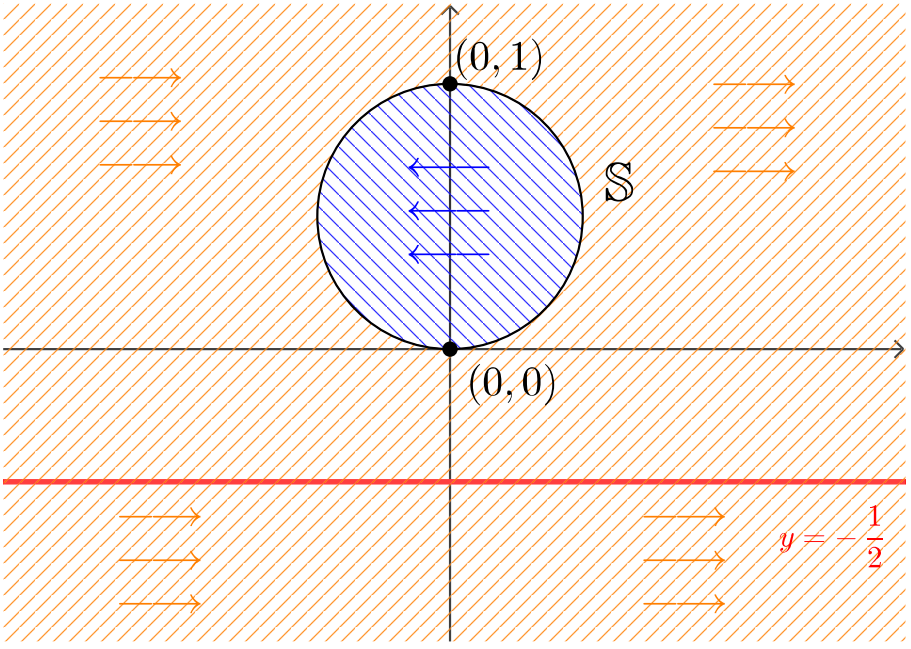
**Übung 14.37.** *Zeigen Sie, dass für jede Lösung $t \mapsto (x(t), y(t))^t$ eines Anfangswertproblems zu* (14.13) *mit Anfangswert $x_0 \in \mathbb{R}$, $y_0 > -\frac{1}{2}$ auch $y(t) > -\frac{1}{2}$ für alle $t$ im Definitionsbereich der Lösung erfüllt ist.*

Nun teilen wir $\mathbb{R}^2$ in mehrere Gebiete auf und untersuchen auf diesen die Vorzeichen von $\dot{x}$ und $\dot{y}$.

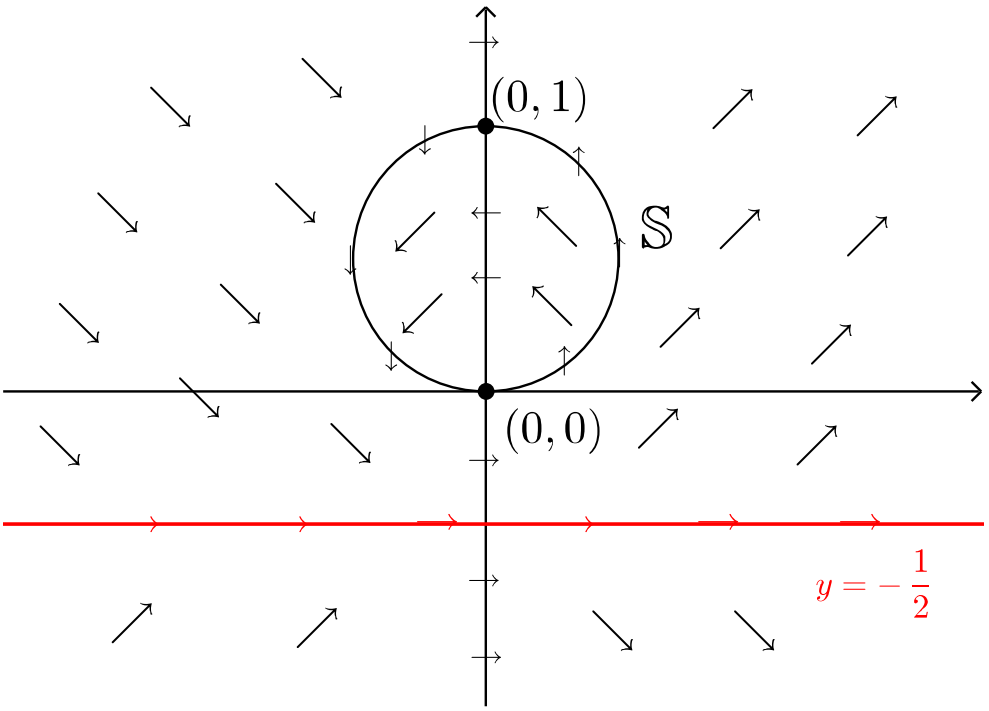
Wir faktorisieren $\dot{y} = x + 2xy = x(1 + 2y)$ und erhalten eine Unterteilung von $\mathbb{R}^2$ in 4 Quadrate, wo $\dot{y}$ jeweils dasselbe Vorzeichen besitzt; siehe folgendes Bild.




Weiter erhalten wir durch quadratische Ergänzung $\dot{x} = -y + x^2 + y^2 = x^2 + (y - \frac{1}{2})^2 - \frac{1}{4}$. Dies ergibt für das Vorzeichen von $\dot{x}$ eine Unterteilung von $\mathbb{R}^2$ in einen Kreis mit Mittelpunkt $(0, \frac{1}{2})^t$ und Radius $\frac{1}{2}$; siehe folgendes Bild.




Wenn wir diese beiden Unterteilungen von $\mathbb{R}^2$ gemeinsam betrachten, erhalten wir die schematische Darstellung in Entsprechung zum Bild 14.2.




Wir wollen nun die Lösung des Anfangswertproblems (14.13) mit Anfangswerten $x(0) = 0$, $y(0) = y_0 \in (0, 1)$ untersuchen. Da $\dot{x}(0) = (y_0 - \frac{1}{2})^2 - \frac{1}{4} < 0$ ist, ist $x(t) < 0$ für alle $t \in (0, \delta]$ für ein gewisses $\delta > 0$. Für diese $t$ gilt weiter $\dot{y} < 0$. Also bewegt sich die Lösung $(x(t), y(t))^t$ nach links und nach unten, bis zu einem Zeitpunkt $t_1 > 0$ (minimal gewählt) $(x(t_1), y(t_1))^t \in \mathbb{S}$ gilt. In der Tat ist $x(t) \leq x(\delta) < 0$ und $y(t) \geq 0$, solange der Kreis nicht erreicht wurde. Des Weiteren gilt dann $\dot{y}(t) \leq x(\delta)(1 + 2y(t)) \leq x(\delta)$. Das heisst, der Kreis $\mathbb{S}$ muss in endlicher Zeit erreicht werden.

Ab dem Zeitpunkt $t_1$ bewegt sich die Lösung nach unten und nach rechts. Wir wollen als nächstes den minimalen Zeitpunkt $t_2 > t_1$ finden, bei dem $y(t_2) = 0$ ist. Da $y(t_1) > 0$ ist und $\dot{y}(t) = x(t)(1 + 2y(t)) < 0$ für alle $t$ mit $x(t) < 0$ gilt, sollte dieses $t_2$ existieren. Doch besteht die Gefahr, dass die Lösung zur konstanten Lösung bei $(0, 0)$ strebt, bevor sie die $x$-Achse erreicht. Hierbei kann aber die Lösung den Kreis $\mathbb{S}$ nicht mehr betreten, da auf dem Kreis das Richtungsfeld nach aussen zeigt.

**Übung 14.38.** *Verwenden Sie die Funktion $H(x, y) = x^2 + y(y-1) = x^2 + (y - \frac{1}{2})^2 - \frac{1}{4}$, um zu zeigen, dass der Abstand der Lösung zum Punkt $(0, \frac{1}{2})^t$ (zumindest in der Nähe des Ursprungs) zunimmt. Schliessen Sie daraus, dass ein minimales $t_2 > t_1$ mit $y(t_2) = 0$ existieren muss.*

Es gibt ein $\eta > 0$ mit $x(t) < 0$ und $\dot{y}(t) < 0$ für alle $t \in (t_2, t_2 + \eta]$. Nun folgt aus $\dot{x}(t) = -y(t) + x^2(t) + y^2(t)$, dass es ein minimales $t_3 > t_2$ gibt mit $x(t_3) = 0$.

Die Argumente bis jetzt haben eigentlich nur das Vorzeichen von gewissen Ableitungen verwendet und sind sehr allgemein einsetzbar. Doch können diese den vermuteten Zufall, dass es periodische Punkte um den Ursprung gibt, nicht erklären. Für diesen Zufall benötigen wir ein globales Argument wie in folgender Übung.

**Übung 14.39.** *Zeigen Sie, dass die Spiegelung um die y-Achse ($x \to -x$) und Zeitumkehr ($t \to -t$) zu demselben Differentialgleichungssystem führt.*

Aus obiger Übung und Eindeutigkeit der Lösung folgt, dass die von uns untersuchte Lösung zu dem Anfangswert $x_0 = 0$, $y_0 \in (0,1)$ die Gleichungen $x(-t) = -x(t)$ und $y(-t) = y(t)$ für alle $t$ erfüllt. Insbesondere gilt $x(-t_3) = 0 = x(t_3)$ und $y(-t_3) = y(t_3)$, womit die Lösung tatsächlich periodisch ist.

Wir bewegen nun die Koordinate $y_0$ unseres Anfangswertes $(0, y_0)^t$ nach 1, womit $t_1$ unbeschränkt grösser wird aber $t_2 - t_1$ und $t_3 - t_2$ beschränkt bleiben (wieso?). Die Punkte $(x(t_3), y(t_3))^t$ in $\{0\} \times (0, -\frac{1}{2})$ besitzen einen Grenzwert $(x_1, y_1)^t \in \{0\} \times (0, -\frac{1}{2}]$ (wieso?). Die Lösung zum Anfangswert $(x_1, y_1)^t$ ist die gesuchte Tropfenfunktion (wieso?).

**Übung 14.40.** *Erklären Sie die letzten drei Behauptungen. (Ein vollständiger Beweis hat natürlich Vorteile, aber eine gute Erklärung ohne formale Details kann auch viel Zeit sparen.)*

## 14.5 Lineare Differentialgleichungssysteme*

**Proposition 14.41** (Maximale Lösungen für lineare Differentialgleichungssysteme)**.** *Sei $I_0 \subseteq \mathbb{R}$ ein offenes Intervall und $A : t \in I_0 \mapsto A_t \in \mathrm{Mat}_{dd}(\mathbb{R})$ stetig. Dann ist die Menge der (maximalen) Lösungen des homogenen Differentialgleichungssystems*

$$x'(t) = A_t x(t)$$

*ein $d$-dimensionaler Teilraum $R$ des Raumes $C^1(I_0, \mathbb{R}^d)$ der $\mathbb{R}^d$-wertigen stetig differenzierbaren Funktion auf $I_0$. Für eine stetige Störfunktion $g : I_0 \to \mathbb{R}^d$ ist die Menge der Lösungen des inhomogenen Differentialgleichungssystems*

$$x'(t) = A_t x(t) + g(t)$$

*von der Form $x_{\text{part}} + R$, wobei $x_{\text{part}} : I_0 \to \mathbb{R}^d$ eine partikuläre Lösung der inhomogenen Differentialgleichungssystems ist. Insbesondere haben Anfangswertprobleme zu jedem $t_0 \in I_0$ eine auf ganz $I_0$ eindeutig bestimmte Lösung.*

**Übung 14.42** (Lineare Differentialgleichungen $d$-ter Ordnung)**.** *Formulieren und beweisen Sie ein Korollar zu Proposition 14.41 für homogene und inhomogene lineare Differentialgleichungen $d$-ter Ordnung*

$$y^{(d)}(t) = f_{d-1}(t)y^{(d-1)}(t) + \ldots + f_0(t)y(t) + g(t),$$

*welches zum Beispiel die Lösungsmenge der Bessel-Differentialgleichung* (10.13) *beschreibt.*

*Beweis.* Wir bemerken zuerst, dass für $U = I_0 \times \mathbb{R}^d$ die rechte Seite $A_t x$ stetig von $(t, x) \in U$ und linear vom Ort $x \in \mathbb{R}^d$ abhängt. Insbesondere hängt $A_t x$ Lipschitz-stetig vom Ort ab und damit sind die Vorraussetzungen vom Existenz- und Eindeutigkeitssatz von Picard-Lindelöf (Theorem 14.23) erfüllt.

Sei $t_0 \in I_0$. Wir verwenden Theorem 14.23 für die Anfangswertsprobleme $x(t_0) = e_k$ zur homogenen Differentialgleichung und erhalten zu jedem Standardbasisvektor $e_k$ von $\mathbb{R}^d$ eine Lösung $x_k : I_k \to \mathbb{R}^d$ mit $x_k(t_0) = e_k$ für $k = 1, \ldots, d$. Hier ist $I_k \subseteq I_0$ das in Theorem 14.23 gefundene maximale Intervall auf dem $x_k$ definiert ist. Wir setzen die einzelnen Lösungen auch zu einer Matrix-wertigen Funktion $L : t \mapsto L_t$ zusammen, wobei $L_t = (x_1(t), \ldots, x_d(t))$ für alle $t$ im gemeinsamen Definitionsbereich.

Wir behaupten $I_1 = I_2 = \ldots = I_d = I_0$ (womit $L$ auf $I_0$ definiert ist) und $\det(L_t) > 0$ für alle $t \in I_0$. Für den Beweis dieser Behauptung setzen wir

$$\begin{aligned}
\alpha &= \inf \{t \in I_1 \cap I_2 \cap \ldots \cap I_d \mid \det(L_s) > 0 \text{ für alle } s \in [t, t_0]\} \\
\beta &= \sup \{t \in I_1 \cap I_2 \cap \ldots \cap I_d \mid \det(L_s) > 0 \text{ für alle } s \in [t_0, t]\} \\
I_{\max} &= (\alpha, \beta) \subseteq I_0 = (a, b),
\end{aligned}$$

und nehmen indirekt $\beta < b$ an. (Der Widerspruchsbeweis im Fall $\alpha > a$ verläuft analog.)

Da $\beta \in I_0$ ist, können wir Theorem 14.23 auch für die Anfangswertprobleme $x(\beta) = e_k$ für $k = 1, \dots, d$ anwenden und erhalten damit Lösungen, die in (möglicherweise verschiedenen) Umgebung von $\beta$ definiert sind. Wir schneiden die Umgebungen um einen gemeinsamen Definitionsbereich $\tilde{I}$ zu erhalten und setzen die einzelnen Lösungen nochmals zu einer Matrix-wertigen Funktion $\tilde{L} : \tilde{I} \to \mathrm{Mat}_{dd}(\mathbb{R})$ zusammen, womit $\big(\tilde{L}(e_k)\big)'(t) = A_t \tilde{L}_t(e_k)$ für $k = 1, \dots, d$ und $\tilde{L}_\beta = I_d$ gilt. Da $\det(\tilde{L}_\beta) = 1$ können wir, wenn nötig, die Umgebung kleiner wählen und und des Weiteren annehmen, dass $\det(\tilde{L}_t) > 0$ für alle $t \in \tilde{I} = (\tilde{\alpha}, \tilde{\beta})$.

Sei nun $t_1 \in I_{\max} \cap \tilde{I}$. Wir definieren für jedes $k \in \{1, \dots, d\}$ die Funktion

$$y_k(t) = \begin{cases} x_k(t) & \text{für } t \in (\alpha, t_1] \\ \tilde{L}_t\big(\tilde{L}_{t_1}^{-1} x_k(t_1)\big) & \text{für } t \in (t_1, \tilde{\beta}), \end{cases}$$

wobei die Definition im zweiten Fall so gewählt wurde, dass $y_k$ bei $t_1$ stetig ist. Da nun $y_k$ die Differentialgleichung im Definitionsbereich sowohl für $t \in (\alpha, t_1)$ als auch für $t \in (t_1, \tilde{\beta})$ erfüllt, folgt aus der Stetigkeit (wie schon im Beweis von Theorem 14.23), dass $y_k$ die Differentialgleichung auch bei $t_1$ erfüllt.

Gemeinsam definieren die Lösungsfunktion $y_1, \dots, y_d$ wieder eine Matrix, die nun auf $(\alpha, \tilde{\beta})$ definiert ist und auch auf $[t_1, \tilde{\beta})$ eine positive Determinante besitzt. Da $\beta < \tilde{\beta}$ widerspricht dies aber der Definition von $\beta$ und es folgt, dass $\beta = b$ gelten muss.

Der Lösungsraum $R$ des homogenen Differentialgleichungssystems besteht nun aus den Funktionen $t \in I_0 \to L_t v$ für beliebige $v \in \mathbb{R}^d$, und ist damit wie behauptet ein $d$-dimensionaler Teilraum von $C^1(I, \mathbb{R}^d)$.

Wir betrachten nun ein stetiges Störglied $g : I_0 \to \mathbb{R}^d$ und nehmen an, dass $x_{\text{part}} : I_{\max} \to \mathbb{R}^d$ eine maximale Lösung zum Anfangswertproblem $x'(t) = A_t x(t) + g(t)$ und $x(t_0) = 0$ ist. Wir bezeichnen den Definitionsbereich wiederum mit $I_{\max} = (\alpha, \beta) \subseteq I_0 = (a, b)$ und nehmen indirekt $\beta < b$ an. Wir wenden nochmals Theorem 14.23 auf das Anfangswertproblem $x'(t) = A_t x(t) + g(t)$ und $x(\beta) = 0$ an. Dies definiert ein Intervall $\tilde{I} = (\tilde{\alpha}, \tilde{\beta}) \ni \beta$ und eine Lösung $y : \tilde{I} \to \mathbb{R}^d$ dieses Anfangswertproblems. Sei wiederum $t_1 \in (\tilde{\alpha}, \beta)$. Dann können wir die allgemeine Lösung des homogenen Differentialgleichungssystems (welches auf ganz $I_0$ existiert) und die partikuläre Lösung $y$ auf $\tilde{I}$ verwenden um eine Lösung $\tilde{y}$ auf $\tilde{I}$ zu dem Anfangswertproblem $x'(t) = A_t x(t) + g(t)$ und $x(t_1) = x_{\text{part}}(t_1)$ zu finden. Da aber nach Theorem 14.23 die maximale Lösung des Anfangswertproblems eindeutig bestimmt ist, erhalten wir einen Widerspruch und $x_{\text{part}}$ muss auf ganz $I_0$ definiert sein.

Die Proposition ergibt sich nun aus dem „linearen Prinzip“, dass die allgemeine Lösung der inhomogenen linearen Differentialgleichung die Summe einer partikulären Lösung der inhomogenen linearen Differentialgleichung und der allgemeinen Lösung der homogenen linearen Differentialgleichung ist. □

VO: 30.05.18

## 14.6 Weitere Lernmaterialien

### 14.6.1 Verwendung des Kapitels

Wie bereits erwähnt ist die Untersuchung von Differentialgleichungen für Anwendungen in vielfältigen Themenbereichen von grosser Bedeutung. Falls aber die betrachtete Differentialgleichung nicht konkret durch bekannte Formeln gelöst werden kann, so muss man sich mitunter mit numerischen Methoden helfen. Für diese oder auch eine theoretische Untersuchung der Lösung ist aber von grosser Wichtigkeit zu wissen, dass es eine Lösung gibt. Denn erst dann kann man versuchen, die Lösung mit weiteren Überlegungen besser zu verstehen oder auch zu verstehen wie nahe die numerische Lösung der eigentlichen Lösung kommt. Der Existenz- und Endeutigkeitssatz von Picard-Lindelöf (Theorem 14.23) ist also aus praktischen und theoretischen Gründen sehr wichtig.

### 14.6.2 Übungen

**Übung.** *(i) Bestimmen Sie die allgemeine Lösung der Differentialgleichung*

$$y'' - 4y' + 4y = \sin(x).$$

*(ii) Bestimmen Sie die Lösung des Anfangswertproblems*

$$y'' - y = x, \quad y(0) = 1, \quad y'(0) = 3.$$

**Übung.** *Finden Sie eine Lösung des Anfangswertproblems*

$$y' - \left(\frac{4}{x} + 1\right) y = x^4$$
$$y(1) = 1$$

*auf dem Intervall* $(0, \infty)$.

**Übung.** *Lösen Sie folgende Differentialgleichungen (Sie müssen Spezialfälle, die durch das Verschwinden von gewissen Ausdrücken im Lösungsverfahren entstehen, nicht weiter untersuchen):*

*(i)* $(x^2 - x)y' = y^2 + y$,

*(ii)* $y' + e^y = 1$,

*(iii)* $xy' = 1 - y^2$,

*(iv)* $y'x^2 = y^2 + yx + x^2$.

### 14.6.3 Lernkarten

Sie können wiederum die Lernkarten oder den Graphen für Ihre Wiederholung der Themen des Kapitels verwenden.

# Anhang A

# Changelog

03. Okt. In Abschnitt 1.4.4 werden nun endliche Mengen definiert und das Schubfachprinzip wird bewiesen.

11. Okt. Abschnitt zu natürlichen Zahlen etwas umgestellt. Wohlordnung von $\mathbb{N}$ wird nun nach der zweiten Induktionsvariante in Abschnitt 2.2.1 bewiesen.

17. Okt. In Abschnitt 3.3.4 wurde eine Permutation von Textteilen vorgenommen, da dies die Idee besser betont. Inhalte sind aber unverändert.

18. Okt. Vereinfachung im Beweis von Satz 2.75 und eine Beweisskizze davor.

29. Okt. Neuer Unterabschnitt 3.8.2.

8. Nov. Transposition von 5.1.1 und 5.1.2, da Folgenstetigkeit zur Berechnung notwendig ist und sowohl Teilfolgen als auch Cauchyfolgen eher theoretisch wichtig sind und besser zusammenpassen. Inhalte sind unverändert. Proposition 5.23 ist neu auch im Skript, wodurch Satz 5.44 einfacher zu beweisen ist.

13. Nov. Monotonie-Beweis in Abschnitt 5.3.2 ist jetzt korrekt.

20. Nov. Nach Korollar 6.29 wurden die Begriffe „Majorante“ und „Minorante“ verkehrt herum definiert; dies ist jetzt behoben.

22. Nov. Verkürzter Beweis vom Produktsatz (Satz 6.36) und neues Beispiel 6.38 zur Veranschaulichung des Cauchy-Produkts.

3. Dez. Beweis von Satz 6.56 ist nun etwas kürzer und eleganter.

7. Dez. Die Liste der Integrale in Abschnitt 7.5.2 hatte ein Vorzeichenfehler bei arcsin.

18. Dez. Abschnitt 8.2.8 fasst die Verwendung der Leibniz-Notation bei Integralberechnungen zusammen.

23. Feb. In Definition 9.14 wurde zweimal und in Klammern „in $X$“ hinzugefügt. Wenn $X$ fest gewählt ist, dann spielt dies keine Rolle. Aber prinzipiell sind offen und abgeschlossen immer relativ zu einem gewählten Grundraum $X$. Im Beweis von Lemma 9.20 waren

zwei Druckfehler: einmal soll $y \in O_Y$ statt $y \in Y$ stehen und $Y \in O \cap Y$ sollte $y \in O \cap Y$ sein. Der Paragraph direkt vor dem Beweis von Lemma 9.20 ist auch neu.

5.März Gegen Ende des Beweises von (5) impliziert (1) sollte „für alle $y \in B_{r/4}(x)$“ statt „$B_{r/2}(x)$“ stehen. Teil des Beweises der weiteren Äquivalenzen wurde als Proposition 9.67 ausgelagert, damit dieser Beweis, der in der Vorlesung besprochen wird, besser von dem zusätzlichen Beweis abgetrennt ist.

12. März Gewisse Fehler in den Übungen des Abschnitts 9.9 wurden ausgebessert.

15. März Übung 10.35 wurde eingefügt.

17. März Stetigkeitsbeweis für Korollar 10.38 fehlte.

21. März Übung 10.53 wurde eingefügt.

23. März Fehler in (10.11) korrigiert und danach ein paar Kommentare eingefügt.

23. März Kleiner Nachtrag zur Dreiecksungleichung für vektorwertige Funktionen – siehe Abschnitt 5.7.6.

28. März Im Beweis von Satz 11.2 war an einer Stelle das Vorzeichen von $f(x+h) - f(x)$ falsch (nämlich wo der Ausdruck das erste Mal unter $h$ steht), der weitere Beweis ändert sich dadurch nicht. Applet 10.54 wurde hinzugefügt.

12. April Beispiel 11.3 hinzugefügt.

13. April Klarere Formulierungen von Definition 11.27 und Korollar 11.28.

20. April In Gleichung (12.6) im Beweis des Lebesgue-Kriteriums wurde fälschlicherweise $f$ anstelle von $f|_K$ geschrieben. In der Tat beschreibt Proposition 9.79 nur die Oszillation von $f$ auf $K$. Der Rest des Beweises bleibt (bis auf eine kleine Änderung) intakt.

23. April Wie von einer Studentin bemerkt, war der in der Vorlesung angedeutete Beweis von Proposition 12.22 allgemeiner als der im Skript durchgeführte Beweis. Die Proposition gilt daher auch für Riemann-integierbare Funktionen. Dies macht ebenso die letzte Aussage von Korollar 12.27 allgemeiner und führte zu einer Änderung und Kürzung von Übung 12.28. Der Abschnitt nach (12.7) wurde neu formuliert, um die kleine Komplikation rund um $\partial Q$ besser zu erklären.

14. Mai Abschnitt 13.2 hat zwei weitere Beispiele, die auch in der Vorlesung angesprochen wurden. Des Weiteren wurden einige Notationen weiter optimiert und der Begriff des Landes $\Phi(K) \subseteq S$ wie in der Vorlesung kurz angesprochen.

16. Mai Beweis von Proposition 13.40 hatte einige Tippfehler, und wird nun in kleineren Schritten erklärt. Vor Übung 13.31 wurde ein Paragraph eingefügt, der die Bedeutung der Positivität der Jacobi-Determinante weiter erklärt.

# Anhang B

# Christmas-Special:

## Ein Crash-Kurs zu Fourier-Reihen*

In diesem kurzen Abschnitt möchten wir das sogenannte „Basler Problem“ lösen und

$$\zeta(2) = \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$$

zeigen. Die Konvergenz der Reihe haben wir bereits in Beispiel 6.14 als eine der ersten Beispiele konvergenter Reihen gesehen, doch haben wir den Wert der Reihe mit den Methoden der Analysis 1 nicht berechnen können. Ähnlich wie bei der Besprechung des Werts der alternierenden harmonischen Reihe oder der Leibniz-Reihe in Abschnitt 8.1 werden wir ein allgemeineres Resultat zeigen, nämlich dass

$$\frac{1}{6} - \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{1}{2\pi^2 n^2} \mathrm{e}^{2\pi \mathrm{i} n x} = \frac{1}{6} - \sum_{n=1}^{\infty} \frac{1}{\pi^2 n^2} \cos(2\pi n x) = x(1-x) \tag{B.1}$$

für alle $x \in [0,1]$. Hier werden wir jedoch sogenannte **Fourier-Reihen** im Beweis verwenden.

Eine weitere Motivation zur Betrachtung von Fourier-Reihen findet man in der Musiktheorie. Was macht den Unterschied der Note $A$ auf der Klanggabel, dem Horn und der Klarinette aus? Alle drei haben die Frequenz 440 Hz. Im Gegensatz zur Klanggabel sind beim Horn und bei der Klarinette abgesehen von der Grundschwingung zu 440 Hz weitere Obertöne vorhanden. Die Amplituden dieser Schwingungen sind für die Klangfarbe verantwortlich. In diesem Sinne ist die mathematische „Definition der Klarinette“ durch

$$\begin{aligned} f(t) = \sin(\omega_1 t) &+ 0.75\sin(3\omega_1 t) + 0.5\sin(5\omega_1 t) + 0.14\sin(7\omega_1 t) \\ &+ 0.5\sin(9\omega_1 t) + 0.12\sin(11\omega_1 t) + 0.17\sin(13\omega_1 t) \end{aligned}$$

gegeben (siehe auch diesen Link). Detailliertere Erklärungen finden sich unter diesem Link. Dieses Experiment des Physik-Departments zeigt verschiedene Arten von Schwingungen, die auf einer Saite erzeugt werden können; weitere solche findet man hier.

Wir kehren nun zur mathematischen Diskussion zurück.

**Definition B.1.** Eine Funktion $f : \mathbb{R} \to \mathbb{C}$ ist $\mathbb{Z}$-**periodisch** (periodisch mit **Periode** 1), falls

$$f(x+n) = f(x)$$

für alle $x \in \mathbb{R}$ und $n \in \mathbb{Z}$.

Die wichtigsten Beispiele von $\mathbb{Z}$-periodischen Funktionen sind die Funktionen

$$x \in \mathbb{R} \mapsto \mathrm{e}^{2\pi \mathrm{i} n x} \in \mathbb{C}$$

für $n \in \mathbb{Z}$. Die $\mathbb{Z}$-periodischen Funktionen sind, wie wir gleich sehen werden, genau die Funktionen auf dem eindimensionalen Torus.

**Definition B.2.** Der **eindimensionale Torus** ist durch

$$\mathbb{T} = {}^{\mathbb{R}}/_{\mathbb{Z}} = \{x + \mathbb{Z} \mid x \in \mathbb{R}\}$$

definiert.

Formal ist der eindimensionale Torus also der Quotient von $\mathbb{R}$ bezüglich der Äquivalenzrelation $\sim$ gegeben durch $x \sim y \iff x - y \in \mathbb{Z}$ für alle $x, y \in \mathbb{R}$ (wieso?). Als Menge kann man $\mathbb{T}$ durch die Bijektion

$$x + \mathbb{Z} \in \mathbb{T} \mapsto \{x\} \in [0, 1)$$

mit dem Intervall $[0, 1)$ identifizieren.

Wie schon erwähnt wurde, gibt es eine eineindeutige Korrespondenz (das heisst, eine Bijektion)

$$\{\text{Funktionen } F : \mathbb{T} \to \mathbb{C}\} \quad \longleftrightarrow \quad \{\mathbb{Z}\text{-periodische Funktionen } f : \mathbb{R} \to \mathbb{C}\}\,.$$

Zum einen definiert eine Funktion $F : \mathbb{T} \to \mathbb{C}$ die $\mathbb{Z}$-periodische Funktion $f : x \in \mathbb{R} \mapsto F(x+\mathbb{Z}) \in \mathbb{C}$. Umgekehrt induziert eine $\mathbb{Z}$-periodische Funktion $f : \mathbb{R} \to \mathbb{C}$ eine wohldefinierte Funktion $F : \mathbb{T} \to \mathbb{C}$, indem man $F(x + \mathbb{Z}) = f(x)$ für $x + \mathbb{Z} \in \mathbb{T}$ setzt. Hierbei war $\mathbb{Z}$-Periodizität fundamental (wieso?). Interessierte können diese Diskussion mit der allgemeineren Diskussion nach Beispiel 1.65 vergleichen. Wir werden im Folgenden $\mathbb{Z}$-periodische Funktionen auf $\mathbb{R}$ und Funktionen auf $\mathbb{T}$ oft miteinander identifizieren.

$$\begin{tikzcd} \mathbb{R} \arrow[rr, "x \mapsto x+\mathbb{Z}"] \arrow[dr, "f"'] & & \mathbb{T} \arrow[dl, "F"] \\ & \mathbb{C} & \end{tikzcd}$$

Im Sinne der obigen Korrespondenz nennen wir eine Funktion $F : \mathbb{T} \to \mathbb{C}$ **stetig**, falls $f : x \in \mathbb{R} \mapsto F(x + \mathbb{Z}) \in \mathbb{C}$ stetig ist. Unter dieser Auffassung ist die Abbildung $x + \mathbb{Z} \in \mathbb{T} \mapsto$

$\{x\} \in [0,1)$ nicht stetig (wieso?). Somit ist $\mathbb{T}$ geometrisches ein anderes Objekt als $[0,1)$. In der Tat ist die Bijektion

$$x + \mathbb{Z} \in \mathbb{T} \mapsto \mathrm{e}^{2\pi \mathrm{i} x} \in \mathbb{S}^1$$

zwischen $\mathbb{T}$ und dem Einheitskreis $\mathbb{S}^1 \subseteq \mathbb{C}$ in beiden Richtungen stetig und in diesem Sinne ist $\mathbb{T}$ ein Kreis.

Wir möchten kurz eine weitere Methode erwähnen, mit der man stetige Funktionen auf dem Torus erhalten kann. Ist $f : [0,1] \to \mathbb{C}$ stetig mit $f(0) = f(1)$, dann definiert

$$\tilde{f} : x \in \mathbb{R} \mapsto f(x - \lfloor x \rfloor)$$

eine stetige Funktion auf $\mathbb{R}$ (wieso?) und es gilt $\tilde{f}(x) = f(x)$ für alle $x \in [0,1]$. Die Funktion $\tilde{f}$ nennt sich auch die $\mathbb{Z}$-**periodische Fortsetzung** von $f$ auf $\mathbb{R}$ und ist eindeutig bestimmt (wieso?).

**Hoffnung.** *$\mathbb{Z}$-periodische Funktionen lassen sich als Linearkombinationen (Superposition) der Funktionen $\mathrm{e}^{2\pi \mathrm{i} n x}$ für $n \in \mathbb{Z}$ oder alternativ $1$, $\cos(nx)$, $\sin(nx)$ für $n \in \mathbb{N}$ (der Grundschwingungen des Kreises) schreiben.*

$$f(x) = \sum_{n \in \mathbb{Z}} a_n \mathrm{e}^{2\pi \mathrm{i} n x} \tag{B.2}$$

*Die Zahl $a_n$ nennt sich der **Fourier-Koeffizient** oder die **Amplitude** zur **Frequenz** $n$. In Anwendungen sieht man häufiger die reellwertigen Funktionen $1$, $\cos(nx)$, $\sin(nx)$ für $n \in \mathbb{N}$. Mathematisch gesehen ist es jedoch einfacher, die komplexwertigen Funktionen $\mathrm{e}^{2\pi \mathrm{i} n x}$ für $n \in \mathbb{Z}$ zu verwenden.*

**Applet B.3** (Einige Fourier-Reihen)**.** *Wir sehen anhand einiger Beispiele wie endliche Fourier-Reihen die vorgegebenen Funktionen approximieren.*

Für welche Funktionen kann (B.2) stimmen? Das hängt davon ab, in welchem Sinn die Konvergenz und die Gleichheit verstanden wird. Wir wollen gleichmässige Konvergenz und Gleichheit für alle $x \in \mathbb{R}$ verwenden. Da die Partialsummen $\sum_{n=-N}^{N} a_n \mathrm{e}^{2\pi \mathrm{i} n x}$ stetig sind, muss auch der gleichmässige Limes $f$ stetig sein (wieso?). Somit können wir eine derartige Dartstellung einer $\mathbb{Z}$-periodischen Funktion nur für stetige Funktionen $f$ erhoffen.

Stetigkeit von $f$ ist aber nicht hinreichend, um gleichmässige Konvergenz in (B.2) zu erzielen. Stetige Differenzierbarkeit der Funktion $f$ ist hierfür hinreichend – dies werden wir hier nicht zeigen. Wir wollen stattdessen zweimalige stetige Differenzierbarkeit annehmen.

**Theorem B.4.** *Angenommen $f : \mathbb{R} \to \mathbb{C}$ ist eine $\mathbb{Z}$-periodische und zweimal stetig differenzierbare Funktion oder angenommen $f$ ist die $\mathbb{Z}$-periodische Fortsetzung von $x(1-x)$ für $x \in [0,1]$. Dann gilt*

$$f(x) = \sum_{n \in \mathbb{Z}} a_n \mathrm{e}^{2\pi \mathrm{i} n x},$$

*wobei die Reihe auf ganz $\mathbb{R}$ gleichmässig konvergiert und die Koeffizienten $a_n$ durch*

$$a_n = \int_0^1 f(x)\mathrm{e}^{-2\pi \mathrm{i} nx}\,\mathrm{d}x$$

*für $n \in \mathbb{Z}$ gegeben sind.*

**Definition B.5.** Für $n \in \mathbb{Z}$ und $f$ wie in Theorem B.4 ist

$$a_n = \int_0^1 f(x)\mathrm{e}^{-2\pi \mathrm{i} nx}\,\mathrm{d}x$$

der $n$-te **Fourier-Koeffizient** von $f$ und $\sum_{n\in\mathbb{Z}} a_n \mathrm{e}^{2\pi \mathrm{i} nx}$ ist die **Fourier-Reihe** von $f$.

Im Vergleich zu Taylor-Reihen in Abschnitt 8.5 ist Theorem B.4 ebenfalls sehr befriedigend – man nimmt zweimalige stetige Differenzierbarkeit an und schon hat man eine gleichmässig konvergente Darstellung der Funktion durch eine Fourier-Reihe. Für solche Funktionen sind die Taylor-Reihen im Allgemeinen nicht einmal definiert.

Wir beginnen den Beweis des Theorems damit, Fourierreihen selbst zu analysieren und diskutieren erst später die Frage der Darstellbarkeit aus Theorem B.4.

**Proposition B.6** (Eindeutigkeit der Fourier-Koeffizienten)**.** *Sei $f(x) = \sum_{n\in\mathbb{Z}} a_n \mathrm{e}^{2\pi \mathrm{i} nx}$ eine gleichmässig konvergente Fourier-Reihe. Dann ist $f$ eine $\mathbb{Z}$-periodische, stetige Funktion und*

$$a_n = \int_0^1 f(x)\mathrm{e}^{-2\pi \mathrm{i} nx}\,\mathrm{d}x.$$

*für alle $n \in \mathbb{Z}$.*

*Beweis.* Nach Annahme strebt $\sum_{n=-N}^{N} a_n \mathrm{e}^{2\pi \mathrm{i} nx}$ gleichmässig gegen $f(x)$. Nach Satz 6.48 ist $f$ stetig. Des Weiteren gilt für $x \in \mathbb{R}$ und $m \in \mathbb{N}$

$$\begin{aligned} f(x+m) &= \lim_{N\to\infty} \sum_{n=-N}^{N} a_n \mathrm{e}^{2\pi \mathrm{i} n(x+m)} \\ &= \lim_{N\to\infty} \sum_{n=-N}^{N} a_n \mathrm{e}^{2\pi \mathrm{i} nx} = f(x), \end{aligned}$$

da die Summe $\sum_{n=-N}^{N} a_n \mathrm{e}^{2\pi \mathrm{i} nx}$ $\mathbb{Z}$-periodisch ist. Somit ist $f$ eine $\mathbb{Z}$-periodische Funktion.

Sei $k \in \mathbb{Z}$. Dann gilt nach Satz 6.49

$$\begin{aligned} \int_0^1 f(x)\mathrm{e}^{-2\pi \mathrm{i} kx}\,\mathrm{d}x &= \int_0^1 \lim_{N\to\infty} \left( \sum_{n=-N}^{N} a_n \mathrm{e}^{2\pi \mathrm{i} nx} \right) \mathrm{e}^{-2\pi \mathrm{i} kx}\,\mathrm{d}x \\ &= \lim_{N\to\infty} \sum_{n=-N}^{N} a_n \int_0^1 \mathrm{e}^{2\pi \mathrm{i}(n-k)x}\,\mathrm{d}x. \end{aligned}$$

Nun berechnet man für $n \neq k$

$$\int_0^1 \mathrm{e}^{2\pi \mathrm{i}(n-k)x}\,\mathrm{d}x = \int_0^1 \cos(2\pi(n-k)x)\,\mathrm{d}x + \mathrm{i}\int_0^1 \sin(2\pi(n-k)x)\,\mathrm{d}x = 0.$$

Somit ist für $N \in \mathbb{N}$ mit $|k| \leq N$

$$\sum_{n=-N}^{N} a_n \int_0^1 \mathrm{e}^{2\pi \mathrm{i}(n-k)x}\,\mathrm{d}x = a_k$$

und daher folgt

$$a_k = \int_0^1 f(x)\mathrm{e}^{-2\pi \mathrm{i}kx}\,\mathrm{d}x,$$

was zu beweisen war. □

**Proposition B.7.** *Falls $f : \mathbb{R} \to \mathbb{C}$ eine $\mathbb{Z}$-periodische, zweimal differenzierbare Funktion ist, dann ist*

$$a_n = \int_0^1 f(x)\mathrm{e}^{-2\pi \mathrm{i}nx}\,\mathrm{d}x = O\left(\tfrac{1}{n^2}\right)$$

*für $n \to \infty$ und $n \to -\infty$. Insbesondere ist $\sum_{n\in\mathbb{Z}} a_n \mathrm{e}^{2\pi \mathrm{i}nx}$ eine gleichmässig konvergente Reihe.*

*Beweis.* Für $n \neq 0$ gilt nach partieller Integration

$$\begin{aligned}
a_n &= \int_0^1 f(x)\mathrm{e}^{-2\pi \mathrm{i}nx}\,\mathrm{d}x \\
&= \underbrace{\left[\frac{f(x)\mathrm{e}^{-2\pi \mathrm{i}nx}}{-2\pi \mathrm{i}n}\right]_0^1}_{=0} - \tfrac{1}{-2\pi \mathrm{i}n}\int_0^1 f'(x)\mathrm{e}^{-2\pi \mathrm{i}nx}\,\mathrm{d}x \\
&= -\underbrace{\left[\frac{f'(x)\mathrm{e}^{-2\pi \mathrm{i}nx}}{(-2\pi \mathrm{i}n)^2}\right]_0^1}_{=0} + \tfrac{1}{(-2\pi \mathrm{i}n)^2}\int_0^1 f''(x)\mathrm{e}^{-2\pi \mathrm{i}nx}\,\mathrm{d}x \\
&= -\tfrac{1}{4\pi^2 n^2}\int_0^1 f''(x)\mathrm{e}^{-2\pi \mathrm{i}nx}\,\mathrm{d}x,
\end{aligned}$$

wobei wir $\mathbb{Z}$-Periodizität von $f$ und $f'$ verwendet haben. Nun gilt aber

$$\left|\int_0^1 f''(x)\mathrm{e}^{-2\pi \mathrm{i}nx}\,\mathrm{d}x\right| \leq \int_0^1 |f''(x)|\,\mathrm{d}x \leq \max_{x\in[0,1]} |f''(x)|,$$

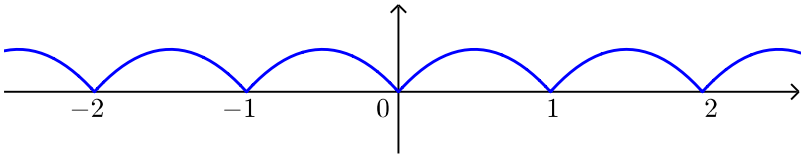
womit $|a_n| \leq C\frac{1}{n^2}$ für eine positive Konstante $C$ und also $a_n = O\left(\frac{1}{n^2}\right)$.

Insbesondere ist für $N \in \mathbb{N}$

$$\left|\sum_{n\in\mathbb{Z}} a_n \mathrm{e}^{2\pi \mathrm{i} n x} - \sum_{n=-N}^{N} a_n \mathrm{e}^{2\pi \mathrm{i} n x}\right| = \left|\sum_{|n|>N} a_n \mathrm{e}^{2\pi \mathrm{i} n x}\right| \leq C \sum_{|n|\geq N} \frac{1}{n^2}$$

für eine positive Konstante $C$ und die Reihe $\sum_{n\in\mathbb{Z}} a_n \mathrm{e}^{2\pi \mathrm{i} n x}$ konvergiert gleichmässig. $\square$

Sei nun $f : \mathbb{R} \to \mathbb{C}$ die periodische Fortsetzung des Polynoms $x(1-x)$ auf $[0,1]$. Konkret ist diese gegeben durch $f(x) = \{x\}\,(1-\{x\})$ für alle $x \in \mathbb{R}$ (wieso?).




Aus obigem Bild ist ersichtlich, dass $f$ nicht zweimal differenzierbar ist, wobei Probleme nur an den Randpunkten erscheinen. Insbesondere kommt Proposition B.7 nicht zur Anwendung. Wir berechnen deswegen die Fourier-Koeffizienten von $f$. Für $n = 0$ ist

$$\begin{aligned} a_0 = \int_0^1 f(x)\,\mathrm{d}x &= \int_0^1 x(1-x)\,\mathrm{d}x \\ &= \left(\frac{x^2}{2} - \frac{x^3}{3}\right)\Big]_0^1 = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}. \end{aligned}$$

Für $n \neq 0$ gilt

$$\begin{aligned} a_n &= \int_0^1 f(x)\mathrm{e}^{-2\pi \mathrm{i} n x}\,\mathrm{d}x = \int_0^1 (x-x^2)\mathrm{e}^{-2\pi \mathrm{i} n x}\,\mathrm{d}x \\ &= \underbrace{\left[\tfrac{1}{-2\pi \mathrm{i} n}(x-x^2)\mathrm{e}^{-2\pi \mathrm{i} n x}\right]_0^1}_{=0} - \tfrac{1}{-2\pi \mathrm{i} n}\int_0^1 (1-2x)\mathrm{e}^{-2\pi \mathrm{i} n x}\,\mathrm{d}x \\ &= -\left[\tfrac{1}{(-2\pi \mathrm{i} n)^2}(1-2x)\mathrm{e}^{-2\pi \mathrm{i} n x}\right]_0^1 + \tfrac{1}{(-2\pi \mathrm{i} n)^2}\underbrace{\int_0^1 (-2)\mathrm{e}^{-2\pi \mathrm{i} n x}\,\mathrm{d}x}_{=0} \\ &= \frac{1}{4\pi^2 n^2}(-1-1) = -\frac{1}{2\pi^2 n^2}. \end{aligned}$$

Insbesondere gilt $a_n = O(\frac{1}{n^2})$ für $n \to \infty$ und $n \to -\infty$ und die Fourier-Reihe von $f$ konvergiert ebenso gleichmässig wie die Fourier-Reihe einer $\mathbb{Z}$-periodischen, zweimal stetig differenzierbaren Funktion. Nach Proposition B.6 ist $\frac{1}{6} - \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{\mathrm{e}^{2\pi \mathrm{i} n x}}{2\pi^2 n^2}$ der einzig mögliche Kandidat für eine Fourierreihe von $f$.

Um Gleichheit zu zeigen, müssen wir zuerst die Faltung definieren.

### Faltung auf dem Torus

Für eine stetige oder „Riemann-integrierbare“ Funktion $f : \mathbb{T} \to \mathbb{C}$ definieren wir

$$\int_{\mathbb{T}} f(x)\,\mathrm{d}x = \int_0^1 f(x)\,\mathrm{d}x,$$

wobei wir $f$ im zweiten Integral als $\mathbb{Z}$-periodische Funktion auf $\mathbb{R}$ auffassen.

**Lemma B.8** (Rotations- und Spiegelungsinvarianz des Integrals auf dem Torus)**.** *Sei $f$ wie oben. Dann gilt für alle $y \in \mathbb{R}$*

$$\int_{\mathbb{T}} f(x+y)\,\mathrm{d}x = \int_{\mathbb{T}} f(x)\,\mathrm{d}x.$$

*Des Weiteren ist*

$$\int_{\mathbb{T}} f(-x)\,\mathrm{d}x = \int_{\mathbb{T}} f(x)\,\mathrm{d}x.$$

Dabei ist $x+y$ für $x = \tilde{x} + \mathbb{Z} \in \mathbb{T}$ und $y \in \mathbb{R}$ durch $(\tilde{x}+y)+\mathbb{Z}$ gegeben. Analog ist $-x$ für $x = \tilde{x} + \mathbb{Z} \in \mathbb{T}$ durch $(-\tilde{x})+\mathbb{Z}$ definiert.

Anschaulich lässt sich obiges Lemma wie folgt erklären. Unter der Auffassung von $\mathbb{T}$ als Einheitskreis $\mathbb{S}^1$ wird die Addition mit $y \in \mathbb{R}$ zu einer Rotation des Kreises mit Winkel $2\pi y$ und die Abbildung $x \in \mathbb{T} \mapsto -x \in \mathbb{T}$ zu einer Spiegelung (wieso?). Nun integriert man $f$ über den Kreis und über die rotierte (respektive gespiegelte) Version des Kreises und erhält beide Male dasselbe.

*Beweis.* Es gilt

$$\begin{aligned}\int_{\mathbb{T}} f(x+y)\,\mathrm{d}x &= \int_0^1 f(x+y)\,\mathrm{d}x = \int_y^{y+1} f(t)\,\mathrm{d}t \\ &= \int_y^n f(t)\,\mathrm{d}t + \int_n^{y+1} f(t)\,\mathrm{d}t,\end{aligned}$$

wobei wir $t = x+y$ substituiert haben und $n \in \mathbb{Z}$ mit $y \le n < y+1$ gewählt haben. Nun substituieren wir nochmals $s = t-(n-1)$ im ersten obigen Integral und $s = t-n$ im zweiten. Damit erhalten wir aus der $\mathbb{Z}$-Periodizität von $f$

$$\begin{aligned}\int_{\mathbb{T}} f(x+y)\,\mathrm{d}x &= \int_{y-(n-1)}^1 f(t)\,\mathrm{d}t + \int_0^{y+1-n} f(t)\,\mathrm{d}t \\ &= \int_{y+1-n}^1 f(t)\,\mathrm{d}t + \int_0^{y+1-n} f(t)\,\mathrm{d}t \\ &= \int_0^1 f(t)\,\mathrm{d}t = \int_{\mathbb{T}} f(x)\,\mathrm{d}x\end{aligned}$$

wie gewünscht. Für die zweite Eigenschaft berechnet man

$$\begin{aligned}\int_{\mathbb{T}} f(x)\,\mathrm{d}x &= \int_0^1 f(x)\,\mathrm{d}x = \int_0^1 f(1-t)\,\mathrm{d}t \\ &= \int_0^1 f(-t)\,\mathrm{d}t = \int_{\mathbb{T}} f(-t)\,\mathrm{d}t\end{aligned}$$

durch die Substitution $x = 1 - t$ und $\mathbb{Z}$-Periodizität von $f$. □

**Definition B.9.** Seien $f, g : \mathbb{R} \to \mathbb{C}$ zwei $\mathbb{Z}$-periodische und stetige Funktionen. Dann definieren wir die Faltung $f * g : \mathbb{R} \to \mathbb{C}$ von $f$ und $g$ durch

$$f * g(x) = \int_{\mathbb{T}} f(t)g(x-t)\,\mathrm{d}t.$$

für $x \in \mathbb{R}$.

Man beachte dabei, dass $f * g$ wieder $\mathbb{Z}$-periodisch ist und somit die Faltung zweier stetigen Funktionen auf dem Torus definiert wurde. Des Weiteren ist $f * g = g * f$ für alle $f, g$ wie oben, das heisst, die Faltung ist kommutativ. In der Tat ist nach Anwendung von Lemma B.8 und der Substitution $s = x - t$

$$\begin{aligned}f * g(x) &= \int_{\mathbb{T}} f(t)g(x-t)\,\mathrm{d}t \\ &= \int_{\mathbb{T}} f(x-s)g(s)\,\mathrm{d}s = g * f(x).\end{aligned}$$

**Übung B.10.** *Zeigen Sie, dass der Ausdruck $f * g$ bilinear (das heisst, linear in $f$ und linear in $g$) ist für $f, g$ wie oben.*

Der Zusammenhang der Faltung zu Fourier-Reihen wird im nächsten Beispiel klarer.

**Beispiel B.11.** *Wir betrachten $g = e_n : x \in \mathbb{R} \mapsto \mathrm{e}^{2\pi \mathrm{i} n x}$ für $n \in \mathbb{Z}$ und $f$ beliebig. Dann ist*

$$(f * e_n)(x) = \int_{\mathbb{T}} f(t)\mathrm{e}^{2\pi \mathrm{i} n(x-t)}\,\mathrm{d}t = \mathrm{e}^{2\pi \mathrm{i} n x}\left(\int_{\mathbb{T}} f(t)\mathrm{e}^{-2\pi \mathrm{i} n t}\,\mathrm{d}t\right) = \mathrm{e}^{2\pi \mathrm{i} n x} a_n$$

*der $n$-te Term der Fourier-Reihe von $f$, wobei $a_n$ der $n$-te Fourier-Koeffizient von $f$ ist.*

Sei $f : \mathbb{T} \to \mathbb{C}$ stetig (oder Riemann-integrierbar). Nach obigem Beispiel und Bilinearität der Faltung gilt für $N \in \mathbb{N}$

$$\left(f * \left(\sum_{n=-N}^{N} e_n\right)\right)(x) = \sum_{n=-N}^{N} a_n \mathrm{e}^{2\pi \mathrm{i} n x}$$

für alle $x \in \mathbb{R}$.

**Definition B.12.** Zu $N \in \mathbb{N}_0$ ist die Abbildung $D_N : \mathbb{T} \to \mathbb{C}$ definiert durch

$$D_N(x) = \sum_{n=-N}^{N} \mathrm{e}^{2\pi \mathrm{i} n x} = \sum_{n=-N}^{N} e_n(x)$$

für alle $x \in \mathbb{T}$ der $N$-te **Dirichlet-Kern**.

Nach obigem gilt, dass das Faltungsprodukt $f * D_N$ genau die $N$-te Partialsumme der Fourier-Reihe von $f$.

**Lemma B.13** (Eine Charakterisierung des Dirichlet-Kerns)**.** *Für $N \in \mathbb{N}_0$ und alle $x \in \mathbb{R}$ ist*

$$D_N(x) = \begin{cases} \frac{\mathrm{e}^{2\pi \mathrm{i}(N+1)x} - \mathrm{e}^{-2\pi \mathrm{i} N x}}{\mathrm{e}^{2\pi \mathrm{i} x} - 1} & \textit{falls } x \notin \mathbb{Z} \\ 2N+1 & \textit{falls } x \in \mathbb{Z} \end{cases}.$$

*Beweis.* Für $x \in \mathbb{Z}$ ist $e_n(x) = 1$ für alle $n$ und somit $D_N(x) = 2N+1$. Wir berechnen für $x \notin \mathbb{Z}$

$$D_N(x) = \sum_{n=-N}^{N} \mathrm{e}^{2\pi \mathrm{i} n x} = \mathrm{e}^{-2\pi \mathrm{i} N x} \sum_{m=-0}^{2N} \left(\mathrm{e}^{2\pi \mathrm{i} x}\right)^m = \mathrm{e}^{-2\pi \mathrm{i} N x} \frac{\mathrm{e}^{2\pi \mathrm{i}(2N+1)x} - 1}{\mathrm{e}^{2\pi \mathrm{i} x} - 1},$$

woraus sich die Aussage im Lemma ergibt. □

Wir wollen eigentlich zeigen, dass $f * D_N(x)$ gegen $f(x)$ strebt für $N \to \infty$ und für $x \in \mathbb{T}$. Wir schlagen aber einen weiteren Umweg ein.

**Definition B.14.** Sei $N \in \mathbb{N}$. Der $N$-te Fejér-Kern $F_N : \mathbb{T} \to \mathbb{C}$ ist gegen durch

$$F_N(x) = \frac{1}{N+1} \sum_{n=0}^{N} D_n(x)$$

für alle $x \in \mathbb{T}$.

Der $N$-te Fejér-Kern ist also das Mittel der ersten $N+1$ Dirichlet-Kerne. Wir werden zeigen, dass $F_N$ „gute“ Eigenschaften hat. (Diese Eigenschaften sind der Grund warum wir mit $F_N$ anstatt mit $D_N$ arbeiten wollen.)

**Proposition B.15** (Eine Charakterisierung des Fejér-Kerns)**.** *Für alle $N \in \mathbb{N}$ und $x \in \mathbb{R}$ gilt*

$$F_N(x) = \begin{cases} \frac{1}{N+1} \frac{\sin^2\left((N+1)\pi x\right)}{\sin^2(\pi x)} & \textit{falls } x \notin \mathbb{Z} \\ N+1 & \textit{falls } x \in \mathbb{Z} \end{cases}$$

*Beweis.* Für $x \in \mathbb{Z}$ ist nach Lemma B.13

$$\begin{aligned} F_N(x) &= \frac{1}{N+1} \sum_{n=0}^{N} (2n+1) = \frac{1}{N+1} \left( (N+1) + 2 \sum_{n=0}^{N} n \right) \\ &= \frac{1}{N+1} \big( (N+1) + N(N+1) \big) = N+1. \end{aligned}$$

wie behauptet. Sei also $x \notin \mathbb{Z}$. Dann gilt nach Lemma B.13

$$\begin{aligned}
F_N(x) &= \frac{1}{N+1} \sum_{n=0}^{N} \frac{\mathrm{e}^{2\pi \mathrm{i}(n+1)x} - \mathrm{e}^{-2\pi \mathrm{i} n x}}{\mathrm{e}^{2\pi \mathrm{i} x} - 1} \\
&= \frac{1}{(N+1)\left(\mathrm{e}^{2\pi \mathrm{i} x} - 1\right)} \left( \mathrm{e}^{2\pi \mathrm{i} x} \sum_{n=0}^{N} \mathrm{e}^{2\pi \mathrm{i} n x} - \sum_{n=0}^{N} \mathrm{e}^{-2\pi \mathrm{i} n x} \right) \\
&= \frac{1}{(N+1)\left(\mathrm{e}^{2\pi \mathrm{i} x} - 1\right)} \left( \mathrm{e}^{2\pi \mathrm{i} x} \frac{\mathrm{e}^{2\pi \mathrm{i}(N+1)x} - 1}{\mathrm{e}^{2\pi \mathrm{i} x} - 1} - \frac{-\mathrm{e}^{2\pi \mathrm{i}(N+1)x} - 1}{\mathrm{e}^{-2\pi \mathrm{i} x} - 1} \right) \\
&= \frac{1}{(N+1)\left(\mathrm{e}^{2\pi \mathrm{i} x} - 1\right)} \left( \mathrm{e}^{2\pi \mathrm{i} x} \frac{\mathrm{e}^{2\pi \mathrm{i}(N+1)x} - 1}{\mathrm{e}^{2\pi \mathrm{i} x} - 1} - \mathrm{e}^{2\pi \mathrm{i} x} \frac{-\mathrm{e}^{2\pi \mathrm{i}(N+1)x} - 1}{1 - \mathrm{e}^{2\pi \mathrm{i} x}} \right) \\
&= \frac{\mathrm{e}^{2\pi \mathrm{i} x}}{(N+1)\left(\mathrm{e}^{2\pi \mathrm{i} x} - 1\right)^2} \left( \mathrm{e}^{2\pi \mathrm{i}(N+1)x} - 2 + \mathrm{e}^{-2\pi \mathrm{i}(N+1)x} \right) \\
&= \frac{1}{N+1} \frac{\left(\mathrm{e}^{\pi \mathrm{i}(N+1)x} - \mathrm{e}^{-\pi \mathrm{i}(N+1)x}\right)^2}{\left(\mathrm{e}^{\pi \mathrm{i} x} - \mathrm{e}^{-\pi \mathrm{i} x}\right)^2} = \frac{1}{N+1} \frac{\sin^2\left((N+1)\pi x\right)}{\sin^2(\pi x)},
\end{aligned}$$

was zu zeigen war. □

**Übung B.16.** *Argumentieren Sie, wieso die Dirichlet-Kerne und die Fejér-Kerne stetig sind. Zeigen Sie damit in Lemma B.13 und Proposition B.15 jeweils den Fall $x \in \mathbb{Z}$ unter Verwendung der Regel von de l'Hôpital (siehe Satz 7.49).*

Mit Proposition B.15 lässt sich insbesondere das asymptotische Verhalten der Fejér-Kerne besser charakterisieren.

**Korollar B.17.** *Die Fejér-Kerne erfüllen folgende Eigenschaften.*

- *Zu $N \in \mathbb{N}$ ist $F_N(x) \geq 0$ für alle $x \in \mathbb{R}$.*
- *Es gilt $F_N(x) \to 0$ für $N \to \infty$ und für alle $x \in \mathbb{R} \setminus \mathbb{Z}$.*
- *Es gilt $F_N(x) \to \infty$ für $N \to \infty$ und für alle $x \in \mathbb{Z}$.*
- *Für alle $\delta \in \left(0, \frac{1}{2}\right)$ konvergiert die Funktionenfolge $F_N$ gleichmässig gegen Null auf $[\delta, 1-\delta]$ für $N \to \infty$.*
- *Für alle $N \in \mathbb{N}$ gilt*

$$\int_0^1 F_N(x)\,\mathrm{d}x = 1.$$

- *Für alle $\delta \in \left(0, \frac{1}{2}\right)$ gilt*

$$\int_\delta^{1-\delta} F_N(x)\,\mathrm{d}x \to 0$$

*für $N \to \infty$.*

*Beweis.* Die ersten drei Eigenschaften folgen unmittelbar aus Proposition B.15. Sei nun ein $\delta \in \left(0, \frac{1}{2}\right)$ gegeben. Dann ist $\alpha = \min_{x\in[\delta,1-\delta]} \sin^2(\pi x) > 0$ und somit für alle $x \in [\delta, 1-\delta]$

$$|F_N(x)| = \frac{1}{N+1} \frac{\sin^2\left((N+1)\pi x\right)}{\sin^2(\pi x)} \leq \frac{1}{N+1}\alpha^{-1} \to 0,$$

woraus die vierte Eigenschaft folgt. Diese impliziert auch die sechste Eigenschaft nach Satz 6.49 über die Vertauschbarkeit eines gleichmässigen Grenzwerts und des Riemann-Integrals. Nach Definition der Dirichlet-Kerne gilt

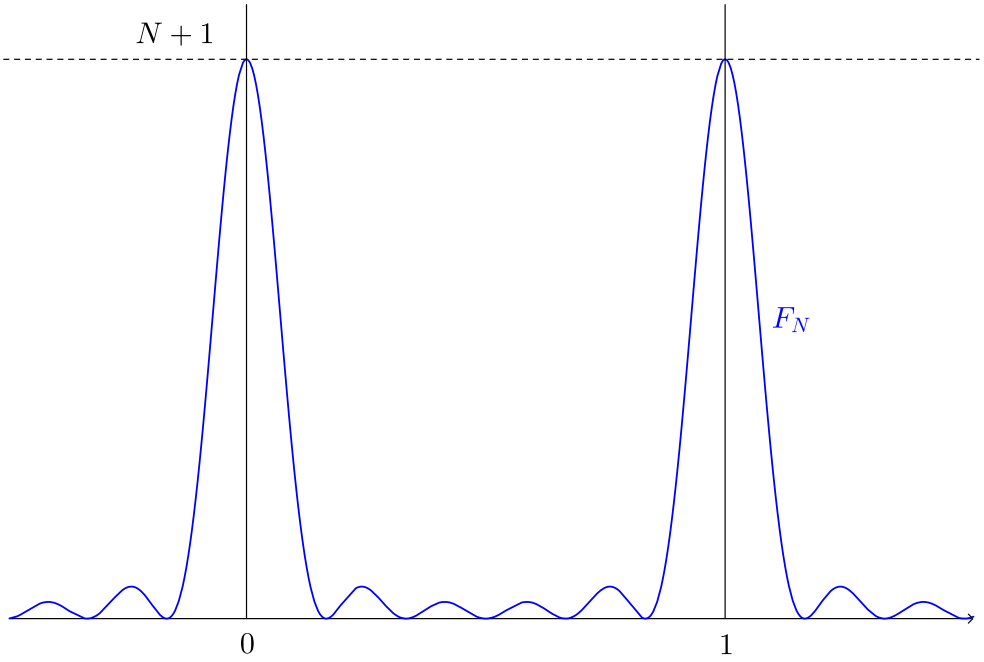
$$\int_{\mathbb{T}} D_n(x)\,\mathrm{d}x = \sum_{k=0}^{n} \int_{\mathbb{T}} \mathrm{e}^{2\pi \mathrm{i} kx}\,\mathrm{d}x = 1.$$

für alle $n \in \mathbb{N}_0$. Damit folgt auch

$$\int_{\mathbb{T}} F_N(x) = \frac{1}{N+1} \sum_{n=0}^{N} \int_{\mathbb{T}} D_n(x)\,\mathrm{d}x = 1$$

für $N \in \mathbb{N}$, was die fünfte Eigenschaft beweist. □

Wir fassen die Erkenntnisse über die Fejér-Kerne aus Proposition B.15 und Korollar B.15 in folgendem Bild zusammen.




Figur B.1: Zwischen den beiden Spitzen von $F_N$ bei 0 und 1 wird $F_N$ immer kleiner für $N \to \infty$ und der Beitrag zu $\int_{\mathbb{T}} F_N(x)\,\mathrm{d}x = 1$ erfolgt vornehmlich um $0 + \mathbb{Z}$.

Der wichtigste Schritt für den Beweis von Theorem B.4 ist der folgende.

**Proposition B.18.** *Sei $f : \mathbb{R} \to \mathbb{C}$ eine $\mathbb{Z}$-periodische, stetige Funktion. Dann konvergiert die Funktionenfolge $f * F_N$ auf $\mathbb{R}$ gleichmässig gegen $f$ für $N \to \infty$.*

*Beweis.* Für $N \in \mathbb{N}$ und $x \in \mathbb{T}$ gilt

$$\begin{aligned} f * F_N(x) - f(x) &= F_N * f(x) - f(x) \\ &= F_N * f(x) - f(x) \int_{\mathbb{T}} F_N(t)\,\mathrm{d}t \\ &= \int_{\mathbb{T}} F_N(t) f(x-t)\,\mathrm{d}t - \int_{\mathbb{T}} F_N(t) f(x)\,\mathrm{d}t \\ &= \int_{\mathbb{T}} F_N(t)\big(f(x-t) - f(x)\big)\,\mathrm{d}t \end{aligned}$$

unter Verwendung der Kommutativität der Faltung, der Gleichheit $\int_{\mathbb{T}} F_N(t)\,\mathrm{d}t = 1$ und der Definition der Faltung. Per Definition des Integrals über dem Torus und wegen der $\mathbb{Z}$-Periodizität von $F_N$ und $f$ können wir schreiben

$$\begin{aligned} f * F_N(x) - f(x) &= \int_{\mathbb{T}} F_N(t)(f(x-t) - f(x))\,\mathrm{d}t \\ &= \int_{-\frac{1}{2}}^{\frac{1}{2}} F_N(t)\big(f(x-t) - f(x)\big)\,\mathrm{d}t, \end{aligned}$$

was bezweckt, dass sich $F_N$ nur um einen Punkt konzentriert (im Sinne von Bild B.1). Sei nun $\varepsilon > 0$. Dann gibt es wegen der gleichmässigen Stetigkeit von $f$ (Satz 3.78) ein $\delta > 0$ (unabhängig von $x$), so dass für alle $x, y \in [-3, 3]$ mit $|x - y| < \delta$ gilt $|f(y) - f(x)| < \varepsilon$. Wir teilen obiges Integral auf und erhalten

$$\begin{aligned} f * F_N(x) - f(x) = &\int_{-\frac{1}{2}}^{-\delta} F_N(t)\big(f(x-t) - f(x)\big)\,\mathrm{d}t \\ &+ \int_{-\delta}^{\delta} F_N(t)\big(f(x-t) - f(x)\big)\,\mathrm{d}t \\ &+ \int_{\delta}^{\frac{1}{2}} F_N(t)\big(f(x-t) - f(x)\big)\,\mathrm{d}t \end{aligned}$$

Nach der Dreiecksungleichung in $\mathbb{R}$ und derjenigen für Riemann-Integrale ist

$$\begin{aligned} |f * F_N(x) - f(x)| \leq &\int_{-\frac{1}{2}}^{-\delta} F_N(t)\big|f(x-t) - f(x)\big|\,\mathrm{d}t \\ &+ \int_{-\delta}^{\delta} F_N(t)\big|f(x-t) - f(x)\big|\,\mathrm{d}t \\ &+ \int_{\delta}^{\frac{1}{2}} F_N(t)\big|f(x-t) - f(x)\big|\,\mathrm{d}t \\ < &\, 2 \max_{x \in \mathbb{R}} |f(x)| \left( \int_{-\frac{1}{2}}^{-\delta} F_N(t)\,\mathrm{d}t + \int_{\delta}^{\frac{1}{2}} F_N(t)\,\mathrm{d}t \right) \\ &+ \varepsilon \int_{-\delta}^{\delta} F_N(t)\,\mathrm{d}t \end{aligned}$$

Nach Korollar B.17 kann man nun $N$ gross genug wählen, so dass der erste Ausdruck oben (inklusive Vorfaktor) kleiner als $\varepsilon$ ist. Da $\int_{-\delta}^{\delta} F_N(t)\,\mathrm{d}t \leq 1$ ist, folgt damit, dass

$$|f * F_N(x) - f(x)| < 2\varepsilon$$

für alle $x \in \mathbb{T}$ und für alle hinreichend grossen $N$. □

**Applet B.19** (Faltung von periodischen Funktionen)**.** *Das Faltungsprodukt mit dem Fejér-Kern mittelt die ursprüngliche Funktion in der Nähe eines variablen Punktes. Wir stellen den Fejér-Kern etwas gestaucht dar, da dieser auf Grund der Konzentration der fast kompletten Gesamtmasse in immer kleineren Teilintervallen ansonsten weit über das Fenster ragen würde.*

*Beweis von Theorem B.4.* Sei $f$ zweimal stetig differenzierbar oder die periodische Fortsetzung des Polynoms $x(1-x)$ auf $[0,1]$. Dann haben wir schon gesehen, dass die Fourier-Reihe von $f$ gleichmässig konvergiert – weil $a_n = O\left(\frac{1}{n^2}\right)$ für $n \to \infty$ und für $n \to -\infty$. Wir müssen noch zeigen, dass der Wert der Fourierreihe von $f$ gleich dem Wert von $f$ ist bei jedem Punkt. Sei also

$$g(x) = \sum_{n\in\mathbb{Z}} a_n \mathrm{e}^{2\pi \mathrm{i} n x}$$

der Wert der Fourierreihe von $f$ bei einem Punkt $x$. Dann gilt also

$$f * D_n(x) \to g(x)$$

für $n \to \infty$ aber

$$f * F_N(x) \to f(x)$$

für $N \to \infty$ nach Proposition B.18. Nach Definition ist aber

$$f * F_N(x) = \frac{1}{N+1} \sum_{n=0}^{N} f * D_n(x)$$

das Cauchy-Mittel der Folge $f * D_n(x)$, welche gegen $g(x)$ konvergiert. Deswegen konvergiert $f * F_N(x)$ gegen $g(x)$ und wir schliessen $g(x) = f(x)$ für alle $x \in \mathbb{T}$ aus der Eindeutigkeit des Grenzwerts. □

**Übung B.20.** *Berechnen Sie den Wert der Reihe* $\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2}$ *und* $\sum_{n=1}^{\infty} \frac{1}{(2n+1)^2}$.

**Übung B.21** (Dreiecksfunktion)**.** *Wir betrachten die periodische Fortsetzung* $f : \mathbb{R} \to \mathbb{R}$ *der Funktion*

$$[0,1] \to \mathbb{R},\ x \mapsto \begin{cases} x & \textit{falls } x \leq \frac{1}{2} \\ 1-x & \textit{falls } x > \frac{1}{2} \end{cases}.$$

*Berechnen Sie die Fourierkoeffizienten $(a_n)_{n\in\mathbb{Z}}$ von $f$ und zeigen Sie, dass $a_n = O(\frac{1}{n^2})$ für $n \to \infty$ und $n \to -\infty$. Verwenden Sie dies dann, um zu zeigen, dass die Fourierreihe von $f$ gleichmässig gegen $f$ konvergiert.*

## Bezug zu weiteren Themen

Die Ideen der Fourier-Theorie auf dem Torus haben viele und weitreichende Verallgemeinerungen auf beispielsweise mehrdimensionalen Tori $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ und führen weiter auch noch zur Theorie der Fourier-Transformation auf $\mathbb{R}^d$.

Des Weiteren wird der Skalar

$$\langle f(x), \mathrm{e}^{2\pi \mathrm{i} nx}\rangle = \int_0^1 f(x)\mathrm{e}^{-2\pi \mathrm{i} nx}\,\mathrm{d}x$$

das innere Produkt der Funktionen $f(x)$ und $\mathrm{e}^{2\pi \mathrm{i} nx}$ genannt. Dieses ist für zwei stetige Funktionen $f, g$ auf $[0, 1]$ durch

$$\langle f, g\rangle = \int_0^1 f(x)\overline{g(x)}\,\mathrm{d}x$$

definiert. Innere Produkte werden sowohl in der Linearen Algebra 2, in der Analysis 2, aber auch in weiterführenden Vorlesungen bei Betrachtung von Fourier-Reihen, Hilberträumen, Riemannschen Metriken und anderen Themen eine zentrale Rolle spielen (siehe zum Beispiel die Vorlesungen „Methoden der mathematischen Physik“ im 2. Jahr, „Funktionalanalysis“ und „Differentialgeometrie“ ab dem 3. Jahr).

Das innere Produkt erlaubt uns folgende einfache geometrische Interpretation der Fourier-Reihen. Die Funktionen $e_n(x) = \mathrm{e}^{2\pi \mathrm{i} nx}$ haben „Länge“ 1 für jedes $n \in \mathbb{Z}$ und der Ausdruck $a_n e_n$ mit $a_n = \langle f, e_n\rangle$, der in der Fourier-Reihe für jedes $n \in \mathbb{Z}$ auftritt, ist die „orthogonale Projektion“ von $f$ auf den eindimensionalen Teilraum $\mathbb{C}e_n$.

Wir möchten des Weiteren erwähnen, dass die von uns gewünschte gleichmässige Konvergenz der Fourier-Reihe nicht der beste Konvergenzbegriff für Fourier-Reihen ist. Wenn man stattdessen den Begriff der Konvergenz im quadratischen Mittel verwendet, dann besitzt jede Riemann-integrierbare Funktion auf dem Torus (und noch viele weitere) eine Fourier-Reihe. Um dies in einem Special erklären zu können, müssten aber Weihnachten und Ostern zusammenfallen.

# Anhang C

# Summer-Special 2018:

## Dynamische Systeme auf Teilmannigfaltigkeiten*

Wir nehmen im Folgenden immer an, dass $M = \{x \in U \mid F(x) = 0\}$ eine Teilmannigfaltigkeit innerhalb einer offenen Menge $U \subseteq \mathbb{R}^n$ ist, welche wie im Theorem über den konstanten Rang (Theorem 11.16) als Niveaumenge einer glatten Funktion $F : U \to \mathbb{R}$ (für einen regulären Wert – hier $0 \in \mathbb{R}^m$) gegeben ist. Um den Anschauungsfaktor und damit auch den Spassfaktor zu maximimieren, setzen wir $m = 1$, $n = 3$ und meist sogar $U = \mathbb{R}^3$ fest.

Wir wollen natürliche Differentialgleichungen auf der Teilmannigfaltigkeit $M$ und deren Tangentialbündel $\mathrm{T}M$ entwerfen und die daraus resultierenden dynamische Systeme (abstrakt) untersuchen.

### C.1 Vektorfelder und Flüsse auf einer Teilmannigfaltigkeit

Wir wollen nun annehmen, dass ein Vektorfeld $f : U \to \mathbb{R}^3$ gegeben ist. Dieses Vektorfeld soll die Bewegungsrichtung und Bewegungsgeschwindigkeit eines Punktes angeben. Doch wird für einen vorgegebenen Punkt $p \in M$ weiter noch verlangt, dass der gesuchte Lösungsweg innerhalb von $M$ bleibt. Wie können wir das Anfangswertswertproblem

$$\begin{aligned} \dot{p} &= f(p) \\ p(0) &= p_0 \end{aligned} \tag{C.1}$$

so modifizieren, dass der Lösungsweg $t \mapsto p(t)$ in $M$ bleibt und abgesehen davon „so gut wie möglich“ noch dem Vektorfeld $f$ folgt?

Die Idee ist hier, einfach $f(p)$ für $p \in M$ durch die Projektion von $f(p)$ auf $\mathrm{T}_pM$ zu ersetzen. Sei also für $p \in U$ und $v \in \mathbb{R}^3$ die orthogonale Projektion $P_p$ durch

$$v \in \mathbb{R}^3 \mapsto P_p(v) = v - \frac{1}{\|\nabla F(p)\|^2} \langle v, \nabla F(p) \rangle \nabla F(p)$$

definiert. Auf Grund der Wahl von $F$ (0 war ein regulärer Wert auf $M$) ist dabei $\|\nabla F(p)\| \neq 0$ für alle $p \in M$, womit wir ohne Beschränkung der Allgemeinheit $\nabla F(p) \neq 0$ für alle $p \in U$ annehmen können. Die Abbildung $P_p$ bildet auf das orthogonale Komplement von $\nabla F(p)$ ab, da

$\langle P_p(v), \nabla F(p)\rangle = 0$ für alle $p \in M$ und $v \in \mathbb{R}^3$ gilt. Insbesondere ist $P_p$ für $p \in M$ die orthogonale Projektion auf $\mathrm{T}_pM$ nach Satz 11.21. Wir wollen also anstelle des Anfangswertproblems in (C.1) das Anfangswertproblem

$$\begin{aligned} \dot{p} &= P_p(f(p)) \\ p(0) &= p_0 \end{aligned} \tag{C.2}$$

betrachten.

Für geeignete Vektorfelder $f$ (wie zum Beispiel glatte Vektorfelder) können wir das Theorem von Picard-Lindelöf (Theorem 14.23) anwenden und erhalten für jedes $p_0 \in M$ einen maximalen Lösungsweg $p : I \to U$. Nun gilt aber für alle $t \in I$

$$(F(p))^{\cdot}(t) = \langle \nabla F(p), \dot{p}(t)\rangle = 0,$$

da $\dot{p}(t) = P_p(F(p))$ orthogonal auf $\nabla F(p(t))$ steht für alle $t \in I$. Auf Grund von $F(p_0) = 0$ erhalten wir also $F(p(t)) = 0$ für alle $t \in I$ oder äquivalenterweise $p(t) \in M$ für alle $t \in I$. Der Lösungsweg $t \mapsto p(t)$ bleibt also wie gewünscht in der Teilmannigfaltigkeit $M$.

Das Anfangswertproblem (C.2) beschreibt gewissermassen die Bewegung eines masselosen Teilchens, das einer Strömung ausgesetzt ist, aber dessen Bewegung durch die Nebenbedingung (gegeben durch $F(p) = 0$) eingeschränkt ist.

Falls wir (C.2) für den Gradienten $\nabla\varphi = f$ einer reellwertigen, glatten Funktion $\varphi$ anwenden, so können wir die so erhaltenen Lösungskurven auch als eine Suche nach einem Maximum von $\varphi$ auf $M$ interpretieren. In der Tat ist in diesem Fall

$$\dot{\varphi}(p) = \langle \nabla\varphi(p), \dot{p}\rangle = \langle \nabla\varphi(p), P_p(\nabla\varphi(p))\rangle \geq 0,$$

da für eine orthogonale Projektion $P : \mathbb{R}^3 \to P(\mathbb{R}^3)$ und einen Vektor $v \in \mathbb{R}^3$ stets $\langle v, P(v)\rangle = \langle P(v), P(v)\rangle \geq 0$ gilt. In anderen Worten ist $t \mapsto \varphi(p(t))$ monoton wachsend (und strikt monoton wachsend, solange $\nabla\varphi(p)$ nicht senkrecht auf $\mathrm{T}_pM$ steht.

Ist nun $M$ zusätzlich kompakt, so strebt $\varphi(p(t))$ gegen einen Kandidaten für ein Maximum von $\varphi$ auf $M$. Es kann dabei allerdings vorkommen, dass $p(t)$ durch ein lokales Maximum oder einen Sattelpunkt „gefangen wird“.

## C.2 Geodätischer Fluss auf dem Tangentenbündel

Wie bereits erwähnt, entspricht (C.2) der Bewegung eines masselosen Teilchens. Wir wollen jetzt die Bewegung eines Massepunktes $p \in M$ untersuchen, wenn nicht nur der Startpunkt, sondern auch die ursprüngliche Geschwindigkeit $(p_0, v_0) \in \mathrm{T}_{p_0}M$ bekannt ist. Weiter soll die Bewegung nach wie vor auf $M$ eingeschränkt sein, aber es sollen sonst keine Kräfte auf den Massepunkt wirken. Die Lösung des gesuchten Anfangswertproblems wird als eine **Geodäte** auf $M$ bezeichnet und beschreibt eine „möglichst gerade Kurve“, die innerhalb $M$ liegt.

Die ungestörte Bewegung eines Teilchens im $\mathbb{R}^3$ mit Ausgangsgeschwindigkeit $(p_0, v_0) \in \mathrm{T}_{p_0}\mathbb{R}^3$ ist natürlich durch das Anfangswertproblem

$$\begin{aligned} \ddot{p} &= 0 \\ p(0) &= p_0 \\ \dot{p}(0) &= v_0 \end{aligned} \tag{C.3}$$

gegeben. Wir wollen wiederum das Anfangswertproblem auf natürliche Weise so modifizieren, dass die Lösungen des modifizierten Problems innerhalb von $M$ bleiben. Der Lösungsweg $p : I \to \mathbb{R}^3$ soll also $F \circ p = 0$ erfüllen. Für die Ableitungen sollen insbesondere die Gleichungen

$$0 = (F \circ p)^{\cdot} = \langle \nabla F(p), \dot{p} \rangle = \sum_{j=1}^{n} \partial_j F(p) \dot{p}_j$$

und

$$\begin{aligned} 0 = (F \circ p)^{\cdot\cdot} &= \sum_{j=1}^{n} \partial_j F(p) \ddot{p}_j + \sum_{j,k=1}^{n} \partial_j \partial_k F(p) \dot{p}_j \dot{p}_k \\ &= \langle \nabla F(p), \ddot{p} \rangle + \dot{p}^t H_F(p) \dot{p} \end{aligned}$$

gelten, wobei $H_F(p) = (\partial_j \partial_k F(p))_{j,k}$ die Hesse-Matrix von $F$ bei $p$ ist. Es ist klar, dass obige Bedingungen noch nicht unsere gewünschte Differentialgleichungen darstellen, da diese von jedem glatten Weg in $M$ erfüllt werden. Sie sind allerdings notwendig für die von uns gewünschte Modifikation von (C.3).

Um die gesuchte Modifikation zu erraten, möchten wir uns kurz in den $\mathbb{R}^2$ setzen und uns die Teilmannigfaltigkeit $\mathbb{S}^1 = M$ vorstellen, welche durch die Gleichung $F(x, y) = x^2 + y^2 - 1 = 0$ gegeben ist. In diesem Fall können wir uns denken, dass der Massepunkt mit Hilfe eines starren Seiles der Länge 1 mit einem Pfeiler im Ursprung verbunden ist. Bei Rotation des Massepunktes entlang des Kreises wirkt ständig eine Kraft durch das straff gespannte Seil auf den Massepunkt. Diese Krafteinwirkung ist stets orthogonal zur Tangente des Punktes an $\mathbb{S}^1$.

Dieses Beispiel drängt die Vermutung auf, dass die gesuchte Differentialgleichung der geodätischen Bahnen auf $M$ verlangen sollte, dass $\ddot{p}$ orthogonal auf $\mathrm{T}_p M$ ist. Gemeinsam mit den bereits diskutierten, notwendigen Bedingungen wollen wir also das Anfangswertproblem

$$\begin{aligned} \ddot{p} &= -\dot{p}^t H_F(p) \dot{p} \|\nabla F(p)\|^{-2} \nabla F(p) \\ p(0) &= p_0 \\ \dot{p}(0) &= v_0 \end{aligned} \tag{C.4}$$

für einen vorgegebenen Anfangswert $(p_0, v_0) \in \mathrm{T}_{p_0} M$ betrachten. In der Tat gilt mit dieser Definition

$$\begin{aligned} (F \circ p)^{\cdot\cdot} &= \langle \nabla F(p), \ddot{p} \rangle + \dot{p}^t H_F(p) \dot{p} \\ &= -\dot{p}^t H_F(p) \dot{p} \left\langle \nabla F(p), \|\nabla F(p)\|^{-2} \nabla F(p) \right\rangle + \dot{p}^t H_F(p) = 0 \end{aligned}$$

sowie

$$\begin{aligned}(F \circ p)^{\cdot}(0) &= \langle \nabla F(p), \dot{p}(0) \rangle = \langle \nabla F(p), v_0 \rangle = 0 \\ (F \circ p)(0) &= F(p_0) = 0.\end{aligned}$$

Daraus folgt, dass die maximale Lösung $p : I \to \mathbb{R}^3$ von (C.4) auch $F(p(t)) = 0$ für alle $t \in I$ erfüllen muss. In anderen Worten liegen Lösungswege von (C.4) auf $M$. Des Weiteren gilt für die Lösung $p$

$$\left(\|\dot{p}\|^2\right)^{\cdot} = \langle \dot{p}, \dot{p} \rangle^{\cdot} = 2 \langle \ddot{p}, \dot{p} \rangle = 0. \tag{C.5}$$

In der Tat wissen wir, dass $F \circ p = 0$ ist, womit nach Ableiten $\langle \nabla F(p), \dot{p} \rangle = 0$ und somit $\dot{p} \in \mathrm{T}_p M$ ist. Auch ist nach Definition des Anfangswertproblems (siehe (C.4)) $\ddot{p}$ ein Vielfaches von $\nabla F(p)$ und insbesondere orthogonal auf $\mathrm{T}_p M$, womit sich (C.5) ergibt.

Hoffentlich lenken obige Formeln nicht zu sehr von der Schönheit und der Natürlichkeit des zugrundeliegenden Gedanken ab. Um letzteres zu betonen und die Formeln etwas weniger in den Vordergrund zu stellen, können wir noch eine formelfreie Vorstellung anbieten. Dazu verbinden wir den Massepunkt $p \in M$ mit zwei Seilen mit den Punkten $p + \nabla F(p)$ und $p - \nabla F(p)$. Die letzteren beiden Punkte statten wir mit genauen Sensoren, starken Raketenantrieben und ausgeklügelten Algorithmen aus, welche bei Bewegung von $p$ jeweils die nötige Kraft auf $p$ in Richtung $\nabla F(p)$ wirken lassen, so dass $p$ eben auf $M$ bleibt. Unser Formalismus in (C.4) beschreibt genau diesen Mechanismus.

**Applet C.1** (Geodätischer Fluss auf einem Doppeltorus)**.** *Es wird der geodätische Fluss auf einer kompakten Fläche visualisiert. Sie können die Bewegung mit dem Pfeil auf dem flachen Teilstück neu starten. Können Sie die Bewegung genau vorhersagen? Wenn ja, dann sollten Sie eine Karriere als Wahrsager versuchen, denn die Bewegung reagiert für lange Zeitspannen sehr sensibel auf kleine Änderungen in den Anfangsbedingungen: Die Bewegung ist „chaotisch“.*

## C.3 Physikalisch relevante Modifikationen

In obiger Diskussion haben wir uns von physikalischer Interpretation leiten lassen, um eine vernünftige Definition der Differentialgleichung der Geodäten von $M$ zu finden. Dennoch haben wir übliche, weitere physikalische Phänomene wie zum Beispiel

- die Schwerkraft,
- Krafteinwirkungen durch Strömungen eines Mediums (zum Beispiel Wind),
- Reibung oder auch
- Magnetfelder

ausser Acht gelassen. Wir wollen kurz andeuten, wie wir durch Modifikation von (C.4) Anfangswertprobleme erhalten können, die diese Phänomene in Betrachtung ziehen und noch

immer Lösungen in $M$ annehmen. In der Tat nimmt jede Lösung von

$$\begin{aligned} \ddot{p} &= -\dot{p}^t H_F(p)\dot{p}\|\nabla F(p)\|^{-2}\nabla F(p) + s(p,\dot{p}) \\ p(0) &= p_0 \\ \dot{p}(0) &= v_0 \end{aligned}$$

für eine glatte Störfunktion $s : \mathrm{T}U \to \mathbb{R}^3$ mit der Eigenschaft $(p, s(p,v)) \in \mathrm{T}_pM$ für alle $(p,v) \in \mathrm{T}_pM$ und für alle Anfangswerte $(p_0, v_0) \in \mathrm{T}_{p_0}M$ nur Werte in $M$ an. Dies verwendet $\langle \nabla F(p), s(p,v)\rangle = 0$ für alle $(p,v) \in \mathrm{T}M$ und dasselbe Argument wie zuvor. Nun kann der Störterm $s$ eine Summe von folgenden Ausdrücken sein:

- (Schwerkraft) Für eine vorgegebene Potentialfunktion $\Psi$ kann der Term $\nabla\Psi$ die Bedeutung der Schwerkraft haben. Die Projektion $(p,v) \mapsto P_p(\nabla\Psi(p))$ auf $\mathrm{T}_pM$ ist der von der Nebenbedingung, dass die Bahn in $M$ liegen soll, erlaubte Anteil der Schwerkraft.
- (Krafteinwirkung) Für ein allgemeines Kraftfeld $f : U \to \mathbb{R}^3$ können wir wiederum $(p,v) \mapsto P_p(f(p))$ erlauben.
- (Reibung) Für eine glatte Funktion $r : \mathrm{T}U \to \mathbb{R}_{\geq 0}$ können wir $(p,v) \mapsto -r(p,v)v$ als Reibungsterm in obiges Anfangswertproblem einbauen.
- (Magnetfeld) Angenommen das Teilchen hat nun eine elektrische Ladung. In diesem Fall könnte eine glatte Funktion $m : U \to \mathbb{R}$ auch als Stärke eines Magnetfelds interpretieren werden. Dementsprechend können wir den Term $(p,v) \mapsto m(p)v \times \nabla F(p)$ in das Anfangswertproblem aufnehmen. Dies wird die Bahn je nach Grösse und Vorzeichen von $m$ mehr oder weniger nach rechts oder nach links ablenken.

Wir bemerken, dass es folgende Erhaltungsgesetze gibt.

- (Nur Schwerkraft) Die Grösse (potentielle plus kinetische Energie) $\Psi(p) + \frac{1}{2}\|\dot{p}\|^2$ ist konstant.
- (Schwerkraft und Reibung) Die Grösse $\Psi(p) + \frac{1}{2}\|\dot{p}\|^2$ nimmt monoton ab.
- (Nur Magnetfeld) Die Geschwindigkeit $\|\dot{p}\|^2$ ist konstant.

**Applet C.2** (Bewegungen auf einem Hyperboloid)**.** *Wir sehen wie sich ein Teilchen auf einem Teil eines Hyperboloids bewegt, wenn wir den Rand der Fläche mit einer elastischen Bande ausstatten. Des Weiteren können wir die Bewegung bei Schwerkraft oder magnetischer Krafteinwirkung betrachten.*

## C.4 Dynamische Systeme

Ein **dynamisches System** besteht aus einem hier immer kompakten, metrischen Raum $X$ und einer Wirkung einer Gruppe, hier immer $\mathbb{R}$, auf $X$. Genauer formuliert wollen wir annehmen, dass wir eine stetige Abbildung

$$\Phi : \mathbb{R} \times X \to X, \quad (t,x) \mapsto \Phi(t,x)$$

haben, die

- $\Phi(0, x) = x$ für alle $x \in X$ und
- $\Phi(s, \Phi(t, x)) = \Phi(s + t, x)$ für alle $s, t \in \mathbb{R}$ und $x \in X$

genügt. Wir sprechen in diesem Fall auch von einem **Fluss** auf $X$. In Worten besagt dann die erste obige Eigenschaft, dass wenn man einen beliebigen Punkt $x \in X$ für Zeit 0 fliessen lässt, der Punkt fixiert bleibt. Die zweite Eigenschaft verlangt, dass der Punkt, den man erhält, wenn man zuerst $x \in X$ für Zeit $t$ fliessen lässt und dann den erhaltenen Punkt $x' = \Phi(t, x)$ für Zeit $s$ fliessen lässt, derselbe ist, wie wenn man auf $x$ direkt den Fluss für Zeit $s + t$ angewendet hätte.

Oft vereinfacht man die Notation auch etwas, indem man den **Fluss für Zeit** $t$ durch $\phi_t = \Phi(t, \cdot) : X \to X$ für alle $t \in \mathbb{R}$ definiert. Dann besagen obige Eigenschaften des Flusses, dass $\phi_0 = \mathrm{id}_X$ sowie $\phi_{s+t} = \phi_s \circ \phi_t$ für alle $s, t \in \mathbb{R}$ gelten soll.

**Beispiel C.3** (Flüsse auf Teilmannigfaltigkeiten)**.** *Seien $M, F, U$ wie zu Beginn dieses Kapitels, wobei wir hier zusätzlich annehmen wollen, dass $M$ kompakt ist. Folgende Beispiele von Flüssen auf $M$, $\mathrm{T}M$ oder auch auf dem Einheits-Tangentialbündel*

$$\mathrm{T}^1 M = \{(p, v) \in \mathrm{T}M \mid \|v\| = 1\}$$

*haben wir bereits gesehen.*

- *Mit dem Anfangswertproblem in* (C.2) *für ein glattes Vektorfeld $f : U \to \mathbb{R}^3$ lässt sich ein Fluss auf $X = M$ angeben. Die maximale Lösung von* (C.2) *zu einem Anfangswert $p_0$ ist auf ganz $\mathbb{R}$ definiert und wir nennen $\Phi(t, p_0) \in M$ den Wert der Lösung für den Anfangswert $p_0$ zum Zeitpunkt $t$.*

  *Um zu sehen, dass die maximale Lösung in der Tat auf ganz $\mathbb{R}$ definiert ist, bemerken wir, dass eine maximale Lösung mit nach oben oder nach unten beschränkten Geltungsbereich nach dem Theorem von Picard-Lindelöf (Theorem 14.23) gegen den Rand $\partial U$ streben sollte, was wegen $\Phi(t, p_0) \in M$ für alle $t$ nicht der Fall ist. Für letzteres Argument haben wir die Kompaktheit von $M$ verwendet, welche garantiert, dass $M$ und $\partial U$ eine positive Distanz voneinander haben.*

  *Die so definierte Abbildung $\Phi : \mathbb{R} \times M \to M,\ (t, p_0) \mapsto \Phi(t, p_0)$ definiert nun einen Fluss. In der Tat ist $p$ nach Proposition 14.32 stetig und erfüllt per Definition $\Phi(0, p_0) = p_0$ für alle $p_0 \in M$. Die zweite Eigenschaft von Flüssen folgt aus der Eindeutigkeit der Lösungen von Anfangswertproblem (siehe wiederum Theorem 14.23), da die Abbildung $s \mapsto \Phi(s + t, p_0)$ das Anfangswertproblem* (C.2) *für den Anfangswert $\Phi(t, p_0)$ löst.*

- *Sei $X = \mathrm{T}^1 M$. Dann definiert das Anfangswertproblem* (C.4) *einen Fluss $p$ auf $\mathrm{T}^1 M$, den **geodätischen Fluss**. Einen Startpunkt $x = (p_0, v_0)$ bildet $\Phi(t, \cdot) = \phi_t$ auf jenen Punkt ab, den man erhält, wenn man der Geodäte mit Ausgangsort und -vektor $(p_0, v_0)$ für Zeit $t$ folgt. Es lässt sich wie im letzten Fall verifizieren, dass dies in der Tat einen Fluss definiert (und insbesondere alle Geodäten für unbeschränkte Zeiten definiert sind).*

*Wir haben hierbei auch verwendet, dass Geodäten Einheitsgeschwindigkeit haben (siehe Gleichung* (C.5)*).*

- *Genauso kann man auf $X = \mathrm{T}^1 M$ für eine glatte Magnetstärke $m : M \to \mathbb{R}$ einen Fluss erhalten, welcher eine Modifikation des geodätischen Flusses darstellt (wie in Abschnitt C.3 beschrieben).*

- *Ebenso können wir Potentialfelder, andere Krafteinwirkungen oder Reibungen erlauben, um auf $X = \mathrm{T}M$ einen Fluss zu definieren. Dieser kann möglicherweise „grosse“ invariante, kompakte Teilmengen besitzen (mehr dazu später). Dabei ist es allerdings wichtig, dass die Störterme zeitunabhängig (autonom) sind, da ansonsten meist kein Fluss im Sinne obiger Definition zustande kommt.*

### C.4.1 Existenz minimaler Teilmengen

Im Folgenden sei $\Phi : \mathbb{R} \times X \to X$ ein Fluss auf einem kompakten metrischen Raum $X$ und $\phi_t = \Phi(t, \cdot)$ der Fluss für Zeit $t$.

**Definition C.4** (Invariante Teilmengen und Limesmenge)**.** Eine Teilmenge $K \subseteq X$ heisst **invariant**, falls für alle $x \in K$ auch $\phi_t(x) \in K$ liegt für alle $t \in \mathbb{R}$ oder in Worten falls die **Bahn** $\{\phi_t(x) \mid t \in \mathbb{R}\}$ jedes Punktes $x \in K$ in $K$ enthalten ist. Des Weiteren ist der **$\omega$-Limes** eines Punktes $x \in X$ definiert durch

$$\omega(x) = \{z \in X \mid \text{es existiert eine Folge } (t_k)_k \text{ mit } t_k \to \infty \text{ und } \phi_{t_k}(x) \to z \text{ für } k \to \infty\}\,.$$

**Lemma C.5.** *Der $\omega$-Limes $\omega(x)$ von $x \in X$ ist eine nicht-leere, abgeschlossene, invariante Teilmenge.*

*Beweis.* Da $X$ kompakt ist, existiert eine Folge $(t_k)_k$ mit $t_k \to \infty$ für $k \to \infty$, so dass die Folge $(\phi_{t_k}(x))_k$ gegen einen Punkt $z \in X$ konvergiert. Also ist $z \in \omega(x)$, womit insbesondere $\omega(x)$ nicht-leer ist.

Wir zeigen nun Invarianz von $\omega(x)$. Sei also $z \in \omega(x)$, $t \in \mathbb{R}$ und $(t_k)_k$ eine Folge mit $t_k \to \infty$ und $\phi_{t_k}(x) \to z$ für $k \to \infty$. Dann gilt

$$\phi_{t+t_k}(x) = \phi_t(\phi_{t_k}(x)) \to \phi_t(z)$$

für $k \to \infty$, also ist $\phi_t(z) \in \omega(x)$ und $\omega(x)$ ist invariant.

Um Abgeschlossenheit des $\omega$-Limes zu zeigen, sei $(z_\ell)_\ell$ eine Folge in $\omega(x)$ mit $z_\ell \to z$ für $\ell \to \infty$. Dann existiert für jedes $\ell$ eine Folge $(t_{\ell,k})_k$ mit $t_{\ell,k} \to \infty$ und $\Phi(t_{\ell,k}, x) \to z_\ell$ für $k \to \infty$. Wir konstruieren nun iterativ eine Folge $(s_k)_k$ in $\mathbb{R}$ wie folgt. Sei $s_1 = t_{1,k_1}$, wobei $k_1$ der Ungleichung $\mathrm{d}(\Phi(t_{1,k_1}, x), z_1) < 1$ genügt. Dann sei $s_2 = t_{2,k_2}$ so gegeben, dass $s_2 > s_1 + 1$ ist und $\mathrm{d}(\Phi(t_{2,k_2}, x), z_2) < \frac{1}{2}$ erfüllt ist. Fährt man so fort, erhält man eine Folge $(s_k)_k$ mit $\mathrm{d}(\Phi(s_k, x), z_k) < \frac{1}{k}$ und $s_k \to \infty$ für $k \to \infty$. Insbesondere gilt

$$\mathrm{d}(\Phi(s_k, x), z) \leq \mathrm{d}(\Phi(s_k, x), z_k) + \mathrm{d}(z_k, z) \to 0$$

für $k \to \infty$, womit $z \in \omega(x)$ ist. Also ist $\omega(x)$ abgeschlossen. □

**Übung C.6.** *Es bezeichne $\psi_x$ die Abbildung $t \mapsto \Phi(t,x) \in X$ für ein $x \in X$, welche die Bahn von $x$ beschreibt. Zeigen Sie, dass*

$$\omega(x) = \bigcap_{t>0} \overline{\psi_x([t,\infty))}$$

*für alle $x \in X$ gilt.*

Als Hauptziel unserer Exkursion in die Theorie der dynamischen Systeme wollen wir zeigen, dass unter einem Fluss auf einem kompakten metrischen Raum stets einen rekurrenten Punkt gibt.

**Definition C.7** (Rekurrenz)**.** Ein Punkt $x \in X$ heisst **rekurrent**, falls $x \in \omega(x)$ ist oder in anderen Worten falls es eine Folge von Zeiten $(t_k)_k$ in $\mathbb{R}$ mit $t_k \to \infty$ für $k \to \infty$ gibt, so dass $\phi_{t_k}(x) \to x$ für $k \to \infty$ gilt.

In Worten ausgedrückt ist ein Punkt rekurrent, falls er beliebig nahe zu sich selbst zurückkehrt. Es wird sich zeigen, dass es im Wesentlichen genauso schwierig ist, eine gewisse stärkere Aussage zu beweisen, die folgenden Begriff benötigt.

**Definition C.8** (Minimalität)**.** Eine nicht-leere invariante abgeschlossene Teilmenge $K \subseteq X$ heisst **minimal**, falls $\omega(x) = K$ für jedes $x \in K$ erfüllt ist.

Jeder Punkt $x$ in einer minimalen Teilmenge $K$ ist dann rekurrent, da $x \in K = \omega(x)$ gilt.

**Satz C.9** (Existenz minimaler Teilmengen)**.** *Sei $\Phi : \mathbb{R} \times X \to X$ ein Fluss auf einem nicht-leeren, kompakten metrischen Raum. Dann existiert eine nicht-leere minimale Teilmenge $K \subseteq X$ für den Fluss $\Phi$. Insbesondere existieren unter $\Phi$ rekurrente Punkte in $X$.*

*Beweis.* Sei $x_1 \in X$ beliebig und $K_1 = \omega(x_1)$. Nach Lemma C.5 ist $K_1$ eine nicht-leere invariante abgeschlossene Teilmenge. Falls $K_1$ minimal ist, so gilt der Satz bereits. Also nehmen wir an, dass $K_1$ nicht minimal ist und können somit einen Punkt $x_2 \in K_1$ wählen, so dass $\omega(x_2) \subseteq K_1$ eine echte Teilmenge ist (die auch nicht-leer, abgeschlossen und invariant ist nach Lemma C.5). Wir iterieren diese Konstruktion und erhalten entweder eine nicht-leere minimale Teilmenge (nach endlich vielen Schritten) oder eine Kette

$$K_1 \supsetneq K_2 \supsetneq K_3 \supsetneq \dots$$

von nicht-leeren invarianten abgeschlossenen Teilmengen von $X$. Da $X$ kompakt ist, gilt das Schachtelungsprinzip (siehe Satz 9.66) und

$$K_\omega = \bigcap_{n=1}^{\infty} K_n$$

ist eine nicht-leere invariante abgeschlossene Teilmenge. Da die Teilmengen $K_n$ für alle $n \in \mathbb{N}$ invariant sind, ist auch $K_\omega$ invariant. (Wieso?) Wir bemerken, dass der Index $\omega$ hier nicht für den $\omega$-Limes steht, sondern für die erste unendliche „Ordinalzahl $\omega$“ aus der Mengentheorie.

Man könnte nun hoffen, dass $K_\omega$ minimal ist. Leider muss dies nicht immer so sein. Wir müssen also obiges Argument wiederholen, um entweder eine nicht-leere minimale Teilmenge (nach weiteren endlich vielen Schritten) oder eine weitere Kette

$$K_\omega \supsetneq K_{\omega+1} \supsetneq K_{\omega+2} \supsetneq \dots$$

von nicht-leeren invarianten abgeschlossenen Teilmengen zu finden. Im zweiten Fall ist

$$K_{\omega\cdot 2} = \bigcap_{n=0}^{\infty} K_{\omega+n}$$

wiederum eine nicht-leere invariante abgeschlossene Teilmenge. Falls $K_{\omega\cdot 2}$ nicht minimal ist, so müssen wir möglicherweise

$$K_{\omega\cdot 2+1}, K_{\omega\cdot 2+2}, \dots, K_{\omega\cdot 3}, K_{\omega\cdot 3+1}, K_{\omega\cdot 3+2}, \dots$$

und so weiter konstruieren.

Eine derartige Induktion über alle „Ordinalzahlen“ der Mengenlehre nennt man auch eine „transfinite Induktion“. Dass diese mit der Konstruktion einer minimalen Teilmenge enden muss, folgt daraus, dass wir ansonsten für eine beliebige Ordinalzahl $\alpha$ eine Injektion von

$$\alpha = \{\beta \mid \beta \text{ Ordinalzahl mit } \beta < \alpha\}$$

in die Potenzmenge $\mathcal{P}(X)$ finden können. Doch gibt es auch Ordinalzahlen der Kardinalität $\mathcal{P}(\mathcal{P}(X))$, was einen Widerspruch liefert.

Mathematiker (abgesehen von Experten in Mengentheorie und mathematischer Logik) verwenden die obige Sprache der transfiniten Induktion nicht so oft. Stattdessen verwendet man häufiger das Lemma von Zorn, welches zum Auswahlaxiom äquivalent ist. Hierfür betrachtet man die nicht-leere Teilmenge

$$\mathcal{G} = \{K \subseteq X \mid K \text{ nicht-leer, invariant, abgeschlossen}\} \subseteq \mathcal{P}(X)$$

und zeigt zuerst, dass für eine beliebige linear geordnete Teilmenge $\mathcal{L} \subseteq \mathcal{G}$ auch

$$\bigcap_{K\in\mathcal{L}} K \in \mathcal{G}$$

gilt. Hierbei ist $\mathcal{L}$ linear geordnet, falls wir für alle $K_1, K_2 \in \mathcal{L}$ eine der Inklusionen $K_1 \subseteq K_2$ oder $K_2 \subseteq K_1$ haben. Der Beweis von $\bigcap_{K\in\mathcal{L}} K \in \mathcal{G}$ verwendet wieder Satz 9.66 (in der stärkeren, nicht abzählbaren Form) und ist völlig analog zum Beweis, dass unser $K_\omega = \bigcap_{n=1}^{\infty} K_n$ von unserer vorherigen Diskussion zu $\mathcal{G}$ gehört.

Nach Überprüfung von $\bigcap_{K\in\mathcal{L}} K \in \mathcal{G}$ für alle linear geordneten Teilmengen $\mathcal{L}$ schliesst man aus dem Lemma von Zorn, dass $\mathcal{G}$ ein minimales Element $K_{\min}$ enthält (das heisst, ein Element mit $K \subseteq K_{\min} \implies K = K_{\min}$).

Nun gilt, dass $K_{\min}$ eine minimale Teilmenge von $X$ ist. Denn für $x \in K_{\min}$ ist $\omega(x) \subseteq K_{\min}$

eine nicht-leere abgeschlossene invariante Teilmenge. Falls also $\omega(x) \neq K_{\min}$ wäre, so hätten wir mit $\omega(x) \in \mathcal{G}$ ein kleineres Element von $\mathcal{G}$ gefunden, was der Wahl von $K_{\min}$ als ein minimales Element widerspricht. $\square$

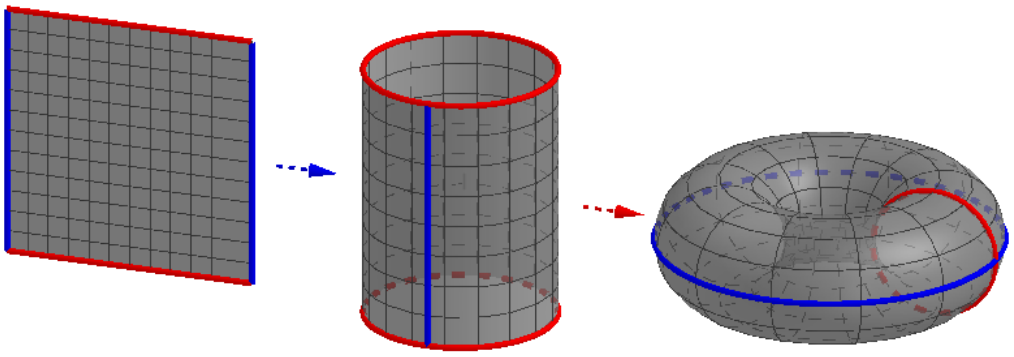
## C.5 Der geodätische Fluss auf (Quotienten) der Ebene

Der geodätische Fluss auf der Ebene $\mathbb{R}^2$ ist gegeben durch

$$(p, v) \in \mathrm{T}^1\mathbb{R}^2 \mapsto (p + tv, v)$$

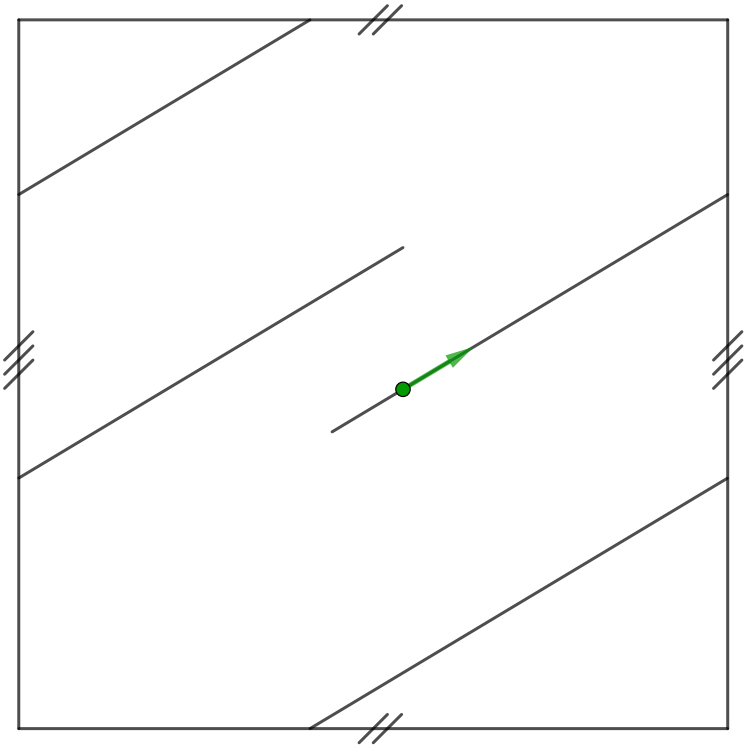
und ist eigentlich ziemlich langweilig. In der Tat bewegt dieser jeden Punkt $p$ bei vorgegebener Richtung $v \in \mathbb{R}^2$ mit $\|v\| = 1$ mit Geschwindigkeit 1 entlang der Geraden $p + \mathbb{R}v$ nach Unendlich. Dabei ist insbesondere der $\omega$-limes für jedes Paar $(p, v) \in \mathrm{T}^1\mathbb{R}^2$ die leere Menge. Dies ist wohlgemerkt kein Widerspruch zu Lemma C.5, da $\mathrm{T}^1\mathbb{R}^2$ kein kompakter Raum ist.

Allerdings wird der geodätische Fluss deutlich interessanter, wenn wir die Ebene zu einem Torus verkleben. Dabei wird das Einheitsquadrat $[0, 1]^2$ wie in folgender Figur zuerst zur Mantelfläche eines Zylinders verbogen und verklebt. Anschliessend wird der obere Kreis mit dem unteren Kreis identifiziert.



Im Gegensatz zu Beispiel 11.18 möchten wir uns den erhaltenen 2-Torus nicht als Teilmannigfaltigkeit des $\mathbb{R}^3$ vorstellen. Bei dieser Vorstellung (welche wir in obigem Bild angewendet haben) werden unsere Geodäten im $\mathbb{R}^2$ nicht auf Geodäten der Teilmannigfaltigkeit abgebildet (wovon man sich zuerst überzeugen müsste). Grund dafür ist, dass der 2-Torus als Teilmannigfaltigkeit des $\mathbb{R}^3$ im Gegensatz zur „flachen“ Ebene $\mathbb{R}^2$ „Krümmung“ aufweist.

Eine bessere Vorstellung erhalten wir, wenn wir die nur mit dem Quadrat $[0, 1]^2$ (mit identifizierten Kanten) arbeiten, ohne den damit erhalten 2-Torus in ein konkretes $\mathbb{R}^n$ einbetten zu wollen (was wohlgemerkt allerdings möglich wäre). Wie bei (eher klassischen) Computer-Spielen kann man das Quadrat (also den Bildschirm) nicht verlassen. Versucht man dies trotzdem, so erscheint man auf der gegenüberliegenden Seite wieder. Dabei ändert sich nur der Ort, aber nicht die Richtung der Geodäte.



Figur C.1: Die Markierungen am Rand deuten die Verklebungen an. Die Geodäte zum Startpunkt $(p, v)$ bleibt auf Grund der den Verklebungen entsprechenden Identifikationen innerhalb des Quadrats.

Wenn Sie eine formalere Definition von obigem wollen, so haben Sie die Wahl zwischen

- der Auffassung des 2-Torus als Quotienten

$$\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2 = \left\{p + \mathbb{Z}^2 \mid p \in \mathbb{R}^2\right\}$$

  womit

$$\begin{aligned} \mathrm{T}\mathbb{T}^2 &= \mathbb{T}^2 \times \mathbb{R}^2, \\ \mathrm{T}^1\mathbb{T}^2 &= \mathbb{T}^2 \times \mathbb{S}^1 = \left\{(p, v) \mid p \in \mathbb{T}^2,\ v \in \mathbb{R}^2 : \|v\| = 1\right\} \end{aligned}$$

  das Tangentialbündel respektive das Einheits-Tangentialbündel ist oder

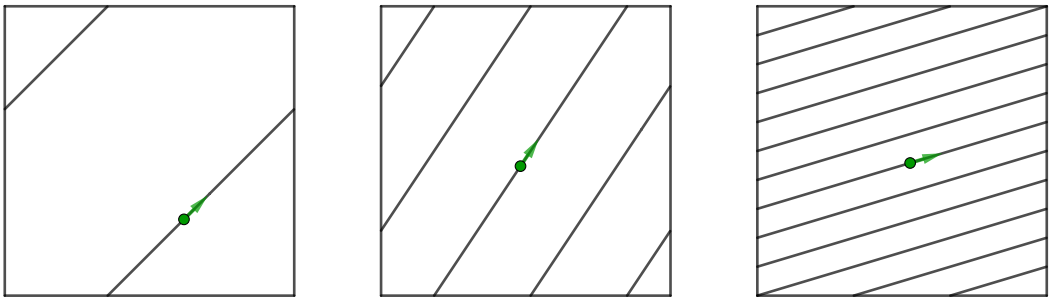
- der eingebetteten Auffassung

$$\mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1 \subseteq \mathbb{R}^4,$$

  welche in Übung 11.19 diskutiert wurde und für welche das Tangentialbündel mit Hilfe von Satz 11.21 beschrieben wird (siehe auch Beispiel 11.23 und Übung 11.24).

Die Dynamik des geodätischen Flusses ist nun definitiv interessanter:

- Für $v_1 = (1, 0)$ oder allgemeiner für Einheitsvektoren $v \in \mathbb{R}^2$, für welche die Linie $\mathbb{R}v$ rational ist (d.h. $\mathbb{R}v \cap \mathbb{Z}^2 \neq \emptyset$), ist jede Bahn periodisch. In anderen Worten existiert für jedes $p \in \mathbb{T}^2$ eine Zeit $T$, so dass $p + Tv = p$ in $\mathbb{T}^2$ ist.



Figur C.2: Darstellung dreier periodischer Bahnen für „rationale Richtungen“.

- Für eine irrationale Richtung $v \in \mathbb{R}^2$ mit $\|v\| = 1$ (also $\mathbb{R}v \cap \mathbb{Z}^2 = \emptyset$) ist die Geodäte $p + \mathbb{R}v$ aufgefaltet in $\mathbb{T}^2$ nicht periodisch und auch nicht abgeschlossen. Stattdessen ist sie dicht und für den $\omega$-Limes von $(p, v)$ gilt

$$\omega((p, v)) = \mathbb{T}^2 \times \{v\}\,,$$

da sich der Richtungsvektor entlang der Geodäte nie ändert. Zum Beweis dieser Aussagen benützt man am besten die Theorie der Fourier-Reihen auf $\mathbb{T}^2$.

## C.6 Ein Quotient einer nicht-euklidschen Ebene

Die **hyperbolische Ebene** $\mathbb{H}^2$ kann auf verschiedene Arten definiert werden; eine davon werden wir unten sehen. Ursprünglich von Interesse war die hyperbolische Ebene unter anderem, da sie alle Axiome der Euklidschen Geometrie erfüllt bis auf das Parallelenaxiom. Sie wurde auch bereits von Gauss untersucht und ist seither in vielen Gebieten der Mathematik wichtig geworden, unter anderem in der Differentialgeometrie, in der Theorie der dynamischen Systeme, in der Theorie der Lie-Gruppen. Man könnte sich aber auch in anderen Gebieten mit weniger direkten Beziehung zur Geometrie wie beispielsweise der Zahlentheorie die hyperbolische Ebene oder ihre Isometriegruppe $\mathrm{SL}_2(\mathbb{R})$ kaum wegdenken.

Wir definieren die hyperbolische Ebene $\mathbb{H}^2$ als die obere Halbebene

$$\mathbb{H}^2 = \{z \in \mathbb{C} \mid \mathrm{Im}(z) > 0\}\,. \tag{C.6}$$

Dabei möchten wir in $\mathbb{H}^2$ die Distanzen aber nicht auf die übliche (Euklidsche) Art und Weise messen, sondern einen Skalierungsfaktor einführen (siehe Abschnitt 10.6.1). Genauer statten wir für jedes $p \in \mathbb{H}^2$ den Tangentialraum $\mathrm{T}_p\mathbb{H}^2 = \{p\} \times \mathbb{R}^2$ mit dem (reellen) inneren Produkt

$$\langle (p, v), (p, w) \rangle_p = \frac{\langle v, w \rangle_2}{\mathrm{Im}(p)^2}$$

für alle $(p, v), (p, w) \in \mathrm{T}_p\mathbb{H}^2$ aus. Die Länge eines (stückweise stetig differenzierbaren) Weges $\gamma : [a, b] \in \mathbb{H}^2$ definiert man nun durch

$$L(\gamma) = \int_a^b \frac{\|\gamma'(t)\|_2}{\mathrm{Im}(\gamma(t))}\,\mathrm{d}t = \int_a^b \sqrt{\langle \gamma'(t), \gamma'(t)\rangle_p}\,\mathrm{d}t.$$

Nimmt man das Infimum aller Wege zwischen zwei Punkten, so lässt sich damit eine Metrik definieren, auf welche wir hier allerdings nicht genauer eingehen wollen (siehe Übung 10.42).

*Bemerkung.* Alternativ kann man die hyperbolische Ebene auch als Teilmannigfaltigkeit $M$ von $\mathbb{R}^6$ auffassen, wobei das innere Produkt auf $\mathrm{T}_pM$ für alle $p \in M$ das Standardskalarprodukt ist (siehe [Bla55]). Insbesondere passt die hyperbolische Ebene in die Diskussion der vorherigen Abschnitte (nach geeigneter Auffassung). Dieses Model ist allerdings nicht gleich praktikabel wie das Model aus (C.6), weswegen wir nicht weiter darauf eingehen wollen.

Nun kann man beispielsweise überprüfen, dass die Kurve

$$\gamma : t \in \mathbb{R} \mapsto x + \mathrm{e}^t\mathrm{i}$$

für alle $x \in \mathbb{H}^2$ eine Geodäte von $\mathbb{H}^2$ darstellt, in dem Sinne, dass jeder beliebige Weg von $\gamma(a)$ nach $\gamma(b)$ für alle $a, b \in \mathbb{R}$ länger als $\gamma|_{[a,b]}$ ist. Darum nennt man die vertikalen Geraden auch Geodäten auf der hyperbolischen Ebene $\mathbb{H}^2$.

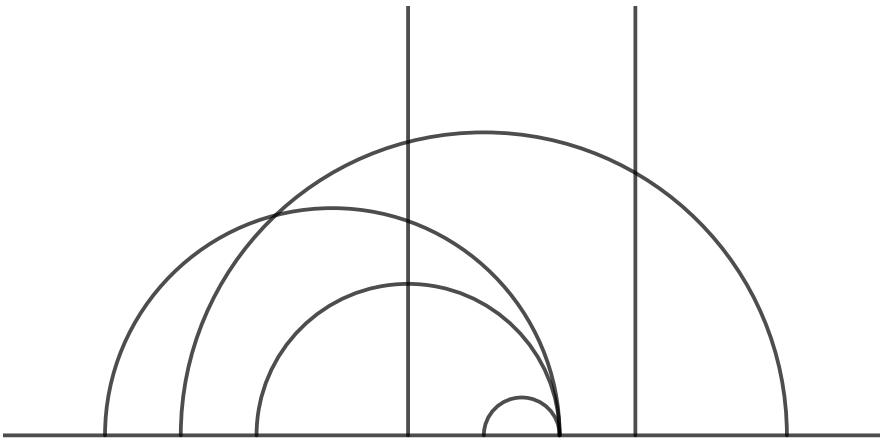
Wir wenden uns nun den Isometrien der hyperbolischen Ebene zu. Es ist vielleicht überraschend, dass die Matrixgruppe $\mathrm{SL}_2(\mathbb{R}) = \{g \in \mathrm{Mat}_2(\mathbb{R}) \mid \det(g) = 1\}$ isometrisch auf $\mathbb{H}^2$ (mittels sogenannten **Möbiustransformationen**) wirkt. Für jedes Element $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{R})$ und $z \in \mathbb{H}^2$ setzt man

$$g.z = \frac{az + b}{cz + d} \in \mathbb{H}^2.$$

Des Weiteren gilt für jeden Weg $\gamma : [a, b] \to \mathbb{H}^2$

$$L(g \circ \gamma) = L(\gamma),$$

womit insbesondere die Aktion von $\mathrm{SL}_2(\mathbb{R})$ die Distanz zwischen je zwei Punkte erhält oder in anderen Worten eben isometrisch wirkt. Auf diese Weise können wir die bereits gefundenen vertikalen Geodäten durch Möbiustransformationen auf eine neue Art von Geodäte auf $\mathbb{H}^2$ wie in folgendem Bild senden.

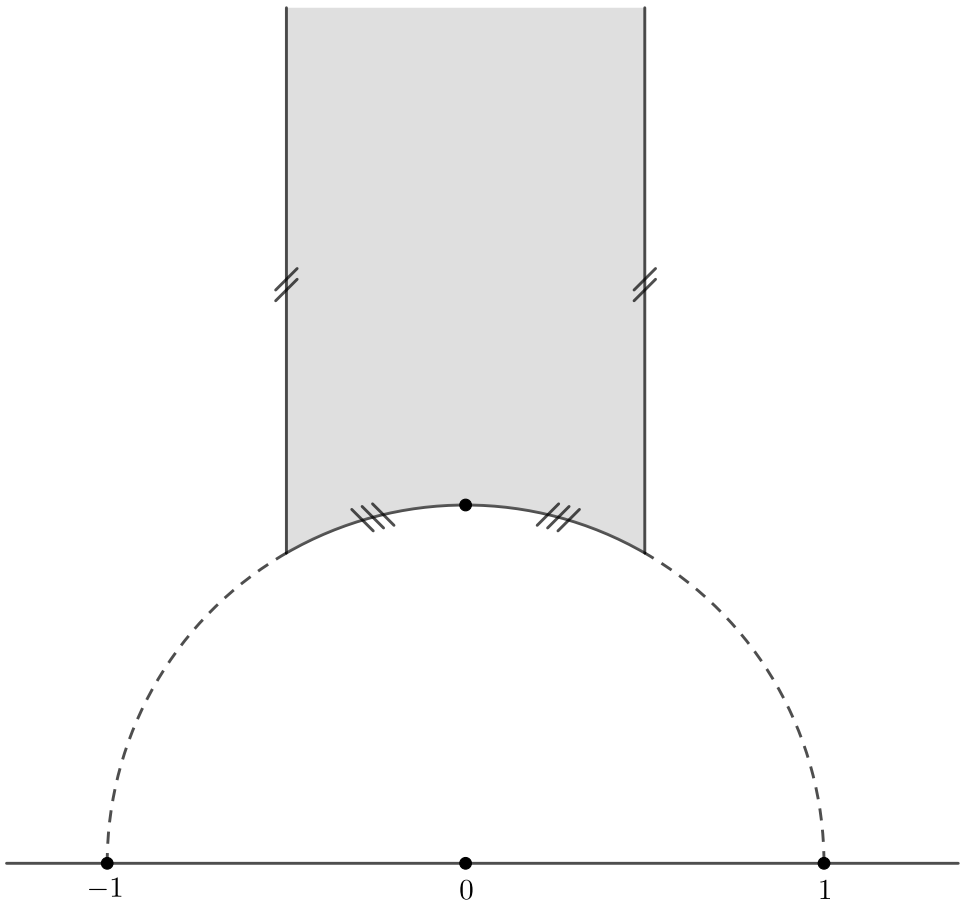


Figur C.3: Einige sich schneidende Geodäten und einige sich nicht schneidende Geodäten der hyperbolischen Ebene.

Viele der obigen Aussagen lassen sich durch konkrete Rechnungen verifizieren, die auch in ihrer Schönheit überzeugen können.

Der geodätische Fluss auf $\mathbb{H}^2$, der für einen vorgegebenen Punkt $p \in \mathbb{H}^2$ und einen Richtungsvektor $(p, v) \in \mathrm{T}^1_p\mathbb{H}^2$ den Punkt $p$ entlang der Geodäte durch $p$ mit Richtung $v$ verschiebt, ist genauso wie schon der geodätische Fluss auf $\mathbb{R}^2$ dynamisch nicht sehr interessant.

Falten und verkleben wir hingegen $\mathbb{H}^2$, so kann man mit den sich ergebenden dynamischen, quantentheoretischen und zahlentheoretischen Fragen und Antworten ohne Probleme eine einjährige Vorlesung (mit sechs Vorlesungsstunden pro Woche) füllen. Beispielsweise kann die hyperbolische Ebene $\mathbb{H}^2$ kann mittels $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ und $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ zu der Modulfläche $M = \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}^2$ verklebt werden.




Der geodätische Fluss auf dem nicht-kompakten Tangentenbündel $\mathrm{T}^1 M$ ist „ergodisch“, womit insbesondere die Bahn zu fast jedem Anfangswert dicht in $\mathrm{T}^1 M$ liegt.

# Anhang D

# Kapitelübergreifende Fragen

## D.1 Wie kann man Konvergenz einer Folge zeigen?

Die Konvergenz komplexwertiger Folgen wurde in Definition 5.2 formuliert und wurde sehr bald für die gesamte Theorie ein unentbehrlicher Begriff. Also ist die Frage, wie man für eine konkret gegebene Folge Konvergenz zeigen kann, äusserst relevant. Allgemeingültig kann man dies jedoch nicht beantworten. Zum Beispiel haben wir in Abschnitt 8.7.1 partielle Integration (!) für die Bestimmung eines Grenzwerts einer Folge und in Abschnitt 8.7.2 Taylor-Approximation auf unendlich vielen Intervallen (!) verwendet. Des Weiteren könnte die Folge das Ergebnis eines näherungsweisen Lösungsverfahren einer Differentialgleichung sein und damit sind möglicherweise die Methoden im Beweis des Existenz- und Eindeutigkeitssatzes von Picard-Lindelöf (Theorem 14.23) relevant. Wir wollen hier also keinen Anspruch auf Vollständigkeit erheben, aber trotzdem einige Möglichkeiten mit der von uns besprochenen Theorie in aufsteigender Schwierigkeit auflisten. Sei also $(a_n)_n$ eine Folge reeller Zahlen.

(i) *Elementare Grenzwerte:* Für rationale Funktionen oder auch Kombinationen von Polynomen und den Ausdrücken in Beispiel 5.33 genügen oft elementare algebraische Umformungen und die Anwendung von Proposition 5.7.

(ii) *Regel von de l'Hôpital:* Existieren zwei differenzierbare Funktion $f, g$, sodass sich $a_n$ als $\frac{f(x_n)}{g(x_n)}$ für jedes $n$ und eine weitere Folge $(x_n)_n$ schreiben lässt, so kann man $\lim_{n\to\infty} a_n$ vielleicht als Spezialfall der Berechnung eines Grenzwerts der Funktion $x \mapsto \frac{f(x)}{g(x)}$ interpretieren. Letzteren kann man oft mit der Regel von de l'Hôpital (Satz 7.49) berechnen.

(iii) *Monotonie:* Falls die Folge $(a_n)_n$ monoton und beschränkt ist, so kann man Satz 5.34 verwenden. Möglicherweise ist $a_n$ nicht monoton, aber „alternierend monoton“ in dem Sinne, als dass $a_{2n} \leq a_{2n+2} \leq a_{2n+3} \leq a_{2n+1}$ für alle $n \in \mathbb{N}$ (siehe auch den Beweis des Leibniz-Kriteriums in Proposition 6.25). Dann lässt sich Satz 5.34 verwenden, um die Existenz der Grenzwerte $\lim_{n\to\infty} a_{2n} \leq \lim_{n\to\infty} a_{2n+1}$ zu zeigen. Es verbleibt also noch die Gleichheit zu verifizieren, damit die Existenz des Grenzwerts $\lim_{n\to\infty} a_n$ folgt.

(iv) *Sandwich:* Für reelle Folgen ist das Sandwich-Lemma (Lemma 5.31) nützlich, falls man eine "etwas kleinere" und eine "etwas grössere Folge" mit demselben Grenzwert finden kann.

(v) *Banachscher Fixpunktsatz:* Falls die Folge rekursiv durch $x_{n+1} = T(x_n)$ für ein gewisses $x_1$ und eine gewisse Abbildung definiert wird, so lässt sich vielleicht der Banachsche Fixpunktsatz in Satz 9.54 verwenden um die Konvergenz der Folge zu zeigen.

(vi) *Interpretation als Riemann-Summe:* Falls $a_n$ durch eine Summe dargestellt ist (oder dargestellt werden kann), so ist $a_n$ vielleicht eine Riemann-Summe zu einem ein- oder mehrdimensionalen Riemann-Integral. Also könnte die Existenz des Grenzwerts (inkl. dessen Wert) aus Satz 5.79 folgen.

(vii) *Newton-Verfahren:* Vielleicht ist die Folge $(a_n)_n$ rekursiv definiert und stellt einen Spezialfall des Newton-Verfahrens in Beispiel 8.65 dar.

(viii) *Umwandlung in eine Reihe:* Falls $(a_n)_n$ durch eine konkrete Formel gegeben ist, kann man anstelle der Folge die Reihe $\sum_{n=1}^{\infty}(a_{n+1}-a_n)$ betrachten. Damit kann man vielleicht eine der Methoden für Reihen anwenden, siehe Kapitel 6.

(ix) *Diagonalisierbarkeit von Matrizen:* Falls die Folge durch Anfangswerte $x_1, x_2, \ldots, x_k$ und eine lineare Rekursion der Form $x_{n+k+1} = a_0x_n + a_1x_{n+1} + \cdots + a_kx_{n+k} + b$ für ein $k \in \mathbb{N}$ und $a_0, \ldots, a_k, b \in \mathbb{R}$ definiert wird, dann lassen sich $x_n, \ldots, x_{n+k}$ auch als Komponenten eines Vektors und der Potenz einer zugehörigen Matrix interpretieren. Diagonalisiert man diese Matrix, so vereinfacht sich das Problem.

(x) *Cauchy-Kriterium:* Die Cauchy-Bedingung in Definition 5.24 ist nach Satz 5.48 eine Charakterisierung der Konvergenz und stellt ein fundamentales theoretisches Hilfsmittel dar. (Vollständigkeit metrischer Räume wird insbesondere genau durch diese Äquivalenz definiert.) Dieses Kriterium hat den Vorteil, dass der Grenzwert der Folge nicht bekannt sein muss, und gleichzeitig den Nachteil, dass man damit den Grenzwert auch nicht bestimmen kann. In der Praxis wird dieses Kriterium allerdings selten angewendet, da obige Methoden (die oft Korollare des Cauchy-Kriteriums sind) einfacher anzuwenden sind.

**Übung D.1.** *Wir wollen hier einige Beispiele zu den obigen Methoden erwähnen, es gibt oft mehr als eine Möglichkeit wie man vorgehen kann.*

*a) Berechnen Sie* $\lim_{n\to\infty} \frac{1}{n} \log \frac{(2n)!}{n!n^n}$.

*b) Berechnen Sie den Grenzwert der Folge, die durch* $x_1 = 1$, $x_2 = 0$, *und* $x_{n+2} = \frac{5}{6}x_{n+1} - \frac{1}{6}x_n + 1$ *für* $n \in \mathbb{N}$ *definiert wird.*

Hinweise: Bei a) kann man die asymptotische Sterling-Formel verwenden, es gibt aber auch die Möglichkeit die Folge in eine Riemann-Summe zu verwandeln. Bei b) kann man eine zugehörige diagonalisierbare Matrix betrachten und entweder diese diagonalisieren oder nach Bestimmung der Eigenwerte den Banachschen Fixpunktsatz anwenden.

## D.2 Warum sind kompakte konvexe Mengen Jordan-messbar?

In der Analysis-VO taucht diese Frage eigentlich nicht auf, da konvexe Mengen für die von uns besprochene Theorie (abgesehen von Korollar 10.17) nicht so wichtig sind. Doch taucht diese Frage mitunter in den Numerik-Vorlesungen auf, weswegen wir hier kurz auf die Frage eingehen wollen.

**Proposition D.2.** *Sei $n \geq 1$. Eine konvexe kompakte Teilmenge $K \subseteq \mathbb{R}^n$ ist Jordan-messbar.*

*Beweis-Skizze.* Wir beweisen die Aussage mittels Induktion nach der Dimension $n$. Für $n = 1$ ist eine konvexe kompakte Menge ein abgeschlossenes beschränktes Intervall, welches in der Tat Jordan-messbar ist.

Angenommen die Aussage stimmt bereits für konvexe kompakte Teilmengen in $\mathbb{R}^n$ und $K \subseteq \mathbb{R}^{n+1}$ ist eine konvexe kompakte Teilmenge. Sei $\pi : \mathbb{R}^{n+1} \to \mathbb{R}^n$ die Projektion auf die ersten $n$ Koordinaten und $L = \pi(K)$. Dann ist $L$ ebenso eine konvexe kompakte Menge und nach Induktionsannahme auch Jordan-messbar. Wir behaupten, dass zwei stetige beschränkte Funktionen $\phi_-, \phi_+ : L^\circ \to \mathbb{R}$ mit

$$\begin{aligned} K &\supseteq \{(x, y) \in L^\circ \times \mathbb{R} \mid \phi_-(x) < y < \phi_+(x)\} \\ K &\subseteq \{(x, y) \in L^\circ \times \mathbb{R} \mid \phi_-(x) \leq y \leq \phi_+(x)\} \cup \partial L \times \mathbb{R} \end{aligned}$$

existieren. Die Behauptung impliziert den Induktionsschritt unter Verwendung von den Methoden im Beweis von Korollar 12.27. Dabei ist es noch hilfreich zu bemerken, dass $\partial J$ eine Jordan-Nullmenge ist.

Für die Behauptung definieren wir

$$\begin{aligned} \phi_+(x) &= \sup \{y \in \mathbb{R} \mid (x, y) \in K\} \\ \phi_-(x) &= \inf \{y \in \mathbb{R} \mid (x, y) \in K\} \end{aligned}$$

für alle $x \in L^\circ$. Für ein festes $x_0 \in L^\circ$ wählen wir ein hinreichend kleines $\delta > 0$ so dass $x_{\pm,k} = x_0 \pm \delta e_k \in L$ für $k = 1, \ldots, n$. Nach Definition von $L$ gibt es zu jeden $x_{\pm,k}$ ein $y_{\pm,k} \in \mathbb{R}$ so dass $(x_{\pm,k}, y_{\pm,k}) \in K$. Unter Verwendung dieser Punkte und dem Punkt $(x_0, \phi_+(x_0))$ lässt sich nun $D = \{y \in \mathbb{R}^n \mid \|x - x_0\|_1 \leq \delta\}$ und mittels Konvexkombinationen auf den verschiedenen Quadranten eine stetige Funktion $\psi : D \to \mathbb{R}$ definieren so dass $(x, \psi(x)) \in K$ für alle $x \in D$. (Es empfiehlt sich hierfür zuerst die Fälle $n = 1, 2$ genauer zu betrachten.) Diese Funktion erfüllt damit $\psi(x) \leq \phi_+$, woraus eine Ungleichung der gewünschten Stetigkeit von $\phi_+$ bei $x_0$ folgt.

Für die zweite Ungleichung betrachtet man stattdessen die „gespiegelte“ Funktion

$$\tilde{\psi} : x \in D \mapsto 2\phi_+(x_0) - \psi(2x_0 - x)$$

und zeigt, dass $\phi_+(x) \leq \tilde{\psi}(x)$ für alle $x \in D$. □

## D.3 Eine Anekdote zur Jordan-Messbarkeit

Nach Veröffentlichung dieses Skripts als eSkript der ETH und dem Verschicken der Einladungen zum zugehörigen Diskussionsforum wurde dort eine Frage zur Substitutionsformel (Satz 12.56) und zur Jordan-Messbarkeit gestellt, die wir einige Tage lang nicht beantworten konnten.

Zufälligerweise fand kurz danach der Departementsausflug des D-MATH statt. Beim Abendessen entwickelte sich dann eine Diskussion über nette Rätsel (wie zum Beispiel die Übung über ganzzahlige Kanten in Abschnitt 12.8.2), Rechnungen wie in Abschnitt 12.7.3, das von John Nash erfundene Spiel Hex oder eben Beweise wie im Sommerspecial. Die Frage aus dem Diskussionsforum kam dabei auch auf, doch war hier im Unterschied zu obigen Themen die Lösung nicht bekannt.

**Frage:** Seien $X, Y \subseteq \mathbb{R}^n$ offen und sei $\Phi : X \to Y$ genügend nett (glatt, $C^1$ oder auch nur Lipschitz stetig). Des Weiteren sei $B \subseteq X$ eine kompakte, Jordan-messbare Teilmenge. Ist dann $\Phi(B)$ auch Jordan-messbar? Welche Voraussetzungen an $\Phi$ benötigt man hierzu?

Für einen Diffeomorphismus $\Phi$ ist dies natürlich Teil von Satz 12.56, wobei in diesem Fall sogar eine Formel für das Volumen von $\Phi(B)$ angegeben werden kann. Formal sollte man in diesem Fall zuerst die Funktion $f = \mathbb{1}_B |\det(\mathrm{D}\Phi^{-1})|$ auf $Y$ für eine Jordan-messbare, kompakte Teilmenge $B \subseteq Y$ betrachten, woraus man erhält, dass $\Phi^{-1}(B)$ Jordan-messbar ist. Da $\Phi$ aber ein Diffeomorphismus, können wir dies auch für $\Phi^{-1} : Y \to X$ anwenden und erhalten die Behauptung.

In der Diskussion beim Abendessen gab es zuerst eine kleine Verwirrung: „Ist dies nicht klar?“, gefolgt von einer Lösung zu Übung **??**. In der Tat ist das Verhalten von Nullmengen unter Abbildungen oft wichtig, weswegen die Frage eben (ähnlich zu der oft zu aggressiven Rechtschreibungsautokorrektur mancher Computerprogramme) anders gehört wurde, als sie gemeint war.

Aufbauend auf dem Result in Übung **??** war der nächste relevante Gedanke in der Diskussion, dass das wahre Problem ist, den Zusammenhang zwischen dem Bild des Randes und dem Rand des Bildes zu verstehen (welche im Allgemeinen nicht gleich sein müssen). Dann erkannten wir, dass dies im eindimensionalen Fall kein Problem darstellt: Falls ein Intervall $B = I$ in der offenen Menge $X \subseteq \mathbb{R}$ liegt, dann ist $\Phi(B)$ wieder ein Intervall und insbesondere Jordan-messbar.

**Übung D.3.** *Beweisen Sie für $n = 1$, dass das Bild einer Jordan-messbaren Teilmenge von $\mathbb{R}$ unter einer Lipschitz stetigen Abbildungen wieder eine Jordan-messbare Menge ist.*

*Hinweis: Approximieren Sie $B$ von Innen und von Aussen durch endliche Vereinigungen von Intervallen.*

Das ist aber leider schon der einzige Fall, den wir mit den Methoden der Analysis-Vorlesung verstehen konnten. Nach einigem Hin und Her, wobei auch der Vorschlag kam, wir sollten doch versuchen in einem Buch wie zum Beispiel „Counterexamples in Analysis“ ([GO03]) die Antwort zu finden, wurden wir immer überzeugter, dass die Antwort in der Tat sogar für eine Lipschitz-Abbildung positiv ist.

Beim Ende des Abendessens hatten wir eine Beweisskizze für $C^1$-Abbildungen, welche auf Sard's Theorem beruhte: Falls $B = Q \subseteq X$ ein abgeschlossener Quader ist und $y \in \partial\Phi(B)$ ist, dann ist $y = \Phi(x)$ für ein $x \in \partial B$ oder für ein $x \in B^\circ$ (da das Bild $\Phi(B)$ kompakt und insbesondere abgeschlossen ist). Nach Übung **??** ist $\Phi(\partial B)$ eine Nullmenge. Aus Satz 11.5 für $x \in B^\circ$ und $y = \Phi(x) \in \partial\Phi(B)$ folgt nun, dass $x$ ein kritischer Punkt von $\Phi$ ist. Sard's Theorem besagt aber, dass die Menge

$$\begin{aligned}&\{y \in \Phi(B) \mid y \text{ ist ein kritischer Wert}\}\\ &\qquad = \{y \in \Phi(B) \mid y = \Phi(x) \text{ für einen kritischen Punkt } x \in B\}\end{aligned}$$

eine Lebesgue-Nullmenge ist. Durch Approximation des Bildes $\Phi(B)$ von Innen und von Aussen durch endliche Vereingungen von Quadern folgt die Behauptung für Jordan-messbare Mengen und $C^1$-Abbildungen.

Für den Lipschitz-Fall $\Phi : X \to Y$ waren wir uns bei der Verabschiedung etwas uneinig darüber, ob wir eine komplette Lösung hatten, aber hier sind die wichtigsten Gedanken: Nach Rademacher's Theorem ist $\Phi$ fast überall differenzierbar und Sard's Theorem sollte eine Verallgemeinerung für Lipschitz-Abbildungen besitzen. Wenn dem so ist, dann sollte die Jordan-Messbarkeit des Bildes $\Phi(Q)$ eines Quaders $Q$ daraus folgen: Für $y \in \partial\Phi(Q)$ gilt $y \in \Phi(\partial Q)$ oder $y$ ist ein kritischer Wert oder $y = \Phi(x)$ für einen Punkt $x \in Q$, bei dem $\Phi$ nicht differenzierbar ist. Da das Bild einer Lebesgue-Nullmenge unter einer Lipschitz-Abbildung nach Übung **??** wieder eine Nullmenge ist und $\partial\Phi(Q)$ kompakt ist, würde wahrscheinlich aus der Verallgemeinerung von Sard's Theorem für Lipschitz-Abbildungen die gewünschte Jordan-Messbarkeit von $\Phi(Q)$ folgen. Der allgemeinere Fall der Jordan-Messbarkeit des Bildes $\Phi(B)$ einer Jordan-messbaren Menge $B \subseteq Y$ folgt dann wieder durch Approximation von Innen und von Aussen.

Was lernen wir aus obiger Anekdote?

- Es ist einfach, in der Mathematik beliebig schwierige Fragen bereits mit sehr grundlegenden Begriffen zu stellen. Die Antworten (falls vorhanden) verwenden vielleicht viel mehr Theorie, als zum Stellen der Frage notwendig war.

- Die genaue Beantwortung der gestellten Frage geht auf jeden Fall über den Stoff der Analyis I- und II-Vorlesungen hinaus, weswegen wir diese hier nicht weiter ausbauen.

- Für die typische Mathematikerin / den typischen Mathematiker sind Fragen zur Jordan-Messbarkeit etwas schwieriger, da dieser Begriff mit wenigen Ausnahmen in den meisten weiterführenden Sätzen durch den Begriff der Lebesgue-Messbarkeit ersetzt wird und wir daher über letzteren häufiger nachdenken.

- Es ist gefährlich in Mathematiker-Kreisen eine mathematische Frage, deren Antwort noch nicht klar ist, bei einer Party zu stellen, denn dann könnten einige Teilnehmer etwas asozial während der Party nur mehr an das Problem denken.

# Anhang E

# Prüfungsrelevant

Wir wollen hier dem Wunsch von vielen Studenten nachkommen und etwas genauer erklären welche Teile der Vorlesung und Übungserien prüfungsrelevant sind. Die Definition von „prüfungsrelevant“ ist hierbei, dass die jeweiligen Themen und Fragen in dieser oder ähnlicher Form bei einer Prüfung auftauchen könnte. Da die Probleme der Serien bewusst eine Bandbreite an Schwierigkeitsgrade enthält, macht es eben Sinn den relevanten Teil der Übungsserien kennzuzeichnen. Die restlichen Fragen in den Serien sind aber meist „lernrelevant“, da sie das Verständnis der diskutierten Themen fördern.

Die Übungen im Skript werden wir nicht auf dieselbe Weise kennzeichnen obwohl viele von ihnen auch prüfungsrelevant sind. Sie sind gerne dazu eingeladen auf dem eSkript diese zu diskutieren. Bitte immer in dem Forum „Public“ kommentieren.

Wie bereits in Vorlesung und Vorlesungswebseite darauf hingewiesen wird die Prüfung zu Analysis I&II aus drei Teilen bestehen, die mit jeweils der gleichen Punkteanzahl bewertet werden:

1. Rechnungen
2. Anwendungen der Theorie
3. Theorie aus der Vorlesung

Dieses Format können Sie bereits in den drei vorherigen (Probe-)Prüfungen einsehen (siehe die Vorlesungswebseite).

## E.1 Rechnungen und Anwendungen der Theorie

Bei den Teilen

1. Rechnungen
2. Anwendungen der Theorie

geht es um Probleme wie in den Übungen. Bei gewissen Übungsserien gab es fast ausschliesslich Probleme vom Typ 1 oder fast ausschliesslich Probleme vom Typ 2, dies wurde durch die

Vorlesung bedingt, die in gewissen Wochen die Theorie weiter ausbaute und erst nach Erreichen eines gewissen Niveaus wieder Rechnungen erlaubte.

Im Rechnungsteil müssen Sie nicht auf die gleiche Weise alle Schritte mathematisch exakt formulieren und beweisen. Zum Beispiel müssen Sie bei trigonometrischer Substitution in einer Rechnung nicht auf das potentielle Problem, dass sich ein uneigentliches Integral ergeben könnte, achten. Oder auch bei Auffinden eines Extremwertes mittels Lagrange Multiplikatoren müssen Sie nicht zuerst zeigen, dass Sie eine Mannigfaltigkeit vorliegen haben, aber es kann natürlich vorkommen, dass sich ein Extremwert bei einem Punkt versteckt, wo die Teilmenge eben keine Teilmannigfaltigkeit ist.

Im Teil Anwendungen der Theorie müssen Sie hingegen mathematisch exakt vorgehen und wenn nötig obige Feinheiten in ihrer Lösung erklären.

### E.1.1 Erstes Semester, Übungen

Die Aufgaben des ersten Semesters finden sich hier. Dabei entsprechen folgende Aufgaben im Stile und Schwierigkeitsgrad in etwa einer Aufgabe der Prüfung:

- Übungsblatt 1 Aufgabe 4
- Übungsblatt 2 Aufgaben 2,3
- Übungsblatt 3 Aufgabe 5
- Übungsblatt 4 Aufgaben 1,3,4,6
- Übungsblatt 5 Aufgaben 4,6
- Übungsblatt 6 Aufgaben 1,2,3,4,6
- Übungsblatt 7 Aufgaben 3,4,5b),6
- Übungsblatt 8 Aufgaben 1-6
- Übungsblatt 9 Aufgaben 1,2,4,5
- Übungsblatt 10 Aufgaben 2-5
- Übungsblatt 11 Aufgaben 2,3,5
- Übungsblatt 12 Aufgaben 1,2,3,5
- Übungsblatt 13 Aufgaben 1-6

### E.1.2 Zweites Semester, Übungen

Die Aufgaben des zweiten Semesters finden Sie hier. Dabei entsprechen folgende Aufgaben im Stile und Schwierigkeitsgrad in etwa einer Aufgabe der Prüfung:

- Übungsblatt 1 Aufgabe 1,2,3,5,6

- Übungsblatt 2 Aufgaben 1,3,4,6
- Übungsblatt 3 Aufgabe 2,4,5a)
- Übungsblatt 4 Aufgaben 1,4
- Übungsblatt 5 Aufgaben 1,2,3,5
- Übungsblatt 6 Aufgaben 1-6
- Übungsblatt 7 Aufgaben 1-4
- Übungsblatt 8 Aufgaben 1,4,6
- Übungsblatt 9 Aufgaben 1,2,3,5a)
- Übungsblatt 10 Aufgaben 1,4
- Übungsblatt 11 Aufgaben 1,2,3,4
- Übungsblatt 12 Aufgaben 1,3,4,6
- Übungsblatt 13 Aufgaben 1,3,4,5
- Übungsblatt 14 Aufgaben 1,2,3a),4
- Übungsblatt 15 Aufgaben 1,2,3,4,6

## E.2 Vorlesung und Theorieteil

Im dritten Teil der Liste

1. Rechnungen
2. Anwendungen der Theorie
3. Theorie aus der Vorlesung

geht es tatsächlich um Definitionen, Sätze, und Beweise aus der Vorlesung, und wie diese zusammengehören.

Wir haben alle Teile, die über den Analysis-Prüfungsstoff hinausgehen, mit einem Stern $*$ markiert, siehe zum Beispiel Kapitel 1 (ihr Ausdruck ist hier nicht notwendigerweise auf dem aktuellen Stand). Insbesondere sind manche der Beweise in der Vorlesung nicht bis ins letzte Detail besprochen worden (siehe zum Beispiel das Ende von Abschnitt 12.6).

Es gibt aber auch Teile des Skripts und der Vorlesung vor allem am Anfang, die auf Grund der Fülle der Themen eine geringere Chance haben in der Prüfung Platz zu finden. Wir werden diese Themen aber nicht kennzeichnen, da diese im Aufbau der Theorie notwendig waren. Des Weiteren sind gewisse Themen im Nachhinein viel einfacher, zum Beispiel ist die Integration von Polynomen in Abschnitt 4.6 nachträglich viel einfacher mit Hilfe des Fundamentalsatzes, analog werden wir nicht nach der Berechnung eines Grenzwertes fragen mit der Auflage die

Regel von de l'Hôpital nicht zu verwenden womit viele der ersten Grenzwertberechnungen erheblich vereinfacht werden.

Aus diesen Gründen ist es beim Lernen für die Prüfung nötig mit einem Überblick zu beginnen und darauf zu achten nicht bei den ersten Themen stecken zu bleiben.

10 Tage vor der Prüfung werden Sie einen Katalog mit Aussagen (Sätze, Propostionen und so weiter) aus der Vorlesung erhalten, die im Theorie-Teil der Prüfung (Teil 3) auftauchen können. Dieser wird aus höchstens 40 Aussagen bestehen. Dabei sollte man beachten, dass wenn beispielsweise Satz 7.48 auf der Liste steht, dass die Addresse 7.48 im Skript völlig irrelevant ist, und dass dann nicht nur nach **Formulierung und Beweis des Erweiterten Mittelwertsatzes** gefragt werden kann, sondern auch verwandte Fragen auftauchen können, wie zum Beispiel:

- Q: Welche Eigenschaften von Funktionen tauchen im Erweiteren Mittelwertsatz auf? Definieren Sie diese.
  A: Stetigkeit: ... und Differenzierbarkeit: ...
- Q: Formulieren Sie das wichtigste Hilfsmittel im Beweis des Erweiteren Mittelwertsatz (ohne Beweis).
  A: Der Satz von Rolle: ....
- Q: Wo war der Erweiterte Mittelwertsatz wichtig?
  A: Zum Beispiel im Beweis von der Regel von de l'Hôpital.
- Q: Ist die Vorraussetzung der Differenzierbarkeit im Inneren in der Formulierung des Erweiterten Mittelwertsatzes bei jedem Punkt notwendig?
  A: Ja, ein Gegenbeispiel ist gegeben durch den Absolutbetrag $f(x) = |x|$ und die Funktion $g(x) = x$ auf dem Intervall $[-1, 1]$.
- Q: Beschreiben Sie in Worten die wichtigste(n) Idee(n) im Beweis des Erweiteren Mittelwertsatzes.
  A: Durch Definition einer geeigneten Hilfsfunktion kann man die Aussage auf den Satz von Rolle zurückführen.

Bei längeren Beweisen kann mitunter auch, nach klarer Formulierung der Vorraussetzungen und der gewünschten (Teil-)Aussage, nach einem Teil des Beweises gefragt werden.

Wie bereits in der Vorlesung erwähnt soll dieser Katolog Ihnen helfen den Stress direkt vor der Prüfung etwas zu reduzieren. Idealerweise verwendet man den Katalog und die letzten 10 Tage zum Festigen des bereits Erlernten und dem allfälligen Schliessen von einigen wenigen Lücken im Wissen. Ein gutes Verständnis der Theorie ist auch daran ersichtlich, dass man weiss warum, zumindest bei den wichtigsten Sätzen, gewisse Vorraussetzungen gemacht wurden (siehe obige Frage nach einem Gegenbeispiel).

Bitte beachten Sie, dass der Katalog nicht den Prüfungsteil „Anwendungen der Theorie" (Teil 2) einschränkt. Genauso ist es möglich, dass im zweiten Teil der Prüfung Themen aus gesternten Abschnitten des Skripts gefragt werden, wobei dann die vorherige Lektüre des jeweiligen Abschnitts nicht vorausgesetzt wird.

Wir möchten an dieser Stelle anmerken, dass an der Winterprüfung 2018/2019 die genau gleichen Rahmenbedingungen gelten werden. Insbesondere werden wir einen Katalog gleicher Länge (nicht denselben) zehn Tage vor der Prüfung zur Verfügung stellen.

### E.2.1 Probeprüfung und Katalog

Als Beispiel davon, wie ein solcher Katalog aussehen könnte, möchten wir Ihnen hier einen Katalog für die Probeprüfung im Stile des Katalogs der Prüfung geben.

- Kapitel 2: Satz 2.15, Satz 2.59, Satz 2.68.
- Kapitel 3: Satz 3.28, Satz 3.59, Satz 3.70.
- Kapitel 4: Proposition 4.12, Satz 4.26.
- Kapitel 5: Proposition 5.17, Satz 5.44, Abschnitt 5.3.2, Satz 5.79.
- Kapitel 6: Proposition 6.16, Proposition 6.25, Satz 6.49.
- Kapitel 7: Proposition 7.40, Satz 7.48.
- Kapitel 8: Theorem 8.2, Lemma 8.32, Satz 8.46, Satz 8.68 (b).
- Kapitel 9: Proposition 9.37, Satz 9.72, Proposition 9.79, Theorem 9.81.
- Kapitel 10: Satz 10.13, Korollar 10.17, Satz 10.36, Satz 10.49.
- Kapitel 11: Satz 11.1, Proposition 11.12, Satz 11.21.
- Kapitel 12: Satz 12.13, Korollar 12.27, Proposition 12.71.
- Kapitel 13: Proposition 13.2, Lemma 13.37.
- Kapitel 14: Proposition 14.15, Proposition 14.25, Proposition 14.41.

Die Nummerierung bezieht sich hier auf die PDF-Version des Skripts vom 1. Juni 2018.

# Anhang F

# Summer-Special 2017:

## Der Igelsatz und der Brouwersche Fixpunktsatz*

Wir wollen hier grösstenteils einer Arbeit [Mil78] von John Milnor (Fields-Medaillenträger von 1962) aus 1978 folgen und zwei interessante geometrisch-topologische Aussagen beweisen, die üblicherweise Teil der algebraischen Topologie und der Differentialtopologie sind. Im Rahmen dieser beiden Themengebiete ergibt sich natürlich auch eine vollständigere Erklärung, als wir hier bewerkstelligen können.

### F.1 Der Satz des gekämmten Igels

Ein (idealisierter) gekämmter Igel besteht aus einem Vektorfeld auf der Kugeloberfläche, welches bei jedem Punkt $p$ die Lage und die Richtung der Stacheln angibt. Hierbei wollen wir aber nur die Situation zulassen, in der die Richtung der Stacheln an einem Punkt $p$ in der Tangentialebene der Kugeloberfläche bei $p$ liegen (der Igel soll ja gekämmt sein). Des Weiteren bestehen wir darauf, dass das Vektorfeld zumindest stetig ist. Wir wollen folgende Aussage beweisen:

Jeder stetig gekämmte Igel hat mindestens einen Glatzpunkt.

Ein „Glatzpunkt“ ist eine Stelle, bei der die Länge der Stachel gleich Null ist; also ein Punkt auf der Kugeloberfläche, wo das Vektorfeld verschwindet.

Eine physikalische Interpretation der obigen Aussage ist durch das „Problem des globalen Windes“ gegeben: Die (Ableitung nach der Zeit der) Windströmungen auf der Erde zu einem gegebenen Zeitpunkt definieren ein stetiges Vektorfeld auf der Erdoberfläche. Die Aussage ist somit, dass es zu jedem Zeitpunkt einen Ort auf der Erde gibt, wo Windstille herrscht.

Für die mathematische Formulierung ist es von Vorteil, nicht nur dreidimensionale Igel und Erden zu betrachten. Sei also $n \geq 2$ und

$$\mathbb{S}^{n-1} = \{p \in \mathbb{R}^n \mid \|p\| = 1\}$$

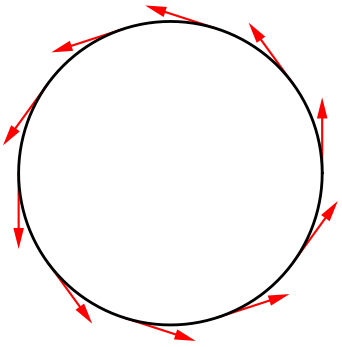
die $(n-1)$-dimensionale Sphäre. Ein **stetiges Tangentenfeld** auf $\mathbb{S}^{n-1}$ ist eine stetige Funktion

$$f : p \in \mathbb{S}^{n-1} \mapsto f(p) \in p^{\perp} \subseteq \mathbb{R}^n,$$

wobei wir $f$ mitunter mit der Abbildung $p \in \mathbb{S}^{n-1} \mapsto (p, f(p)) \in \mathrm{T}_p\mathbb{S}^{n-1}$ identifizieren. Wir sagen, dass das Tangentenfeld **nirgends verschwindet**, falls $f(p) \neq 0$ für alle $p \in \mathbb{S}^{n-1}$ gilt. Der folgende Satz wurde von Poincaré formuliert und von L. E. J. Brouwer (1881-1966) im Jahr 1910 bewiesen, siehe [Bro11].

**Theorem F.1** (Igelsatz, engl. Hairy Ball Theorem)**.** *Sei $n \geq 2$. Es gibt genau dann ein nirgends verschwindendes, stetiges Tangentenfeld auf der Sphäre $\mathbb{S}^{n-1}$, wenn $n$ gerade ist.*

Die „erste Häfte“ des Beweises ist relativ einfach und für $n = 2$ durch folgendes Bild gegeben.



Allgemeiner können wir für $n = 2k$ zu $k \in \mathbb{N}$ ein nirgends verschwindendes Tangentenfeld $f : \mathbb{S}^{n-1} \to \mathrm{T}\mathbb{S}^{n-1}$ durch

$$f(p) = (-p_2, p_1, -p_4, p_3, \ldots, -p_{2k}, p_{2k-1})$$

für $p = (p_1, p_2, p_3, p_4, \ldots, p_{2k-1}, p_{2k}) \in \mathbb{S}^{n-1}$ definieren.

Die „zweite Hälfte“ ist erheblich schwieriger, da wir alle möglichen stetigen Tangentenfelder auf $\mathbb{S}^{n-1}$ betrachten müssen. Wir werden den Beweis in Abschnitt F.5 auf folgende Übung zurückführen.

**Übung F.2.** *Zeigen Sie, dass die Funktion $t \in [0,1] \mapsto \left(\sqrt{1+t^2}\right)^n$ genau dann durch ein Polynom gegeben ist, wenn $n \in \mathbb{N}$ gerade ist.*

## F.2 Der Brouwersche Fixpunktsatz

Der folgende Satz wurde 1910 von Hadamard [Had10] und Brouwer [Bro11] bewiesen.

**Theorem F.3** (Brouwerscher Fixpunktsatz)**.** *Sei $m \in \mathbb{N}$. Eine stetige Funktion $T : \mathbb{D}^m \to \mathbb{D}^m$ auf der abgeschlossenen Einheitskugel $\mathbb{D}^m = \overline{B_1(0)} \subseteq \mathbb{R}^m$ hat stets einen Fixpunkt (also ein $p \in \mathbb{D}^m$ mit $T(p) = p$).*

Der Spezialfall $m = 1$ ist ein Korollar eines Satzes aus dem ersten Semester.

**Übung F.4.** *Beweisen Sie Theorem F.3 im Fall $m = 1$.*

Für $m \geq 2$ gibt es genügend anschauliche Erklärungen. Zum Beispiel gibt es beim Umrühren eines Getränkes stets einen Punkt der Flüssigkeit, der sich vor und nach dem Umrühren an derstelben Stelle befindet. Eine weitere Veranschaulichung erhält man, wenn man zwei gleich grosse Kreisscheiben aus Papier ausschneidet, eine der beiden auf willkürliche Art und Weise zusammenfaltet und anschliessend auf die andere Kreisscheibe legt (ohne über den Rand zu reichen). Hier besagt obiger Fixpunktsatz, dass es einen Punkt in der gefalteten Kreisscheibe gibt, der nach der Faltung in der neuen Kreisscheibe die gleichen Koordinaten besitzt wie vor der Faltung.

## F.3 Beweis des Brouwerschen Fixpunktsatzes via dem Igelsatz

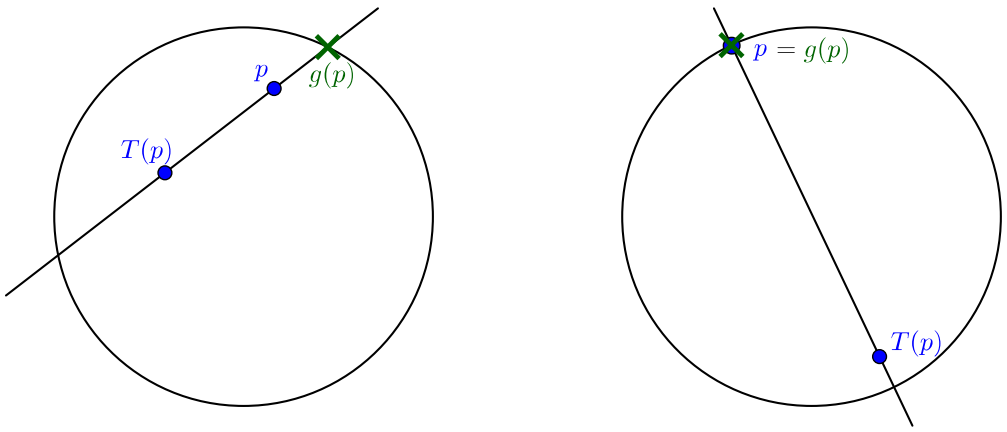
In diesem Abschnitt möchten wir unter Annahme des Igelsatzes (Theorem F.1) den Brouwerschen Fixpunktsatz (Theorem F.3) beweisen. Hierfür aber auch für den Beweis im nächsten Unterabschnitt wird folgendes Lemma nützlich sein.

**Lemma F.5** (Igelperücke, Retraktion)**.** *Sei $m \geq 2$ und $T : \mathbb{D}^m \to \mathbb{D}^m$ eine stetige Abbildung. Falls $T$ keine Fixpunkte besitzt, dann gibt es ein nirgends verschwindendes stetiges Vektorfeld $g$ auf $\mathbb{D}^m$ mit Werten in $\mathbb{S}^{m-1}$ und $g(p) = p$ für alle $p \in \mathbb{S}^{m-1}$. Falls des Weiteren $T$ eine glatte Fortsetzung auf einer offenen Obermenge von $\mathbb{D}^m$ hat, so gilt dies auch für $g$.*

Wir können uns das Vektorfeld als eine vorgekämmte Igelperücke vorstellen, und werden das Vektorfeld unten auch auf diese Weise verwenden um damit eine Kämmung der oberen Halbspäre zu erzielen, die am Rand alle Stacheln nach unten zeigen lässt. Ganz genau können wir uns die Kämmung der Perücke nicht vorstellen, da diese eigentlich nicht existieren sollte.

Dieser Typ an Abbildung wird in der Algebraischen Topologie auch eine Retraktion genannt, da sie gewissermassen die Einheitskugel $\mathbb{D}^m$ auf stetige Art und Weise auf die Sphäre $\mathbb{S}^{m-1}$ zurückzieht.

Wir wollen zuerst für $m = 2$ ein geometrisches Argumente für die Konstruktion von $g$ geben: Da $T$ keinen Fixpunkt besitzt, können wir für jeden Punkt eine eindeutig bestimmte Gerade durch $p$ und $T(p)$ finden, welche den Kreis $\mathbb{S}^1 = \partial\mathbb{D}^2$ in genau zwei Punkten schneidet; siehe folgendes Bild.



Zu $p \in \mathbb{D}^2$ definieren wir $g(p) \in \mathbb{S}^1$ als jenen Schnittpunkt der Geraden durch $p$ und $T(p)$ mit dem Einheitskreis $\mathbb{S}^1$, der auf der Seite von $p$ liegt. In Formeln definieren wir die Richtung

$$v(p) = \frac{1}{\|p - T(p)\|}\big(p - T(p)\big),$$

in die wir gehen wollen und den Punkt

$$g(p) = p + \Big(\sqrt{1 - \|p\|^2 + \langle p, v(p)\rangle^2} - \langle p, v(p)\rangle\Big)v(p)$$

Eine Rechnung zeigt nun, dass $g(p) \in \mathbb{S}^1$ für $p \in \mathbb{D}^2$ den oben beschrieben Punkt definiert und dass $g(p) = p$ für alle $p \in \mathbb{S}^1$. Diese Definitionen sind auch für beliebige $m \geq 2$ sinnvoll, doch wollen wir ein etwas anderes Argument ausführen.

*Beweis.* Sei $T : \mathbb{D}^m \to \mathbb{D}^m$ stetig so dass $T(p) \neq p$ für alle $p \in \mathbb{D}^m$ gilt. Wir definieren für $p \in \mathbb{D}^m$ den Vektor

$$f(p) = p - \frac{1 - \langle p, p\rangle}{1 - \langle p, T(p)\rangle}T(p).$$

Da $\langle p, T(p)\rangle \leq \|p\|\|T(p)\| \leq 1$ für alle $p \in \mathbb{D}^m$ ist und Gleichheit nur gelten kann, wenn $\|p\| = 1$ und $T(p) = p$ zutrifft, definiert dies ein stetiges Vektorfeld auf $\mathbb{D}^m$. Des Weiteren ist $f(p) \neq 0$ für alle $p \in \mathbb{D}^m$. Dies folgt direkt aus der Definition, wenn $p$ und $T(p)$ linear unabhängig sind. Sind $p$ und $T(p)$ linear abhängig, so gilt $\langle p, p\rangle\, T(p) = \langle p, T(p)\rangle\, p$ und

$$\begin{aligned} f(p) &= \frac{1}{1 - \langle p, T(p)\rangle}\big((1 - \langle p, T(p)\rangle)p - (1 - \langle p, p\rangle)T(p)\big) \\ &= \frac{1}{1 - \langle p, T(p)\rangle}(p - T(p)) \neq 0, \end{aligned}$$

da per Annahme $T$ keinen Fixpunkt hat. Schlussendlich bemerken wir noch, dass für $p \in \mathbb{D}^m$ mit $\|p\| = 1$ die Gleichung

$$f(p) = p$$

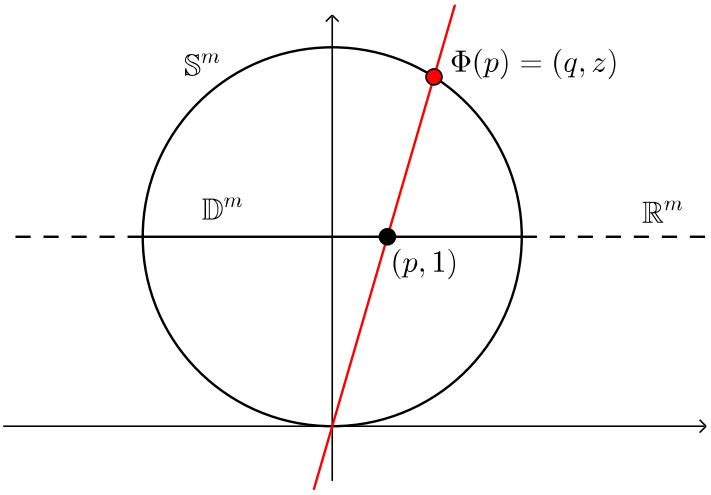
gilt. Schlussendlich definieren wir $g(p) = \|f(p)\|^{-1}f(p)$ für alle $p \in \mathbb{D}^m$, was nun die gewünschten Eigenschaften erfüllt.

Falls $T$ auf einer offenen Obermenge von $\mathbb{D}^m$ definiert und glatt ist, so können wir obige Formeln auch auf einer (möglicherweise etwas kleineren) offenen Obermenge von $\mathbb{D}^m$ für die Definition einer glatten Abbildung $g$ verwenden. □

Informell besteht der Beweis des Brouwerschen Fixpunktsatzes mittels des Igelsatzes aus dem „Zusammenfügen von zwei Igelperücken“.

*Beweis von Theorem F.3.* Wir betrachten zuerst den Fall $m = 2k$ für $k \geq 1$ und setzen $n = m + 1$. Nach dem Igelsatz (Theorem F.1) gibt es somit kein nirgends verschwindendes Tangentenfeld $\mathbb{S}^m \subseteq \mathbb{R}^n$.

Wir wollen das „hypothetische Vektorfeld“ $g$ von Lemma F.5 nun von $\mathbb{D}^m$ auf die nördliche Hemisphäre von $\mathbb{S}^m$ verschieben. Zur Vereinfachung der Notation verschieben wir die Sphäre $\mathbb{S}^m$, so dass $e_n$ der Mittelpunkt von $\mathbb{S}^m$ ist, das heisst, $\mathbb{S}^m = \{q \in \mathbb{R}^n \mid \|q - e_n\| = 1\}$. Wir betrachten nun die **stereographische Projektion** $\Phi : \mathbb{R}^m \to \mathbb{S}^m$ als Parametrisierung der oberen Hemisphäre, welche wie in folgendem Bild gegeben ist.




Eine kurze Rechnung zeigt, dass die stereographische Projektion durch

$$\Phi : p \in \mathbb{R}^m \to \left(\tfrac{2}{1+\|p\|^2}p, \tfrac{2}{1+\|p\|^2}\right) \in \mathbb{S}^m$$

gegeben ist. Die Abbildung $\Psi$, welche durch

$$\Psi(q, z) = \tfrac{1}{z}q \in \mathbb{R}^m$$

für $(q, z) \in \mathbb{R}^n$ mit $z \neq 0$ definiert ist, erfüllt $\Psi \circ \Phi = \mathrm{Id}$. Da sowohl $\Phi$ als auch $\Psi$ differenzierbar sind, erhalten wir mit der Kettenregel (Satz 10.13) $\mathrm{D}_{\Phi(p)}\Psi \cdot \mathrm{D}_p\Phi = I_m$ und damit hat $\mathrm{D}_p\Phi$ für alle $p \in \mathbb{R}^m$ trivialen Kern. Insbesondere ist $\{\Phi(p)\} \times \mathrm{D}_p\Phi(\mathbb{R}^m) = \mathrm{T}_{\Phi(p)}\mathbb{S}^m$ (wieso?).

Wir definieren für $(q, z) \in \mathbb{S}^{n-1}$ mit $z \geq 1$ nun das Tangentenfeld

$$f(q, z) = \big(\mathrm{D}_{\Psi(q,z)}\Phi\big)g(\Psi(q, z)).$$

In Worten verwenden wir also den nicht-verschwindenden Vektor $g(p)$ bei dem Punkt $p \in \mathbb{D}^m$, der $(q, z)$ entspricht, gemeinsam mit der Ableitung $\mathrm{D}_p\Phi$ der Parametrisierung, um das Vektorfeld bei $(q, z)$ zu definieren. (Dies ist eine Standardprozedur, wenn man ein Vektorfeld auf einer Karte via einer Parametrisierung auf ein Stück einer Teilmannigfaltigkeit schieben möchte.)

Für $p \in \partial\mathbb{D}^m$ mit $\|p\| = 1$ gilt des Weiteren $g(p) = p$. Um daraus $f(p, 1)$ zu berechnen, betrachten wir den Weg $\gamma : t \mapsto tp$ und erhalten

$$f(p, 1) = (\Phi \circ \gamma)'(1) = \tfrac{\mathrm{d}}{\mathrm{d}t}\left(\tfrac{2t}{1+t^2}p, \tfrac{2t}{1+t^2}\right)\Big|_{t=1} = (0, -1).$$

Also zeigt das Tangentenfeld am Äquator direkt nach unten. Wir können $f$ auf ganz $\mathbb{S}^m$ fortsetzen, indem wir die obere Hemisphäre und das Tangentenfeld nach unten spiegeln, aber anschliessend das Tangentenfeld auf der unteren Hemisphäre mit $-1$ multiplizieren. Dies bewirkt, dass die Tangentenfelder auf der unteren und der oberen Hemisphäre beide am Äquator jeweils den Wert $(0, -1)$ annehmen. Wir können diese also zu einem nirgends verschwindenden Tangentenfeld $f$ auf $\mathbb{S}^m$ zusammenfügen. Dies widerspricht dem Igelsatz (Theorem F.1). Also muss $T$ einen Fixpunkt besitzen.

Wir betrachten nun den Fall, wo $m$ ungerade ist. Sei also $m = 2k + 1$ für $k \geq 1$ und $T : \mathbb{D}^m \to \mathbb{D}^m$ eine stetige Abbildung. Wir definieren nun eine stetige Abbildung

$$\tilde{T} : \mathbb{D}^{m+1} \to \mathbb{D}^{m+1}$$

durch

$$\tilde{T}(p, z) = (T(p), 0)$$

für $p \in \mathbb{D}^m$ und $z \in [-1, 1]$ mit $(p, z) \in \mathbb{D}^{m+1}$. Nach obigem besitzt $\tilde{T}$ einen Fixpunkt $(p, z) = \tilde{T}(p, z) = (T(p), 0)$ und daher gilt dasselbe auch für $T$. □

**Übung F.6.** *Für gerade Dimensionen $m$ hat obiger Beweis gezeigt, dass die Retraktion aus Lemma F.5 nicht existieren kann. Verwenden Sie nun den Satz von Brouwer (Theorem F.3) um dies für alle $m \in \mathbb{N}$ zu zeigen.*

*Hinweis: Betrachten Sie die Abbildung $T(p) = p - \delta g(p)$ für $p \in \mathbb{D}^m$ und ein geeignet gewähltes $\delta > 0$.*

## F.4 Beweis des Brouwerschen Fixpunktsatzes via dem Divergenzsatz im Zwei- und Dreidimensionalen

Wir wollen hier einen zweiten Beweis des Brouwerschen Fixpunktsatz geben, der anstatt des Igelsatzes die Divergenzsätze verwendet.

**Lemma F.7** (Reduktion auf glatte Abbildungen)**.** *Für den Beweis von Theorem F.3 genügt es für jedes $m \in \mathbb{N}$ glatte Abbildungen $T$ zu betrachten, die auf einer Obermenge von $\mathbb{D}^m$ definiert sind und $T(\mathbb{D}^m) \subseteq \mathbb{D}^m$ erfüllen.*

*Beweis.* Sei $m \geq 1$ und $T : \mathbb{D}^m \to \mathbb{D}^m$ eine stetige Abbildung. Wir wollen annehmen, dass jede glatte Abbildung $\tilde{T}$, welche auf einer offenen Obermenge von $\mathbb{D}^m$ definiert ist, Werte in $\mathbb{R}^m$ hat und $\tilde{T}(\mathbb{D}^m) \subseteq \mathbb{D}^m$ erfüllt, einen Fixpunkt besitzt und daraus ableiten, dass auch $T$ einen Fixpunkt besitzt. Um ausgehend von $T$ derartige glatte Abbildungen $\tilde{T}$ zu finden, bedienen wir uns der Faltung im Abschnitt 13.6.1, welche wir auf die stetige Fortsetzung

$$\overline{T} : p \in \mathbb{R}^m \mapsto \begin{cases} T(p) & \text{falls } p \in \mathbb{D}^m \\ T(\|p\|^{-1}p) & \text{falls } p \notin \mathbb{D}^m \end{cases}$$

der Abbildung $T$ anwenden wollen.

Wir definieren $\tilde{T}_k = \psi_k * \overline{T}$, wobei $\psi_k : \mathbb{R}^m \to \mathbb{R}_{\geq 0}$ wie in Proposition 13.57 eine glatte Abbildung mit kompaktem Träger in $B_{\frac{1}{k}}(0)$ und mit $\int_{\mathbb{R}^m} \psi_k \,\mathrm{d}\,\mathrm{vol} = 1$ ist. Nach Proposition 13.55 ist $\tilde{T}_k$ für jedes $k \geq 1$ glatt. Nach Proposition 9.77 und Proposition 13.57 strebt $\tilde{T}_k$ eingeschränkt auf $\mathbb{D}^m$ gleichmässig gegen $T$ für $k \to \infty$. Des Weiteren gilt für $p \in \mathbb{R}^m$

$$\|\tilde{T}_k(p)\| = \left\| \int_{\mathbb{R}^m} \psi(p-v)\overline{T}(v) \,\mathrm{d}\,\mathrm{vol}(v) \right\| \leq \int_{\mathbb{R}^m} \psi(p-v) \underbrace{\|\overline{T}(v)\|}_{\leq 1} \,\mathrm{d}\,\mathrm{vol}(v) \leq 1,$$

das heisst $\tilde{T}_k$ bildet $\mathbb{D}^m$ wieder nach $\mathbb{D}^m$ ab.

Per Annahme existiert für jedes $k \in \mathbb{N}$ ein Fixpunkt $p_k \in \mathbb{D}^m$ der Abbildung $\tilde{T}_k$. Da $\mathbb{D}^m$ kompakt ist, existiert eine konvergente Teilfolge $(p_{k_\ell})_\ell$ der Folge $(p_k)_k$ mit Grenzwert $q \in \mathbb{D}^m$. Auf Grund der gleichmässigen Konvergenz folgt nun, dass

$$q = \lim_{\ell \to \infty} p_{k_\ell} = \lim_{\ell \to \infty} \tilde{T}_{k_\ell}(p_{k_\ell}) = \lim_{\ell \to \infty} T(p_{k_\ell}) = T(q),$$

was zu beweisen war. □

Wie schon angekündigt, möchten wir nun den Brouwerschen Fixpunktsatz mittels Divergenzsatz und Lemma F.5 beweisen.

*Beweis von Theorem F.3 für $m = 2$.* Sei nun $T$ eine glatte Abbildung auf einer offenen Obermenge von $\mathbb{D}^2$ mit Werten in $\mathbb{R}^2$ so dass $T(\mathbb{D}^2) \subseteq \mathbb{D}^2$. Wir wollen annehmen, dass $T$ keinen Fixpunkt in $\mathbb{D}^2$ besitzt und werden daraus einen Widerspruch ableiten.

Ausgehend von $g = \left(\begin{smallmatrix} g_1 \\ g_2 \end{smallmatrix}\right)$ wie in Lemma F.5 definieren wir das Vektorfeld

$$f = g_1 \begin{pmatrix} \partial_2 g_2 \\ -\partial_1 g_2 \end{pmatrix}$$

und erhalten

$$\begin{aligned} \operatorname{div}(f) &= \partial_1(g_1 \partial_2 g_2) - \partial_2(g_1 \partial_1 g_2) = \partial_1 g_1 \partial_2 g_2 + g_1 \partial_1 \partial_2 g_2 - \partial_2 g_1 \partial_1 g_2 - g_1 \partial_2 \partial_1 g_2 \\ &= \det \begin{pmatrix} \partial_1 g_1 & \partial_2 g_1 \\ \partial_1 g_2 & \partial_2 g_2 \end{pmatrix} = \det(\mathrm{D}g). \end{aligned}$$

Da aber $g$ nur Werte in $\mathbb{S}^1$ annimmt, folgt aus dem Satz über die inverse Abbildung (Satz 11.5), dass $\det(\mathrm{D}g) = 0$ gilt. Daher ist $f$ divergenzfrei und es gilt

$$\int_{\mathbb{S}^1} f \cdot \mathrm{d}\mathbf{n} = \int_{\mathbb{D}^2} \operatorname{div}(f) \,\mathrm{d}\,\mathrm{vol} = 0.$$

Das Flussintegral von $f$ ist aber gleich

$$\begin{aligned}
\int_{\mathbb{S}^1} f \cdot \mathrm{d}\mathbf{n} &= \int_0^{2\pi} \left\langle f\begin{pmatrix}\cos(t)\\ \sin(t)\end{pmatrix}, \begin{pmatrix}\cos(t)\\ \sin(t)\end{pmatrix}\right\rangle \mathrm{d}t \\
&= \int_0^{2\pi} \cos(t) \left\langle \begin{pmatrix}\partial_2 g_2(\cos(t),\sin(t))\\ -\partial_1 g_2(\cos(t),\sin(t))\end{pmatrix}, \begin{pmatrix}\cos(t)\\ \sin(t)\end{pmatrix}\right\rangle \mathrm{d}t \\
&= \int_0^{2\pi} \cos(t) \left\langle \begin{pmatrix}\partial_1 g_2(\cos(t),\sin(t))\\ \partial_2 g_2(\cos(t),\sin(t))\end{pmatrix}, \begin{pmatrix}-\sin(t)\\ \cos(t)\end{pmatrix}\right\rangle \mathrm{d}t \\
&= \int_0^{2\pi} \cos(t) \frac{\mathrm{d}}{\mathrm{d}t}\Big(\underbrace{g_2\begin{pmatrix}\cos(t)\\ \sin(t)\end{pmatrix}}_{=\sin(t)}\Big) \mathrm{d}t \\
&= \int_0^{2\pi} \cos^2(t)\, \mathrm{d}t = \pi.
\end{aligned}$$

Dieser Widerspruch zeigt, dass $T$ doch einen Fixpunkt in $\mathbb{D}^2$ besitzen muss. □

*Beweis-Idee von Theorem F.3 für $m = 3$.* Für eine Abbildung $T : \mathbb{D}^3 \to \mathbb{D}^3$ können wir wiederum $g$ wie in Lemma F.5 verwenden. Das Vektorfeld $f$ definiert man in diesem Fall durch

$$f = g_1 \begin{pmatrix}\partial_2 g_2 \partial_3 g_3 - \partial_3 g_2 \partial_2 g_3\\ \partial_3 g_2 \partial_1 g_3 - \partial_1 g_2 \partial_3 g_3\\ \partial_1 g_2 \partial_2 g_3 - \partial_2 g_2 \partial_1 g_3\end{pmatrix}.$$

Eine Rechnung zeigt nun wiederum, dass $\operatorname{div}(f) = \det(\mathrm{D}g) = 0$. Andererseits gilt aber $g(p) = p$ für alle $p \in \mathbb{S}^2$ und man kann dies benützen um

$$\int_{\mathbb{S}^2} f \cdot \mathrm{d}\mathbf{n} = \operatorname{vol}(\mathbb{D}^3) \neq 0$$

zu zeigen, wobei man den drei-dimensionalen Divergenzsatz (für $\tilde{f}$ definiert durch $\tilde{g} = I$) verwenden kann um die Rechnung abzukürzen. Dies führt wieder zu einem Widerspruch und damit muss der Brouwer Fixpunktsatz auch für $m = 3$ gelten. □

*Bemerkung.* Die Definition der Vektorfelder in den obigen Beweisen für $m = 2$ und $m = 3$ ist in unserer Darstellung etwas mysteriös. Verwenden wir stattdessen Differentialformen, so kann man in allen Dimensionen $m \geq 2$ mittels

$$\omega = g_1\, \mathrm{d}g_2 \wedge \ldots\, \mathrm{d}g_m$$

einen einheitlichen Ausdruck angeben, welcher die Rolle des Vektorfeldes $f$ übernimmt. Gemeinsam mit der entsprechenden Formulierung des Satzes von Stokes kann man dann einen Beweis des Brouwerschen Fixpunktsatz in allen Dimensionen angeben.

## F.5 Beweis des Igelsatzes

Im Folgenden sei $n \geq 2$.

**Lemma F.8** (Glatte Version auf Kugelschale)**.** *Falls es ein stetiges nirgends verschwindendes Tangentenfeld auf $\mathbb{S}^{n-1}$ gibt, dann existiert auch eine glatte Funktion $f : \mathbb{R}^n \to \mathbb{R}^n$, so dass $\|f(p)\| = \|p\|$ und $f(p) \perp p$ für alle $p$ in der Kugelschale*

$$K = \{q \in \mathbb{R}^n \mid 1 \leq \|q\| \leq 2\}\,.$$

*Beweis.* Sei $g : \mathbb{S}^{n-1} \to \mathbb{R}^n$ ein stetiges nirgends verschwindendes Tangentenfeld auf $\mathbb{S}^{n-1}$. Wir definieren das Vektorfeld $\tilde{g} : \mathbb{R}^n \to \mathbb{R}^n$ durch

$$p \in \mathbb{R}^n \mapsto \tilde{g}(p) = \begin{cases} 0 & \text{falls } p = 0, \\ \|p\| g(\|p\|^{-1} p) & \text{falls } p \neq 0. \end{cases}$$

Dies definiert eine stetige Fortsetzung von $g$ auf ganz $\mathbb{R}^n$. Als nächstes definieren wir

$$\tilde{f} = \psi * \tilde{g}$$

mit Hilfe einer glatten Funktion $\psi$ mit kompaktem Träger in $B_\delta(0)$ wie in Proposition 13.57. Nach Proposition 13.55 ist $\tilde{f}$ glatt und nach Proposition 13.57 können wir erreichen, dass $\tilde{f}$ gleichmässig nahe an $\tilde{g}$ ist, wenn wir $\delta > 0$ geeignet wählen.

Wir definieren nun eine weitere Abbildung $\overline{f}$ durch

$$\overline{f}(p) = \tilde{f}(p) - \left\langle \tilde{f}(p), p \right\rangle p$$

für alle $p \in \mathbb{R}^n$. Insbesondere erhalten wir, dass $\overline{f}(p)$ für $p \in \mathbb{S}^{n-1}$ gleichmässig und beliebig nahe an $\tilde{g}(p) - \langle \tilde{g}(p), p \rangle\, p = g(p) \neq 0$ ist. Da $\mathbb{S}^{n-1}$ kompakt ist, ist

$$\min \left\{ \|g(p)\| \mid p \in \mathbb{S}^{n-1} \right\} > 0$$

und wir können $\delta$ klein genug wählen, so dass $\overline{f}$ auf $\mathbb{S}^{n-1}$ ebenfalls nirgends verschwindet.

Sei nun $\psi_1 : \mathbb{R} \to \mathbb{R}$ eine glatte Funktion mit $\psi_1(t) = 1$ für alle $t \in [1, 2]$ und mit Träger in $[\frac{1}{2}, 3]$ (welche man aus Satz 13.60 für $K = [1, 2]$ und $O_1 = (\frac{1}{2}, 3)$ erhält). Dann erfüllt das Vektorfeld $f$ gegeben durch

$$f(p) = \begin{cases} 0 & \text{falls } p = 0 \\ \frac{\psi_1(\|p\|) \cdot \|p\|}{\|\overline{f}(\|p\|^{-1} p)\|} \overline{f}(\|p\|^{-1} p) & \text{falls } p \neq 0 \end{cases}$$

für $p \in \mathbb{R}^n$ alle Aussagen des Lemmas. □

**Proposition F.9** (Diffeomorphismus zwischen Kugelschalen)**.** *Seien $f : \mathbb{R}^n \to \mathbb{R}^n$ und die Kugelschale $K$ wie in Lemma F.8. Es existiert ein $T > 0$, so dass für alle $t \in (0, T)$ die Funktion $\Phi_t$ definiert durch*

$$p \in B_3(0) \mapsto \Phi_t(p) = p + t f(p) \in \Phi_t(B_3(0))$$

*ein Diffeomorphismus ist. Des Weiteren ist $\Phi_t$ orientierungserhaltend und $\Phi_t(K) = \sqrt{1+t^2}\,K$ ist die Kugelschale mit innerem Radius $\sqrt{1+t^2}$ und äusserem Radius $2\sqrt{1+t^2}$.*

*Beweis.* Für $t \in \mathbb{R}$ ist die Ableitung von $\Phi_t$ wie in der Proposition für alle $p \in B_3(0)$ durch

$$\mathrm{D}_p\Phi_t = I + t\,\mathrm{D}_p f$$

gegeben. Wir wählen $T = \frac{1}{2}M^{-1}$, wobei

$$M = \max_{p\in\overline{B_3(0)}} \|\,\mathrm{D}_p f\|,$$

und erhalten, dass $\mathrm{D}_p\Phi_t$ für alle $t \in (0,T)$ und alle $p \in \overline{B_3(0)}$ auf Grund von $\|t\,\mathrm{D}_p f\| \leq \frac{1}{2}$ invertierbar ist. Für $p, q \in \overline{B_3(0)}$ gilt des Weiteren

$$\|f(p) - f(q)\| \leq M\|p-q\|$$

und daher

$$\|\Phi_t(p) - \Phi_t(q)\| = \|p - q + t(f(p) - f(q))\| \geq \|p-q\| - tM\|p-q\| \geq \tfrac{1}{2}\|p-q\|.$$

Dies zeigt, dass $\Phi_t$ auf $B_3(0)$ injektiv ist und damit nach Korollar 11.8 einen Diffeomorphismus $\Phi_t : B_3(0) \to \Phi_t(B_3(0))$ definiert.

Sei nun $p \in K$. Dann gilt $\|f(p)\| = \|p\|$ und $f(p) \perp p$ nach Lemma F.8. Daher folgt

$$\|\Phi_t(p)\| = \|p + tf(p)\| = \sqrt{\|p\|^2 + t^2\|f(p)\|^2} = \sqrt{1+t^2}\,\|p\|.$$

Dies zeigt die Inklusionen

$$\Phi_t(r\mathbb{S}^{n-1}) \subseteq \sqrt{1+t^2}\,r\mathbb{S}^{n-1} \tag{F.1}$$

für alle $r \in [1,2]$ sowie

$$\Phi_t(K) \subseteq \sqrt{1+t^2}\,K.$$

Um die Gleichheit $\Phi_t(K) = \sqrt{1+t^2}\,K$ zu zeigen, bemerken wir zuerst, dass für $r \in (1,2)$ die Menge $\Phi_t(r\mathbb{S}^{n-1})$ eine relativ offene Teilmenge von $\sqrt{1+t^2}\,r\mathbb{S}^{n-1}$ ist. In der Tat ist $\Phi_t(K^\circ)$ eine offene Teilmenge von $\mathbb{R}^n$, da $\Phi_t$ ein Diffeomorphismus ist. Wegen (F.1) für alle $r \in (1,2)$ gilt dann, dass

$$\Phi_t(r\mathbb{S}^{n-1}) = \sqrt{1+t^2}\,r\mathbb{S}^{n-1} \cap \Phi_t(K^\circ)$$

in $\sqrt{1+t^2}\,r\mathbb{S}^{n-1}$ relativ offen ist. Da aber $\mathbb{S}^{n-1}$ kompakt ist, ist nach Satz 9.72 dass Bild $\Phi_t(r\mathbb{S}^{n-1})$ auch kompakt und somit abgeschlossen. Da $\mathbb{S}^{n-1}$ jedoch zusammenhängend ist (was zum Beipiel mit $n$-dimensionalen Kugelkoordinaten gezeigt werden kann), muss jede

nicht-leere offene und abgeschlossene Teilmenge die ganze Menge sein und es folgt

$$\Phi_t(r\mathbb{S}^{n-1}) = \sqrt{1+t^2}\, r\mathbb{S}^{n-1}$$

für alle $r \in (1,2)$. Daher gilt $\Phi_t(K^\circ) = \sqrt{1+t^2}\, K^\circ$. Da $K = \overline{K^\circ}$ kompakt ist, folgt mit

$$\Phi_t(K) \supseteq \Phi_t(K^\circ) = \sqrt{1+t^2}\, K^\circ,$$

dass $\Phi_t(K) = \sqrt{1+t^2}\, K$ wie behauptet. □

**Korollar F.10** (Volumenformel)**.** *Sei $n \geq 1$ und sei $K$ die Kugelschale $\{p \in \mathbb{R}^n \mid 1 \leq \|p\| \leq 2\}$. Angenommen es gibt ein stetiges nirgends verschwindendes Tangentenfeld auf $\mathbb{S}^{n-1}$. Dann gibt es ein $T > 0$, so dass das Volumen von $\sqrt{1+t^2}\, K$ ein Polynom in $t \in (0,T)$ (und Grad kleiner oder gleich $n$) ist.*

*Beweis.* Nach Annahme gibt es ein Vektorfeld $f : \mathbb{R}^n \to \mathbb{R}^n$ wie in Lemma F.8. Seien $T > 0$ und $\Phi_t$ für $t \in (0,T)$ wie Proposition F.9. Wir berechnen das Volumen von $\sqrt{1+t^2}\, K = \Phi_t(K)$ durch Substitution (siehe Satz 12.56). Dazu definieren wir $X = B_3(0)$, $Y = \Phi_t(B_3(0))$ und $f = \mathbb{1}_{\Phi_t(K)}$ und erhalten

$$\operatorname{vol}(\Phi_t(K)) = \int_Y \mathbb{1}_{\Phi_t(K)} \,\mathrm{d}\operatorname{vol} = \int_X \mathbb{1}_K \det(\mathrm{D}_p\Phi_t)\,\mathrm{d}\operatorname{vol}(p) = \int_K \det(I + t\,\mathrm{D}_p f)\,\mathrm{d}\operatorname{vol}(p)$$

Expandieren wir $\det(I + t\,\mathrm{D}_p f)$, so erhalten wir ein Polynom in $t$ mit Koeffizienten, die stetig von $p \in K$ abhängen. Durch Integration über $p \in K$ folgt das Korollar. □

*Beweis des Igelsatzes (Theorem F.1).* Sei $n = 2k+1$ ungerade. Angenommen es existiert ein stetiges, nirgends verschwindendes Tangentenfeld auf $\mathbb{S}^{n-1}$. Aus Korollar F.10 folgt nun, dass $\operatorname{vol}(\sqrt{1+t^2}\, K)$ ein Polynom in $t \in (0,T)$ ist. Wenden wir die lineare Substitution in Proposition 12.49 an, so ergibt sich stattdessen

$$\operatorname{vol}\left(\sqrt{1+t^2}\, K\right) = (\sqrt{1+t^2})^n \operatorname{vol}(K) = \sqrt{1+t^2}\,(1+t^2)^k \operatorname{vol}(K).$$

Da dies aber kein Polynom in $t \in (0,T)$ ist, erhalten wir einen Widerspruch und der Igelsatz folgt. □

## F.6 Erweiterungen und Anwendungen

Der Brouwersche Fixpunktsatz hat ebenso wie der Banachsche Fixpunktsatz zahlreiche Anwendungen, auch ausserhalb der Analysis und der Mathematik. Zum Beispiel wurde er von John Nash in 1950 für den Beweis des Nash Gleichgewichts [Nas50], [Nas51] verwendet, welches für Spieltheorie und Ökonomie von zentraler Bedeutung ist. Obwohl sich die Ökonomie weiterentwickelt hat, sind auch manche neuere Gleichgewichtssätze Korollare des Brouwerschen Fixpunktsatzes.

Eine amüsantere (aber irrelevantere) Anwendung ist durch folgende Beschreibung gegeben. Angenommen ein Boot fährt auf dem Zürichsee und macht mehrere starke Kurven. Sie kennen

den Weg des Bootes schon zu Beginn und wollen den ganzen Weg über stehend schlafen. Dann gibt es eine ideale Orientierung und Neigung zum Boden, so dass Sie ohne weiter auf ihre Haltung zu achten bis zum Ende der Bootsfahrt nicht umfallen. Dies folgt gewissermassen aus dem zweidimensionalen Satz von Brouwer. Sei $p \in \mathbb{D}^2$ ihre Ausgangsneigung. Die Schwerkraft und jede Kurve bewirken über die Bootsreise gewisse Änderungen der Neigung. Die Punkte in $\mathbb{S}^1 = \partial\mathbb{D}^2$ entsprechen einem „in einer bestimmten Richtung Hinlegen oder Fallen“. Wir wollen davon ausgehen, dass man sich nach dem Hinfallen nicht mehr bewegt. Dadurch erhält man eine Abbildung $f : \mathbb{D}^2 \mapsto \mathbb{D}^2$, die am Rand mit der Identität übereinstimmt und welche die Ausgangsneigung auf die Endneigung abbildet. Nach dem Beweis in Abschnitt F.3 muss ein derartiges Vektorfeld eine Nullstelle besitzen. Also gibt es eine Ausgangsneigung so dass man am Ende der Reise genau senkrecht steht.

Wir erwähnten in Abschnitt 13.1.6 den Jordanschen Kurvensatz und betrachteten den Fall von glatten Wegen. Für stetige Funktionen ist der Satz aber überraschend schwierig zu beweisen; einer der kürzesten Beweise verwendet den Brouwerschen Fixpunktsatz als zentrales Hilfsmittel, siehe [Mae84].

Eine weitere Anwendung besteht aus einem sehr kurzen Beweis des allgemeinen Satz von Perron-Frobenius (siehe auch Abschnitt 9.4.2): Wenn $A \in \mathrm{Mat}_{n,n}(\mathbb{R})$ eine nichtnegative Matrix ist so dass jede Spalte ungleich Null ist, dann gibt es einen Eigenvektor $v$ mit nichtnegativen Eintragungen. Hierzu setzt man $S = \{v \in [0,1]^n \mid \|v\|_1 = 1\}$ und $T(v) = \|Av\|_1^{-1} Av \in S$ für alle $v \in S$. Da es topologisch keinen Unterschied zwischen dem Simplex $S$ und der Kreisscheibe $\mathbb{D}^{n-1}$ gibt, besagt der Brouwersche Fixpunktsatz in diesem Fall, dass es einen Eigenvektor von $A$ in $S$ gibt.

Der Brouwersche Fixpunktsatz wurde in vielen Richtungen weiter verallgemeinert. Der Kakutani Fixpunktsatz betrachtet mengenwertige Funktionen auf kompakten konvexen Teilmengen von $\mathbb{R}^n$ und hat neben mathematischen Anwendungen ebenso Anwendungen in Spieltheorie und Ökonomie.

Das Satz von Borsuk-Ulam [Bor33] nimmt eine andere Richtung und besagt, dass es zu jeder stetigen Funktion $f : \mathbb{S}^m \to \mathbb{R}^m$ einen Punkt $p \in \mathbb{S}^m$ mit $f(p) = f(-p)$ gibt. Also gibt es zum Beispiel zu jedem Zeitpunkt einen Punkt $p$ auf der Erde, so dass an der genau gegenüberliegenden Stelle auf der Erdoberfläche (also in $-p$) die gleiche Temperatur und der gleiche Luftdruck herrscht wie bei $p$.

Zu guter Letzt wollen wir noch kurz eine Verschärfung vom Igelsatz erwähnen: Für welche $m$-dimensionale Mannigfaltigkeiten $M$ gibt es $m$ stetige nirgends verschwindendende Tangentenfelder, so dass die Werte dieser Tangentenfelder bei jedem Punkt eine Basis des Tangentialraums bilden? (Man spricht in diesem Fall von einem trivialen Tangentenbündel, da damit das Tangentenbündel in einem geeigneten Sinn zu $M \times \mathbb{R}^m$ isomorph ist.) Spezieller wollen wir fragen, für welche $m \geq 2$ dies für die $m$-dimensionalen Sphären $\mathbb{S}^m$ zutrifft. Die Antwort mag etwas überraschen, denn wie von Kervaire [Ker58] (1927-2007, ETH-Absolvent) gleichzeitig und getrennt von Bott und Milnor [BM58] in 1958 gezeigt wurde, ist dies genau für $m = 1$, $m = 3$, und $m = 7$ der Fall. Diese Fälle sind mit gewissen algebraischen Strukturen (genauer den komplexen Zahlen, den Hamiltonsche Quaternionen und den Oktonionen) verknüpft, doch gibt es keine weiteren Sphären mit dieser Eigenschaft.

Wir hoffen, dass diese Sätze Ihnen klar veranschaulichen, dass die Mathematik des letzten Jahrhundert auch sehr interessante und mitunter überraschend relevante Aussagen enthält, die auf Grund des benötigten Vorwissen allerdings nicht in dem gleichen Ausmass bekannt sind wie die Differential- und Integralrechnung.

# Liste der Symbole


$\mathcal{P}(X)$, Potenzmenge einer Menge $X$, 20

$\mathbb{1}_A$, charakteristische Funktion einer Teilmenge $A$ einer Menge, 22

$\mathbb{R}$, Menge der reellen Zahlen, 68

$\mathbb{N}$, Menge der natürlichen Zahlen, 80

$\mathbb{N}_0$, Menge der nicht-negativen ganzen Zahlen, 84

$\mathbb{Z}$, Menge der ganzen Zahlen, 85

$\mathbb{Q}$, Menge der rationalen Zahlen, 86

$\mathbb{C}$, Menge der komplexen Zahlen, 89

$\mathbb{K}[T]$, Polynomring über einem Körper $\mathbb{K}$ in der Variablen $T$, 136

$\mathcal{S}_n$, Menge der Permutationen auf der Menge $\{1, \ldots, n\}$ (symmetrische Gruppe), 142

$\mathcal{F}(D)$, Menge der reellwertigen Funktionen auf einer Menge $D$, 151

$C(D)$, Menge der stetigen reellwertigen Funktionen auf der Teilmenge $D \subseteq \mathbb{R}$, 157

$\mathcal{F}_{\mathbb{C}}(D)$, Menge der komplexwertigen Funktionen auf einer Menge $D$, 160

$\mathcal{TF}([a, b])$, Menge der Treppenfunktionen auf einem Intervall $[a, b]$ mit Endpunkten $a < b$, 182

$\mathcal{R}([a, b])$, Menge der Riemann-integrierbaren Funktionen auf einem Intervall $[a, b]$ mit Endpunkten $a < b$, 185

$C^k(I)$, Menge der $k$-mal stetig differenziebaren Funktionen auf einem Intervall $I$ mit Endpunkten $a < b$ für $k \in \mathbb{N}$ oder $k = \infty$, 355

$\mathrm{Mat}_{d,d}(\mathbb{K})$, Menge der $d \times d$-Matrizen über einem Körper $\mathbb{K}$, 756

$\mathrm{GL}_d(\mathbb{K})$, Gruppe der invertierbaren $d \times d$-Matrizen über einem Körper $\mathbb{K}$, 756

# Literaturverzeichnis


[A'C03] A'Campo, N.: *A natural construction for the real numbers.* arXiv preprint, 2003.

[AE06] Amann, H. und Escher, J.: *Analysis I.* Grundstudium Mathematik. Birkhäuser Basel, 3. Auflage, 2006.

[Apo83] Apostol, T.: *A proof that Euler missed: Evaluating $\zeta(2)$ the easy way.* The Mathematical Intelligencer, 5:59–60, 1983.

[Beu09] Beutelspacher, A.: *Das ist o.B.d.A trivial.* Vieweg+Teubner Verlag, 2009.

[Bla55] Blanuša, D.: *Über die Einbettung hyperbolischer Räume in euklidische Räume.* Monatsh. Math., 59:217–229, 1955.

[Bla03] Blatter, C.: *Analysis I*, 2003. `https://people.math.ethz.ch/~blatter/dlp.html`.

[BM58] Bott, R. und Milnor, J.: *On the parallelizability of spheres.* Bull. Amer. Math. Soc, 64(8789):87–89, 1958.

[Boo47] Boole, G.: *The mathematical analysis of logic.* Philosophical library, 1847.

[Bor33] Borsuk, K.: *Drei Säetze über die $n$-dimensionale euklidische Sphäre.* Fundamenta Mathematicae, 1(20):77–90, 1933.

[Bro11] Brouwer, L.E.J.: *Über die Abbildungen von Mannigfaltigkeiten.* Mathematische Annalen, 71(1):97–115, 1911.

[Can95] Cantor, G.: *Beiträge zur Begründung der transfiniten Mengenlehre.* Mathematische Annalen, 46(4):481–512, 1895.

[Coo49] Coolidge, J. L.: *The story of the binomial theorem.* The American Mathematical Monthly, 56(3):147–157, 1949.

[GO03] Gelbaum, B. und Orsted, J.: *Counterexamples in Analysis.* Courier Corporation, 2003.

[Had10] Hadamard, J.: *Sur quelque applications de l'indice de Kronecker.* in Jules Tannery: Introduction à la théorie des fonctions d'une variable, 2:433–477, 1910.

[Hil93] Hilbert, D.: *Über die Transzendenz der Zahlen* e *und* $\pi$. Mathematische Annalen, 43:216–219, 1893.

[Ker58] Kervaire, M.: *On the non-parallelizability of the $n$-sphere for $n > 7$.* Proceedings of the National Academy of Sciences, 44(3):280–283, 1958.

[Lin94] Lindelöf, E.: *Sur l'application des méthodes d'approximations successives à l'étude des intégrales réelles des équations différentielles ordinaires.* Journal de mathématiques pures et appliquées, 10(4):117–128, 1894.

[Mae84] Maehara, R.: *The Jordan Curve Theorem Via the Brouwer Fixed Point Theorem.* The American Mathematical Monthly, 91(10):641–643, 1984.

[Mil78] Milnor, J.: *Analytic Proofs of the „Hairy Ball Theorem" and the Brouwer Fixed Point Theorem.* The American Mathematical Monthly, 85(7):521–524, 1978.

[Nas50] Nash, J.: *Equilibrium points in $n$-person games.* Proceedings of the National Acadamy of Sciences, 36(1):48–49, 1950.

[Nas51] Nash, J.: *Non-cooperative games.* Annals of Mathematics, 54(2):286–295, 1951.

[Niv47] Niven, I.: *A simple proof that $\pi$ is irrational.* Bull. Amer. Math. Soc., 53(6):509, 1947.

[Pea90] Peano, G.: *Sur une courbe, qui remplit toute une aire plane.* Mathematische Annalen, 36(1):157–160, 1890.

[RU08] Remmert, R. und Ullrich, P.: *Elementare Zahlentheorie.* Grundstudium Mathematik. Birkhäuser Basel, 3. Auflage, 2008.

[Rus03] Russell, B.: *The principles of mathematics.* WW Norton & Company, 1903.

[Smu78] Smullyan, R.: *What is the name of this book?* Prentice-Hall, 1978.

[SS12] Schichl, H. und Steinbauer, R.: *Einführung in das mathematische Arbeiten.* Springer Spektrum, 2. Auflage, 2012.

[Wal04] Walter, W.: *Analysis 1.* Springer-Verlag Berlin Heidelberg, 7. Auflage, 2004.

[Web92] Weber, H.: *Leopold Kronecker.* Jahresbericht der Deutschen Mathematiker-Vereinigung, 2, 1892.

[Zag90] Zagier, D.: *A one-sentence proof that every prime $p \equiv 1 \mod 4$ is a sum of two squares.* Amer. Math. Monthly, 97(2):144, 1990.

[Zor07] Zorich, V.: *Analysis II.* Springer-Verlag, 2007.