Eidgenössische Technische Hochschule Zürich

Begleitendes Skript zur Linearen Algebra I HS2020

Menny Akka Ginosar

2. Januar 2021

Inhaltsverzeichnis

Kapitel 0

Einführung

In dieser Einführung werden wir (mittels linearer Algebra) ein Problem lösen, das mich als Jugendlicher genervt hat.

In der Mittelschule haben wir Folgen kennengelernt. Es gibt die arithmetische Folge: Für zwei Zahlen a, d definiert man

$$\begin{cases} a_0 &= a \\ a_n &= a_{n-1} + d, \quad n \geq 1 \end{cases}.$$

Hier kann man fast problemlos eine Formel für das n-te Folgenglied finden. Konkret heisst das, dass wir einen Ausdruck für das n-te Folgenglied finden, der nur von n abhängt. Der Ausdruck ist gegeben durch $a_n = a + nd$ für $n \geq 0$ und wenn wir wollten, könnten wir problemlos a_{12345} auch ohne Taschenrechner in weniger als einer halben Minute berechnen. Eine ähnliche Geschichte haben wir in der Mittelschule mit der geometrischen Folge gehabt: Seien a, q zwei Zahlen, dann definieren wir

$$\begin{cases} a_0 &= a \\ a_n &= a_{n-1}q, \quad n \geq 1 \end{cases}.$$

In diesem Fall ist eine Formel für a_n mit n als Variable durch $a_n = aq^n$ für $n \geq 0$ gegeben.

Für mich hat der interessante Teil des Stoffs hier aufgehört. Man „beweist“ dann in der Regel viele Identitäten, die fast tautologisch sind, betrachtet die Summenfolge (das ist vielleicht doch interessant). Man kann jedoch viele interessante Fragen zum Beispiel über arithmetische Folgen stellen. Hier sind einige:

1. Sei $a_n = a + nd$ eine arithmetische Folge. Enthält die Folge unendlich viele Primzahlen? Das heisst, gibt es unendlich viele $n \in \mathbb{N}$, so dass a_n eine Primzahl ist?

   Angenommen a, d sind nicht teilerfremd (d.h. es gibt $l \geq 1$, so dass l sowohl a als auch d teilt), dann kann die Folge höchstens eine Primzahl enthalten.

Theorem 0.1 (Dirichlet, 1837). *Seien $a, d \in \mathbb{N}, a \geq 1, d > 1$ teilerfremd. Dann enthält die Folge $a_n = a + nd$ unendlich viele Primzahlen*[1].

Für den Beweis des Satzes brauchen Sie mehr oder weniger einen Bachelor in Mathematik (dies könnte zum Beispiel das Thema Ihrer Bachelorarbeit sein). Falls Sie nicht bis dahin warten wollen, dann können Sie hier [11] anfangen.

2. Wegen Dirichlets Theorem (Theorem 0.1) könnten wir uns fragen, wann die erste Primzahl in einer arithmetischen Folge vorkommt. Im Jahre 1944 hat Yuri Vladimirovich Linnik, ein Mathematiker, der meine eigene Forschung stark beeinflusst, bewiesen, dass es Konstanten $c, L > 0$ gibt, so dass die erste Primzahl in einer arithmetischen Folge $a + nd$ mit a, d teilerfremd (und $1 \leq a < d$) kleiner als cd^L ist. Nehmen Sie sich ein bisschen Zeit, um diese Aussage richtig zu verstehen. Hier ist eine Vermutung[2], dass obiges Theorem von Linnik mit $c = 1, L = 2$ gilt. Mehr dazu hier.

3. Jetzt können wir uns fragen, ob wir eine „Kette" von Primzahlen in irgendeiner arithmetischen Folge finden können. Was genau damit gemeint ist, können sie hier nachlesen.

Diese Fragen und Theoreme sind schön, aber deren Lösungen sind nicht direkt mit linearer Algebra verbunden. Um näher zur linearen Algebra zu kommen, betrachten wir eine andere berühmte Folge, die Fibonacci-Folge. Diese ist folgendermassen definiert:

$$\begin{cases} a_0 &= 0 \\ a_1 &= 1 \\ a_n &= a_{n-1} + a_{n-2}, \quad n \geq 2 \end{cases} \tag{0.1}$$

Wenn wir diese mit den oben definierten Folgen vergleichen, stellt sich sofort die Frage, ob wir einen Ausdruck/eine Formel für das n-te Glied der Fibonacci-Folge finden können. Es ist eigentlich nicht offensichtlich, dass es so eine Formel überhaupt gibt, aber als Jugendlicher habe ich gehört, dass es eine Formel für das n-te Glied der Fibonacci-Folge geben soll. Es konnte mir jedoch niemand sagen, wie diese Formel aussieht oder wie man sie findet (in diesen uralten Zeiten gab es noch kein Wikipedia...). Wir werden jetzt diese Formel zusammen finden und dabei fast allen Begriffen und Themen von diesem Kurs (zumindest vom ersten Semester) begegnen. Der erste Schritt in diesem Abenteuer ist sehr unintuitv; wir verkomplizieren die Sache. Wir haben keine Ahnung

[1]Zum Beispiel gibt es unendlich viele Primzahlen der Form $1 + 4k$ für $k \in \mathbb{N}$ und unendlich viele der Form $3 + 4k$ für $k \in \mathbb{N}$.

[2]Eine Vermutung ist eine Aussage, von welcher die mathematische Gemeinschaft glaubt, dass sie richtig ist, aber die bis anhin noch niemand beweisen konnte.

wie wir dieses Problem lösen können und betrachten dazu noch unendlich viele (verwandte) Probleme, die wir ebenfalls nicht lösen können. . . ziemlich verrückt! Um diese neuen Probleme einführen zu können, entwickeln wir ein bisschen „Sprache“. Das heisst, wir führen einige Begriffe und Definitionen ein. Wir werden jetzt Folgen als Objekte betrachten und sie mit schönen kalligrafischen Buchstaben bezeichnen. Zum Beispiel schreiben wir: Sei $\mathcal{F}$ die Fibonacci-Folge, die in (0.1) definiert ist.

Definition 0.2. Wir sagen, dass wir eine Folge *gut kennen*, wenn wir eine Formel für ihr n-tes Folgenglied gefunden haben.

Zum Beispiel kennen wir alle arithmetischen und geometrischen Folgen gut. Wir können jetzt unser Ziel so ausdrücken: Wir möchten gerne $\mathcal{F}$ gut kennen. Wie versprochen, machen wir die Sache komplizierter:

Definition 0.3. Seien $a, b \in \mathbb{R}$ zwei reelle Zahlen. Wir definieren die Folge $\mathcal{F}_{a,b}$ mittels der Rekursion

$$\begin{cases} a_0 = a \\ a_1 = b \\ a_n = a_{n-1} + a_{n-2}, \text{ für } n \geq 2 \end{cases}$$

Zum Beispiel gilt $\mathcal{F} = \mathcal{F}_{0,1}$. Also können wir unser Ziel so ausdrücken: Wir möchten $\mathcal{F}_{0,1}$ gut kennen. Der Schlüssel für die Lösung ist nun das verrückte neue Ziel:

Neues Ziel: Wir möchten $\mathcal{F}_{a,b}$ für alle $a, b \in \mathbb{R}$ gut kennen.

Wie kann es sein, dass wir mehr Hoffnung für dieses neue Ziel haben als für das ursprüngliche Ziel? Zunächst klingt das verrückt, da das neue Ziel so aussieht wie unendlich viele Varianten des alten Problems. Das neue Problem ist ein Raum von Problemen und hier liegt eigentlich der Hund begraben. Dieser Raum hat eine gewisse Struktur und wir können diese Struktur ausnutzen. Bevor wir anfangen, geben wir noch eine Definition an:

Definition 0.4. Eine Folge $\mathcal{A}$ heisst eine *Fibonacci-Folge*, wenn es $a, b \in \mathbb{R}$ gibt, so dass $\mathcal{A} = \mathcal{F}_{a,b}$. Den Raum aller Fibonacci-Folgen nennen wir **Fib**.

Übung 0.5. Zeigen Sie, dass eine Folge $(a_0, a_1, \ldots)$ eine Fibonacci-Folge ist genau dann, wenn $a_n = a_{n-1} + a_{n-2}$ für alle $n \geq 2$.

Es gilt

$$\mathbf{Fib} = \{\mathcal{F}_{a,b} \mid a, b \in \mathbb{R}\}$$

und unser Ziel ist es, alle Elemente von **Fib** gut zu kennen.

0.1 Welche Struktur hat Fib?

Wir nehmen zwei Folgen $\mathcal{F}_1 = (a_0, a_1, a_2, a_3, \ldots)$, $\mathcal{F}_2 = (b_0, b_1, b_2, b_3, \ldots)$ und formen ihre komponentenweise Addition

$$\mathcal{F}_1 + \mathcal{F}_2 := (a_0 + b_0, a_1 + b_1, a_2 + b_2, a_3 + b_3, \ldots).$$

Behauptung: Seien $\mathcal{F}_1$ und $\mathcal{F}_2$ zwei Fibonacci-Folgen, dann ist auch $\mathcal{F}_1 + \mathcal{F}_2$ eine Fibonacci Folge. Wir bezeichnen die Glieder von $\mathcal{F}_1 + \mathcal{F}_2$ als $\mathcal{F}_1 + \mathcal{F}_2 = (c_0, c_1, c_2, c_3, \ldots)$. Laut Übung 2.20 müssen wir lediglich zeigen, dass

$$c_n = c_{n-1} + c_{n-2}$$

für alle $n \geq 2$ gilt. Beweisen wir es: Es gilt

$$c_n = a_n + b_n$$

und da $\mathcal{F}_1, \mathcal{F}_2 \in \textbf{Fib}$ haben wir

$$a_n + b_n = (a_{n-1} + a_{n-2}) + (b_{n-1} + b_{n-2}) = (a_{n-1} + b_{n-1}) + (a_{n-2} + b_{n-2}) = c_{n-1} + c_{n-2}.$$

Daher gilt in der Tat $c_n = c_{n-1} + c_{n-2}$.

Übung 0.6. Mit der Notation von Definition 0.3 erklären Sie für sich selbst, dass das obige Argument Folgendes zeigt:

$$\mathcal{F}_{a,b} + \mathcal{F}_{c,d} = \mathcal{F}_{a+c,b+d}.$$

Kurz gesagt, können wir zwei Elemente von **Fib** addieren. Der Fibonacci-Folgen Raum **Fib** hat also eine Addition. Ausser der Addition haben wir noch eine andere Operation in **Fib**: eine Multiplikation mit einem Skalar. In diesem Zusammenhang ist „Skalar“ einfach ein komischer Name für eine reelle Zahl. Später in diesem Kurs möchten wir andere „Typen“ von Zahlen betrachten, die einen sogenannten Körper formen (vielleicht haben Sie schon vom Körper der reellen/komplexen Zahlen oder von endlichen Körpern gehört) und ein Skalar ist dann einfach ein Element eines Körpers. Im Moment ist jedenfalls ein Skalar einfach ein Synonym für eine reelle Zahl. Also, sei $\alpha \in \mathbb{R}$ ein Skalar und $\mathcal{A} = (a_0, a_1, a_2, \ldots)$. Wir definieren die Multiplikation von $\mathcal{A}$ mit dem Skalar $\alpha \in \mathbb{R}$ als

$$\alpha\mathcal{A} := (\alpha a_0, \alpha a_1, \alpha a_2, \ldots).$$

Übung 0.7. Argumentieren sie ähnlich wie im Fall der Addition, um folgende Behauptung zu zeigen: Sei $\alpha \in \mathbb{R}$ und $\mathcal{A} \in \textbf{Fib}$. Dann gilt auch $\alpha\mathcal{A} \in \textbf{Fib}$. Zeigen Sie des

Weiteren, dass $\alpha\mathcal{F}_{a,b} = \mathcal{F}_{\alpha a,\alpha b}$.

Wir haben also gesehen, dass der Raum **Fib** Addition und Multiplikation mit einem Skalar hat. Wie hilft uns das, unsere Aufgabe $\mathcal{F}_{0,1}$ gut zu kennen, zu lösen? Geduld bringt Rosen...

Bemerken Sie zuerst das Folgende: Wenn man $\mathcal{F}_1 = (a_0, a_1, \ldots)$ und $\mathcal{F}_2 = (b_0, b_1, \ldots)$ gut kennt, dann kennt man auch $\mathcal{F}_1 + \mathcal{F}_2$ gut. In der Tat, wenn man eine Formel für a_n und b_n hat, dann hat man auch eine Formel für das n-te Folgenglied. Nämlich, wenn $a_n = f(n)$ und $b_n = g(n)$ Formeln für das n-Glied sind, dann ist $f(n)+g(n)$ eine Formel für $\mathcal{F}_1 + \mathcal{F}_2$. Ähnlicherweise, wenn man eine Folge $\mathcal{A} \in$ **Fib** gut kennt, dann kennt man auch $\alpha\mathcal{A}$ für alle $\alpha \in \mathbb{R}$ gut. Allgemeiner, zeigen Sie:

Übung 0.8. Wenn wir $\mathcal{F}_1, \ldots, \mathcal{F}_k$ gut kennen, dann kennen wir

$$\alpha_1\mathcal{F}_1 + \ldots + \alpha_k\mathcal{F}_k \tag{0.2}$$

für alle $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ gut.

Ausdrücke der Form (0.2) heissen lineare Kombinationen von Folgen.

Anders gesagt, kann man mittels der Struktur von **Fib** das Wissen über einige Elemente von **Fib** auf andere Elemente von **Fib** übertragen.

Das ist alles sehr schön, aber es gibt kein einziges Element von **Fib**, das wir gut kennen! Das ist eigentlich nicht wahr. Es gibt doch ein Element von **Fib**, das wir gut kennen.

Übung 0.9. Finden Sie es, bevor Sie weiter lesen.

Die Fibonacci-Folge $\mathcal{F}_{0,0}$ kennen wir sehr gut: Wenn $\mathcal{F}_{0,0} = (a_0, a_1, \ldots)$, dann ist $a_n = 0$ für alle $n \in \mathbb{N} \cup \{0\}$. Heisst das, dass wir jetzt mittels der Struktur von **Fib** viele andere Folgen gut kennen? Leider nicht, denn $\mathcal{F}_{0,0} + \mathcal{F}_{0,0}$ oder $\alpha\mathcal{F}_{0,0}$ ergeben leider keine neuen Folgen... Wir bekommen einfach $\mathcal{F}_{0,0}$ wieder und wieder. Also, wenn wir nur $\mathcal{F}_{0,0}$ gut kennen, können wir mittels Addition und Skalarmultiplikation nicht weiterkommen. Wie viele Folgen sollten wir kennen, um alle Elemente von **Fib** gut zu kennen?

Nehmen wir an, dass wir eine Fibonacci-Folge $\mathcal{F}_{a,b}$ mit a und b beide nicht Null gut kennen. Dann kennen wir mittels Skalarmultiplikation $\mathcal{F}_{\alpha a,\alpha b}$ für alle $\alpha \in \mathbb{R}$ gut.

Übung 0.10. Zeigen Sie, dass die Menge $\{\mathcal{F}_{\alpha a,\alpha b} \mid \alpha \in \mathbb{R}\} \subseteq$ **Fib** unter der Annahme $(a, b) \neq (0, 0)$ unendlich viele Elemente enthält, aber trotzdem niemals gleich **Fib** ist.

Diese Übung zeigt, dass wir zumindest zwei Fibonacci-Folgen gut kennen müssen, um alle Elemente von **Fib** gut zu kennen. Existieren also zwei Folgen in **Fib**, die wir gut kennen und wodurch wir dann alle Elemente von **Fib** gut kennen? Könnte man also zwei Folgen $\mathcal{A}, \mathcal{B} \in \mathbf{Fib}$ finden, so dass jede Folge eine lineare Kombination von $\mathcal{A}$ und $\mathcal{B}$ ist?

Wenn wir zum Beispiel $\mathcal{F}_{0,1}$ und $\mathcal{F}_{1,0}$ gut kennen, dann kennen wir alle Elemente von **Fib** gut: Wir können ein allgemeines Element $\mathcal{F}_{a,b}$ von **Fib** folgendermassen schreiben

$$\mathcal{F}_{a,b} = a\mathcal{F}_{1,0} + b\mathcal{F}_{0,1}$$

und daraus folgt eine Formel für $\mathcal{F}_{a,b}$ aus Formeln für $\mathcal{F}_{0,1}$ und $\mathcal{F}_{1,0}$.

Dies ist nochmals sehr schön, aber wir sind wieder bei unserem ursprünglichen Problem angelangt! Wir kennen $\mathcal{F}_{0,1}$ noch nicht gut! Im nächsten Abschnitt werden wir zwei andere Folgen gut kennenlernen und mit diesen kann man auch jedes andere Element von **Fib** mit Skalarmultiplikation und Addition erreichen.

Bemerkung 0.11. Die Tatsachen, dass

(1) zwei Folgen reichen,

(2) eine Folge nicht reicht (Übung 0.10),

(3) in jeder Menge von drei Folgen es eine Folge gibt, die wir weglassen können ohne die Menge der „erreichbaren Folgen“ zu ändern,

sind verbunden mit der Aussage, dass der Raum **Fib** Dimension 2 hat. Das ist keine Überraschung. Dimension ist fast gleichbedeutend mit der Anzahl „Freiheitsgraden“ des Raums. Überlegen Sie sich, wieso **Fib** genau 2 Freiheitsgrade von reellen Zahlen hat. In anderen Worten, **Fib** ist eine Ebene, in der jeder Punkt eine Folge darstellt.

0.2 Vorkenntnisse und Symmetrie

Wie kommt man auf eine neue Idee für die Lösung eines Problems? Normalerweise durch die Nutzung von Vorkenntnissen, die für die Lösung relevant sind oder durch das Erkennen, dass das Problem eine gewisse Symmetrie hat, die wir benutzen können. Vielleicht denken Sie, dass unser Problem gar keine Symmetrie enthält. Immerhin ist dies kein geometrisches Problem...

0.2.1 Vorkenntnisse

Bevor wir erklären, welche Symmetrie dieses Problem dennoch geniesst und wieso diese Symmetrie der Schlüssel zur Lösung ist, benutzen wir unsere Vorkenntnisse um Fibonacci-Folgen zu finden, die wir gut kennen.

Übung 0.12. Zeigen Sie, dass **Fib** keine arithmetische Folge enthält (ausser $\mathcal{F}_{0,0}$).

Die Übung zeigt, dass wir mit arithmetischen Folgen nicht vorankommen können. Wie wäre es mit geometrischen Folgen? Kann eine Folge der Form $(a, aq, aq^2, aq^3, \ldots)$ eine Fibonacci-Folge sein? Der Einfachheit halber versuchen wir zuerst $a = 1$, das heisst mit einer Folge der Form

$$\mathcal{G}_q = (1, q, q^2, \ldots)$$

mit $q \neq 0$. Die Folge $\mathcal{G}_q$ ist eine Fibonacci-Folge, genau dann, wenn

$$q^n = q^{n-1} + q^{n-2} \tag{0.3}$$

für alle $n \geq 2$ gilt. Da $q \neq 0$, können wir (0.3) durch q^{n-2} dividieren und daher ist (0.3) äquivalent zu

$$q^2 = q + 1. \tag{0.4}$$

Dies können wir mittels anderer Vorkenntnisse lösen, der Mitternachtsformel[3]! Gleichung (0.4) gilt genau dann, wenn

$$q = \frac{1 \pm \sqrt{5}}{2}.$$

Es ist üblich die Zahl

$$\frac{1+\sqrt{5}}{2} \approx 1.618033\ldots \quad \text{(goldener Schnitt)}$$

mit φ zu bezeichnen und die Zahl

$$\frac{1-\sqrt{5}}{2} \approx -0.618033\ldots \quad \text{(konjugierter goldener Schnitt)}$$

mit ψ. Fassen wir diesen Teil zusammen: die Folgen

$$\mathcal{G}_\varphi = (1, \varphi, \varphi^2, \varphi^3, \ldots) \quad \text{und} \quad \mathcal{G}_\psi = (1, \psi, \psi^2, \psi^3, \ldots)$$

sind beides Elemente von **Fib**, die wir gut kennen! Dies ist wunderbar, aber können wir mittels Addition und Skalarmultiplikation von $\mathcal{G}_\varphi$ und $\mathcal{G}_\psi$, die ursprüngliche Folge $\mathcal{F}_{0,1}$ ausdrücken? Ja!

Übung 0.13. Verifizieren Sie, dass

$$\frac{1}{\varphi - \psi}\mathcal{G}_\varphi + \frac{1}{\psi - \varphi}\mathcal{G}_\psi = \mathcal{F}_{0,1}.$$

Bemerkung 0.14. Um Übung 0.13 zu lösen, müssen Sie wahrscheinlich ein lineares Gleichungssystem lösen. Lineare Gleichungssysteme liegen im Herzen der linearen Algebra.

[3] Endlich eine Motivation zum Lösen einer quadratischen Gleichung!

Eines unserer ersten Themen wird ein Algorithmus sein, um lineare Gleichungssysteme zu lösen, die Gauss-Elimination.

Übung 0.13 gibt eine Formel für das n-te Glied von $\mathcal{F}_{0,1} = (a_0, a_1, a_2, \ldots)$:

$$a_n = \frac{1}{\varphi - \psi}\varphi^n + \frac{1}{\psi - \varphi}\psi^n = \frac{\varphi^n - \psi^n}{\varphi - \psi} = \frac{(\frac{1+\sqrt{5}}{2})^n - (\frac{1-\sqrt{5}}{2})^n}{\sqrt{5}}.$$

Wir sind am Ziel angekommen!

Übung 0.15. Sei $\mathcal{F}_{0,1} = (F_0, F_1, F_2, \ldots)$. Unter der Annahme, dass $\lim_{n\to\infty} \frac{F_n}{F_{n-1}}$ existiert, berechnen Sie diesen Grenzwert. Dies gibt eine andere Motivation geometrische Folgen mit q als den Wert dieses Grenzwert zu betrachten.

0.2.2 Symmetrie

Man könnte sagen, dass dies nur ein Glückstreffer war. Wie konnten wir wissen, dass geometrische Folgen hilfreich sein würden? Das ist eine berechtigte Frage. Wie vorhin erwähnt, könnte man diese geometrischen Folgen entdecken, wenn man die Symmetrie des Raums **Fib** betrachtet. Was aber bedeutet Symmetrie in diesem Zusammenhang? Wenn X irgendein geometrischer Raum ist, dann ist eine Symmetrie von X eine Abbildung[4]

$$T : X \to X,$$

die alle/einige der geometrischen Eigenschaften von X erhält, wie beispielsweise Distanz, Winkel usw. Wenn $T : X \to X$ eine Symmetrie ist, dann ist die Menge der Fixpunkte

$$\mathrm{Fix}(T) = \{x \in X : T(x) = x\}$$

normalerweise eine interessante Menge zum Betrachten. In anderen Worten ist ein Fixpunkt ein Element, welches auf sich selbst abgebildet wird. Um den Symmetriebegriff auf allgemeinere Räume zu verallgemeinern, könnten wir das Folgende definieren: Sei X ein Raum mit einer gewissen Struktur. Eine Symmetrie von X ist eine Abbildung $T : X \to X$, die die Struktur von X erhält/respektiert.

Wenn $X = \mathbf{Fib}$ bekommen wir das Folgende:

Definition 0.16. Eine Abbildung $T : \mathbf{Fib} \to \mathbf{Fib}$ heisst eine Symmetrie von **Fib**, wenn für $\mathcal{A}, \mathcal{B} \in \mathbf{Fib}$ und $\alpha \in \mathbb{R}$ gilt, dass

$$T(\mathcal{A} + \mathcal{B}) = T(\mathcal{A}) + T(\mathcal{B}) \quad \text{und} \quad T(\alpha\mathcal{A}) = \alpha T(\mathcal{A}). \tag{0.5}$$

[4]Eine *Abbildung* ist ein anderer Name für eine Funktion. Das ist eine Regel, die jedem Element von X ein bestimmtes Element von X zuordnet. Dies und andere Grundlagen werden wir zusammen mit der Analysisvorlesung bald sorgfältig definieren.

Die Anforderungen in (0.5) sind das, was wir meinen mit „T respektiert die Struktur von **Fib**“.

Überlegen wir zuerst, welche Abbildungen $T : X \to X$ wir überhaupt kennen. Hier sind drei Beispiele, die nicht so interessant sind:

1. Die Identitäts-Abbildung

$$\begin{aligned}\mathrm{Id} : \mathbf{Fib} &\to \mathbf{Fib}\\ \mathcal{A} &\mapsto \mathcal{A},\end{aligned}$$

   die „nichts“ macht.

2. Die Skalarmultiplikations-Abbildung: Für $\alpha \in \mathbb{R}$ definieren wir

$$\begin{aligned}\mathrm{M}_\alpha : \mathbf{Fib} &\to \mathbf{Fib}\\ \mathcal{A} &\mapsto \alpha\mathcal{A}.\end{aligned}$$

3. Die Vektoradditions/Folgenadditions-Abbildung: Für $\mathcal{B} \in \mathbf{Fib}$ definieren wir

$$\begin{aligned}\mathrm{A}_\mathcal{B} : \mathbf{Fib} &\to \mathbf{Fib}\\ \mathcal{A} &\mapsto \mathcal{A} + \mathcal{B}.\end{aligned}$$

Übung 0.17. Zeigen Sie, dass Id und M_α für alle $\alpha \in \mathbb{R}$ die Anforderungen in (0.5) erfüllen. Zeigen Sie des Weiteren, dass $\mathrm{A}_\mathcal{B}$ für $\mathcal{B} \neq \mathcal{F}_{0,0}$ die Anforderungen in (0.5) **nicht** erfüllt.

Wie gesagt, sind diese Symmetrien nicht besonders interessant, vielleicht weil sie nicht mit der Tatsache verbunden sind, dass es sich bei **Fib** um einen Raum von Folgen handelt. Diese Abbildungen existieren für jeden Raum, der Addition und Skalarmultiplikation hat.

Übung 0.18. Bevor Sie weiter lesen, versuchen Sie eine interessante andere Abbildung $T : \mathbf{Fib} \to \mathbf{Fib}$ zu finden, welche berücksichtigt, dass es sich um einen Folgen-Raum handelt? (Hinweis: Ich behaupte: Wenn man $\mathcal{F}_{1,0}$ gut kennt, dann kennt man auch $\mathcal{F}_{0,1}$ gut. Wieso?)

Bei Folgen-Räumen gibt es die Verschiebungsabbildung

$$\begin{aligned}S : \mathbf{Fib} &\to \mathbf{Fib}\\ (a_0, a_1, a_2, \ldots) &\mapsto (a_1, a_2, \ldots).\end{aligned}$$

Übung 0.19. (1) Verifizieren Sie, dass $S(\mathcal{A}) \in \mathbf{Fib}$ für $\mathcal{A} \in \mathbf{Fib}$.

(2) Verifizieren Sie, dass S die Anforderungen in (0.5) erfüllt.

Diese Verschiebungsabbildung ist schon eine interessante Symmetrie, die mit Folgen-Räumen verbunden ist. Daher fragen wir uns, welche Fixpunkte S hat:

Übung 0.20. Zeigen Sie, dass $\mathcal{F}_{0,0}$ der einzige Fixpunkt von S ist.

Gut, die Fixpunkte von S sind nicht so interessant. Es stellt sich heraus, dass für Abbildungen, die (0.5) erfüllen, die Menge von Fixpunkten häufig nicht interessant ist. Die Forderung $S(\mathcal{A}) = \mathcal{A}$ ist einfach zu einschränkend, um interessante Folgen zu bekommen. Daher betrachtet man eine schwächere Anforderung (die auch mit der Skalarmultiplikation verbunden ist).

Definition 0.21. Sei $T : \mathbf{Fib} \to \mathbf{Fib}$ eine Symmetrie. Eine Folge $\mathcal{A} \neq \mathcal{F}_{0,0}$ heisst eine *Eigenfolge* von T, wenn es ein $\alpha \in \mathbb{R}$ gibt, so dass

$$T(\mathcal{A}) = \alpha A.$$

Der Skalar α heisst der *Eigenwert* von $\mathcal{A}$.

Wir möchten also alle Eigenfolgen der Symmetrie S finden[5]. Dafür nehmen wir an, dass $\mathcal{A} = (a_0, a_1, a_2, \ldots) \neq F_{0,0}$ eine Eigenfolge mit Eigenwert α ist. Dann gilt $S(\mathcal{A}) = \alpha A$ oder in anderen Worten

$$(a_1, a_2, a_3, \ldots) = \alpha(a_0, a_1, a_2, \ldots) = (\alpha a_0, \alpha a_1, \alpha a_2, \ldots).$$

Also gilt $a_n = \alpha a_{n-1}$ für alle $n \geq 1$. Daher hat $\mathcal{A}$ die Form

$$\mathcal{A} = (a_0, a_0\alpha, a_0\alpha^2, a_0\alpha^3, \ldots). \tag{0.6}$$

Das heisst, dass $\mathcal{A}$ eine geometrische Folge ist. Welche α (und welche a_0) kommen also in Frage? Dies haben wir schon in Abschnitt 0.2.1 gesehen, aber leiten wir es nochmals her: Aus (0.6) folgt insbesondere, dass

$$a_n = \alpha^n a_0$$

für alle $n \geq 0$. Da $\mathcal{A} \in \mathbf{Fib}$ gilt auch $a_2 = a_1 + a_0$ und zusammen bekommen wir

$$\alpha^2 a_0 = a_2 = a_1 + a_0 = \alpha a_0 + a_0 = (\alpha + 1)a_0. \tag{0.7}$$

Aus $\mathcal{A} \neq \mathcal{F}_{0,0}$ folgt, dass $a_0 \neq 0$. (Wieso? Finden Sie eine überzeugende Erklärung.) Daher ist (0.7) äquivalent zu $\alpha^2 = \alpha + 1$. Sieht das bekannt aus? Dies ist gerade die

[5] Für die fortgeschrittenen Leser bemerken wir das Folgende. Eigenfolgen sind eigentlich Fixpunkte bezüglich der Wirkung von S auf dem projektiven Raum $\mathbb{P}(\mathbf{Fib})$.

Gleichung (0.4). Dies bedeutet: Wenn $\mathcal{A}$ eine Eigenfolge mit Eigenwert α ist, dann gilt $\alpha^2 = \alpha + 1$ beziehungsweise $\alpha = \psi$ oder $\alpha = \varphi$. Das heisst, $\mathcal{A}$ ist eine geometrische Folge mit $\alpha = \varphi$ oder $\alpha = \psi$. Der Einfachheit halber wählen wir $a_0 = 1$. Wir bekommen die zwei Folgen, die wir vorher erraten haben

$$\mathcal{G}_\varphi = (1, \varphi, \varphi^2, \varphi^3, \ldots) \quad \text{und} \quad \mathcal{G}_\psi = (1, \psi, \psi^2, \psi^3, \ldots).$$

<u>Zusammenfassung</u> : Die n-te Fibonacci-Zahl F_n ist durch

$$\frac{\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n}{\sqrt{5}}$$

gegeben.

Allgemeiner zeigen Sie:

Übung 0.22. Finden Sie eine Formel für das n-te Glied von $\mathcal{F}_{a,b}$, die von a, b und n abhängt.

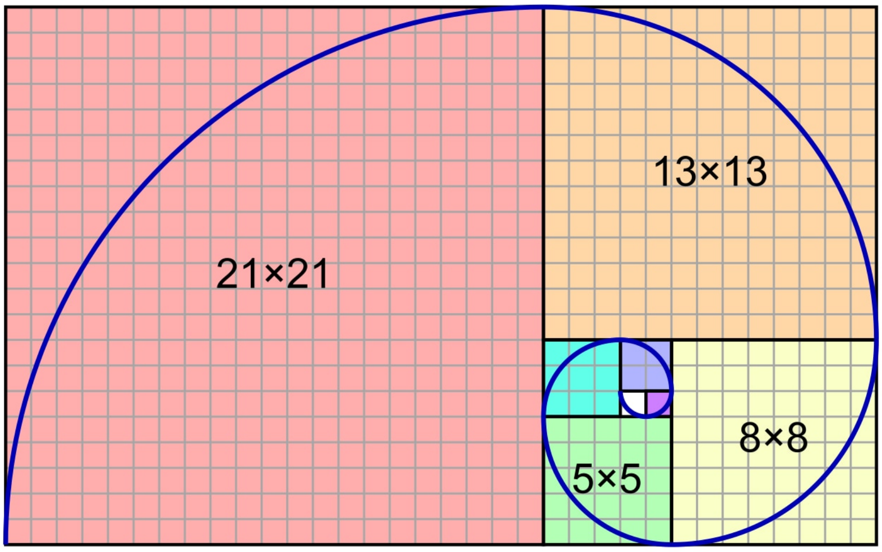
Es gibt noch viele Sachen in diesem Zusammenhang zu sagen, aber wir müssen irgendwann mit dem Kurs anfangen. Hier ist eine kleine Übung als Dessert:

Übung 0.23. Zeigen Sie, dass

$$\varphi = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \ldots}}}.$$

Was genau gemeint ist mit dieser Schreibweise und wie dies mit der Bewegung der Planeten zusammenhängt, können Sie mich fragen oder hier [4, §1.1] nachlesen.

Bemerkung 0.24. Auf die Idee zu dieser Einführung bin ich gekommen, als ich Übungsstunden an der Hebrew University gehalten habe und gefragt wurde, wieso lineare Algebra Spass machen soll. Dies habe ich durch diese Frage mit der Welt geteilt. Viele der Antworten dort sind sehr interessant und lohnen sich, darüber nachzudenken!



Figur 0.1: Ihr nächstes T-Shirt?

Kapitel 1

Grundlagen

1.1 Mengen und Abbildungen

Dieses Kapitel haben wir sowohl in der Linearen Algebra wie auch in der Analysis besprochen. Sie können dieses Kapitel im Fischer [6] als Wiederholung lesen.

1.2 Gruppen

Teile von diesem Kapitel haben Sie in der Analysis behandelt (die Definition von Gruppen, Ringen und Körper). Für uns ist vor allem die Definition von Körpern relevant. In der Vorlesung Algebra I im dritten Semester werden sie noch mehr Gruppen- und Ring-Theorie behandeln. In diesem Kapitel geben wir nur einige Definitionen und einige Beispiele.

Definition 1.1 (Analysis 24. Sept.)**.** Eine *Gruppe* $(G, e, \cdot)$ ist eine Menge G versehen mit einem Element $e \in G$ (das neutrale Element) und einer Verknüpfung $\cdot : G \times G \to G$ mit folgenden Eigenschaften:

a) Für alle $a, b, c \in G$ gilt $(a \cdot b) \cdot c = a \cdot (b \cdot c)$. ($\cdot$ ist assoziativ)

b) Für alle $g \in G$ gilt $e \cdot g = g \cdot e = g$. (daher der Name „neutrales Element“)

c) Zu jedem $g \in G$ gibt es ein $g' \in G$ mit $g \cdot g' = g' \cdot g = e$. ($g'$ nennen wir ein inverses Element von g)

Beachten Sie, dass a) der Eigenschaft G1 im Fischer [6] entspricht und b), c) entsprechen G2. Wir schreiben oft zur Vereinfachung ab statt $a \cdot b$ für Elemente $a, b \in G$, falls klar ist, welche Operation gemeint ist.

Folgende Proposition haben Sie auch in der Analysis gesehen (vgl. auch Fischer [6, Kap. 1.2.3]):

Proposition 1.2. *1) Das neutrale Element ist eindeutig bestimmt (d.h. falls für $e, e' \in G$ gilt, dass $eg = ge$ und $e'g = ge'$ für alle $g \in G$, dann folgt stets $e = e'$).*

2) Das inverse Element von G ist eindeutig bestimmt. (Beachten Sie, dass wir dadurch a^{-1} für das inverse Element von a schreiben können.)

3) Für alle $a, b \in G$ gilt $(a^{-1})^{-1} = a$ und $(ab)^{-1} = b^{-1}a^{-1}$.

4) Für alle $a, b, c \in G$ gilt $ab = ac$ genau dann wenn $b = c$ und $ba = ca$ genau, dann wenn $b = c$.

Beweis. Für 1) und 2) verweisen wir auf die Analysis Vorlesung (21. Sept.). Wir beweisen jetzt 3):

Beachten Sie, dass $(a^{-1})^{-1}$ das einzige Element $b \in G$ ist mit der Eigenschaft

$$a^{-1}b = ba^{-1} = e.$$

Da a auch diese Eigenschaft hat, folgt nach 2), dass $a = (a^{-1})^{-1}$. Für den zweiten Teil bemerken Sie, dass

$$(ab)(b^{-1}a^{-1}) = a(bb^{-1})a^{-1} = aea^{-1} = aa^{-1} = e$$

und analog folgt $(b^{-1}a^{-1})(ab) = e$. Nach 2) folgt wiederum, dass $b^{-1}a^{-1} = (ab)^{-1}$. Teil 4) der Proposition beweisen Sie in Serie 4. □

Definition 1.3. Eine Gruppe $(G, e, \cdot)$ heisst *abelsch* oder *kommutativ*, falls $a \cdot b = b \cdot a$ für alle $a, b \in G$ gilt.

Bemerkung 1.4. Wenn eine Gruppe G abelsch[1] ist, dann verwenden wir fast immer die additive Schreibweise und schreiben $(G, e, +)$ für die Gruppe. Also schreiben wir insbesondere $g_1 + g_2$ statt $g_1 g_2$ und $-g$ statt g^{-1} für $g_1, g_2, g \in G$.

Beispiel 1.5. Wir betrachten einige Beispiele von Gruppen.

- $(\mathbb{Z}, 0, +)$ ist eine Gruppe.
- $(\mathbb{N} \cup \{0\}, 0, +)$ ist keine Gruppe, da jedes $n \in \mathbb{N}$ kein inverses Element hat.
- $(\mathbb{Q}, 0, +)$ ist eine Gruppe.
- Allgemeiner ist $(K, 0, +)$ für jeden Körper K eine Gruppe. (Die Definition eines Körpers folgt noch.)

[1] Der Begriff „abelsch“ ist zurückzuführen auf den Mathematiker Niels Abel. Dieser Begriff ist das einzige Beispiel eines nach einer Person benannten mathematischen Begriffs, der ohne Grossbuchstabe geschrieben wird. So grundlegend ist dieser Begriff also!

- Sei X eine Menge und
  $\mathrm{Abb}(X) = \{f : X \to X \mid f \text{ ist bijektiv}\}$. Dann ist $(\mathrm{Abb}(X), \mathrm{Id}_X, \circ)$ eine Gruppe. Hier ist $\circ$ die Verkettung und $\mathrm{Id}_X : X \to X, x \mapsto x$ die Identitätsabbildung. (Dieses Beispiel haben Sie schon in der Analysis gesehen).

- Wenn $X = \{1, \ldots, n\}$ ist, dann schreiben wir $S_n = \mathrm{Abb}(X)$ und somit ist $(S_n, \mathrm{Id}_X, \circ)$ nach dem obigen Beispiel eine Gruppe. Diese Gruppe benutzen wir später, wenn wir Determinanten besprechen (Kapitel 3 im Fischer [6]).

1.2.1 Wichtiges Beispiel: Restklassen modulo n

Dieses Beispiel ist zentral für uns. Es wird uns als Beispiel für Gruppen, Ringe und für gewisse n auch für Körper dienen. Um das „Objekt“ der Restklassen zu definieren, betrachten wir die folgende Äquivalenzrelation auf $\mathbb{Z}$: Sei $n \in \mathbb{N}$. Wir definieren für $a, b \in \mathbb{Z}$

$$a \sim_n b \iff (b-a) \text{ ist durch } n \text{ teilbar} \iff \exists l \in \mathbb{Z} : (b-a) = ln.$$

Bemerkung 1.6. Wenn n aus dem Zusammenhang klar ist, dann schreiben wir einfach $\sim$ statt $\sim_n$.

Übung 1.7. Beweisen Sie, dass $\sim_n$ eine Äquivalenzrelation ist. (Reflexivität und Symmetrie sind fast direkt, nur für die Transitivität muss man kurz überlegen.)

Übung 1.8. Schreiben Sie die Äquivalenzklassen für $\sim_2, \sim_3, \sim_4, \ldots$ auf, bis Sie verstanden haben, was die Äquivalenzklassen von $\sim_n$ für ein allgemeines $n \in \mathbb{N}$ sind. Wie viele Äquivalenzklassen gibt es für $\sim_1$?

Wir lösen Übung 1.8 für $\sim_2$ (jedoch ohne Erklärung!):

$$[0]_{\sim_2} = \{\ldots, -6, -4, -2, 0, 2, 4, \ldots\} = \text{die Menge der geraden Zahlen in } \mathbb{Z}$$
$$[1]_{\sim_2} = \{\ldots, -5, -3, -1, 1, 3, 5, \ldots\} = \text{die Menge der ungeraden Zahlen in } \mathbb{Z}$$

Also sehen wir, dass $a \sim_2 b$ ein ausgeklügelter (aber präziser) Weg ist zu sagen, dass a und b entweder beide gerade sind oder a und b beide ungerade sind.[2]

[2]Merken Sie sich: Es ist in der Tat präziser die obige Definition mit $\sim_2$ zu verwenden. Entweder ... oder ... hat keine ganz eindeutige Bedeutung in der Alltagssprache. Einige ETH-Dozenten verbieten die Nutzung von „entweder“ sogar! Ich erlaube es Ihnen, aber nur wenn es klar ist, was Sie meinen!

Was passiert nun, wenn man beispielsweise eine gerade Zahl und eine ungeraden Zahl addiert? Man bekommt immer eine ungerade Zahl! Es gilt

$$\begin{align} \text{gerade} + \text{gerade} &= \text{gerade} \tag{1.1}\\ \text{gerade} + \text{ungerade} &= \text{ungerade} \tag{1.2}\\ \text{ungerade} + \text{gerade} &= \text{ungerade} \tag{1.3}\\ \text{ungerade} + \text{ungerade} &= \text{gerade.} \tag{1.4} \end{align}$$

Übung 1.9. Zeigen Sie, dass (1.1) der Menge $\mathbb{Z}/\sim_2 = \{[0]_{\sim_2}, [1]_{\sim_2}\}$ die Struktur einer Gruppe gibt mit $[0]_{\sim_2}$ als neutralem Element. Die Additions-Tabelle sieht so aus:

$+$	$[0]_{\sim_2}$	$[1]_{\sim_2}$
$[0]_{\sim_2}$	$[0]_{\sim_2}$	$[1]_{\sim_2}$
$[1]_{\sim_2}$	$[1]_{\sim_2}$	$[0]_{\sim_2}$

Dies möchten wir auch in einem ausgeklügelten Weg schreiben, und zwar für $\sim_n$. Zuerst geben wir aber die Antwort zu Übung 1.8: Für $\sim_n$ gibt es n verschiedene Äquivalenzklassen, die man durch „Restklassen“ darstellen kann:

$$\begin{aligned} [0]_{\sim_n} = 0 + n\mathbb{Z} :=& \{0 + nl \mid l \in \mathbb{Z}\}\\ =& \{\text{alle ganzen Zahlen mit Rest 0 bei Division mit } n\}\\ [1]_{\sim_n} = 1 + n\mathbb{Z} :=& \{1 + nl \mid l \in \mathbb{Z}\}\\ =& \{\text{alle ganzen Zahlen mit Rest 1 bei Division mit } n\}\\ &\vdots\\ [n-1]_{\sim_n} = (n-1) + n\mathbb{Z} :=& \{(n-1) + nl \mid l \in \mathbb{Z}\}\\ =& \{\text{alle ganzen Zahlen mit Rest } (n-1) \text{ bei Division mit } n\}. \end{aligned}$$

Allgemeiner gilt $[a]_{\sim_n} = a + n\mathbb{Z}$. Wir schreiben $\overline{a}$ für $[a]_{\sim_n}$ (unter der Annahme, dass n aus dem Zusammenhang klar ist). Also ist $\overline{a} = a + n\mathbb{Z} = [a]_{\sim_n}$. Des Weiteren schreiben wir $a \equiv b \,\text{modulo}\, n$ oder $a \equiv b \mod n$ oder $a \equiv b\ (n)$, falls $a \sim_n b$.

Beispiel 1.10. Sei $n = 3$. Dann ist beispielsweise

- $\overline{1} = \overline{4} = \overline{-2} = \overline{10} = \overline{3k+1} \quad \text{für } k \in \mathbb{Z}$
- $\overline{0} = \overline{3} = \overline{-9} = \overline{3k} \quad \text{für } k \in \mathbb{Z}$
- $2 \equiv 5 \equiv -1 \equiv 11 \mod 3$

Obiges Beispiel zeigt, dass jede Restklasse viele Darstellungen als $\overline{a}$ hat.

Definition 1.11 (vgl. Fischer [6, Kap. 1.2.7])**.** Sei $n \in \mathbb{N}$. Wir schreiben $\mathbb{Z}/n\mathbb{Z}$ für die Menge der Äquivalenzklassen von $\sim_n$, das heisst[3]

$$\mathbb{Z}/n\mathbb{Z} = \mathbb{Z}/\sim_n = \{\overline{0}, \overline{1}, \ldots, \overline{(n-1)}\}.$$

Wir möchten gerne zwei Elemente von $\mathbb{Z}/n\mathbb{Z}$ addieren können. (Denken Sie an $\mathbb{Z}/n\mathbb{Z}$ für $n = 2$ und die Addition von geraden/ungeraden Zahlen oben.)

Definition 1.12. Seien $\overline{a}, \overline{b} \in \mathbb{Z}/n\mathbb{Z}$. Wir definieren

$$\overline{a} + \overline{b} := \overline{a+b}. \tag{1.5}$$

Ist das eine gute Definition? Ist es „erlaubt", diese Addition einfach so zu definieren? Da wir Addition von *Restklassen/Äquivalenzklassen* durch eine Auswahl von Repräsentationen definieren[4], müssen wir beweisen, dass diese Definition nicht von dieser Wahl abhängt. In der Mathematik sagen wir, dass wir zeigen müssen, dass die Addition in (1.5) <u>wohl definiert</u> ist. Dies ist am besten erklärt, in dem man den Beweis sieht:

Lemma 1.13. *Die Addition in (1.5) ist wohl definiert.*

Beweis. Seien $a, a' \in \mathbb{Z}$ mit $\overline{a} = \overline{a'}$ (d.h. a und a' sind zwei möglicherweise verschiedene Repräsentationen derselben Restklasse) und $b, b' \in \mathbb{Z}$ mit $\overline{b} = \overline{b'}$. Wir müssen zeigen, dass

$$\overline{a+b} = \overline{a'+b'}.$$

Das ist einfach: Laut Annahme existieren $l, k \in \mathbb{Z}$ mit

$$a - a' = ln \quad \text{und} \quad b - b' = kn.$$

Daraus folgt, dass

$$(a+b) - (a'+b') = (a-a') + (b-b') = ln + kn = (l+k)n.$$

Also gilt $\overline{a+b} = \overline{a'+b'}$ und daher ist die Addition wohl definiert (sie hängt nicht von der Wahl der Repräsentanten ab). Dies beendet den Beweis. □

Theorem 1.14 (vgl. Fischer [6, Kap. 1.2.7])**.** *Die Menge $\mathbb{Z}/n\mathbb{Z}$ zusammen mit der obigen Addition ist eine abelsche Gruppe.*

Beweis. Seien $\overline{a}, \overline{b}, \overline{c} \in \mathbb{Z}/n\mathbb{Z}$. Wir müssen vier Eigenschaften überprüfen:

[3] Die erste Gleichung gilt nach Definition der Quotientenmenge und die zweite laut Übung 1.8.

[4] Für $n = 3$ ist es nach obiger Definition zum Beispiel a priori nicht klar, dass $\overline{1} + \overline{3} = \overline{1} + \overline{-9}$, vgl. Beispiel 1.10.

- Assoziativität folgt aus Assoziativität in $\mathbb{Z}$:
  $(\overline{a}+\overline{b})+\overline{c} = \overline{a+b}+\overline{c} = \overline{(a+b)+c} = \overline{a+(b+c)} = \overline{a}+\overline{b+c} = \overline{a}+(\overline{b}+\overline{c})$.

- $\overline{0}$ ist das neutrale Element: $\overline{0}+\overline{a} = \overline{0+a} = \overline{a} = \overline{a+0} = \overline{a}+\overline{0}$.

- $\overline{-a}$ ist die Inverse von $\overline{a}$: $\overline{a}+\overline{-a} = \overline{a+(-a)} = \overline{0} = \overline{(-a)+a} = \overline{-a}+\overline{a}$.

- Die Gruppe ist abelsch, da $\mathbb{Z}$ abelsch ist: $\overline{a}+\overline{b} = \overline{a+b} = \overline{b+a} = \overline{b}+\overline{a}$.

□

Bemerkung 1.15. Wie bei jeder abelschen Gruppe bedeutet Subtraktion Addition der Inversen: Erinnern Sie sich, dass $\overline{-b}$ die Inverse von $\overline{b}$ ist. Dann gilt: $\overline{a}-\overline{b} = \overline{a}+\overline{-b} = \overline{a+(-b)} = \overline{a-b}$.

1.2.2 Untergruppen, Homomorphismen und Isomorphismen

Dieses Kapitel vergleicht sich mit Kapitel 1.2.6 im Fischer [6]. Wir besprechen kurz einige Begriffe der Gruppen-Theorie.

Definition 1.16. Sei $(G,\cdot)$ eine Gruppe. Eine Teilmenge $H \subseteq G$ heisst eine *Untergruppe* (von G), falls das Folgende gilt:

- H ist nicht leer.

- Für alle $h_1, h_2 \in H$ gilt $h_1h_2 \in H$.

- Für alle $h \in H$ gilt $h^{-1} \in H$.

Beispiel 1.17. Wir betrachten einige Beispiele von Untergruppen.

- $\mathbb{Z} \subseteq (\mathbb{Q},+)$ ist eine Untergruppe von $(\mathbb{Q},+)$.

- $\{0\} \subseteq \mathbb{Z}$ ist eine Untergruppe von $\mathbb{Z}$.

- Für jedes $m \in \mathbb{N}$ ist die Menge $m\mathbb{Z} = \{ml \mid l \in \mathbb{Z}\}$ ist eine Untergruppe von $\mathbb{Z}$.

- Sei $(G,e,\cdot)$ eine Gruppe. Dann sind $\{e\}$ und G selbst zwei Untergruppen von G.

- $\{\overline{0},\overline{2}\} \subseteq \mathbb{Z}/4\mathbb{Z} = \{\overline{0},\overline{1},\overline{2},\overline{3}\}$ ist eine Untergruppe von $\mathbb{Z}/4\mathbb{Z}$.

Bemerkung 1.18. Ist $H \subseteq (G,\cdot)$ eine Untergruppe, so ist $(H,\cdot)$ mit der Verknüpfung $\cdot$ aus G wieder eine Gruppe:

Assoziativität der Verknüpfung $\cdot$ auf H folgt, da $H \subseteq G$. Ausserdem gilt, dass $h^{-1} \in H$ für $h \in H$ per Definition von Untergruppen. Da $h, h^{-1} \in H$ folgt aus der Definition von Untergruppen auch, dass $e = hh^{-1} \in H$ und e hat die Eigenschaft $eg = ge = g$ für alle $g \in G$. Da $H \subseteq G$ gilt dies also insbesondere für alle $h \in H$. Also ist $(H,\cdot)$ in der Tat eine Gruppe.

Um gewisse strukturelle Ähnlichkeiten zwischen verschiedenen Gruppen auszudrücken, brauchen wir den Begriff von Homomorphismus und Isomorphismus. Allgemeiner definiert man in der Mathematik (insbesondere in der abstrakten Algebra) eine Klasse von Objekten mit gewisser Struktur und betrachtet dann alle Abbildungen zwischen diesen Objekten, welche die Struktur der Objekte „erhalten/respektieren". Kürzer gesagt:

Definition 1.19 (vgl. Fischer [6, Kap. 1.2.6])**.** Seien $(G, \cdot)$ und $(H, *)$ zwei Gruppen. Eine Abbildung $\varphi : G \to H$ ist ein *Homomorphismus*, falls für alle $g_1, g_2 \in G$ gilt $\varphi(g_1 \cdot g_2) = \varphi(g_1) * \varphi(g_2)$[5]. Ein bijektiver Homomorphismus nennen wir einen *Isomorphismus*.

Lemma 1.20. *Sei $\varphi : (G, e_G, \cdot) \to (H, e_H, *)$ ein Homomorphismus. Dann gilt:*

(a) $\varphi(e_G) = e_H$, wobei e_G und e_H jeweils die neutralen Elemente von G und H sind.

(b) $\varphi(a^{-1}) = \varphi(a)^{-1}$ für alle $a \in G$.

Beweis. (a) Es gilt $\varphi(e_G) = \varphi(e_G \cdot e_G) = \varphi(e_G) * \varphi(e_G)$. Wir multiplizieren beide Seiten von links mit $\varphi(e_G)^{-1}$ und erhalten:

$$e_H = \varphi(e_G)^{-1} * \varphi(e_G) = \varphi(e_G)^{-1} * \varphi(e_G) * \varphi(e_G) = e_H * \varphi(e_G) = \varphi(e_G).$$

(b) Sei $a \in G$. Nach (a) gilt $\varphi(a) * \varphi(a^{-1}) = \varphi(a \cdot a^{-1}) = \varphi(e_G) = e_H$ und $\varphi(a^{-1}) * \varphi(a) = \varphi(a^{-1}a) = \varphi(e_G) = e_H$. Aus der Eindeutigkeit der Inversen folgt nun (b).

□

Bemerkung 1.21. Wie wir in Bemerkung 1.4 gesagt haben, verwenden wir normalerweise additive Schreibweise für abelsche Gruppen. In additiver Schreibweise sieht Definition 1.19 so aus: Seien $(G, +_G)$ und $(H, +_H)$ zwei (abelsche) Gruppen. Eine Abbildung $\varphi : G \to H$ ist ein Homomorphismus, falls $\varphi(g_1 +_G g_2) = \varphi(g_1) +_H \varphi(g_2)$ gilt.

Beispiel 1.22. (1) Sei $m \in \mathbb{N}$. Die Abbildung

$$\begin{aligned} \varphi_m : (\mathbb{Z}, +) &\to (\mathbb{Z}, +) \\ a &\mapsto ma \end{aligned}$$

ist ein Homomorphismus: $\varphi_m(a+b) = m(a+b) = ma+mb = \varphi_m(a)+\varphi_m(b)$. Diese Abbildung ist genau, dann ein Isomorphismus, wenn $m = 1$.

(2) Sei $n \in \mathbb{N}$. Die Abbildung

$$\begin{aligned} \pi_n : \mathbb{Z} &\to \mathbb{Z}/n\mathbb{Z} \\ a &\mapsto \overline{a} \end{aligned}$$

[5] Beachten Sie, dass $\cdot$ die Multiplikation in G ist und $*$ die Multiplikation in H!

ist ein Homomorphismus: $\pi_n(a+b) = \overline{a+b} = \overline{a} + \overline{b} = \pi_n(a) + \pi_n(b)$. Bemerken Sie, dass die ersten zwei Additionszeichen die Addition in $\mathbb{Z}$ bezeichnen und die letzten zwei die Addition in $\mathbb{Z}/n\mathbb{Z}$.

Übung 1.23. Zeigen Sie, dass $\mathrm{Bild}(\varphi_n) = \pi_n^{-1}(\overline{0})$, und dass diese Menge eine Untergruppe von $\mathbb{Z}$ ist.

1.3 Ringe und Körper

Unsere Hauptmotivation für dieses Kapitel ist es Körper zu definieren. Diese Kapitel vergleicht sich mit Kapitel 1.3 im Fischer [6]. Als erster Schritt definieren wir Ringe:

1.3.1 Ringe

Definition 1.24. Eine Menge R zusammen mit zwei Verknüpfungen

$$\begin{aligned} + : R \times R \to R, &\quad (a,b) \mapsto a+b \quad (\mathit{Addition}) \\ \cdot : R \times R \to R, &\quad (a,b) \mapsto a \cdot b \quad (\mathit{Multiplikation}) \end{aligned}$$

heisst ein *Ring*, falls:

(R1) $(R,+)$ ist eine abelsche Gruppe.

(R2) $\cdot$ ist assoziativ.

(R3) Für alle $a,b,c \in R$ gilt $a \cdot (b+c) = a \cdot b + a \cdot c$ und $(a+b) \cdot c = a \cdot c + b \cdot c$.

Des Weiteren definieren wir:

- R heisst *kommutativ*, falls für alle $a,b \in R$ gilt $a \cdot b = b \cdot a$.
- R heisst *Ring mit Eins*, falls es ein Element $1 \in R$ gibt, so dass $1 \cdot a = a \cdot 1 = a$ für alle $a \in R$ gilt. Das Element $1 \in R$ mit dieser Eigenschaft heisst *Einselement*.

Wir schreiben oft zur Vereinfachung ab statt $a \cdot b$ für Elemente $a,b \in R$, falls klar ist, welche Operation gemeint ist.

Beispiel 1.25. $(\mathbb{Z},+,\cdot)$ ist ein kommutativer Ring (mit Eins), $1 \in \mathbb{Z}$ ist das Einselement.

Beispiel 1.26 (Restklassen)**.** Wir betrachten $\mathbb{Z}/n\mathbb{Z}$ für $n \in \mathbb{N}$. Wir definieren für $\overline{a}, \overline{b} \in \mathbb{Z}/n\mathbb{Z}$:

$$\overline{a} \cdot \overline{b} = \overline{a \cdot b}. \tag{1.6}$$

Man überprüft (ähnlich wie für die Addition zuvor), dass die Multiplikation in (1.6) wohl definiert ist (d.h. sie hängt nicht von der Wahl der Repräsentanten ab). (Machen Sie das!) Mit dieser Definition ist $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ ein (endlicher) kommutativer Ring mit Eins; das Einselement ist $\overline{1}$. Dies ermöglicht die sogenannte modulo Arithmetik, welche äusserst nützlich ist für viele Anwendungen in der Zahlen-Theorie (siehe Spass-Aufgabe in Serie 4). Aus Zeitgründen verzichten wir jedoch auf mehr Details. Hier sind die Multiplikations-Tabellen für $n = 4, 5$:

$\cdot$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{0}$	$\overline{2}$
$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

$n = 4$

$\cdot$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{4}$	$\overline{1}$	$\overline{3}$
$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{1}$	$\overline{4}$	$\overline{2}$
$\overline{4}$	$\overline{0}$	$\overline{4}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

$n = 5$

Wenn wir eine Ring R mit Eins haben, könnten wir uns nun fragen, ob es auf R eine Gruppenstruktur bezüglich Multiplikation gibt. In diesem Fall wäre $1 \in R$ das neutrale Element. Ein Problem, das wir dabei haben ist das Element $0 \in R$:

Lemma 1.27 (vgl. Fischer [6, Kap. 1.3.1])**.** *Für alle $a \in R$ haben wir $a \cdot 0 = 0 \cdot a = 0$.*

Beweis. Wir beweisen, dass $0 \cdot a = 0$. Der Beweis $0 \cdot a = 0$ ist sehr ähnlich.

Da $0 + 0 = 0$ ist, haben wir wegen der Distributivität, dass

$$0 \cdot a = (0 + 0) \cdot a = 0 \cdot a + 0 \cdot a.$$

Wir addieren die additive Inverse von $0 \cdot a$ auf beide Seiten, benutzen Assoziativität und erhalten

$$0 = 0 \cdot a - 0 \cdot a = (0 \cdot a + 0 \cdot a) - 0 \cdot a = 0 \cdot a + (0 \cdot a - 0 \cdot a) = 0 \cdot a + 0 = 0 \cdot a.$$

$\square$

Wir können natürlich die Frage ändern und uns fragen, ob es auf $R \setminus \{0\}$ eine Gruppenstruktur bezüglich Multiplikation gibt. Wir sehen jedoch schnell, dass 0 nicht das einzige Hindernis ist/war:

$$\begin{aligned} &\bullet \text{ In } \mathbb{Z}/4\mathbb{Z} \text{ ist } \overline{2} \cdot \overline{2} = \overline{0}, \text{ aber } \overline{2} \neq \overline{0}. \\ &\bullet \text{ In } \mathbb{Z}/6\mathbb{Z} \text{ ist } \overline{2} \cdot \overline{3} = \overline{0}, \text{ aber } \overline{2} \neq \overline{0} \text{ und } \overline{3} \neq \overline{0}. \end{aligned} \tag{1.7}$$

Dies zeigt, dass für $R = \mathbb{Z}/6\mathbb{Z}$ oder $R = \mathbb{Z}/4\mathbb{Z}$ die Menge $R \setminus \{\overline{0}\}$ keine multiplikative Gruppe sein kann, da die Menge nicht einmal abgeschlossen ist bezüglich Multiplikation (siehe (1.7)). Man kann überprüfen, dass man in $\mathbb{Z}/2\mathbb{Z}, \mathbb{Z}/3\mathbb{Z}$ und $\mathbb{Z}/5\mathbb{Z}$ keine Beispiele wie in (1.7) findet. Um genauer zu erklären, was passiert, brauchen wir eine Definition:

Definition 1.28. Ein Ring R heisst *nullteilerfrei*, wenn für alle $a, b \in R$ mit $ab = 0$ folgt, dass $a = 0$ oder $b = 0$. In Quantoren:

$$\forall a, b \in R : ab = 0 \Rightarrow (a = 0) \vee (b = 0). \tag{1.8}$$

Bemerkung 1.29. Es ist auch nützlich sich die Kontraposition von (1.8) zu merken:

$$\forall a, b \in R : (a \neq 0) \wedge (b \neq 0) \Rightarrow ab \neq 0.$$

Wir definieren zuerst Körper und fragen uns dann, wann genau unser Lieblings-Beispiel $\mathbb{Z}/n\mathbb{Z}$ ein nullteilerfreier Ring ist.

1.3.2 Körper

Kurz gesagt ist ein Körper ein Ring mit Eins, wo jedes von Null verschiedene Element eine multiplikative Inverse hat.

Definition 1.30 (vgl. Fischer [6, Kap. 3.3])**.** Eine Menge K (oder $(K, +, \cdot, 0, 1)$) zusammen mit zwei Verknüpfungen

$$\begin{aligned} + &: K \times K \to K, \quad (a, b) \mapsto a + b \quad \text{(Addition)} \\ \cdot &: K \times K \to K, \quad (a, b) \mapsto a \cdot b \quad \text{(Multiplikation)} \end{aligned}$$

heisst *Körper*, wenn Folgendes gilt:

(1) K zusammen mit der Addition ist eine abelsche Gruppe (0 ist ihr neutrales Element, $-a$ die Inverse von $a \in K$).

(2) Sei $K^\times := K \setminus \{0\}$. Dann ist $(K^\times, \cdot, 1)$ eine abelsche Gruppe. Das heisst, für alle $a, b \in K^\times$ ist $ab \in K^\times$ und für alle $a \in K^\times$ existiert ein $b \in K^\times$ mit $ab = 1$. In diesem Fall schreiben wir $b = a^{-1} =: \frac{1}{a}$ und allgemeiner schreiben wir $\frac{c}{d} := cd^{-1}$.

(3) Es gelten die Distributivgesetze:

$$\begin{aligned} a \cdot (b + c) &= a \cdot b + a \cdot c \\ (b + c) \cdot a &= b \cdot a + c \cdot a \end{aligned}$$

In anderen Worten, K ist auch ein Ring.

Beachten Sie, dass wir auch hier oft ab statt $a \cdot b$ schreiben für $a, b \in K$.

Bemerkung 1.31. Man könnte auf die Definition einer Gruppe und eines Rings verzichten und direkt einen Körper wie in nachfolgender Definition definieren.

Definition 1.32 (alternative direkte Definition)**.** *Ein Körper ist ein Tupel* $(K, +, \cdot, 0, 1)$ *bestehend aus einer Menge* K *mit zwei Abbildungen*

$$+ : K \times K \to K; (x, y) \mapsto x + y$$

$$\cdot : K \times K \to K; (x, y) \mapsto x \cdot y$$

und ausgezeichneten Elementen $0, 1 \in K$*, so dass die Körperaxiome gelten:*

(K1) $\forall x, y, z \in K : x + (y + z) = (x + y) + z$ *(Assoziativität der Addition)*

(K2) $\forall x, y \in K : x + y = y + x$ *(Kommutativität der Addition)*

(K3) $\forall x \in K : 0 + x = x$ *(Neutrales Element der Addition)*

(K4) $\forall x \in K \, \exists x' \in K : x + x' = 0$ *(Inverses Element der Addition)*

(K5) $\forall x, y, z \in K : x \cdot (y \cdot z) = (x \cdot y) \cdot z$ *(Assoziativität der Multiplikation)*

(K6) $\forall x \in K : 1 \cdot x = x$ *(Neutrales Element der Multiplikation)*

(K7) $\forall x \in K \setminus \{0\} \, \exists x' \in K : x' \cdot x = 1$ *(Inverses Element der Multiplikation)*

(K8) $\forall x, y, z \in K : x \cdot (y + z) = x \cdot y + x \cdot z$ *und*
$\forall x, y, z \in K : (y + z) \cdot x = y \cdot x + z \cdot x$ *(Distributivität)*

(K9) $1 \neq 0$ *(Nichttrivialität)*

(K10) $\forall x, y \in K : x \cdot y = y \cdot x$ *(Kommutativität der Multiplikation)*

Die Definition einer Gruppe kann später vielleicht doch hilfreich sein, wenn wir Determinanten besprechen und gewisse Matrix-Gruppen erwähnen. Im Moment ist es wichtig, sich Folgendes zu merken: Unser erstes wichtiges Thema sind Vektorräume über Körper. Körper spielen jetzt also eine wichtige Rolle, Gruppen und Ringe (noch) nicht.

Lemma 1.33. *Wir haben folgende Folgerungen aus der Definition eines Körpers:*

(a) Es gilt $1 \neq 0$*. Das heisst, jeder Körper hat mindestens zwei Elemente.*

(b) Es gilt $0 \cdot a = a \cdot 0 = 0$*.*

(c) Falls $a \cdot b = 0$, dann ist $a = 0$ oder $b = 0$. Anders gesagt, ist jeder Körper nullteilerfrei.

(d) Wir haben $a \cdot (-b) = -(a \cdot b)$ und $(-a) \cdot (-b) = a \cdot b$.

(e) Falls $x \cdot a = y \cdot a$ mit $a \in K^\times$ und $x, y \in K$, dann folgt $x = y$.

Beweis. (a) Da $K^\times$ eine Gruppe ist, ist $1 \in K^\times$. Das Element 0 ist hingegen per Definition nicht in $K^\times$. Daher ist $1 \neq 0$.

(b) Dies haben wir zuvor schon bewiesen.

(c) Die Tatsache, dass $K^\times$ eine Gruppe ist, impliziert insbesondere, dass $K^\times$ abgeschlossen ist bezüglich Multiplikation. Das heisst, aus $a, b \in K^\times$ folgt $ab \in K^\times$. Dies ist genau die Kontraposition der Aussage in (c).

(d) Bemerken Sie, dass

$$ab + a(-b) = a(b + (-b)) = a \cdot 0 = 0,$$

wobei die erste Gleichheit wegen Distributivität gilt, die zweite da $-b$ die Inverse von b ist und die letzte Gleichheit gilt wegen (b). Daraus folgt (Wieso? Eindeutigkeit der Inversen), dass

$$a(-b) = -(ab). \tag{1.9}$$

Für die zweite Aussage verwenden wir (1.9):

$$(-a)(-b) \overset{(1.9)}{=} -(-(a)b) = -(b(-a)) \overset{(1.9)}{=} -(-(ba)) = -(-(ab)) = ab, \tag{1.10}$$

wobei die letzte Gleichheit wegen Proposition 1.2 gilt.

(e) Hier geben wir nur einen Hinweis: Multiplizieren Sie die Gleichheit $xa = ya$ mit $a^{-1} \in K^\times$ von der rechten Seite.

Andere Folgerungen sehen sie in der Serie/Übungsstunde. □

Beispiel 1.34. Sie haben in der Analysis schon die Körper $\mathbb{Q}, \mathbb{R}$ und $\mathbb{C}$ gesehen. Diese Körper sind auch für die lineare Algebra sehr wichtig. Falls Sie Fragen haben, dann sollten Sie uns fragen!

Wir kehren zu unserem Restklassen-Beispiel zurück, um einige endliche Körper zu bestimmen.

Lemma 1.35. *Der Ring $\mathbb{Z}/n\mathbb{Z}$ ist genau dann nullteilerfrei, wenn n eine Primzahl ist.*

Beweis. „$\Rightarrow$“ Hier verwenden wir die Kontraposition. Angenommen n ist nicht prim, dann existieren $1 < k, l < n$, sodass $n = kl$[6]. Damit gilt in $\mathbb{Z}/n\mathbb{Z}$:

$$\overline{0} = \overline{n} = \overline{k} \cdot \overline{l},$$

aber $\overline{k} \neq 0$ und $\overline{l} \neq 0$ (Wieso?). Daher ist $\mathbb{Z}/n\mathbb{Z}$ nicht nullteilerfrei. „$\Leftarrow$“ Nehmen wir jetzt an, dass n eine Primzahl ist. Seien $\overline{k}, \overline{l} \in \mathbb{Z}/n\mathbb{Z}$ mit

$$\overline{k} \cdot \overline{l} = \overline{0}.$$

Daraus folgt, dass ein $r \in \mathbb{Z}$ existiert mit $kl = rn$ und insbesondere teilt die Primzahl n das Produkt kl. Dies impliziert[7], dass ((n teilt k) oder (n teilt l)), was äquivalent ist zu ($\overline{k} = 0$ oder $\overline{l} = 0$). Dies zeigt, dass $\mathbb{Z}/n\mathbb{Z}$ nullteilerfrei ist. $\square$

Bemerkung 1.36. Wir möchten jetzt nicht gross darauf eingehen; nur als Bemerkung: Wir sagen, dass ein Element in einem nullteilerfreien Ring R zwei verschiedene Eigenschaften haben kann:

- Man sagt, dass $p \in R$ prim ist, falls p teilt ab für $a, b \in R$ stets impliziert, dass p teilt a oder p teilt b. In Quantoren ausgedrückt:

$$\forall a, b \in R : p \text{ teilt } ab \Rightarrow p \text{ teilt } a \vee p \text{ teilt } b.$$

- Man sagt, dass $p \in R$ irreduzibel ist, falls p kein Produkt von zwei Elementen in R ist, die nicht invertierbar sind. In Quantoren ausgedrückt:

$$\forall a, b \in R : p = ab \Rightarrow a \in R^\times \vee b \in R^\times.$$

In $\mathbb{Z}$ sind diese beiden Eigenschaften äquivalent. Die bekannte Tatsache, dass es in $\mathbb{Z}$ eine Primfaktorzerlegung gibt ist äquivalent zu der Aussage, dass in $\mathbb{Z}$ Primelemente und irreduzible Elemente gleich sind.

Lemma 1.37. *Sei R ein nullteilerfreier kommutativer Ring mit 1, so dass $1 \neq 0$. Falls R endlich vielen Elementen besitzt, dann ist R ein Körper.*

Beweis. Sei R ein Ring wie in den Voraussetzungen des Lemmas. Um zu zeigen, dass R ein Körper ist, genügt es zu zeigen, dass jedes $a \in R \setminus \{0\}$ eine multiplikative Inverse hat (Wieso?). Der Trick ist die folgende „Multiplikation-mit-a-Abbildung“ zu betrachten:

$$\begin{aligned} m_a : R \setminus \{0\} &\to R \setminus \{0\} \\ b &\mapsto m_a(b) := ab \end{aligned}$$

[6] Hier benutzt man, dass Primzahlen in $\mathbb{Z}$ irreduzibel sind, siehe Bemerkung 1.36.
[7] Hier verwendet man, dass Primzahlen prim sind, siehe Bemerkung 1.36.

Da R nullteilerfrei ist, ist m_a definiert (d.h. die Zielmenge von m_a ist in der Tat in $R \setminus \{0\}$). Ausserdem folgt die Injektivität aus der Nullteilerfreiheit: Seien $b, b' \in R \setminus \{0\}$ mit $m_a(b) = ab = ab' = m_a(b')$. Dann folgt (aus der Inverse und Distributivität), dass $a(b - b') = 0$. Da $a \neq 0$, folgt aus der Nullteilerfreiheit, dass $b - b' = 0$. Dies zeigt die Injektivität.

Wir haben noch nicht die Annahme, dass $|R| < \infty$ (und daher $|R \setminus \{0\}| < \infty$) benutzt. Aus dem Schubfachprinzip folgt: Weil $|R \setminus \{0\}| < \infty$ und m_a injektiv ist, ist m_a auch surjektiv. Bemerken Sie nun zuletzt, dass $1 \in R \setminus \{0\}$ (per Annahme) und daher existiert ein $b \in R$ mit $ab = m_a(b) = 1$. Dies zeigt, dass a eine Inverse besitzt, was wir zeigen wollten. □

Korollar 1.38. *Sei p eine Primzahl. Dann ist $\mathbb{Z}/p\mathbb{Z}$ ein Körper mit p Elementen. Das Nullelement ist $\overline{0}$ und das Einselement ist $\overline{1}$.*

Bemerkung 1.39. Der Körper[8] $\mathbb{Z}/p\mathbb{Z}$ für p eine Primzahl ist zentral im Gebiet der Algebra und Zahlentheorie. In Serie 5 haben Sie einige lustige Aufgaben dazu.

Übung 1.40. Wir haben eigentlich gezeigt, dass a eine rechtsseitige Inverse besitzt. Zeigen Sie, dass eine rechtsseitige Inverse auch eine linksseitige Inverse ist. (vgl. Fischer [6, 1.2.3, b)])

Definition 1.41. Sei $(K, 0, 1)$ ein Körper. Für $n \in \mathbb{N}$ definieren[9] wir

$$n \cdot 1 = \underbrace{1 + \ldots + 1}_{n\text{-Mal}} \in K.$$

Die *Charakteristik* von K ist dann definiert als

$$\operatorname{char}(K) = \begin{cases} 0, & \text{falls } n \cdot 1 \neq 0 \text{ für alle } n \in \mathbb{N} \\ \min\{n \mid n \cdot 1 = 0\}, & \text{sonst} \end{cases}.$$

Beachten Sie, dass $\operatorname{char}(K) \in \mathbb{N} \cup \{0\}$!

Beispiel 1.42. Wir berechnen einige Charakteristiken für uns schon bekannte Körper:

- Für $\mathbb{Q}, \mathbb{R}$ und $\mathbb{C}$ gilt $\operatorname{char}(\mathbb{Q}) = \operatorname{char}(\mathbb{R}) = \operatorname{char}(\mathbb{C}) = 0$.
- Für $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ gilt $\operatorname{char}(\mathbb{F}_p) = p$. (Wieso?)

Man kann mit einem ähnlichen Argument wie in Lemma 1.35 zeigen, dass die Charakteristik eines Körpers 0 oder eine Primzahl $p \in \mathbb{N}$ ist.

[8] Normalerweise bezeichnet man diesen Körper mit $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$.
[9] Hier ist eine Induktion/rekursive Definition versteckt.

1.4 Polynome

Die letzte Vorbereitung, bevor wir endlich anfangen können (mit linearer Algebra), ist es den Raum von Polynomen über einem Körper einzuführen. In der Mittelschule betrachtet man Polynome als Funktionen, aber wir möchten Polynome als formale Objekte betrachten. Dies bereitet in der Regel einigen Studenten Mühe. Daher erklären wir unten den subtilen Unterschied zwischen Polynomen als formalen Objekten und ihren entsprechenden Funktionen. Zuerst aber definieren wir Polynome:

Grob gesagt, ist ein Polynom ein Ausdruck, der durch Konstanten und Symbole (Variablen oder Unbestimmte genannt) mittels Addition, Multiplikation und Potenzieren (mit nicht-negativen ganzen Potenzen) gebildet werden kann. Zwei solcher Ausdrücke, welche durch anwenden von Kommutativität, Assoziativität und Distributivität von Addition und Multiplikation ineinander überführt werden können, definieren dasselbe Polynom. Nun die formale Definition:

Definition 1.43. Sei K ein Körper und x eine Unbestimmte (unbekannte Variable)[10]. Wir nennen eine formalen Ausdruck der Form

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n, \tag{1.11}$$

wobei $n \in \mathbb{N}$ und $a_i \in K$ für alle $0 \leq i \leq n$, ein *Polynom* über K. Die Menge

$$K[x] = \{f(x) \text{ wie in } (1.11),\ \text{mit } n \in \mathbb{N} \text{ und } a_i \in K \text{ für alle } 0 \leq i \leq n\}$$

heisst der *Polynomraum* über K (mit Unbestimmter x).

Bemerkung 1.44. Um ganz präzis zu sein, müssen wir sagen, dass ein Polynom über K eindeutig bestimmt ist durch eine Folge $\{a_k\}_{k=0}^{\infty}$ mit $a_k \in K$ für alle $k \geq 0$, so dass ein $n \in \mathbb{N}$ existiert mit $a_k = 0$ für alle $k > n$. Zum Beispiel entspricht das Polynom $1 - x^2$ der Folge $(1, 0, -1, 0, 0, \ldots)$ und die Polynome $1 - x^2$, $1 + 0 \cdot x - x^2$ und $1 - x^2 + 0 \cdot x^7$ sind alle gleich. Diese Bemerkung ist vielleicht sehr wichtig für unser „Computer-Programm“ (bestehend aus formalen Definitionen und Theoremen), das wir langsam am Schreiben sind, aber jeder von uns versteht schon, wenn zwei Polynome gleich sind durch die (impräzise) Erklärung vor der Definition.

Hier sind noch einige Definitionen und Notationen, die wir brauchen werden:

Definition 1.45. Das Polynom mit $a_k = 0$ für alle $k \in \mathbb{N} \cup \{0\}$ heisst das *Nullpolynom* und wird einfach mit $0 \in K[x]$ bezeichnet. Normalerweise schreiben wir $f \in K[x]$ statt

[10] Denken Sie bei x als Platzhalter, oder besser gesagt, als sogenanntes „freies Element “: „frei“ im Sinn, dass es nichts mit K zu tun hat, und „Element“ im Sinn, dass wir Potenzen davon bilden können, wie zum Beispiel $x^n = \underbrace{x \cdot \ldots \cdot x}_{n\text{-Mal}}$.

$f(x) \in K[x]$. Der Grad von $f \in K[x]$ wie in (1.11) definieren wir als

$$\deg(f) = \begin{cases} -\infty, & \text{falls } f = 0 \\ \max\{k \in \mathbb{N} \cup \{0\} \mid a_k \neq 0\}, & \text{sonst} \end{cases}.$$

Manchmal verwenden wir das Summenzeichen, um Polynome kürzer zu schreiben. Dabei definieren wir $x^0 = 1$ und $x^1 = x$:

$$\sum_{k=0}^{n} a_k x^k = a_0 x^0 + a_1 x^1 + \ldots + a_n x^n = a_0 + a_1 x + \ldots + a_n x^n.$$

Der höchste Koeffizient, der ungleich Null ist, heisst der *Leitkoeffizient* (oder der führende Koeffizient). Wenn der Leitkoeffizient 1 ist, dann heisst das Polynom *normiert.* Polynome der Form $a_l x^l$ für ein $l \in \mathbb{N}$ heissen *Monome*. Wir können zwei Polynome $f, g \in K[x]$ addieren und multiplizieren, indem wir die üblichen Regeln der Kommutativität, Assoziativität und Distributivität sowie Potenzregeln benutzen. Zum Beispiel in $\mathbb{Q}[x]$: Seien

$$f = 3x^2 + 4x + 2 \quad \text{und} \quad g = x^3 - 3x + 4$$

Dann ist

$$\begin{aligned} f + g &= (3x^2 + 4x + 2) + (x^3 - 3x + 4) = x^3 + 3x^2 + x + 6, \\ f \cdot g &= (3x^2 + 4x + 2) \cdot (x^3 - 3x + 4) = 3x^5 + 4x^4 - 7x^3 + 10x + 8. \end{aligned}$$

Übung 1.46. Für $n \in \mathbb{N} \cup \{0\}$ definieren wir die Regel $-\infty + n = -\infty$. Seien $f, g \in K[x]$ nicht Null. Zeigen Sie (unter Benutzung der Regel), dass

$$\deg(fg) = \deg(f) + \deg(g).$$

Gilt

$$\deg(f + g) = \max(\deg(f), \deg(g))?$$

Wenn ja, beweisen Sie es. Wenn nein, überlegen Sie sich, in welchen Fällen es nicht gilt und in welchen Fällen es doch gilt!

Bemerkung 1.47. Wie Sie oben sehen, erlauben wir uns die Freiheit Polynome nicht immer mit aufsteigenden Potenzen zu schreiben.

Bemerkung 1.48. Unser „Computer-Programm“ ist verwirrt. Hier ist eine präzise Definition der Addition und Multiplikation: Seien $f, g \in K[x]$ mit entsprechenden Folgen $\{a_k\}_{k=0}^{\infty}$ und $\{b_k\}_{k=0}^{\infty}$.

- Wir definieren $f+g$ als das eindeutige Polynom, das der Folge $\{c_k\}_{k=0}^{\infty}$ mit $c_k = a_k + b_k$ entspricht. Auch für $\{c_k\}_{k=0}^{\infty}$ gilt, dass $c_k \in K$ für alle $k \geq 0$, und dass ein $M \in \mathbb{N}$ existiert mit $c_k = 0$ für alle $k \geq M$. (Wieso haben wir diesen Satz geschrieben? Wieso existiert ein solches M?)

- Die Formel für die Multiplikation sieht etwas seltsam aus in dieser „Folgen-Schreibweise". Dies erklärt wieso die Darstellung in (1.11) besser ist, obwohl sie weniger präzis ist. Wir definieren $f \cdot g$ als das eindeutige Polynom, das der Folge $\{c_k\}_{k=0}^{\infty}$ entspricht, wobei

$$c_k := a_0 b_k + a_1 b_{k-1} + \ldots + a_k b_0 = \sum_{i=0}^{k} a_i b_{k-i} = \sum_{\substack{0 \leq i,j \leq k \\ i+j=k}} a_i b_j. \tag{1.12}$$

(Üben Sie ihr Verständnis des Summenzeichens, indem Sie (1.12) gut verstehen).

Definition 1.49. Elemente von $\mathbb{Q}[x]$ (resp. $\mathbb{R}[x], \mathbb{C}[x]$) heissen *rationale Polynome* (resp. *reelle Polynome*, *komplexe Polynome*).

Definition 1.50 (Teilbarkeit)**.** Seien $f, g \in K[x]$. Wir sagen, f *teilt* g und schreiben $f \mid g$, falls ein $h \in K[x]$ existiert, so dass $g = f \cdot h$.

Ziele von diesem Abschnitt : (1) Der Unterschied zwischen Polynomen und Polynomfunktionen zu verstehen.

(2) Wir möchten in $K[x]$ dividieren können, aber das geht nicht immer. Der Ersatz ist „Division mit Rest in $K[x]$".

(3) Nullstellen in K für Polynome in $K[x]$ besprechen. Insbesondere beweisen wir: $\lambda \in K$ ist eine Nullstelle von f genau dann, wenn f durch $(x-\lambda)$ geteilt wird.

(4) Fakten über Nullstellen von komplexen und reellen Polynomen besprechen.

1.4.1 Unterschied zwischen Polynomen und Polynomfunktionen

Die Unbestimmte x in $K[x]$ wartet darauf, dass jemand kommt und etwas in sie einsetzt. In diesem Sinn kann man bei x an einen Platzhalter denken. Mann kann verschiedene Dinge in x einsetzen: Die „üblichen Verdächtigen" sind Elemente von K. Wir werden später jedoch auch Matrizen und Funktionen in x einsetzen.[11]

[11] Fortgeschrittene Aussage (d.h. man kann dies sorglos ignorieren): Es ergibt sicherlich Sinn Elemente einer beliebigen K-Algebra einzusetzen.

Betrachten wir was passiert, wenn wir Elemente von K einsetzen. Wir definieren die „Auswertungs-Abbildung“ (evaluation auf Englisch) durch

$$\begin{aligned} \mathrm{ev}_K : K[X] &\to \mathrm{Abb}(K,K) = \{\text{Funktionen von } K \text{ nach } K\}, \\ f &\mapsto \mathrm{ev}_K(f), \end{aligned}$$

wobei

$$\begin{aligned} \mathrm{ev}_K(f) : K &\to K, \\ \lambda &\mapsto f(\lambda) = a_0 + a_1\lambda + \ldots + a_n\lambda^n \end{aligned}$$

für $f = a_0 + a_1x + \ldots + a_nx^n$.

Bemerkung 1.51. Wenn keine Verwechslungsgefahr entsteht, schreiben wir ev statt ev_K.

Wir möchten zwischen $f \in K[x]$ und $\mathrm{ev}_K(f)$ unterscheiden. Der Grund ist, dass ev_K im Allgemeinen nicht injektiv ist!

Beispiel 1.52. Sei $K = \mathbb{F}_3$. Betrachten Sie $f = x^3 - x \in \mathbb{F}_3[x]$ und $g = 0 \in \mathbb{F}_3[x]$. Wir behaupten, dass $\mathrm{ev}(f) = \mathrm{ev}(g)$:

Natürlich ist $\mathrm{ev}(g)$ die konstante Abbildung, die alles auf $\overline{0} \in \mathbb{F}_3$ abbildet. Für $\mathrm{ev}(f)$ müssen wir etwas berechnen:

$$\begin{aligned} \mathrm{ev}(f)(\overline{0}) &= \overline{0}^3 - \overline{0} = \overline{0} \\ \mathrm{ev}(f)(\overline{1}) &= \overline{1}^3 - \overline{1} = \overline{0} \\ \mathrm{ev}(f)(\overline{2}) &= \overline{2}^3 - \overline{2} = \overline{0}. \end{aligned}$$

Also ist in der Tat $\mathrm{ev}(f) = \mathrm{ev}(g)$.

Übung 1.53. Finden Sie $f \neq 0 \in \mathbb{F}_p[x]$ mit $\mathrm{ev}_{\mathbb{F}_p}(f) = 0$, wobei 0 für die Nullfunktion in $\mathrm{Abb}(\mathbb{F}_p, \mathbb{F}_p)$. (Hinweis: In Beispiel 1.52 gilt $x^3 - x = (x - \overline{0})(x - \overline{1})(x - \overline{2})$. Wieso kann man diesen Trick in einem endlichen Körper benutzen?)

Man kann zeigen (wir zeigen in diesem Abschnitt Teile davon): Sei K ein Körper mit unendlich vielen Elementen. Dann ist $\mathrm{ev}_K : K[x] \to \mathrm{Abb}(K,K)$ injektiv.

1.4.2 Division mit Rest in $K[x]$

Lemma 1.54 (Division mit Rest in $K[x]$)**.** *Sei $f \in K[x]$ ein beliebiges Polynom und $0 \neq g \in K[x]$. Dann gibt es Polynome $q, r \in K[x]$, so dass*

$$f = qg + r$$

mit $\deg(r) < \deg(g)$ *oder* $r = 0$.

Wir werden sehen, dass die Polynome q und r durch die Bedingungen in Lemma 1.54 eindeutig bestimmt sind. Des Weiteren bemerken wir, dass der Beweis von Lemma 1.54 einen Algorithmus angibt, wie man q und r finden kann[12].

Beweis von Lemma 1.54. Da $g \neq 0$, gilt $\deg(g) \neq -\infty$. Falls $\deg(g) = 0$, dann ist $g = b_0 \in K \setminus \{0\}$ ein Skalar ungleich Null. Daher können wir $q = b_0^{-1} f$ und $r = 0$ setzen für $f \in K[x]$ und es gilt $f = qg + r$, wie gewünscht. Wir nehmen jetzt an, dass $\deg(g) \geq 1$. In diesem Fall beweisen wir die Existenz von q und r mittels Induktion über $\deg(f)$:

Für die Induktionsverankerung nehmen wir an, dass $\deg(f) < \deg(g)$. (Beachten Sie, dass dies insbesondere den Fall $\deg(f) = 0$ abdeckt.) Wir wählen $q = 0$ und $r = f$ und erhalten $f = qg + r$.

Für den Induktionsschritt nehmen wir nun an, dass $\deg(f) =: m \geq n := \deg(g)$ und die Aussage des Lemmas für alle Polynome mit Grad kleiner als f gilt. Da $\deg(f) = m$ und $\deg(g) = n$, sind f und g von der Form

$$f(x) = a_m x^m + a_{m-1} x^{m-1} + \ldots + a_0$$

mit $a_m \neq 0$ und

$$g(x) = b_n x^n + b_{n-1} x^{n-1} + \ldots + b_0$$

mit $b_n \neq 0$. Betrachten Sie das Polynom

$$f_1(x) = f(x) - \left(\frac{a_m}{b_n}\right) x^{m-n} g(x). \tag{1.13}$$

Beachten Sie, dass die Leitkoeffizienten von $f(x)$ und $(\frac{a_m}{b_n}) x^{m-n} g(x)$ gleich sind und sich somit auslöschen. Also haben wir entweder $\deg(f_1) < \deg(f)$ oder $f_1 = 0$.

Falls $f_1 = 0$ ist, dann haben wir

$$f(x) = \underbrace{\left(\frac{a_m}{b_n} x^{m-n}\right)}_{q(x)} g(x) + \underbrace{0}_{r(x)},$$

wie gewünscht. Falls $f_1 \neq 0$, dann gilt $\deg(f_1) < \deg(f)$. Also können wir die Induktionsannahme auf f_1 anwenden und finden $q_1, r \in K[x]$ mit

$$f_1(x) = q_1(x) g(x) + r(x) \tag{1.14}$$

[12] Dieser Abschnitt folgt [3, §5.2].

und mit $\deg(r) < \deg(g)$. Wir setzen (1.14) in (1.13) ein und erhalten

$$f(x) = f_1(x) + \left(\frac{a_m}{b_n}\right)x^{m-n}g(x) = \underbrace{\left(q_1(x) + \frac{a_m}{b_n}x^{m-n}\right)}_{q(x)}g(x) + r(x),$$

wie gewünscht. Dies schliesst den Induktionsschritt ab und beweist somit das Lemma. □

Übung 1.55. Beweisen Sie die Eindeutigkeit der Polynome q und r in Lemma 1.54.

In der Praxis benutzt man oft die sogenannte „schriftliche Division“, um Division mit Rest durchzuführen. Bei der Division von $x^3 + x^2 + 3x - 1$ durch $x^2 - 1$ erhalten wir beispielsweise $x + 1$ als Quotient und $4x$ als Rest:

$$\begin{array}{l}
(x^3 + x^2 + 3x - 1) : (x^2 - 1) = x + 1 \\
\underline{-x^3 + x} \\
\qquad x^2 + 4x - 1 \\
\qquad \underline{-x^2 + 1} \\
\qquad\qquad 4x
\end{array}$$

Die Art und Weise wie man solch eine schriftliche Division aufschreibt, ist von Land zu Land unterschiedlich. In jedem Fall ist jedoch die Idee, dass man in jedem Schritt versucht die Leitkoeffizienten auszulöschen. Dies ist auch genau die Idee im Beweis von Lemma 1.54.

Übung 1.56. Teilen Sie $x^7 + x^6 - x^5 - x^4 + x^3 + x^2 - x - 1$ durch $x^2 + x + 1$.

1.4.3 Nullstellen

Eine wichtige Konsequenz der Polynomdivision mit Rest ist, dass wir Nullstellensuche und Polynomfaktorisierung zueinander in Beziehung setzen können. Wir sagen, dass $a \in K$ eine Nullstelle von $0 \neq f \in K[x]$ ist, falls $f(a) = 0$.

Korollar 1.57 (Linearfaktorzerlegung)**.** *Sei $a \in K$ und $f \in K[x]$ ein von Null verschiedenes Polynom. Dann ist a eine Nullstelle von f genau dann, wenn $(x - a)$ ein Teiler von f in $K[x]$ ist.*

Beweis. Sei $g(x) = x - a$. Wir wenden Division mit Rest an und erhalten

$$f(x) = q(x)(x - a) + r(x)$$

mit $\deg(r) < \deg(g) = 1$, so dass r eine Konstante sein muss. Wir beweisen nun die Aussage: Wir nehmen zuerst an, dass $f(a) = 0$. Wir setzen für x in obiger Gleichung a

ein und erhalten $r = 0$. Also gilt

$$f(x) = q(x)(x-a) \tag{1.15}$$

und $(x-a)$ ist ein Teiler von f. Für die Umkehrung nehmen wir an, dass $(x-a)$ ein Teiler von f ist und (1.15) gilt. In diesem Fall ist $f(a) = q(a)(a-a) = 0$. $\square$

Sei $a \in K$. Das Polynom $(x-a)$ kann mehrmals ein Polynom $P \in K[x]$ teilen. Dann spricht man von *mehrfache Nullstelle*. Hier ist eine präzisere Definition:

Definition 1.58. Sei $p \in K[x]$ mit $\deg(p) > 0$. Für $a \in K$ definieren wir die *Vielfachheit von a bei p* durch

$$\mu(p \mid a) := \max\{r \in \mathbb{N} \cup \{0\} \mid \exists g \in K[x] \text{ mit } p = (x-a)^r g\}.$$

Beachten Sie: $\mu(p \mid a)$ ist definiert für jedes $a \in K$, und ist einfach 0, wenn a keine Nullstelle von p ist.

Korollar 1.59 (Anzahl Nullstellen)**.** *Für $f \in K[x]$ von Grad $n \in \mathbb{N} \cup \{0\}$ existieren höchstens n verschiedene Zahlen $a \in K$ mit $f(a) = 0$.*

Beweis. Wir beweisen die Aussage per Induktion über $\deg(f)$.

Falls $\deg(f) = 0$ ist, dann entspricht f einer von Null verschiedenen Konstanten und hat keine Nullstellen in K.

Angenommen es gilt $\deg(f) > 0$ und es existiert eine Nullstelle $a \in K$ (sonst sind wir fertig, wieso?). Also gilt nach der Linearfaktorzerlegung (Korollar 1.57), dass ein Polynom $g \in K[x]$ existiert mit

$$f(x) = (x-a)g(x).$$

Wir benutzen, dass $\deg(g_1 g_2) = \deg(g_1) + \deg(g_2)$ ist für beliebige von Null verschiedene Polynome $g_1, g_2 \in K[x]$, um zu sehen, dass $\deg(g) = \deg(f) - 1$ ist. Nach Induktionsannahme (angewendet auf g) hat g höchstens $n-1$ Nullstellen. Eine Nullstellen von f ist entweder a oder eine Nullstelle von g. Also hat f höchstens n Nullstellen und das Korollar folgt. $\square$

1.4.4 Polynome über $\mathbb{R}$ und $\mathbb{C}$

Kurze Einführung und der Fundamentalsatz der Algebra

Wir betrachten ganz schnell die Zahlenbereich-Erweiterungen, die Sie in Analysis besprochen haben und jedes Mal nennen wir einen Grund, wieso wir diese bestimmte Erweiterung brauchen können.

- Von $\mathbb{N}$ nach $\mathbb{Z}$: Weil wir $x + 5 = 3$ lösen wollen.
- Von $\mathbb{Z}$ nach $\mathbb{Q}$: Weil wir $2x + 3 = 0$ lösen wollen.
- Von $\mathbb{Q}$ nach $\mathbb{R}$: Weil wir $x^2 - 2 = 0$ lösen wollen.
- Von $\mathbb{R}$ nach $\mathbb{C}$: Weil wir $x^2 + 1 = 0$ lösen wollen.

Der Leser sollte sich fragen: Wohin jetzt? Eigentlich gibt es viele Antworten[13]. In Anbetracht von Polynome sind wir allerdings schon am Ziel! Dies ist der Inhalt des folgenden viel zelebrierten Satzes.

Theorem 1.60 (Fundamentalsatz der Algebra)**.** *Sei $P \in \mathbb{C}[x]$ mit $\deg(P) > 0$. Dann existiert $\lambda \in \mathbb{C}$ mit $P(\lambda) = 0$.*

Mittels Korollar 1.57 folgt:

Korollar 1.61. *Sei $P \in \mathbb{C}[x]$ mit $\deg(P) = n > 0$ und Leitkoeffizient $a \in \mathbb{C}$. Dann existieren $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$ paarweise verschieden, so dass*

$$P = a(x - \lambda_1)^{l_1} \cdot \ldots \cdot (x - \lambda_k)^{l_k} = a \prod_{i=1}^{k} (x - \lambda_i)^{l_i} \tag{1.16}$$

und $l_1 + \ldots + l_k = \sum_{i=1}^{k} l_i = n$ gilt.

Beweis-Idee. Solche Aussagen beweist man mit Induktion über $\deg(P)$. Der Induktionsschritt ist dank Theorem 1.60 machbar. Hier ist kurz die Idee: Theorem 1.60 impliziert für $\deg(P) > 0$, dass es ein $\lambda \in \mathbb{C}$ gibt mit $P(\lambda) = 0$. Letzteres impliziert wiederum, dass $(x - \lambda)$ ein Teiler ist von P, also können wir $P = (x - \lambda)\tilde{P}$ schreiben mit $\deg(\tilde{P}) = n - 1$. (Versuchen Sie einen schönen Induktions-Beweis zu schreiben, Forum-Wieso.) $\square$

Dies bedeutet, dass wir keine weiteren Erweiterungen von Zahlen mehr brauchen, falls wir Polynom-Gleichungen lösen möchten: Alle Lösungen sind schon in $\mathbb{C}$. Ausdrücke wie in (1.16) nennt man eine Zerlegung von P in lineare Faktoren.

Reelle Polynome

In $\mathbb{R}$ ist die Sache etwas komplizierter: Es gibt Polynome $P \in \mathbb{R}[x]$ mit $\deg(P) > 1$, die keine Nullstellen über $\mathbb{R}$ haben, das heisst es gibt kein $\lambda \in \mathbb{R}$ mit $P(\lambda) = 0$.

Zum Beispiel ist $P = x^2+1$ ein solches Polynom. Schlechter kann es jedoch eigentlich nicht werden, wie wir jetzt erklären.

[13] Eine Antwort sind die Quaternionen.

Lemma 1.62. *Sei $p \in \mathbb{R}[x] \subseteq \mathbb{C}[x]$ und $\lambda \in \mathbb{C}$ eine Nullstelle von P. Dann ist $\overline{\lambda}$ auch eine Nullstelle von P. (Hier steht $\overline{\lambda}$ für die komplexe Konjugierte von λ.)*

Beweis. Sei $P(x) = a_n x^n + \ldots + a_1 x + a_0$. Laut Annahme gilt

$$P(\lambda) = a_n \lambda^n + \ldots + a_1 \lambda + a_0 = 0. \tag{1.17}$$

Erinnern Sie sich an die Eigenschaften der komplexen Konjugation aus der Analysis: Für alle $\alpha, \beta \in \mathbb{C}$ gilt:

- $\overline{\alpha \cdot \beta} = \overline{\alpha} \cdot \overline{\beta}$,
- $\overline{\alpha + \beta} = \overline{\alpha} + \overline{\beta}$,
- $\overline{\alpha} = \alpha$ genau dann, wenn $\alpha \in \mathbb{R}$.

Wir konjugieren Gleichung (1.17) und benutzen obige Eigenschaften, so dass wir

$$\begin{aligned}
\overline{P(\lambda)} = \overline{a_n \lambda^n + \ldots + a_1 \lambda + a_0} &= \overline{a_n}\,\overline{\lambda^n} + \ldots + \overline{a_1}\,\overline{\lambda} + \overline{a_0} \\
&= \underbrace{a_n \overline{\lambda}^n + \ldots + a_1 \overline{\lambda} + a_0}_{P(\overline{\lambda})} \\
&= \overline{0} \\
&= 0.
\end{aligned}$$

Also ist $P(\overline{\lambda}) = 0$, was wir zeigen wollten. □

Folgendes Lemma ist eine stärkere Version von Lemma 1.62.

Lemma 1.63. *Sei $\lambda \in \mathbb{C} \setminus \mathbb{R}$ und $q_\lambda := (x - \lambda)(x - \overline{\lambda})$. Dann gilt $q_\lambda \in \mathbb{R}[x]$, $\deg(q_\lambda) = 2$ und q_λ hat keine reelle Nullstelle.*

Beweis. Wir lösen die Klammer auf

$$q_\lambda := (x - \lambda)(x - \overline{\lambda}) = x^2 - (\lambda + \overline{\lambda})x + \lambda\overline{\lambda}.$$

Da $\overline{\lambda + \overline{\lambda}} = \overline{\lambda} + \lambda = \lambda + \overline{\lambda}$ und $\overline{\lambda \cdot \overline{\lambda}} = \overline{\lambda} \cdot \overline{\overline{\lambda}} = \overline{\lambda} \cdot \lambda = \lambda \cdot \overline{\lambda}$ gilt, folgt $q_\lambda \in \mathbb{R}[x]$. Aus der Definition von P folgt, dass $\deg(P) = 2$. Für die letzte Aussage beachten Sie, dass λ und $\overline{\lambda}$ zwei Nullstellen von P sind, die nicht in $\mathbb{R}$ sind. (Dies gilt, weil $\lambda \in \mathbb{C} \setminus \mathbb{R}$ impliziert $\overline{\lambda} \in \mathbb{C} \setminus \mathbb{R}$. Wieso gilt die Implikation?) Aus Korollar 1.61 folgt nun, dass P keine anderen Nullstellen haben kann, also insbesondere auch keine reelle Nullstelle. □

Jetzt können wir eine stärkere Version von Lemma 1.62 beweisen:

Lemma 1.64. *Sei $P \in \mathbb{R}[x]$ und $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Dann ist $\mu(P \mid \lambda) = \mu(P \mid \overline{\lambda})$.*

Beweis. Wir beweisen das Lemma per Induktion über $\mu(P \mid \lambda)$.

Falls $\mu(P \mid \lambda) = 0$, dann folgt $\mu(P \mid \overline{\lambda}) = 0$ aus Lemma 1.62. (Wieso? $\lambda = \overline{\overline{\lambda}}$.)

Wir nehmen jetzt an, dass $\mu(\tilde{P} \mid \tilde{\lambda}) = n$ genau dann, wenn $\mu(\tilde{P} \mid \overline{\tilde{\lambda}}) = n$ für alle $\tilde{P} \in \mathbb{R}[x]$ mit Nullstelle $\tilde{\lambda} \in \mathbb{C} \setminus \mathbb{R}$ und $\mu(P \mid \lambda) = n+1$. Per Definition von $\mu(P \mid \lambda)$ ist $P(\lambda) = 0$ und nach Lemma 1.62 gilt auch $P(\overline{\lambda}) = 0$. Nach Korollar 1.57 sind $(x - \lambda)$ und $(x - \overline{\lambda})$ Teiler von P. Da $\lambda \in \mathbb{C} \setminus \mathbb{R}$ ist, gilt $\lambda \neq \overline{\lambda}$ und daher ist

$$q_\lambda = (x - \lambda)(x - \overline{\lambda}) \in \mathbb{R}[x]$$

ein Teiler von P. Also können wir $P = q_\lambda P'$ für ein $P' \in \mathbb{R}[x]$ schreiben und per Definition der Vielfachheit folgt

$$\mu(P \mid \lambda) = \mu(P' \mid \lambda) + 1 \quad \text{und} \quad \mu(P \mid \overline{\lambda}) = \mu(P' \mid \overline{\lambda}) + 1. \tag{1.18}$$

Laut Annahme ist $\mu(P \mid \lambda) = n+1$ und daher ist $\mu(P' \mid \lambda) = n$. Die Induktionsannahme impliziert dann $\mu(P' \mid \overline{\lambda}) = n$. Aus (1.18) folgt dann $\mu(P \mid \overline{\lambda}) = n + 1$, was wir zeigen wollten. □

Korollar 1.65. *Für jedes $P \in \mathbb{R}[x]$ existieren $k, l \in \mathbb{N} \cup \{0\}$ und $\eta_1, \dots, \eta_k \in \mathbb{R}$ und $\lambda_1, \dots, \lambda_l \in \mathbb{C} \setminus \mathbb{R}$, so dass $\{\eta_1, \dots, \eta_k, \lambda_1, \dots, \lambda_l\, \overline{\lambda_1}, \dots, \overline{\lambda_l}\}$ alle Nullstellen von P sind und $\deg(P) = k + 2l$*

In anderen Worten sagt Korollar 1.65, dass die Nullstellen, die nicht reell sind in Paaren λ und $\overline{\lambda}$ vorkommen.

Korollar 1.66. *Sei $P \in \mathbb{R}[x]$ mit Leitkoeffizient $a \in \mathbb{R}$ und $k, l \in \mathbb{N} \cup \{0\}$, $\eta_1, \dots, \eta_k \in \mathbb{R}$, $\lambda_1, \dots, \lambda_l \in \mathbb{C} \setminus \mathbb{R}$ wie in Korollar 1.65. Dann gilt*

$$P(x) = a(x - \eta_1) \cdot \ldots \cdot (x - \eta_k) \cdot q_{\lambda_1} \cdot \ldots \cdot q_{\lambda_l} = \prod_{i=1}^{k} (x - \eta_i) \prod_{i=1}^{l} q_{\lambda_i}$$

und dieses Produkt kann man nicht weiter in $\mathbb{R}[x]$ zerlegen.

Changelog: Kapitel 1

- 14.10: Im Beweis von Lemma 1.54 wurde der Exponent in (1.13) und darauf folgend zu x^{m-n} geändert.
- 25.10: Im Beweis von Lemma 1.54 wurde im Induktionsschritt $b_0 \neq 0$ zu $b_n \neq 0$ geändert.
- 25.10: In Definition 1.43 wurde die Bedingung in $K[x]$ zu $0 \leq i \leq n$ geändert.
- 09.11: Im Beweis von Lemma 1.64 wurden einige Typos korrigiert.
- 20.12: In Definition 1.45 wurde $\max\{k \in \mathbb{N} \mid a_k \neq 0\}$ zu $\max\{k \in \mathbb{N} \cup \{0\} \mid a_k \neq 0\}$ korrigiert.
- 01.01: Im Beweis von Lemma 1.63 wurde $x^2 - (\lambda - \overline{\lambda})x + \lambda\overline{\lambda}$ zu $x^2 - (\lambda + \overline{\lambda})x + \lambda\overline{\lambda}$ korrigiert.
- 01.01: In Korollar 1.65 wurden die Nullstellen $\overline{\lambda_1}, \ldots, \overline{\lambda_l}$ hinzugefügt.
- 01.01: In Korollar 1.66 wurde der Leitkoeffizient $a \in \mathbb{R}$ hinzugefügt.

Kapitel 2

Vektorräume

Wir besprechen endlich die Hauptobjekte der linearen Algebra, Vektorräume. In unserer Fibonacci-Einführung haben wir schon die grobe Definition von Vektorräumen gesehen: Mengen mit einer gewissen Addition und Skalarmultiplikation. Seither sind wir schon viel besser ausgerüstet mit der Sprache der abstrakten Mathematik. Daher definieren wir:

Definition 2.1. Ein *Vektorraum* über einem Körper K ist eine Menge $(V, +, \cdot)$ mit zwei Verknüpfungen

$$\begin{aligned} + &: V \times V \to V, (v_1, v_2) \mapsto v_1 + v_2 \\ \cdot &: K \times V \to V, (a, v) \mapsto a \cdot v = av, \end{aligned}$$

so dass folgende Axiome gelten:

(V1) $\forall v_1, v_2, v_3 \in V : v_1 + (v_2 + v_3) = (v_1 + v_2) + v_3$

(V2) $\exists 0 = 0_V \in V \, \forall v \in V : 0 + v = v$

(V3) $\forall v \in V \, \exists v' \in V : v + v' = 0$

(V4) $\forall v_1, v_2 \in V : v_1 + v_2 = v_2 + v_1$

(V5) $\forall a, b \in K, v \in V : a \cdot (b \cdot v) = (a \cdot b) \cdot v$

(V6) $\forall v \in V : 1 \cdot v = v$

(V7) $\forall a \in K, v_1, v_2 \in V : a \cdot (v_1 + v_2) = a \cdot v_1 + a \cdot v_2$

(V8) $\forall a_1, a_2 \in K, v \in V : (a_1 + a_2) \cdot v = a_1 \cdot v + a_2 \cdot v$

Wie Sie sehen können, benutzt Definition 2.1 nicht direkt das Konzept einer Gruppe. Wir bemerken lediglich: Die Axiome (V1) bis (V3) implizieren, dass $(V, 0, +)$ eine Gruppe ist und (V4) sagt, dass V eine abelsche Gruppe ist. Das Axiom (V5) beschreibt die Kompatibilität der Skalarmultiplikation mit der Multiplikation im Körper[1] K, in (V6) ist 1 die multiplikative Identität im Körper K und die Axiome (V7) und (V8) beschreiben schlussendlich Distributivitätsregeln.

Die Elemente eines Vektorraums V über einem Körper K nennen wir Vektoren und die Elemente des Körpers K heissen Skalare. Ein Vektorraum über $\mathbb{R}$ (bzw. über $\mathbb{C}$) heisst ein reeller (bzw. ein komplexer) Vektorraum. Genau wie bei Körpern können wir gewisse „Arithmetik-Regeln“ aus den Axiomen herleiten.

Lemma 2.2. *Sei V ein Vektorraum über einem Körper K. Dann gilt:*

(a) Der Nullvektor in (V2) ist eindeutig und wird mit 0_V bezeichnet, wenn Verwechslungsgefahr besteht.

(b) Die additive Inverse in (V3) ist eindeutig und wird mit $-v$ bezeichnet. Damit bedeutet $w - v := w + (-v)$ für $w \in V$.

(c) Für alle $v \in V$ gilt $0 \cdot v = 0$ (oder präziser: $0_K \cdot v = 0_V$).

(d) Für alle $a \in K$ gilt $a \cdot 0 = 0$ (oder präziser: $a \cdot 0_V = 0_V$).

(e) Für alle $v \in V$ gilt $-1 \cdot v = -v$.

(f) Für alle $v \in V$ gilt $-(-v) = v$.

(g) Für alle $a \in K, v \in V$ mit $av = 0$ folgt, dass $a = 0$ oder $v = 0$.

(h) Assoziativität gilt auch für die Summe von n Elementen.

Beweis. Wir lassen fast alle Teile des Lemmas als Übungen, da wir ähnliche Beweise schon mit Körpern gemacht haben. Bemerken Sie, dass (a), (b) und (f) direkt aus Tatsachen über Gruppen folgen, die wir schon bewiesen haben. (Wieso?) Zum Spass beweisen wir (c): Sei $v \in V$, dann gilt

$$0 \cdot v = (0 + 0) \cdot v \overset{(V8)}{=} 0 \cdot v + 0 \cdot v. \tag{2.1}$$

Wir addieren $-0 \cdot v$ auf beide Seiten und erhalten

$$0 \overset{(V3)}{=} 0 \cdot v + (-0 \cdot v) = (0 \cdot v + 0 \cdot v) + (-0 \cdot v) \overset{(V1)\&(V3)}{=} 0 \cdot v + 0 = 0 \cdot v.$$

Vergleichen Sie dies mit dem Beweis von Lemma 1.27.

[1] In (V5) ist $(a \cdot b)$ die Multiplikation im Körper und $(b \cdot v)$ bzw. $a \cdot (b \cdot v)$ ist die Verknüpfung $\cdot : K \times V \to V$ in Definition 2.1!

Für (g) brauchen wir eine neue Idee: Sei $a \in K$ und $v \in V$. Wir nehmen an, dass $a \neq 0$ und $av = 0$. Da K ein Körper ist, gibt es $a^{-1} \in K$ und

$$v \overset{(*)}{=} 1 \cdot v = (a^{-1}a) \cdot v \overset{(*)}{=} a^{-1}(av) = a^{-1} \cdot 0 \overset{(c)}{=} 0.$$

(Erklären Sie für sich selbst, welche Axiome jeweils in $(*)$ benutzt wurden und wie die „Aussagen-Logik" in diesem Beweis funktioniert!) Für (h) bemerken wir, dass man diese Aussage mit Induktion beweisen kann, wobei man (V1) für den Induktionsschritt verwendet. □

Wir geben jetzt (nur) ein Beispiel, welches sowohl einfach als auch zentral ist[2]. Danach definieren wir Untervektorräume und geben viele Beispiele von Vektorräumen und ihren Untervektorräumen.

Beispiel 2.3 (Der Koordinatenraum K^n)**.** Sei K ein Körper und $n \in \mathbb{N}$. Als eine ziemlich direkte Verallgemeinerung der René Descartes Ideen definieren wir den *Koordinatenraum* (über K) durch

$$K^n = \{(a_1, \ldots, a_n) \mid a_i \in K, 1 \leq i \leq n\}$$

Die Addition auf K^n ist durch komponentenweise Addition definiert: Für $v = (a_1, \ldots, a_n)$ und $w = (b_1, \ldots, b_n) \in K^n$ definieren wir

$$v + w = (a_1, \ldots, a_n) + (b_1, \ldots, b_n) := (a_1 + b_1, \ldots, a_n + b_n). \tag{2.2}$$

Hier ist die Addition auf der linken Seite die Addition in K^n, die wir gerade definieren und die Addition auf der rechten Seite ist die Addition in K. In anderen Worten kann man (2.2) auch so schreiben:

$$v +_{K^n} w = (a_1, \ldots, a_n) +_{K^n} (b_1, \ldots, b_n) := (a_1 +_K b_1, \ldots, a_n +_K b_n).$$

Wir benutzen aber die Schreibweise (2.2) , um den Leser zu zwingen sich immer wieder zu fragen, was genau was bedeutet. Skalarmultiplikation ist ebenfalls komponentenweise definiert: Für $a \in K$ ist

$$a \cdot v = a \cdot (a_1, \ldots, a_n) := (aa_1, \ldots, aa_n).$$

Mit diesen Definitionen folgen alle Axiome der Vektorraumstruktur auf K^n aus den entsprechenden Axiomen der Körperstruktur auf K. (Verifizieren Sie einige/alle! Beispielsweise ist $0 := (0, \ldots, 0) \in K^n$ der Nullvektor[3].)

[2] Später werden Sie wohl überrascht sein, wie zentral dieses Beispiel ist!

[3] Auch hier könnte man $0_{K^n} = (0_K, \ldots, 0_K)$ schreiben, aber dann müssten die Leser gar nicht denken, oder?

Bemerkung 2.4. Streng genommen sind die folgenden Vektorräume nicht gleich

$$K^n_{Zeil} := \{(x_1, \ldots, x_n) \mid x_i \in K\} \quad \text{und} \quad K^n_{Spal} := \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mid x_i \in K \right\},$$

obwohl wir bei beiden meistens an K^n denken. Wenn der Unterschied nicht wichtig ist, schreiben wir einfach K^n und benutzen die Notation von K^n_{Zeil}, da sie angenehmer ist zum Schreiben. Manchmal ist dieser Unterschied für uns aber dennoch wichtig. Zum Beispiel müssen wir später bei K^n an K^n_{Spal} denken, wenn wir das Produkt von Matrizen definieren.

Der Koordinatenraum heisst auch der *n-dimensionale Koordinatenraum über K*. Der Leser könnte schon jetzt zustimmen, dass K^n n Dimensionen hat. Was genau damit gemeint ist, müssen wir noch erklären, da wir die Dimension eines Vektorraums noch nicht definiert haben. Die Überraschung wird sein, dass K^n in einem gewissen Sinn der einzige Vektorraum über K mit Dimension n ist! Präziser gesagt, jeder andere Vektorraum über K mit Dimension n ist isomorph zu K^n. Um dort jedoch anzugelangen, müssen wir mit der Plauderei aufhören und weiterreisen…

Definition 2.5. Sei V ein Vektorraum über K. Eine Teilmenge $W \subseteq V$ heisst ein *Untervektorraum* von V, falls das Folgende gilt:

(UVR1) W ist nicht leer.

(UVR2) Für alle $w_1, w_2 \in W$ ist $w_1 + w_2 \in W$.

(UVR3) Für alle $a \in K, w \in W$ ist $aw \in W$.

Übung 2.6. (a) Sei $0 \in V$ der Nullvektor und (UVR1') die Aussage $0 \in W$. Zeigen Sie, dass ((UVR1') und (UVR2) und (UVR3)) äquivalent ist zu ((UVR1) und (UVR2) und (UVR3)). In Aussagenlogik:

$$\text{(UVR1')} \wedge \text{(UVR2)} \wedge \text{(UVR3)} \iff \text{(UVR1)} \wedge \text{(UVR2)} \wedge \text{(UVR3)}.$$

(b) Zeigen Sie, dass die Aussage

$$\forall a_1, a_2 \in K \ \forall w_1, w_2 \in W : a_1 w_1 + a_2 w_2 \in W$$

äquivalent ist zu

$$\text{(UVR2)} \wedge \text{(UVR3)}.$$

Lemma 2.7. *Sei V ein Vektorraum über K und $W \subseteq V$ ein Untervektorraum. Dann ist W auch ein Vektorraum über K mit der induzierten Addition und Skalarmultiplikation von V.*

Beweis. Der Beweis besteht aus der direkten Verifizierung aller Axiome mittels (UVR1), (UVR2) und (UVR3) und der Benutzung von $W \subseteq V$. Beweisen wir zum Beispiel das Axiom (V2) für W: Aus Übung 2.6 folgt, dass $0_V \in W$. Wir behaupten, dass 0_V auch der Nullvektor von W ist: Sei $w \in W$. Da $W \subseteq V$ ist, gilt $0_V + w = w$ (aus (V2) für V). Dies zeigt (V2) für W. Wir überlassen die anderen Axiome dem Leser zum Überprüfen. □

Bemerkung 2.8. Lemma 2.7 ist eigentlich äquivalent zu der Definition 2.5. Präziser gesagt: Man könnte umgekehrt einen Untervektorraum folgendermassen definieren: Eine Teilmenge $W \subseteq V$ eines Vektorraums V heisst ein Unterverktorraum, wenn W ein Vektorraum mit der induziereten Addition und Skalarmultiplikation aus V ist. Dann könnte man beweisen, dass dies (UVR1)-(UVR3) impliziert. Zusammen mit Lemma 2.7 zeigt dies die Äquivalenz der beiden Definitionen.

2.1 Ein Haufen Beispiele und ein bisschen Theorie

In diesem Unterkapitel geben wir einen Haufen Beispielen und ein kleines Theorie-Intermezzo.

Triviale Beispiele

Triviale Beispiele sind eine sehr wichtige Art von Beispielen! Diese können helfen das „Kleingedruckte“ einer Definition zu verstehen.

Beispiel 2.9. Sei V ein Vektorraum. Dann sind $\{0_V\}$ und V selbst Untervektorräume.

Bemerkung 2.10. Sie haben diese Woche interessante Multiple Choice Fragen im Zusammenhang mit diesem Beispiel: Entscheiden Sie, ob die folgenden Aussagen wahr oder falsch sind.

- Jeder Vektorraum hat zwei verschiedene Untervektorräume.
- Jeder Körper hat zwei verschiedene Elemente.
- Jeder Vektorraum hat zwei verschiedene Elemente.

Bitte bearbeiten Sie auch die Multiple Choice Serie!

Untervektorräume von K^n

Beispiel 2.11. Sei $b \in K$. Die Menge

$$U_b = \{(x_1, x_2, x_3) \mid x_1 - x_2 + x_3 = b\} \subseteq K^3$$

ist ein Untervektorraum von K^3 genau dann, wenn $b = 0$. Wir zeigen nur „⇒“ mit Kontraposition: Falls $b \neq 0$ und $(x_1, x_2, x_3), (y_1, y_2, y_3) \in U_b$, dann ist

$$(z_1, z_2, z_3) = (x_1, x_2, x_3) + (y_1, y_2, y_3)$$

nicht in U_b, da

$$z_1 - z_2 + z_3 = b + b \neq b.$$

(Beachten Sie: Die letzte Ungleichung gilt in jedem Körper K. Wieso?)

Theorie-Intermezzo

Wir möchten das letzte Beispiel verallgemeinern und dabei ein grundlegendes Beispiel für Untervektorräume von K^n besprechen. Wir wiederholen die Definition einer Matrix über einem allgemeinen Körper K:

Definition 2.12. Sei K ein Körper und seien $m, n \in \mathbb{N}$. Eine $m \times n$ *Matrix*

$$A = (a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$$

über K ist eine rechteckige Anordnung von Elementen in K mit m Zeilen und n Spalten. Die Skalare a_{ij} heissen die *Einträge* der Matrix A. Wir schreiben $M_{m\times n}(K)$ für die Menge aller $m \times n$ Matrizen über K.

Später auf Seite 51 und in der Übungsstunde werden wir sehen, dass $M_{m\times n}(K)$ auch ein wichtiger Vektorraum ist. Im Moment möchten wir aber nur eine Multiplikation zwischen $m \times n$ Matrizen und Vektoren (geschrieben als Spaltenvektoren) in

$$K^n_{Spal} := \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mid x_i \in K \right\}$$

definieren. Sei $A = (a_{ij})_{ij} \in M_{m\times n}(K)$ und $v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n_{Spal}$.[4] Wir definieren

$$Av = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ a_{21}x_1 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix}. \tag{2.3}$$

[4] Zur Vereinfachung der Notation schreiben wir oft $A = (a_{ij})_{ij}$ statt $A = (a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$.

Bemerken Sie, dass $Av \in K^m_{Spal}$. Man zeigt mit der Definition in (2.3), dass für alle $v, w \in K^n_{Spal}$ und für alle $a \in K$ Folgendes gilt:

$$A(v+w) = Av + Aw \quad \text{und} \quad A(av) = aA(v). \tag{2.4}$$

Definition 2.13. Seien $A \in M_{m\times n}(K), x \in K^n (= K^n_{Spal})$ und $b \in K^m (= K^m_{Spal})$. Eine Gleichung der Form $Ax = b$ heisst ein *lineares Gleichungssystem* über K.

Bemerkung 2.14. Ein solches lineares Gleichungssystem können Sie mittels Gauss-Elimination über K lösen, genau wie Sie dies in der Serie und in der Übungsstunde gemacht haben.

Jetzt sind wir bereit, eine Verallgemeinerung von Beispiel 2.11 zu geben.

Beispiel 2.15 (Lösungen von linearen Gleichungssystemen)**.** Sei $A \in M_{m\times n}(K)$ und $b \in K^m (= K^m_{Spal})$. Dann ist

$$L = L_{A,b} = \{x \in K^n \mid Ax = b\} \subseteq K^n$$

ein Untervektorraum genau dann, wenn $b = 0 \in K^m$.

Beweis. „$\Leftarrow$“: Sei $b = 0 = 0_{K^m}$. Wir überprüfen die Axiome eines Untervektorraums:

- $0 = 0_{K^n} \in L$, da $A0_{K^n} = 0_{K^m}$.
- Seien $v_1, v_2 \in L$ und $a_1, a_2 \in K$. Laut (2.4) gilt, dass $A(a_1v_1 + a_2v_2) = a_1Av_1 + a_2Av_2 = a_10_{K^m} + a_20_{K^m} = 0_{K^m}$.

Daher ist L ein Untervektorraum. Für „$\Rightarrow$“ ist der Beweis sehr ähnlich zu dem Beweis in Beispiel 2.11. Merken Sie sich lediglich, dass in K^n auch gilt, dass $0 \neq b \in K^m$ die Ungleichung $b + b \neq b$ impliziert. (Wieso?) $\square$

Dieses Beispiel ist wirklich grundlegend: Wir werden bald zeigen, dass jeder Untervektorraum von K^n die Form $L_{A,b}$ für gewisse A und b hat. Dies ist auch der Grund dafür, dass man durch Gauss-Elimination die ganze lineare Algebra auf K^n erhalten kann. Dies allein wäre aber nicht sehr schön...

Bemerkung 2.16. Wir werden später noch sehen, dass $L_{A,b}$ auch viel Struktur hat, wenn $b \neq 0$. (Hinweis: Es gilt $L_{A,b} = s + L_{A,0} = \{s + v \mid v \in L_{A,0}\}$ für irgendein $s \in L_{A,b}$. Daher ist $L_{A,b}$ ein „verschobener“ Untervektorraum.)

Hier ist ein einfaches Beispiel, das wir später nochmals betrachten:

Beispiel 2.17. Die Menge $W = \{(x_1, \ldots, x_6) \mid x_1 + \ldots + x_6 = 0\} \subseteq K^6$ ist ein Untervektorraum.

Beweis. Es gilt $W = L_{A,0}$ für $A = (1\ 1 \cdots 1) \in M_{1,6}(K)$. $\square$

Folgenräume

In diesem (Unter-)Kapitel stellen wir einen Zusammenhang zu unserer Fibonacci-Einführung her. Sei K ein Körper. Wir definieren den *Folgenraum* K^∞ durch

$$K^\infty = \{(a_i)_{i=1}^\infty \mid a_i \in K,\ i = 1, 2, \ldots\}.$$

Wie in der Einführung definieren wir die Addition und Skalarmultiplikation folgendermassen:

$$(a_1, a_2, \ldots) + (b_1, b_2, \ldots) := (a_1 + b_1, a_2 + b_2, \ldots), \tag{2.5}$$

$$a \cdot (a_1, a_2, \ldots) := (aa_1, aa_2, \ldots). \tag{2.6}$$

Behauptung: K^∞ mit der Addition und Skalarmultiplikation in (2.5) ist ein Vektorraum.

Übung 2.18. Überprüfen Sie diese Behauptung, bis sie für Sie so offensichtlich wird wie die Tatsache, dass K^n ein Vektorraum ist.

Bemerkung 2.19. Hier indizieren wir die Folgen mit den natürlichen Zahlen $\mathbb{N}$, so dass die Folgen in K^∞ also beim Index 1 starten. Wir könnten natürlich auch die Indizes aus $\mathbb{N} \cup \{0\}$ wählen wie in der Einführung oder allgemeiner könnten wir bei einer beliebigen Zahl in $\mathbb{Z}$ starten.

Eine ähnliche, aber andere Konstruktion, sind zweiseitige Folgen: Sei

$$\begin{aligned} K_{-\infty}^\infty &= \{(a_i)_{i\in\mathbb{Z}} \mid a_i \in K,\ i \in \mathbb{Z}\} \\ &= \{(a_i)_{i=-\infty}^\infty \mid a_i \in K,\ i \in \mathbb{Z}\} \\ &= \{(\ldots, a_{-1}, a_0, a_1, \ldots) \mid a_i \in K\}. \end{aligned}$$

Dies ist auch ein Vektorraum.

Untervektorräume von Folgenräumen

Zuerst, um auf unsere Einführung zurück zu kommen, betrachten wir **Fib**:

Beispiel 2.20. Sei $\mathbf{Fib} = \{(a_i)_{i=1}^\infty \mid a_n = a_{n-1} + a_{n-2}$ für alle $n \geq 3\}$. Dann ist **Fib** ein Untervektorraum von K^∞. Allgemeiner gibt uns jede lineare Rekursions-Regel einen Untervektorraum. Zum Beispiel, seien $\alpha, \beta \in K$ zwei fixierte Zahlen. Dann ist

$$U_{\alpha,\beta} = \{(a_i)_{i=1}^\infty \mid a_n = \alpha a_{n-1} + \beta a_{n-2} \text{ für alle } n \geq 3\}$$

ein Untervektorraum von K^∞. Den Beweis davon haben wir eigentlich schon in der Einführung gesehen!

Übung 2.21. Ist $W = \{(a_i)_{i=1}^\infty \mid a_n = a_{n-1}^2 \text{ für alle } n \geq 2\} \subseteq K^\infty$ ein Untervektorraum? (Hinweis: Die Antwort hängt von etwas ab! (Siehe Forum-Wieso))

Beispiel 2.22. Für jedes $n \in \mathbb{N}$ definieren wir

$$W_n := \{(a_i)_{i=1}^\infty \mid a_i = 0 \text{ für alle } i > n\} \subseteq K^\infty.$$

Es gilt $W_n \subseteq W_m$ genau dann, wenn $n \leq m$. Wir definieren $W_\infty := \bigcup_{n \geq 1} W_n$.

Übung 2.23. Beweisen Sie, dass $W_\infty = \{(a_i)_{i=1}^\infty \mid \exists M \in \mathbb{N}\ \forall i > M : a_i = 0\}$ ist, und dass W_∞ ein Untervektorraum ist. Der Raum W_∞ heisst *Raum der Folgen mit endlichem Träger*[5].

Wir verallgemeinern diese Familie von Beispielen weiter.

Funktions-Räume

Wie immer sei K ein Körper, und sei S eine nicht leere Menge. Wir definieren

$$K^S = \{f : S \to K \mid f \text{ ist eine Funktion}\}.$$

Eine andere Notation ist also $K^S = \mathrm{Abb}(S, K)$. Wir möchten eine Addition und Skalarmultiplikation auf K^S definieren, um K^S mit einer Vektorraumstruktur zu versehen. Dazu kann man die Addition und Multiplikation auf K benutzen: Wir definieren für $f, g \in K^S$ und $a \in K$:

$$\begin{aligned} (f +_{K^S} g)(x) &= f(x) +_K g(x), \\ (a \cdot_{K^S} f)(x) &= a \cdot_K f(x). \end{aligned} \tag{2.7}$$

In anderen Worten ist $f + g$ die Funktion

$$\begin{aligned} f + g : S &\to K \\ x &\mapsto f(x) + g(x) \end{aligned}$$

und in ähnlicher Weise ist $a \cdot f$ die Funktion

$$\begin{aligned} a \cdot f : S &\to K \\ x &\mapsto a \cdot f(x). \end{aligned}$$

Bemerkung 2.24. Wir waren nett zu den Lesern und haben in (2.7) geschrieben, um welche Addition und Multiplikation es sich jeweils handelt. Auch die zweite Erklärung

[5]„Endlicher Träger“ bezieht sich auf die Tatsache, dass jeweils nur endlich viele Komponenten jedes Elements von W_∞ ungleich Null sind.

für die Definition in (2.7) war ein „Gefallen“ von unserer Seite. Immer alles genau zu erklären ist ein Bärendienst (und auch mühsam. . .). Deshalb und da die Leser eine gewisse Unabhängigkeit und Fähigkeit zur „Kompilierung“ eines mathematischen Texts entwickeln müssen, werden wir dies von jetzt an normalerweise vermeiden.

Nun können die Leser also überprüfen, dass K^S ein Vektorraum ist. (Der Beweis ist ähnlich zu dem Beweis zu K^n und folgt im Grossen und Ganzen aus den Eigenschaften des Körpers K.)

Beispiel 2.25. Sei $s_0 \in S$. Dann ist

$$\{f \in K^S \mid f(s_0) = 0\}$$

ein Untervektorraum von K^S. Seien $k \in \mathbb{N}$, $s_1, \ldots, s_k \in S$ und $\alpha_1, \ldots, \alpha_k \in K$. Dann ist auch

$$\left\{f \in K^S \mid \sum_{i=1}^{k} \alpha_i f(s_i) = 0\right\}$$

ein Untervektorraum. In ähnlicher Weise (aber doch auf eine etwas andere Art) betrachten wir

$$\mathrm{Ann}(S') = \{f \in K^S \mid f(s) = 0 \text{ für alle } s \in S'\}$$

für $S' \subseteq S$. Dann ist $\mathrm{Ann}(S')$ ein Untervektorraum. Darf $S' = \emptyset$ sein? Wenn ja, was ist $\mathrm{Ann}(\emptyset)$? Was ist $\mathrm{Ann}(S)$?

Um weiter zu spielen, könnten wir definieren:

$$W = \{f \in K^S \mid f(s) = 0 \text{ für alle ausser endliche viele } s \in S\}.$$

Ist W ein Untervektorraum von K^S? Alle diese Fragen sind in einem „Forum-Wieso“.

Beispiel 2.26. Wenn S nicht nur irgendeine Menge ist, sondern noch eine andere Struktur hat, dann können wir oft Untervektorräume von K^S mittels dieser Struktur definieren. Zum Beispiel sei $S = [0, 1]$ das Intervall $\{x \in \mathbb{R} \mid 0 \leq x \leq 1\}$. Dann ist

$$\{f \mid f \text{ stetig auf } [0, 1]\} \subseteq K^{[0,1]}$$

ein Untervektorraum[6] von $K^{[0,1]}$.

Bemerkung 2.27. Für $S = \{1, \ldots, n\}$ ist K^S sehr ähnlich zu K^n, für $S = \mathbb{N}$ ist K^S sehr ähnlich zu K^∞ und für $S = \mathbb{Z}$ ist K^S sehr ähnlich zu $K^\infty_{-\infty}$. Um „sehr ähnlich“ präzise auszudrücken, brauchen wir einige Begriffe, die wir später definieren. In der Zwischenzeit können sich die fleissigen Leser selbst überzeugen, dass es zwischen den erwähnten Mengen oben eine Bijektion gibt.

[6] Stetigkeit werden Sie in der Analysis in ein paar Wochen behandeln.

Beispiel 2.28 (Ein Beispiel aus der Analysis)**.** Sei $K = \mathbb{R}$. Die Lösungen einer sogenannten linearen Differentialgleichung bilden einen Untervektorraum von $\mathbb{R}^{\mathbb{R}}$. Zum Beispiel ist

$$\{f : \mathbb{R} \to \mathbb{R} \mid f' = f\}$$

ein Untervektorraum von $\mathbb{R}^{\mathbb{R}}$. Mit $f' = f$ ist insbesondere implizit gemeint, dass die Ableitung von f existiert.[7]

Polynomräume

Sei K ein Körper. Auf $K[x]$ haben wir die Addition schon in Bemerkung 1.48 definiert. Wir definieren jetzt die Skalarmultiplikation auf $K[x]$. Seien $a \in K$ und $f = a_0 + a_1 x + \ldots + a_n x^n \in K[x]$, dann definieren wir

$$a \cdot f = a \cdot (a_0 + a_1 x + \ldots + a_n x^n) := a a_0 + a a_1 x + \ldots + a a_n x^n.$$

Mit dieser Addition und Skalarmultiplikation ist $K[x]$ ein Untervektorraum (von K^K).

Beispiel 2.29. Sei $h \in \mathbb{N} \cup \{0, -\infty\}$. Dann ist

$$K[x]_h = \{p \in K[x] \mid \deg(p) \leq h\}$$

ein Untervektorraum sowohl von $K[x]$ als auch von K^K. Ist $\{p \in K[x] \mid \deg(p) = h\}$ auch ein Untervektorraum?

Beispiel 2.30. Die Menge $\left\{p \in K[x]_5 \mid p = a_0 + \ldots + a_5 x^5 \text{ mit } \sum_{i=0}^{5} a_i = 0\right\}$ ist ein Untervektorraum von $K[x]_5$.

Sehen Sie eine Ähnlichkeit zwischen Beispiel 2.17 und Beispiel 2.30?

Matrizenräume

In $M_{m\times n}(K)$ kann man eine Addition und Skalarmultiplikation definieren, um eine Vektorraumstruktur zu erhalten. Wie das genau geht und Beispiele von Untervektorräumen zu diesem Vektorraum besprechen Sie in der Übungsstunde!

Die Potenzmenge als Vektorraum

Bevor wir die Theorie weiterentwickeln, besprechen wir ein Beispiel, bei dem die Vektorraumstruktur nicht ganz so offensichtlich ist.

Beispiel 2.31. Sei X eine Menge. Wir möchten eine Vektorraumstruktur über $\mathbb{F}_2$ auf der Potenzmenge $P(X)$ definieren. Nehmen Sie sich einige Minuten Zeit, um sich zu

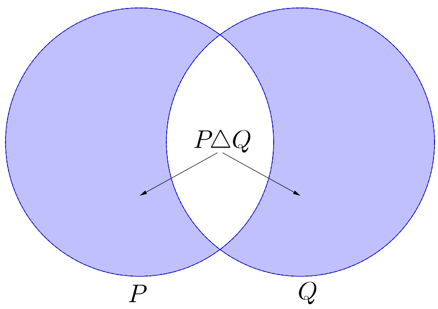
[7] Im Beispiel bezeichnet f' die Ableitung von f, mehr dazu werden Sie später in der Analysis sehen.

überlegen, was die Addition sein könnte. Dann überlegen Sie sich, was der Nullvektor ist. Wenn Sie dann wissen, was der Nullvektor ist, dann wissen Sie, wie die Skalarmultiplikation definiert sein muss: Laut den Axiomen eines Vektorraums und Lemma 2.2 gilt in jedem Vektorraum V

$$1 \cdot v = v \quad \text{und} \quad 0 \cdot v = 0$$

für alle $v \in V$. Aber $\mathbb{F}_2$ enthält nur 0 und 1! Zur Addition: Es gibt eine interessante symmetrische Operation auf $P(X)$, nämlich die symmetrische Differenz:
Für $A, B \in P(X)$ sei

$$A \triangle B = (A \cup B) \setminus (A \cap B).$$



Figur 2.1: Die symmetrische Differenz $P \triangle Q$.

Mann kann zeigen, dass für alle $A, B, C \in P(X)$ Folgendes gilt

$$A \triangle B = B \triangle A \quad \text{und} \quad (A \triangle B) \triangle C = A \triangle (B \triangle C).$$

Wenn wir $\triangle$ als Addition auf $P(X)$ nehmen, was wäre dann der Nullvektor? Wir suchen $S \subseteq X$ mit $A \triangle S = A$ für alle $A \in P(X)$. Eine kurze Überlegung führt uns dazu, dass

$$A \triangle \emptyset = A.$$

für alle $A \in P(X)$ gilt. Also, wenn $(P(X), \triangle, \cdot)$ einen Vektorraum über $\mathbb{F}_2 = \{\overline{0}, \overline{1}\}$ bilden soll, dann muss gelten

$$\overline{0} \cdot A = \emptyset \quad \text{und} \quad \overline{1} \cdot A = A.$$

Übung 2.32. Zeigen Sie, dass $(P(X), +, \cdot)$ ein Vektorraum über $\mathbb{F}_2$ ist.

Bemerkung 2.33. Wir empfehlen diese Übung mit allen Details zu lösen - siehe Serie 6!

2.2 Zurück zur Theorie

Nun da wir so viele Beispiele von Vektorräumen gesehen haben, ist es höchste Zeit zurück zur Theorie zu kehren. Jetzt müssen Sie ihren „Computer-Compiler“ immer benutzen: Wenn wir einen neuen Begriff definieren, wählen Sie eines der vorherigen Beispiele und überprüfen Sie mittels ihrem „Computer-Compiler“, ob der neue Begriff auf ihr gewähltes Beispiel zutrifft. Wir werden definieren:

- Der Span[8] einer Menge $S \subseteq V$, wobei V ein Vektorraum ist;
- Linearkombinationen;
- lineare Unabhängigkeit (und daher auch lineare Abhängigkeit);
- Basen und Dimensionen;
- Summen und (innere) direkte Summen.

2.2.1 Span und Linearkombinationen

In diesem Unterkapitel bezeichnet K immer einen Körper und V einen Vektorraum über K. Sei $S \subseteq V$ eine Teilmenge (die nicht unbedingt ein Untervektorraum ist). Wir können uns zwei Fragen stellen:

1. Welches ist der kleinste Untervektorraum von V, der S enthält?
2. Was bekommen wir, wenn wir die Menge aller Elemente betrachten, die wir mittels Addition und Skalarmultiplikation von Elementen aus S erreichen können?

Die erste Frage können wir mit diesem einfachen, aber wichtigen, Lemma beantworten:

Lemma 2.34. *Sei V ein Vektorraum über K, I eine Indexmenge und $(W_i)_{i\in I}$ eine Familie von Untervektorräumen. Dann ist*

$$W = \bigcap_{i\in I} W_i$$

ein Untervektorraum.

Beweis. Sei W wie im Lemma. Dann gilt:

- Der Nullvektor 0 liegt in W, da $0 \in W_i$ für alle $i \in I$.

[8]Dieser Begriff kommt ursprünglich aus dem Englischen. Im Deutschen kann man auch Erzeugnis sagen. Wir werden jedoch in der Regel den englischen Begriff verwenden.

- Seien $a_1, a_2 \in K$ und $v_1, v_2 \in W$. Dann sind $v_1, v_2 \in W_i$ für alle $i \in I$ per Definition von W. Da W_i ein Untervektorraum ist für alle $i \in I$, ist $a_1v_1 + a_2v_2 \in W_i$. Daher ist $a_1v_1 + a_2v_2 \in W$.

Dies zeigt, dass W in der Tat ein Untervektorraum von V ist. □

Definition 2.35 (Erste Definition des Spans)**.** Sei $S \subseteq V$. Wir definieren den *Span* (auch *Spann*, oder *lineare Hülle* genannt) $\mathrm{Sp}(S)$ von S durch

$$\mathrm{Sp}(S) := \bigcap_{W \in \mathcal{N}} W,$$

wobei $\mathcal{N} = \{W \mid S \subseteq W,\ W \text{ ein Untervektorraum von } V\}$.

Eine direkte Folgerung ist:

Lemma 2.36. *Sei $S \subseteq V$ eine Teilmenge. Dann ist $\mathrm{Sp}(S)$ der kleinste Untervektorraum, der S enthält. Das heisst, es gilt:*

(1) $\mathrm{Sp}(S)$ ist ein Untervektorraum.

(2) Wenn $S \subseteq W$ für $W \subseteq V$ ein Untervektorraum, dann ist $\mathrm{Sp}(S) \subseteq W$.

Beweis. (1) Dies folgt aus Lemma 2.34.

(2) Dies folgt aus der Definition von $\mathrm{Sp}(S)$ als Durchschnitt aller Untervektorräume, die S enthalten.

Dies beweist das Lemma. □

Das ist schön und gut, gibt uns aber keine „konkrete“ Beschreibung von $\mathrm{Sp}(S)$. Zu diesem Zweck definieren wir:

Definition 2.37. Sei $n \in \mathbb{N}$ und seien $a_1, \ldots, a_n \in K$ und $v_1, \ldots, v_n \in V$. Ein Vektor $v \in V$ der Form

$$v = a_1v_1 + \ldots + a_nv_n = \sum_{i=1}^{n} a_iv_i$$

heisst eine *Linearkombination von $v_1, \ldots, v_n$ (über K)*. Die Skalare $a_1, \ldots, a_n$ heissen die *Koeffizienten* der Linearkombination.

Bemerkung 2.38 (Wichtige Bemerkung)*.* Am Anfang von Definition 2.37 steht $n \in \mathbb{N}$. Das heisst bei einer Linearkombination kommen (nur) endlich viele Elemente in der Summe vor![9]

[9]Nachdem Sie mehr Analysis gelernt haben, könnte man auch Linearkombinationen von unendlich vielen Elementen betrachten, siehe zum Beispiel Fourierreihen.

Lemma 2.39. *Sei $S \subseteq V$ eine (nicht leere) Teilmenge. Dann gilt*

$$\begin{aligned}\mathrm{Sp}(S) &= \{a_1v_1 + \ldots + a_nv_n \mid n \in \mathbb{N}, a_i \in K, v_i \in S \textit{ für alle } 1 \leq i \leq n\} \\ &= \{\textit{alle Linearkombinationen von Vektoren aus } S\}.\end{aligned}$$

Bemerkung 2.40. Die Menge S kann unendlich sein. Nichtsdestotrotz benutzt jede Linearkombination von S nur endlich viele Elemente von S.

Beweis von Lemma 2.39. Sei

$$\widetilde{\mathrm{Sp}}(S) := \{a_1v_1 + \ldots + a_nv_n \mid n \in \mathbb{N}, a_i \in K, v_i \in S \text{ für alle } 1 \leq i \leq n\}.$$

Wir zeigen:

(a) $\widetilde{\mathrm{Sp}}(S)$ ist ein Untervektorraum.

(b) Für jeden Untervektorraum $W \subseteq V$ mit $S \subseteq W$ gilt $\widetilde{\mathrm{Sp}}(S) \subseteq W$.

Daraus folgt $\widetilde{\mathrm{Sp}}(S) = \mathrm{Sp}(S)$, was wir zeigen wollten. (Wieso?) Wir beweisen zuerst (a). Seien v und w zwei Linearkombinationen von Vektoren aus S, das heisst, $v, w \in \widetilde{\mathrm{Sp}}(S)$ und wir können dementsprechend schreiben

$$v = a_1v_1 + \ldots + a_nv_n \quad \text{und} \quad w = b_1w_1 + \ldots + b_mw_m,$$

wobei $n, m \in \mathbb{N}$, $a_i, b_i \in K$ und $v_i, w_j \in S$ für alle $1 \leq i \leq n$ und $1 \leq j \leq m$. Seien $\alpha, \beta \in K$. Dann ist

$$\alpha v + \beta w = \alpha a_1v_1 + \ldots + \alpha a_nv_n + \beta b_1w_1 + \ldots + \beta b_mw_m$$

auch eine Linearkombination von Vektoren aus S (mit $m + n \in \mathbb{N}$ Vektoren aus S). Daher ist $\alpha v + \beta w \in \widetilde{\mathrm{Sp}}(S)$ für alle $\alpha, \beta \in K$. Ausserdem, da $S \neq \emptyset$, sei $v \in S$.[10] Dann ist $0_K \cdot v$ eine Linearkombination von Vektoren aus S und daher ist $0_K \cdot v = 0_V \in \widetilde{\mathrm{Sp}}(S)$. Dies zeigt (a). Für Teil (b): Man zeigt mittels Induktion (unter Verwendung von UVR2 und UVR3 aus Definition 2.5)[11], dass für jeden Untervektorraum W Folgendes gilt: Falls $S \subseteq W$ ist, dann ist jede Linearkombination von Vektoren aus S in W enthalten. Das heisst, dass $\widetilde{\mathrm{Sp}}(S) \subseteq W$. □

Zwischenzeit-Zusammenfassung : Wir haben oben eigentlich gezeigt, dass $\mathrm{Sp}(S)$ zwei äquivalente Definitionen hat:

(1) Definition 2.35:

$$\mathrm{Sp}(S) := \bigcap_{W \in \mathcal{N}} W,$$

[10] Siehe Bemerkung 2.45.

[11] Zum Beispiel ist UVR3 die Basis der Induktion.

wobei $\mathcal{N} = \{W \mid S \subseteq W, W \text{ ein Untervektorraum von } V\}$.

(2) $\mathrm{Sp}(S) = \{\text{alle Linearkombinationen von Vektoren in } S\}$.

Sie können jetzt sowohl (1) als auch (2) als Definitionen des Spans $\mathrm{Sp}(S)$ betrachten.

Bevor wir Beispiele besprechen, geben wir noch drei kurze Definitionen und eine Bemerkung.

Definition 2.41 (Notation)**.** Seien $n \in \mathbb{N}$ und $v_1, \ldots, v_n \in V$. Wir schreiben

$$\mathrm{Sp}(v_1, \ldots, v_n) := \mathrm{Sp}(\{v_1, \ldots, v_n\}).$$

Übung 2.42. Zeigen Sie:

- $\mathrm{Sp}(v) = \{\alpha v \mid \alpha \in K\}$.
- $\mathrm{Sp}(v, w) = \{\alpha v + \beta w \mid \alpha, \beta \in K\}$.
- Allgmeiner: $\mathrm{Sp}(v_1, \ldots, v_n) = \{\alpha_1 v_1 \ldots + \alpha_n v_n \mid \alpha_i \in K \text{ für alle } 1 \leq i \leq n\}$.

(Ist dies nicht einfach die Definition? Hier ist es schwieriger zu wissen, was genau man zeigen soll als es zu zeigen.)

Definition 2.43. Wir sagen, dass V von $S \subseteq V$ *erzeugt* wird, falls $\mathrm{Sp}(S) = V$. In diesem Fall sagt man auch, dass S erzeugend (für V) ist oder, dass S den Vektorraum V aufspannt. Die Menge S nennt man dann ein *Erzeugendensystem* von V. Allgemeiner benutzt man dieselbe Terminologie wenn $\mathrm{Sp}(S) = W$ und $W \subseteq V$ ein Untervektorraum ist: Zum Beispiel sagt man, dass S den Untervektorraum W *erzeugt* (oder S spannt W auf etc.).[12]

Definition 2.44. Ein Vektorraum V heisst *endlich-dimensional*, falls es $S \subseteq V$ gibt, so dass $|S| < \infty$ und $\mathrm{Sp}(S) = V$.

Bemerkung 2.45. Wenn wir die Definition 2.35 von $\mathrm{Sp}(S)$ benutzen, dann ist $\mathrm{Sp}(\emptyset) = \{0_V\}$. (Wieso?) Wenn wir $\mathrm{Sp}(\emptyset)$ auch durch die Linearkombinations-Definition definieren wollen, könnten wir sagen, dass die „leere Summe" immer 0_V ist. Das heisst, dass $\sum_{x\in\emptyset} x := 0$. Oder wir deklarieren einfach $\mathrm{Sp}(\emptyset) = \{0_V\}$. In jedem Fall definieren/folgern wir jedoch, dass

$$\mathrm{Sp}(\emptyset) = \{0_V\}.$$

Dies macht einige Formulierungen später einfacher.

[12] Auf Englisch würde man sagen: S spans W.

2.2.2 Beispiele

Beispiel 2.46 (Geometrisches Intermezzo)**.** Wir betrachten $\mathbb{R}^2$. Welche Untervektorräume gibt es in $\mathbb{R}^2$ und wie können wir sie erzeugen? Zuerst einmal gibt es $\{0_{\mathbb{R}^2}\} = \{(0,0)\}$. Für diesen Untervektorraum gilt $\mathrm{Sp}(\emptyset) = \mathrm{Sp}(0_{\mathbb{R}^2}) = \{0_{\mathbb{R}^2}\}$.

Sei nun $0 \neq v = (a,b) \in W \subseteq \mathbb{R}^2$, wobei $W \subseteq \mathbb{R}^2$ ein Untervektorraum ist. Dann wissen wir, dass $\mathrm{Sp}(v) = \{\alpha v \mid \alpha \in \mathbb{R}\} \subseteq W$. Geometrisch heisst das: $\mathrm{Sp}(v)$ ist die Gerade G_v durch v und den Ursprung $(0,0)$. Wir haben also gezeigt, dass $v \in W$ die Inklusion $G_v \subseteq W$ impliziert. Ausserdem sehen wir, dass

$$\mathrm{Sp}(v) = \mathrm{Sp}(w) \iff G_v = G_w \iff \exists\, \alpha \neq 0 : w = \alpha v$$

(Wieso?). Da $\mathrm{Sp}(v)$ immer ein Untervektorraum ist, zeigt dies auch, dass alle Geraden $\{G_v \mid 0 \neq v \in \mathbb{R}^2\}$ Untervektorräume von $\mathbb{R}^2$ sind.

Um die letzte Möglichkeit für einen Untervektorraum von $\mathbb{R}^2$ zu finden, könnte man jetzt Folgendes zeigen:

Übung 2.47. Sei W ein Untervektorraum von $\mathbb{R}^2$, $G_v \subseteq W$ für $v \neq 0$ und $w \in W$ mit $w \notin G_v$. Dann gilt $\mathrm{Sp}(v,w) = \mathbb{R}^2$.

Es gilt zum Beispiel, dass $\mathrm{Sp}((1,0),(0,1)) = \mathbb{R}^2$. Also ist $\mathbb{R}^2$ insbesondere endlichdimensional.
Zusammenfassung : $\mathbb{R}^2$ hat folgende Untervektorräume:

(1) $\{(0,0)\}$;

(2) G_v für $0 \neq v \in \mathbb{R}^2$;

(3) $\mathbb{R}^2$.

Oder geometrisch gesagt: (1) Der Ursprung der Ebene; (2) Alle Geraden in der Ebene; (3) Die Ebene selbst.

Bemerkung 2.48. Dies gilt über jedem Körper K, aber wir „verlieren" die geometrische Intuition. Ausserdem finden wir später eine ähnliche Beschreibung von allen Untervektorräumen von $\mathbb{R}^n$ (oder K^n).[13]

Beispiel 2.49 (Gauss und lineare Vektoren)**.** Zuerst ein Beispiel in $\mathbb{Q}^2$:
Seien $v = (1,3), w = (7,73) \in \mathbb{Q}^2$. Dann ist $(11,137)$ eine Linearkombination von v und w:

$$-3 \cdot (1,3) + 2 \cdot (7,73) = (11,137).$$

Aber wie könnten wir die Skalare -3 und 2 finden? Oder wie können wir bestimmen, ob es solche Skalare überhaupt gibt? Mit Gauss!

[13] Vgl. Bemerkung 2.16.

Lemma 2.50. *Sei* $A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in M_{m\times n}(K)$ *und* $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$. *Wir definieren die Spaltenvektoren* $v_1, \ldots, v_n$ *in* K^m *durch die Spalten von* A*:*

$$A = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix}.$$

Dann gilt

$$Ax = x_1 \cdot \begin{pmatrix} | \\ v_1 \\ | \end{pmatrix} + x_2 \cdot \begin{pmatrix} | \\ v_2 \\ | \end{pmatrix} + \cdots + x_n \cdot \begin{pmatrix} | \\ v_n \\ | \end{pmatrix} = \sum_{i=1}^{n} x_i v_i.$$

Beweis. Mit Definition 2.3 ist dies nur eine kurze Rechnung. □

Seien $v_1, \ldots, v_n \in K^m$ und $w \in K^m$. Wir wollen wissen, ob w eine Linearkombination von $v_1, \ldots, v_n$ ist. Anders gesagt, wollen wir entscheiden, ob $w \in \mathrm{Sp}(v_1, \ldots, v_n)$ ist. Dazu betrachten wir die Matrix

$$A = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix},$$

deren Spalten die Vektoren $v_1, \ldots, v_n$ sind. Jetzt betrachten wir das lineare Gleichungssystem[14]

$$Ax = w \quad \text{für } x \in K^n. \tag{$\star$}$$

Aus Lemma 2.50 folgt:

$$\begin{aligned} (\star) \text{ hat eine Lösung } x = (x_1, \ldots, x_n) &\iff Ax = x_1 v_1 + \ldots + x_n v_n = w \\ &\iff w \in \mathrm{Sp}(v_1, \ldots, v_n). \end{aligned}$$

Mehr dazu in der Übungsstunde.

Bemerkung 2.51. Eine „Baby-Version“ der linearen Algebra bespricht die Vektorräume $\mathbb{R}^n$ und vor allem $\mathbb{R}^2$ und $\mathbb{R}^3$. Man sollte diese einfachen Fälle nicht unterschätzen und versuchen, alle Begriffe am Anfang in diesen Vektorräumen zu verstehen. So entwickelt man ein geometrisches Verständnis, welches auch in Vektorräumen „ohne Geometrie“ nützlich ist. Spielen Sie mit Beispielen, um dieses geometrische Verständnis zu entwickeln. Sie können auch die schöne Video-Serie von 3Blue1Brown anschauen. Im jetzigen Zusammenhang ist besonders dieses Video relevant.

[14] Anders gesagt, machen wir Gauss mit $(A \mid w)$.

Beispiel 2.52. Seien
$$\varphi = \frac{1+\sqrt{5}}{2} \quad \text{und} \quad \psi = \frac{1-\sqrt{5}}{2}.$$
Dann sind
$$v = \big(1, \varphi, \varphi^2, \ldots\big) \quad \text{und} \quad w = \big(1, \psi, \psi^2, \ldots\big)$$
Elemente von **Fib** aus dem Beispiel 2.20. (Wieso?) In der Einfürung haben wir gezeigt, dass $\mathrm{Sp}(v, w) = \mathbf{Fib}$. Daher ist **Fib** endlich-dimensional. Das heisst, dass man jede Folge in **Fib** als Linearkombination von v und w schreiben kann. Im Sinne der Definition aus der Einführung heisst das, dass wir jedes Element von **Fib** gut kennen, da wir v und w gut kennen!

Beispiel 2.53 (Standard-Basen)**.** Es gibt gewisse „Standard-Wege", um die Vektorräume, die wir bisher besprochen haben, zu erzeugen. Hier hat das Wort Standard keine präzise Bedeutung: man meint damit einfach, dass viele Leute diesen Weg benutzen, um diese Vektorräume zu erzeugen. Wir überlassen es dem Leser zu überprüfen, dass die angegebenen Mengen tatsächlich Erzeugendensysteme sind:

(1) Seien $e_1 = (1, 0, \ldots, 0), e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, \ldots, 0, 1)$ Vektoren in K^n. Es gilt $K^n = \mathrm{Sp}(e_1, \ldots, e_n)$.

(2) Für $K[x]_n$ gilt $K[x]_n = \mathrm{Sp}(x^0, x^1, \ldots, x^n)$. Ausserdem gilt
$$K[x] = \mathrm{Sp}(x^0, x^1, \ldots) := \mathrm{Sp}\left(\left\{x^i \mid i \in \mathbb{N} \cup \{0\}\right\}\right).$$
Die letzte Tatsache lässt sich ebenfalls einfach überprüfen. Man muss sich lediglich daran erinnern, dass eine Linearkombination immer nur endlich viele Summanden hat!

(3) Seien $m, n \in \mathbb{N}$. Für $1 \le i \le m, 1 \le j \le n$ sei $E_{ij} = (a_{kl}) \in M_{m\times n}(\mathbb{R})$ die Matrix mit
$$a_{kl} = \begin{cases} 1, & \text{falls } k = i \text{ und } l = j \\ 0, & \text{sonst} \end{cases}.$$
Dann gilt
$$M_{m\times n}(\mathbb{R}) = \mathrm{Sp}(\{E_{ij} \mid 1 \le i \le m, 1 \le j \le n\})$$
Also ist zum Beispiel
$$M_{2\times 2} = \mathrm{Sp}\left(\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\right).$$

(4) Sei X eine Menge mit $|X| = n$. Dann gilt

$$P(X) = \mathrm{Sp}(\{\text{alle einelementigen Teilmengen von } X\}).$$

Bemerkung 2.54. Einige Autoren nennen die Basen in Beispiel 2.53 „kanonisch“ (vgl. Fischer [6, 1.5.1]). Dies ist ein Missbrauch des Wortes kanonisch. Nichts ist kanonisch an diesen Basen. Wir werden mehr dazu sagen, wenn wir ein Beispiel angeben, wo etwas wirklich kanonisch ist. Bitte vermeiden Sie es dieses Wort in Zusammenhang mit den Basen aus Beispiel 2.53 zu verwenden.

Übung 2.55. Nehmen Sie sich Zeit die Aussagen in Beispiel 2.53 zu überprüfen.

Übung 2.56. (1) Sei $i \in \mathbb{N}$. Wir definieren $e_i \in K^\infty$ als die Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_i = 1$ und $a_j = 0$ für alle $j \neq i$. Gilt $\mathrm{Sp}(\{e_i \mid i \in \mathbb{N}\}) = K^\infty$? Wenn nicht, haben wir diesen Untervektorraum schon zuvor erwähnt?

(2) Sei X eine unendliche Menge. Gilt

$$P(X) = \mathrm{Sp}(\{\text{alle einelementigen Teilmengen von } X\})?$$

Wenn nicht, wie könnten Sie diesen Span in Worten beschreiben?

2.2.3 Lineare Unabhängigkeit und die Definition einer Basis

Einführung

Ein alternativer Titel dieses Abschnitts könnte sein: „Wie kann ich einen Vektorraum beherrschen/beschreiben/kontrollieren?“ oder weniger romantisch: „Wie könnte ich einen Vektorraum mit Koordinaten versehen?“ Dieselben Fragen könnten wir bezüglich Untervektorräumen stellen. Dies ist aber unnötig, da jeder Untervektorraum auch ein Vektorraum ist.

Also, was meinen wir mit „einen Vektorraum mit Koordinaten versehen“? Nehmen wir an, dass V ein Vektorraum über K ist und $V = \mathrm{Sp}(v_1, \ldots, v_n)$. Dann können wir jeden Vektor als Linearkombination von $v_1, \ldots, v_n$ schreiben: Sei $v \in V$. Dann existieren $a_1, \ldots, a_n \in K$ mit

$$v = a_1v_1 + \ldots + a_nv_n. \tag{2.8}$$

Wenn diese Beschreibung eindeutig ist, dann könnte man bei den Skalaren

$$a_1, \ldots, a_n$$

an die „Koordinaten" von v in der Beschreibung von v als Linearkombination von $v_1, \dots, v_n$ denken. Wenn jeder Vektor eine eindeutige Beschreibung als Linearkombination von $v_1, \dots, v_n$ hat, könnten wir durch $v_1, \dots, v_n$ jeden Vektor in V mit „Koordinaten" versehen. Dadurch könnten wir den Vektorraum V beschreiben/kontrollieren. Wenn dies der Fall ist, dann sagen wir, dass $v_1, \dots, v_n$ eine *Basis* von V ist. Wir möchten also verstehen, welche Bedingung wir an $v_1, \dots, v_n$ stellen müssen, damit wir wissen, dass jeder Vektor eine eindeutige Beschreibung als Linearkombination von $v_1, \dots, v_n$ hat. Wir werden bald sehen, dass wir dies (überraschenderweise) schon durch eine mögliche Beschreibung des Nullvektors verstehen können.

Lineare Unabhängigkeit

Wie immer bezeichnet V einen Vektorraum über einem Körper K.

Definition 2.57. Sei $n \in \mathbb{N}$. Eine endliche Menge[15] $\{v_1, \dots, v_n\} \subseteq V$ ist *linear unabhängig*, falls aus $a_1v_1 + \ldots + a_nv_n = 0$ für $a_1, \dots, a_n \in K$ stets folgt, dass $a_1 = 0, \dots, a_n = 0$. In Quantoren ausgedrückt:

$$\forall\, a_1, \dots, a_n \in K : a_1v_1 + \ldots + a_nv_n = 0 \;\Rightarrow\; a_1 = 0, \dots, a_n = 0.$$

Später werden wir in diesem Fall auch schreiben: die Liste $v_1, \cdots, v_n$ ist linear unabhängig (Siehe Bemerkung 2.71). Ausserdem ist die leere Menge $\emptyset$ linear unabhängig per Konvention.

Bemerkung 2.58. Der Nullvektor hat immer die folgende Beschreibung als Linearkombination von $v_1, \dots, v_n$:

$$0 \cdot v_1 + \ldots + 0 \cdot v_n = 0. \tag{2.9}$$

Daher sagt Definition 2.57: $v_1, \dots, v_n$ sind linear unabhängig genau dann, wenn (2.9) die einzige Beschreibung des Nullvektors als Linearkombination von $v_1, \dots, v_n$ ist.

Definition 2.59. Eine nichtleere Menge $S \subseteq V$ heisst *linear unabhängig*, falls jede endliche nichtleere Teilmenge von S linear unabhängig ist.

Beispiel 2.60. (1) Die Menge $\{e_1, \dots, e_n\} \subseteq K^n$ ist linear unabhängig. (Wieso?)

(2) Die Menge $\{e_1, e_2, \dots\} \subseteq K^\infty$ ist linear unabhängig. (Wieso?)

(3) Die Menge $\{1, x, x^2, \dots, x^n\}$ ist linear unabhängig in $K[x]_n$. Ausserdem ist die Menge $\{1, x, x^2, \dots\}$ linear unabhängig in $K[x]$. (Wieso?)

Übung 2.61. (1) Beweisen Sie: Die Menge $\{v\} \subseteq V$ ist linear unabhängig genau dann, wenn $v \neq 0$.

[15] Da $0 \notin \mathbb{N}$ ist, ist diese Menge nicht die leere Menge.

(2) Wann sind zwei Vektoren in $\mathbb{R}^2$ linear unabhängig?

(3) Versuchen Sie drei linear unabhängige Vektoren in $\mathbb{R}^3$ zu finden!

Definition 2.62. Eine Menge heisst *linear abhängig*, falls sie nicht linear unabhängig ist.

Definition 2.63. Die *triviale Linearkombination* von $v_1, \ldots, v_n$ ist

$$0 = 0 \cdot v_1 + \ldots + 0 \cdot v_n.$$

Daher gilt: $v_1, \ldots, v_n$ ist linear unabhängig genau dann, wenn die triviale Linearkombination die einzige Möglichkeit ist 0 als Linearkombination von $v_1, \ldots, v_n$ zu schreiben.

Lineare (Un-)Abhängigkeit ist so wichtig, dass es sich lohnt die Definition von linearer Abhängigkeit explizit hinzuschreiben.

Definition 2.64. Eine nichtleere Menge $S \subseteq V$ heisst *linear abhängig*, wenn es paarweise verschiedene $v_1, \ldots, v_n \in S$ gibt und $a_1, \ldots, a_n \in K$, die nicht alle Null sind, so dass

$$a_1 v_1 + \ldots + a_n v_n = 0.$$

Bemerkung 2.65. Nehmen Sie sich Zeit, um zu überprüfen, dass Definition 2.64 die Negation der Definition 2.59 ist. Falls Sie dies verwirrend finden, führen Sie sich folgende Idee vor Augen: $v_1, \ldots, v_n$ sind linear abhängig, wenn der Nullvektor nicht nur die triviale Beschreibung $0 \cdot v_1 + \ldots + 0 \cdot v_n$ hat, sondern auch eine andere „nicht-triviale“ Beschreibung.

Jetzt sind wir bereit für den Begriff der Basis, welcher die gewünschte Eindeutigkeit aus der Einführung herbeiführt.

Lemma 2.66. *Sei S linear unabhängig und $\emptyset \neq S' \subseteq S$. Dann ist auch S' linear unabhängig.*

Beweis. (Wieso?) □

Definition 2.67. Eine Menge $S \subseteq V$ heisst eine *Basis* (für V), falls S linear unabhängig ist und V von S erzeugt wird.

Proposition 2.68. *Eine Menge $S \subseteq V$ ist eine Basis von V genau dann, wenn jedes $v \in V$ in einer eindeutigen Weise als Linearkombination von Vektoren aus S geschrieben werden kann.*

Beweis. Der Einfachheit halber nehmen wir an, dass $|S| < \infty$ ist. Der Beweis für $|S| = \infty$ ist ziemlich ähnlich.

„$\Leftarrow$“: Nehmen wir an, dass S die Eigenschaft hat, dass jedes $v \in V$ in einer eindeutigen Weise als Linearkombination von Vektoren aus S geschrieben werden kann. Dies impliziert direkt, dass $\mathrm{Sp}(S) = V$. Sei $S = \{v_1, \ldots, v_n\}$. Da 0 als triviale Linearkombination $0 = 0 \cdot v_1 + \ldots + 0 \cdot v_n$ geschrieben werden kann, folgt aus der Eindeutigkeit, dass für $a_1, \ldots, a_n \in K$ mit

$$a_1 v_1 + \ldots + a_n v_n = 0$$

$a_1 = \ldots = a_n = 0$ folgen muss. Dies zeigt, dass S linear unabhängig ist und daher eine Basis ist.

„$\Rightarrow$“: Diese Richtung beweisen wir mittels Kontraposition. Nehmen wir an, dass $v \in V$ auf zwei verschiedene Weisen als Linearkombination von Vektoren aus S geschrieben werden kann: Angenommen

$$v = a_1 v_1 + \ldots a_n v_n \quad \text{und} \quad v = b_1 v_1 + \ldots + b_n v_n$$

mit $a_1, \ldots, a_n, b_1, \ldots, b_n \in K$, so dass ein $1 \leq i_0 \leq n$ existiert mit $a_{i_0} \neq b_{i_0}$. Dann ist

$$\begin{aligned} 0 = v - v &= (a_1 v_1 + \ldots + a_n v_n) - (b_1 v_1 + \ldots + b_n v_n) \\ &= (a_1 - b_1) v_1 + \ldots + (a_n - b_n) v_n, \end{aligned}$$

aber $a_{i_0} - b_{i_0} \neq 0$ und daher ist $S = \{v_1, \ldots, v_n\}$ linear abhängig. Insbesondere ist S keine Basis. □

Bemerkung 2.69. Proposition 2.68 zeigt: Eine Basis eines Vektorraums gibt uns, wie in der Einführung gewünscht, „eindeutige Koordinaten“ um die Vektoren aus dem Vektorraum zu beschreiben.

Beispiel 2.70. Alle Mengen in Beispiel 2.53 (Standard-Basen) sind Basen der jeweiligen Vektorräume.

Bevor wir einige Beispiele geben, wollen wir besser verstehen, wie Basen in endlich-dimensionalen Vektorräumen aussehen. Das ruft nach einem neuen Abschnitt!

2.3 Endlich-dimensionale Vektorräume

Dieses Kapitel ist das schönste in der Geschichte der Linearen Algebra I, das wir erzählen wollen. Wir werden sehen, dass wir mittels der Begriffe linearer (Un-)Abhängigkeit und Span ein starkes Struktur-Theorem für endlich-dimensionale Vektorräume über einem Körper K beweisen können. Teile dieser Präsentation sind aus dem Buch von Halmos [7] und [5, Kap. 2]. Im Folgenden ist V ein Vektorraum über K, der endlich-dimensional ist (siehe Definition 2.44).

Bemerkung 2.71. In diesem Abschnitt ist es besser an „Listen von Vektoren“ zu denken. Das heisst, wir stellen uns die Menge $\{v_1, \ldots, v_m\} \subseteq V$ als geordnete Menge von V vor, wobei die Ordnung durch die Nummerierung gegeben ist. Daher schreiben wir häufig $v_1, \ldots, v_m$ statt $\{v_1, \ldots, v_m\}$ und nennen $v_1, \ldots, v_m$ eine *Liste*. Mit der *Länge* einer Liste meinen wir einfach m. (Der Fall $m = 0$ ist erlaubt, wenn die Liste leer ist!)

Lemma 2.72. *Seien $v_1, \ldots, v_m$ linear abhängig in V. Dann existiert ein $1 \leq j \leq m$ mit*

(a) $v_j \in \mathrm{Sp}(v_1, \ldots, v_{j-1})$ *und*

(b) $\mathrm{Sp}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m) = \mathrm{Sp}(v_1, \ldots, v_m)$.

Beweis. Da $v_1, \ldots, v_m$ linear abhängig sind, existieren $a_1, \ldots, a_m \in K$, so dass

$$a_1v_1 + \ldots + a_mv_m = 0$$

und nicht alle $a_i = 0$ sind für $1 \leq i \leq m$. Sei $j = \max\{i \mid a_i \neq 0\}$. Dann gilt

$$a_1v_1 + \ldots + a_jv_j = 0.$$

Wir lösen nach a_jv_j auf und dividieren durch $a_j(\neq 0)$:

$$v_j = \frac{-a_1}{a_j}v_1 + \ldots + \frac{-a_{j-1}}{a_j}v_{j-1}. \tag{2.10}$$

Dies zeigt (a). Für (b) betrachten wir $v \in \mathrm{Sp}(v_1, \ldots, v_m)$. Dann ist

$$v = a_1v_1 + \ldots + a_mv_m$$

für $a_1, \ldots, a_m \in K$. Man setzt nun (2.10) für v_j ein und sieht, dass v auch eine Linearkombination von $\{v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m\}$ ist. Dies zeigt die Behauptung. (Wieso?) □

Bemerkung 2.73. Was passiert im Beweis, falls $j = 1$? Dann ist der Span von $v_1, \ldots, v_{j-1}$ gleich $\mathrm{Sp}(v_1, \ldots, v_{j-1}) = \mathrm{Sp}(\emptyset) = \{0\}$. Das impliziert, dass $v_j = v_1 = 0$. Dann gelten in der Tat (a) und (b).

Lemma 2.74. *Angenommen wir haben in Lemma 2.72 zusätzlich die Bedingung, dass $v_1, \ldots, v_k$ für $k < m$ linear unabhängig sind. Dann gilt $k < j$.*

Beweis. Angenommen $j \leq k$, dann gilt $v_j \in \mathrm{Sp}(v_1, \ldots, v_{j-1})$ beziehungsweise

$$v_j = a_1v_1 + \ldots + a_{j-1}v_{j-1}$$

für $a_1, \ldots, a_{j-1} \in K$. Dann ist jedoch

$$0 = a_1v_1 + \ldots + a_{j-1}v_{j-1} + (-1)v_j$$

eine nicht-triviale Linearkombination von $v_1, \ldots, v_j$, was im Widerspruch zur linearen Unabhängigkeit von $v_1, \ldots, v_k$ steht. □

Lemma 2.75. *Seien* $w_1, \ldots, w_n \in V$ *mit* $\mathrm{Sp}(w_1, \ldots, w_n) = V$ *und* $v \in V$. *Dann sind* $v, w_1, \ldots, w_n$ *linear abhängig.*

Beweis. Der Beweis funktioniert gleich wie im vorherigen Lemma: Da es $a_1, \ldots, a_n \in K$ gibt mit $v = a_1 w_1 + \ldots + a_n w_n$, gilt $0 = a_1 w_1 + \ldots + a_n w_n - v$. Dies zeigt die lineare Abhängigkeit von $\{v, w_1, \ldots, w_n\}$. □

Lemma 2.76. *Seien* $v_1, \ldots, v_n \in V$ *mit* $\mathrm{Sp}(v_1, \ldots, v_n) = V$ *und seien* $u_1, \ldots, u_m \in V$ *linear unabhängig. Dann gilt* $m \leq n$.

Beweis. Wir werden m Etappen durchführen. In der ersten Etappe werden wir u_1 mit einem der Vektoren $\{v_1, \ldots, v_n\}$ „austauschen“, und zwar wie folgt:

Etappe 1: Wir betrachten die Menge

$$u_1, v_1, \ldots, v_n \tag{2.11}$$

von $n+1$ Vektoren. Laut Lemma 2.75 ist diese Menge linear abhängig. Laut Lemma 2.72 können wir einen der Vektoren in (2.11) „wegwerfen“ ohne den Span zu ändern. Laut Lemma 2.74 ist dieser „unbrauchbare“ Vektor nicht u_1, da $\{u_1\}$ linear unabhängig ist. (Streng genommen, müssen wir hier Lemma 2.66 benutzen, um zu argumentieren, dass $\{u_1\}$ linear unabhängig ist.) Wir bekommen eine neue Liste der Länge n, in welcher u_1 der erste Vektor ist und die anderen $n-1$ Vektoren Elemente von $\{v_1, \ldots, v_n\}$ sind. Des Weiteren spannt diese Liste V auf.

Etappe $1 < j \leq m$: (Wir empfehlen dem Leser an $j = 2$ zu denken.) Aus der letzten Etappe haben wir eine Liste der Länge n, die V aufspannt, deren erste $j-1$ Vektoren $u_1, \ldots, u_{j-1}$ sind und deren letzte $n-(j-1)$ Vektoren in $\{v_1, \ldots, v_n\}$ enthalten sind. Der Lesbarkeit halber führen wir eine Umnummerierung durch und schreiben diese Liste als

$$u_1, \ldots, u_{j-1}, w_1, \ldots, w_{n-(j-1)}.$$

Wir betrachten jetzt die Liste

$$u_1, \ldots, u_{j-1}, u_j, w_1, \ldots, w_{n-(j-1)}.$$

Auch hier gilt wegen Lemma 2.75, dass diese Liste linear abhängig ist. Laut Lemma 2.72 können wir einen Vektor dieser Liste wegwerfen ohne den Span zu ändern und wegen Lemma 2.74 muss dieser Vektor einer der Vektoren $w_1, \ldots, w_{n-(j-1)}$ sein, da $u_1, \ldots, u_j$ linear unabhängig sind. Wir führen nacheinander die m Etappen durch.

Am Ende von Etappe m bekommen wir eine Liste

$$u_1, \ldots, u_m, w'_1, \ldots, w'_{n-m}$$

der Länge n, die alle Vektoren $u_1, \dots, u_m$ enthält. Daher ist $n \geq m$, was wir zeigen wollten. □

Bemerkung 2.77. Wir wir gleich sehen werden, spielt Lemma 2.76 eine zentrale Rolle. Es ist eng verbunden mit dem sogenannten „Steinitz'schen Austauschsatz“ (vgl. Fischer [6, 1.5.4]).

Wir sind jetzt bereit für eine der Säulen der linearen Algebra:

Theorem 2.78. *Sei V ein endlich-dimensionaler Vektorraum über einem Körper K. Dann hat V eine Basis mit endlicher Länge. Ausserdem hat jede Basis von V die gleiche Länge.*

Beweis der zweiten Aussage in Theorem 2.78. Wir beweisen zunächst die zweite Aussage. Seien $\mathcal{A}$ und $\mathcal{B}$ zwei Basen von V. Dann sind $\mathcal{A}$ und $\mathcal{B}$ endliche Mengen, weil V endlich-dimensional ist. Da $\mathcal{A}$ erzeugend und $\mathcal{B}$ linear unabhängig ist, folgt aus Lemma 2.76, dass $|\mathcal{B}| \leq |\mathcal{A}|$. Wir können jedoch die Rollen von $\mathcal{A}$ und $\mathcal{B}$ vertauschen! Da $\mathcal{B}$ erzeugend und $\mathcal{A}$ linear unabhängig ist, gilt auch $|\mathcal{A}| \leq |\mathcal{B}|$, also $|\mathcal{A}| = |\mathcal{B}|$. □

Für den Beweis der Existenz einer Basis beweisen wir zunächst ein separates Lemma, da dessen Inhalt und der Beweis davon später wichtig sein werden für uns.

Lemma 2.79. *Sei $v_1, \dots, v_n$ ein Erzeugendensystem. Dann enthält $v_1, \dots, v_n$ eine Basis.*

Beweis. Kurz gesagt (und ohne einen expliziten Algorithmus anzugeben) könnte man das Lemma beweisen, indem man Vektoren aus $v_1, \dots, v_n$ nacheinander weg wirft bis man eine linear unabhängige Liste bekommt mit demselben Span. Und daher wäre diese neue Liste dann eine Basis. Wir möchten den Beweis allerdings etwas algorithmischer besprechen. Wir geben einen Algorithmus mit n Etappen an, der in der Etappe j entscheidet, ob v_j ein Teil der Basis sein wird.

<u>Etappe 1</u>: Wenn $v_1 = 0$ ist, werfen wir v_1 weg. Wenn nicht, dann behalten wir v_1.

<u>Etappe $2 \leq j \leq n$</u>: Falls $v_j \in \mathrm{Sp}(v_1, \dots, v_{j-1})$, so werfen wir v_j weg. Falls nicht, behalten wir v_j.

Nach diesen n Etappen bekommen wir eine Liste. Mit einer Umnummerierung schreiben wir diese Liste als $v_1, \dots, v_m$. Bemerken Sie, dass $\mathrm{Sp}(v_1, \dots, v_m) = V$, da wir in jeder Etappe j nur Vektoren, die im Span von Vektoren in $v_1, \dots, v_{j-1}$ enthalten sind, weggeworfen haben. Ausserdem behaupten wir, dass $v_1, \dots, v_m$ linear unabhängig sind, da $v_1, \dots, v_m$ die Bedingungen des folgenden Lemmas 2.80 erfüllen. □

Lemma 2.80. *Seien $w_1, \dots, w_l \in V$, so dass $w_j \notin \mathrm{Sp}(w_1, \dots, w_{j-1})$ für alle $1 \leq j \leq l$ ist. Dann sind $w_1, \dots, w_l$ linear unabhängig.*

Beweis. Falls die Aussage des Lemmas nicht stimmt, folgt aus Lemma 2.72 (a), dass es ein $1 \leq j \leq l$ gibt mit $w_j \in \mathrm{Sp}(w_1, \dots, w_{j-1})$, im Widerspruch zur Annahme. □

Beweis der ersten Aussage in Theorem 2.78. Nach Definition 2.44 eines endlich-dimensionalen Vektorraums existiert eine Erzeugendensystem für V. Nach Lemma 2.79 enthält dieses Erzeugendensystem eine Basis. Die erste Aussage in Theorem 2.78 folgt somit. Dies beendet den Beweis von Theorem 2.78. □

Theorem 2.78 ermöglicht folgende zentrale Definition:

Definition 2.81. Sei V ein endlich-dimensionaler Vektorraum. Wir definieren die *Dimension* von V als $\dim V = n$, wobei $n \in \mathbb{N} \cup \{0\}$ die Länge einer Basis von V ist.

Beachten Sie, dass dank Theorem 2.78 die Dimension in der Tat ein Element von $\mathbb{N} \cup \{0\}$ und *wohl-definiert* (das heisst, sie hängt nicht von der Wahl der Basis ab) ist. Im nächsten Kapitel werden wir sogenannte Isomorphismen zwischen Vektorräumen definieren. Hier ist ein Spoiler: Bis auf Isomorphismus gibt es nur einen Vektorraum über K der Dimension n. Mit der Hilfe des nächsten Beispiels, zeigt dies, dass jeder Vektorraum über K mit $\dim V = n$ isomorph zu K^n ist.

Beispiel 2.82. Wie wir in Beispiel 2.70 gesehen haben, sind die Mengen aus Beispiel 2.53 Basen der jeweiligen Vektorräume. Dies zeigt:

(1) $\dim K^n = n$;

(2) $\dim K[x]_n = n + 1$;

(3) $\dim M_{m\times n}(K) = mn$;

(4) für $|X| = n$ ist $\dim P(X) = n$ (vgl. Beispiel 2.31).

Zu Theorem 2.78 gibt es ein analoges Theorem auch im Fall, wenn V unendlich-dimensional ist.

Theorem 2.83 (Hamel Basis)**.** *Jeder Vektorraum über K hat eine Basis. Je zwei Basen haben die gleiche Kardinalität.*

Wir werden dieses Theorem nicht beweisen, nicht weil es schwierig ist, sondern weil:

(1) Wir werden schon genug Spass haben mit endlich-dimensionalen Vektorräumen.

(2) Hamel-Basen in unendlich-dimensionalen Vektorräumen sind normalerweise nicht interessant: unendlich-dimensionale Vektorräume haben normalerweise zusätzliche Strukturen (zum Beispiel topologische und analytische Strukturen) und man ist normalerweise interessiert an Basen, die diese Strukturen „auffassen“. Das beste Beispiel dazu ist die Theorie der Fourierreihen.

(3) Die Existenz einer Hamel-Basis ist äquivalent zum Auswalaxiom (bzw. zum Lemma von Zorn). (Trotzdem lohnt es sich vielleicht die Existenz einer Basis mit dem Lemma von Zorn zu beweisen, was nicht schwierig ist.)

Bevor wir diesen Abschnitt abschliessen, beweisen wir noch ein nützliches Lemma für später. In gewissem Sinne (vgl. Abschnitt 2.3.1) ist dieses Lemma dual zu Lemma 2.79.

Lemma 2.84. *Jede linear unabhängige Teilmenge von V kann man zu einer Basis von V erweitern.*

Beweis. Sei $u_1, \ldots, u_l$ eine linear unabhängige Liste in V und $v_1, \ldots, v_m$ eine Basis von V. Betrachten Sie die Liste

$$u_1, \ldots, u_l, v_1, \ldots, v_m,$$

welche sicher V erzeugt. Man macht die gleichen Etappen wie im Beweis von Lemma 2.79. Die Vektoren $u_1, \ldots, u_l$ werden dabei nicht weggeworfen (Wieso?). Daher bekommt man am Ende nach $l + m$ Etappen eine Liste, die mit $u_1, \ldots, u_l$ anfängt, linear unabhängig ist und V aufspannt. Das heisst, die erhaltene Liste ist eine Basis von V, die $u_1, \ldots, u_l$ enthält. Dies zeigt die Aussage des Lemmas. □

2.3.1 Gleichgewicht

Vielleicht hat es Ihre Mutter Ihnen schon gesagt: Gleichgewicht ist das wichtigste im Leben![16]

In der Mathematik hat das Gleichgewicht eine andere Rolle: Wie wir jetzt erklären werden, führt es uns zu nützlichen und interessanten Objekten.

Maximale linear unabhängige Mengen und minimale Erzeugendensysteme

Je grösser eine Teilmenge $S \subseteq V$ ist, desto grösser ist die Wahrscheinlichkeit, dass sie erzeugend ist. Je kleiner eine Teilmenge $S \subseteq V$ ist, desto grösser ist die Wahrscheinlichkeit, dass sie linear unabhängig ist. Anders gesagt, gibt es für eine Menge S einen Wettbewerb zwischen den Eigenschaften „erzeugend“ und „linear unabhängig“: „erzeugend“ zieht nach oben und „linear unabhängig“ zieht nach unten. Der folgende Satz zeigt, dass die „Gleichgewichts-Punkte“ Mengen der Grösse $\dim V$ sind.

Satz 2.85. *Sei V ein Vektorraum mit $\dim V = n$, wobei $n \in \mathbb{N} \cup \{0\}$. Für eine Liste $v_1, \ldots, v_n \in V$ der Länge n sind die folgenden Aussagen äquivalent:*

(1) Die Vektoren $v_1, \ldots, v_n$ sind linear unabhängig.

(2) Die Vektoren $v_1, \ldots, v_n$ sind erzeugend.

(3) Die Vektoren $v_1, \ldots, v_n$ bilden eine Basis.

[16]Falls nicht, dann haben sie es sicher von Karate Kid gelernt, oder?!

Beweis. Wir zeigen zunächst, dass (1) die Aussage in (3) impliziert: Laut Lemma 2.84 kann die linear unabhängige Liste $v_1, \ldots, v_n$ durch die Annahme in (1) zu einer Basis erweitert werden. Aber laut Theorem 2.78 hat jede Basis Länge n. Daher ist $v_1, \ldots, v_n$ schon eine Basis.

Die Implikationen (3) $\Rightarrow$ (1) und (3) $\Rightarrow$ (2) folgen aus der Definition einer Basis, vgl. Definition 2.67.

Um zu zeigen, dass die Aussage (2) die Aussage (3) impliziert, seien $v_1, \ldots, v_n$ ein Erzeugendensystem. Laut Lemma 2.79 enthält dieses System eine Basis. Da aber alle Basen Länge n haben, ist $v_1, \ldots, v_n$ schon eine Basis. □

Bemerkung 2.86. Es lohnt sich Satz 2.85 noch anders zu formulieren: (1) impliziert: Jede maximal linear unabhängige Menge ist eine Basis. (2) impliziert: Jede minimal erzeugende Menge ist eine Basis. (Wieso? Überlegen Sie sich, wieso das eine Umformulierung des Satzes 2.85 ist.)

Korollar 2.87 (Fehlendes Gleichgewicht)**.** *Sei* $n = \dim V < \infty$*.*

(a) Eine Menge $v_1, \ldots, v_k$ *mit* $k < n$ *ist nicht erzeugend.*

(b) Eine Menge $v_1, \ldots, v_k$ *mit* $n < k$ *ist linear abhängig.*

Beweis. (a) Falls die Aussage in (a) nicht gilt, dann enthält die Menge laut Lemma 2.79 eine Basis mit Länge kleiner als n.

(b) Falls die Aussage in (b) nicht gilt, dann können wir laut Lemma 2.84 die Menge zu einer Basis mit Länge grösser als n erweitern.

Dies beweist das Korollar. □

Bemerkung 2.88 (Gauss und Basen in K^m)*.* Genau wie wir nach Lemma 2.50 den Begriff Span in K^m mit Gauss-Elimination verbinden konnten, können wir auch die Begriffe der linearen (Un-)Abhängigkeit und Basis mit der Gauss'schen Elimination verbinden.

Aus Lemma 2.50 folgt, dass $v_1, \ldots, v_n \in K^m$ linear unabhängig sind genau dann, wenn

$$\begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \in K^m$$

nur die triviale Lösung $x_1 = \ldots = x_n = 0$ hat, was man mit Gauss-Elimination bestimmen kann. Für eine Charakterisierung einer Basis durch ein lineares Gleichungssystem, könnten wir Proposition 2.68 benutzen. (Wieso und wie genau?) Aber wir sind jetzt viel schlauer: $v_1, \dots, v_n \in K^m$ ist eine Basis von K^m, falls $n = m$ und

$$\begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \in K^n$$

nur die triviale Lösung hat, was man mit Gauss-Elimination überprüfen kann.

Beispiel 2.89. (1) Je drei Vektoren in $\mathbb{R}^2$ (oder K^2) sind linear abhängig.

(2) Je vier Polynome in $\mathbb{F}_7[x]_2$ (oder $K[x]_2$) sind linear abhängig.

(3) Man hat keine Chance den Raum $M_{2\times 3}(\mathbb{C})$ (als Vektorraum über $\mathbb{C}$) mit 5 Matrizen aufzuspannen.

(4) Zwei Vektoren $v = (a, b)$ und $w = (c, d)$ sind eine Basis für K^2, falls

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

nur die Lösung $x = y = 0$ hat.

(5) Die Vektoren

$$\begin{aligned} f_1 &= (1, 0, \dots, 0) \\ f_2 &= (1, 1, 0, \dots, 0) \\ &\vdots \\ f_n &= (1, 1, \dots, 1) \end{aligned}$$

in K^n bilden eine Basis von K^n.

(6) Je $n + 1$ (vom Nullpolynom verschiedene) Polynome von verschiedenem Grad in $K[x]_n$ bilden eine Basis von $K[x]_n$ (Forum-Wieso?).

(7) Sei V ein Vektorraum. Per Konvention ist die leere Menge $\emptyset$ linear unabhängig. Ausserdem gilt $\mathrm{Sp}(\emptyset) = \{0_V\}$. Daher ist die leere Menge eine Basis des Untervektorraums $\{0_V\}$ und es gilt $\dim\{0_V\} = 0$.

(8) In der Übungsstunde: Gibt es eine Basis $v_1, \dots, v_4$ von $\mathbb{R}[x]_3$, so dass kein v_i ein quadratisches Polynom[17] ist?

[17] Ein quadratisches Polynom ist ein Polynom von Grad 2.

(9) Sei $\mathcal{A}$ eine Menge von Teilmengen von $\{1, \ldots, n\}$ mit $|A| < n$. Dann kann man nicht jede Teilmenge von $\{1, \ldots, n\}$ mittels der symmetrischen Differenz von Mengen aus $\mathcal{A}$ schreiben. Das heisst $\mathrm{Sp}(\mathcal{A}) \neq P(\{1, \ldots, n\})$.

Es kann gut sein, dass Ihnen die Gültigkeit eines der obigen Beispiele unklar ist. Fragen Sie im Forum nach!

2.3.2 Basen von Vektorräumen ohne Gauss'sche Elimination

Wie findet man eine Basis für einen endlich-dimensionalen Vektorraum oder einen Untervektorraum? Auch hier ist V ein Vektorraum über einem Körper K. Laut Theorem 2.78 hat jeder Vektorraum eine Basis. Als Korollar erhalten wir:

Korollar 2.90. *Jeder Untervektorraum hat eine Basis.*

Beweis. Laut Lemma 2.7 ist jeder Untervektorraum auch ein Vektorraum mit der auf den Untervektorraum eingeschränkten Addition und Skalarmultiplikation. Theorem 2.78 zeigt die Existenz einer Basis. □

Was unklar bleibt, ist wie lang die Basis eines Untervektorraums sein darf!

Proposition 2.91. *Seien V ein endlich-dimensionaler Vektorraum und U ein Untervektorraum. Dann ist U auch endlich-dimensional und* $\dim U \leq \dim V$. *Ausserdem gilt* $\dim U = \dim V$ *genau dann, wenn $U = V$ ist.*

Beweis. Sei $S \subseteq U$ eine Basis von U. Wir behaupten, dass $|S| \leq \dim V$. Falls dies nicht gilt, dann enthält S eine Menge S' mit $\dim V + 1$ Vektoren, die linear unabhängig sind, da S linear unabhängig ist. Laut Lemma 2.84 kann man S' zu einer Basis von V erweitern. Dann bekommt man eine Basis von V, deren Länge grösser ist als $\dim V$, was ein Widerspruch zu Theorem 2.78 ist.

Wir zeigen noch die zweite Aussage: Für „$\Leftarrow$" gibt es nichts zu beweisen. Für „$\Rightarrow$" sei $n = \dim U = \dim V$ und $\mathcal{B} = \{v_1, \ldots, v_n\}$ eine Basis für U. Da $\mathcal{B}$ eine Liste von linear unabhängigen Vektoren in V der Länge $\dim V$ ist, ist $\mathcal{B}$ auch eine Basis für V (laut Satz 2.85) und daher ist $V = \mathrm{Sp}(\mathcal{B}) = U$. Die Proposition folgt. □

Wir kommen also zu der Frage, wie man eine Basis eines Untervektorraums findet. Mittels Korollar 2.90 können wir uns auf Vektorräume konzentrieren.

Erste Methode: Die Beweise benutzen

Die obigen Beweise sind eigentlich algorithmisch, obwohl wir nicht versucht haben die Algorithmen effizient zu wählen. Nichtsdestotrotz führen sie zu einer Antwort, wie wir jetzt kurz erklären:

Top-Down-Methode: Sei $V = \mathrm{Sp}(v_1, \ldots, v_n)$. Wie könnten wir eine Basis für V finden?

- Falls $v_1, \ldots, v_n$ linear unabhängig sind, ist $v_1, \ldots, v_n$ schon eine Basis.
- Falls $v_1, \ldots, v_n$ linear abhängig sind, können wir den Algorithmus im Beweis von Lemma 2.79 benutzen, um Vektoren aus $v_1, \ldots, v_n$ „weg zuwerfen“, ohne den Span zu ändern, bis wir eine linear unabhängige Liste bekommen.

Bottom-Up-Methode: Sei V ein endlich-dimensionaler Vektorraum. Man wählt $v \in V$ mit $v \neq 0$. (Wenn dies nicht möglich ist, dann ist $V = \{0\}$.) Jetzt verwendet man Lemma 2.80: Man wählt $v_2 \notin \mathrm{Sp}(v_1)$ und dann $v_3 \notin \mathrm{Sp}(v_1, v_2)$ bis man dies nicht weiterführen kann, da man den ganzen Vektorraum aufgespannt hat. Die erhaltene Liste ist laut Lemma 2.80 eine linear unabhängige Liste, die V aufspannt (da man keine Vektoren mehr hinzufügen konnte). Das heisst, die erhaltene Liste ist eine Basis. Grundsätzlich ist dies genau das, was im Beweis von Lemma 2.79 passiert.

Die erste Methode oben ist hauptsächlich von einem theoretischen Standpunkt aus interessant. Die zweite Methode benutzt die Gauss'sche Elimination und ist viel effizienter. Es lohnt sich aber, dies von einem allgemeineren Blickwinkel aus zu betrachten.

2.3.3 Zeilen- und Spaltenraum und die Gauss'sche Elimination

Definition 2.92. Sei $A \in M_{m\times n}(K)$ und seien $u_1, \ldots, u_m \in K^n$ die Zeilen der Matrix und $v_1, \ldots, v_n \in K^m$ die Spalten der Matrix. Wir definieren

$$\begin{aligned}\mathrm{ZR}(A) &= \mathrm{Zeilenraum}(A) = \mathrm{Sp}(u_1, \ldots, u_m) \subseteq K^n\\ \mathrm{SR}(A) &= \mathrm{Spaltenraum}(A) = \mathrm{Sp}(v_1, \ldots, v_n) \subseteq K^m.\end{aligned}$$

Lemma 2.93. *Seien $A, B \in M_{m\times n}(K)$, so dass B durch Zeilenoperationen aus A entstanden ist. Dann gilt* $\mathrm{ZR}(A) = \mathrm{ZR}(B)$.

Bemerkung 2.94. Vorsicht! Normalerweise gilt nicht, dass $\mathrm{SR}(A) = \mathrm{SR}(B)$, wenn man Zeilenoperationen verwendet. Die Überraschung ist jedoch, dass dann trotzdem $\dim \mathrm{SR}(A) = \dim \mathrm{SR}(B)$ gilt!

Beweis. Dieser Beweis ist ähnlich zum Beweis, welchen wir im Zusatzskript [2] zu $\mathrm{Lös}(A, b) = \mathrm{Lös}(B, b)$ gegeben haben. Seien $u_1, \ldots, u_m \in K^n$ die Zeilen von A.

$L_i \leftrightarrow L_j$: Zeilen vertauschen hat keinen Einfluss auf den Span:

$$\mathrm{Sp}(u_1, \ldots, u_i, \ldots, u_j, \ldots, u_m) = \mathrm{Sp}(u_1, \ldots, u_{i-1}, u_j, u_{i+1}, \ldots, u_{j-1}, u_i, u_{j+1}, \ldots, u_m)$$

$\lambda L_i \to L_i$: Hier muss man zeigen, dass für alle $0 \neq \lambda \in K$ gilt:

$$\mathrm{Sp}(u_1, \ldots, u_i, \ldots, u_m) = \mathrm{Sp}(u_1, \ldots, \lambda u_i, \ldots, u_m).$$

(Wieso? Zeigen Sie es!)

$\underline{L_i + \lambda L_j \to L_i \ (i \neq j):}$ Hier muss man zeigen:

$$\mathrm{Sp}(u_1, \ldots, u_m) = \mathrm{Sp}(u_1, \ldots, u_{i-1}, u_i + \lambda u_j, u_{i+1}, \ldots, u_m).$$

Die Inklusion „$\supseteq$“ folgt aus der Tatsache, dass $u_i + \lambda u_j$ eine Linearkombination von $u_1, \ldots, u_m$ ist. Die Inklusion „$\subseteq$“ folgt aus der Tatsache, dass

$$u_i = (u_i + \lambda u_j) - \lambda u_j$$

eine Linearkombination von $u_1, \ldots, u_{i-1}, u_i + \lambda u_j, u_{i+1}, \ldots, u_m$ ist. □

Definition 2.95. Seien $n, m \in \mathbb{N}$ und $A \in M_{m\times n}(K)$. Wir definieren den *Spaltenrang* und den *Zeilenrang* von A durch

$$\begin{aligned} \mathrm{Spaltenrang}(A) &:= \dim \mathrm{SR}(A), \\ \mathrm{Zeilenrang}(A) &:= \dim \mathrm{ZR}(A). \end{aligned}$$

Bemerkung 2.96. Wir werden später sehen, dass für jede Matrix A die Gleichung

$$\mathrm{Spaltenrang}(A) = \mathrm{Zeilenrang}(A) \tag{2.12}$$

gilt. Dies ist sicher nicht offensichtlich: Immerhin sind $\mathrm{SR}(A)$ und $\mathrm{ZR}(A)$ im Allgemeinen Untervektorräume von verschiedenen Räumen! Das folgende Lemma zeigt (2.12) für den Spezialfall, dass A in Zeilenstufenform ist.

Lemma 2.97. *Sei $B \in M_{m\times n}(K)$ eine Matrix in Zeilenstufenform. Dann sind die Zeilen, die keine Nullzeilen sind, eine Basis des Zeilenraums* $\mathrm{ZR}(B)$. *Insbesondere gilt* $\dim \mathrm{ZR}(B) = \mathrm{Rang}(B)$. *Ähnlich gilt: Die Spalten, die die Pivots enthalten, sind eine Basis des Spaltenraums. Insbesondere gilt Spaltenrang(B) = Zeilenrang(B).*

Beweis. Ohne Beschränkung der Allgemeinheit nehmen wir an, dass alle Pivots 1 sind. Seien $a_{1j_1}, \ldots, a_{rj_r}$ die Einträge der Pivots (das heisst $j_1 < j_2 < \ldots < j_r$) und $u_1, \ldots, u_r$ die entsprechenden Zeilen.

$\underline{u_1, \ldots, u_r \text{ sind linear unabhängig:}}$ Sei

$$\sum_{i=1}^{r} \lambda_i u_i = 0. \tag{2.13}$$

Da in der j_1-ten Spalte ausser dem Pivot $a_{1j_1} = 1$ nur Nullen stehen, ist die j_1-te Koordinate des Vektors $\sum_{i=1}^{r} \lambda_i u_i \in K^n$ gleich λ_1. Also folgt aus (2.13), dass $\lambda_1 = 0$ sein muss. Die j_2-te Koordinate von $\sum_{i=1}^{r} \lambda_i u_i \in K^n$ ist dann (weil $\lambda_1 = 0$) gleich λ_2. Aus (2.13) folgt wiederum, dass $\lambda_2 = 0$. Mit Induktion zeigt man: Falls $\lambda_1 = \ldots = \lambda_k = 0$ für $k < r$, dann ist die j_{k+1}-te Koordinate von $\sum_{i=1}^{r} \lambda_i u_i \in K^n$ gleich λ_{k+1}. Aus (2.13) folgt, dass $\lambda_{k+1} = 0$. Daher gilt $\lambda_1 = \ldots = \lambda_r = 0$, was die lineare Unabhängigkeit von $u_1, \ldots, u_r$ zeigt. Da alle anderen Zeilen Nullzeilen sind, folgt $\mathrm{ZR}(B) = \mathrm{Sp}(u_1, \ldots, u_r)$ und daher ist $u_1, \ldots, u_r$ eine Basis von $\mathrm{ZR}(B)$.

Den Beweis der zweiten Aussage (bezüglich der Basis des Spaltenraums) ist sehr ähnlich und wir überlassen ihn den Lesern. Die Gleichheit $\mathrm{Spaltenrang}(B) = \mathrm{Zeilenrang}(B)$ folgt. □

Dies beweist auch eine wichtige Tatsache zur Gauss'schen Elimination, die wir vorher nicht bewiesen haben.

Erinnerung: Für eine Matrix A ist $\mathrm{Rang}(A)$ die Anzahl der Pivots einer Matrix in Zeilenstufenform, die durch Zeilenoperationen auf A entstanden ist.

Korollar 2.98. *Der Rang einer Matrix A ist wohl-definiert. Das heisst, der Rang hängt nicht von den Zeilenoperationen ab, die zur Stufenform führen.*

Beweis. Wieso? Dieses „Wieso?“ hat für einmal eine Lösung am Ende des Kapitels. □

Korollar 2.99 (Algorithmus zur Berechnung einer Basis aus einem Erzeugendensystem in K^n)**.** *Sei $U = \mathrm{Sp}(u_1, \ldots, u_m) \subseteq K^n$. Um eine Basis für U zu finden, schreibt man $u_1, \ldots, u_m$ als Zeilen einer Matrix $A \in M_{m\times n}(K)$. Mit Zeilenoperationen führt man A zu einer Matrix B Zeilenstufenform über. Dann sind die Zeilen von B, die keine Nullzeilen sind, eine Basis von U.*

Wie findet man eine Basis für den Spaltenraum? Man könnte statt Zeilenoperationen ganz analog Spaltenoperationen durchführen, Stufenform bezüglich Spalten und so weiter definieren. Man könnte aber auch die Rollen von Zeilen und Spalten vertauschen. Dieser Vorgang kommt mehrmals vor, daher gibt man ihm einen Namen:

Definition 2.100 (Transposition)**.** Sei $A = (a_{ij})_{ij} \in M_{m\times n}(K)$. Wir definieren die *Transponierte* von A als die Matrix

$$A^T = (b_{ij})_{\substack{1\leq i\leq n\\ 1\leq j\leq m}} \in M_{n\times m}(K)$$

mit $b_{ij} = a_{ji}$.

Sei $1 \leq k \leq n$. Dann gelten folgende Eigenschaften:

- Die k-te Zeile von A^T ist die k-te Spalte von A.
- Die k-te Spalte von A^T ist die k-te Zeile von A.
- $(A+B)^T = A^T + B^T$.
- $(\lambda A)^T = \lambda A^T$.

- $(A^T)^T = A$.
- $\mathrm{SR}(A) = \mathrm{ZR}(A^T)$.
- $\mathrm{ZR}(A) = \mathrm{SR}(A^T)$.
- Spaltenoperationen auf A sind dasselbe wie Zeilenoperationen auf A^T.

Bemerkung 2.101. Die Addition und Skalarmultiplikation von Matrizen haben Sie in der Übungsstunde definiert.

Satz 2.102. *Seien $m, n \in \mathbb{N}$ und $A \in M_{m\times n}(K)$. Es gilt*

$$\textit{Spaltenrang}(A) = \textit{Zeilenrang}(A).$$

Wir werden den kompletten Beweis erst in den nächsten Kapiteln durchführen (und bis dann nicht verwenden!), geben dafür aber schon eine Beweisskizze mit Gauss'scher Elimination, die man auch zu einem echten Beweis „upgraden“ kann.

<u>Magie:</u> Wie man mit Gauss'scher Elimination gleichzeitig Basen für vier verschiedene Vektorräume berechnen kann!

Beispiel 2.103. Sei

$$A = \begin{pmatrix} -3 & 6 & -1 & 1 \\ 1 & -2 & 2 & 3 \\ 2 & -4 & 5 & 8 \end{pmatrix}.$$

Wir bringen die Matrix in (Zeilen-)Stufenform. In jedem Schritt bezeichnet w_i die i-te Zeile der vorhergehenden Matrix.

$$A \xrightarrow{w_1 \leftrightarrow w_2} \begin{pmatrix} 1 & -2 & 2 & 3 \\ -3 & 6 & -1 & 1 \\ 2 & -4 & 5 & 8 \end{pmatrix} \xrightarrow{w_2 + 3w_1 \to w_2} \begin{pmatrix} 1 & -2 & 2 & 3 \\ 0 & 0 & 5 & 10 \\ 2 & -4 & 5 & 8 \end{pmatrix}$$
$$\xrightarrow{w_3 - 2w_1 \to w_3} \begin{pmatrix} 1 & -2 & 2 & 3 \\ 0 & 0 & 5 & 10 \\ 0 & 0 & 1 & 2 \end{pmatrix} \xrightarrow{w_2 \leftrightarrow w_3} \begin{pmatrix} 1 & -2 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 5 & 10 \end{pmatrix}$$
$$\xrightarrow{w_3 - 5w_2 \to w_3} \begin{pmatrix} 1 & -2 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{w_1 - w_2 \to w_1} \begin{pmatrix} 1 & -2 & 1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Zusammenfassend haben wir aus Zeilenoperationen die Stufenform

$$A \xrightarrow{\text{Zeilenoperationen}} \begin{pmatrix} 1 & -2 & 1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{2.14}$$

erhalten.

Korollar 2.99 besagt, dass $\{(1,-2,1,1),(0,0,1,2)\}$ eine Basis des Zeilenraums $\mathrm{ZR}(A)$ ist. Wenn man jetzt eine Basis des Spaltenraums $\mathrm{SR}(A)$ finden will, gibt es zwei Methoden:

Methode für die Fleissigen: Ohne viel zu denken, betrachtet man

$$A^T = \begin{pmatrix} -3 & 1 & 2 \\ 6 & -2 & -4 \\ -1 & 2 & 5 \\ 1 & 3 & 8 \end{pmatrix}$$

und führt die Gauss'sche Elimination mit Zeilenoperationen durch:

$$A^T \xrightarrow{\text{Zeilenoperationen}} \begin{pmatrix} 1 & 0 & \frac{1}{5} \\ 0 & 1 & \frac{13}{5} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Daher gilt, dass

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ \frac{1}{5} \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \frac{13}{5} \end{pmatrix} \right\}$$

eine Basis des Spaltenraums $\mathrm{SR}(A)$ ist. (Oder falls Sie etwas gegen Brüche haben, ist

$$\left\{ \begin{pmatrix} 5 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 5 \\ 13 \end{pmatrix} \right\}$$

auch eine Basis. (Wieso?))

Methode für Faule: Hier ist der Trick, mit welchem man eine Basis für $\mathrm{SR}(A)$ und $\mathrm{ZR}(A)$ gleichzeitig finden kann.

Betrachten Sie das Eliminationsverfahren in (2.14). Für eine Matrix $M \in M_{3\times 4}(K)$ schreiben wir $M^{(1)}, M^{(2)}, M^{(3)}, M^{(4)}$ für ihre Spalten. Haben Sie bemerkt, dass

$$A^{(2)} = -2A^{(1)}? \tag{2.15}$$

Bemerken Sie auch, dass die lineare Relation in (2.15) auch für alle anderen Matrizen im Eliminationsprozess gilt. Allgemein gilt: Zeilenoperationen erhalten alle linearen Relationen zwischen den Spalten. Insbesondere Aussagen wie

$$\text{die zweite Spalte } = -2 \text{ (die erste Spalte)}$$

sind invariant unter Zeilenoperationen. Das heisst, Zeilenoperationen ändern die Gültigkeit solcher Aussagen nicht.

Am Ende des Eliminationsprozesses in (2.14) steht die Matrix

$$B := \begin{pmatrix} 1 & -2 & 1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Laut Lemma 2.97 ist $\{$die erste Spalte von B, die dritte Spalte von $B\}$ eine Basis von $\mathrm{SR}(B)$. Diese Aussage ist auch invariant unter Zeilenoperationen. Das heisst, die Aussage ändert sich nicht, wenn wir Zeilenoperationen verwenden. Daher gilt

$$\{\text{die erste Spalte von } A, \text{ die dritte Spalte von } A\}$$

ist eine Basis von $\mathrm{SR}(A)$ oder

$$\left\{ \begin{pmatrix} -3 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 5 \end{pmatrix} \right\}$$

ist eine Basis des Spaltenraums $\mathrm{SR}(A)$ von A. Achtung:

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}$$

ist aber keine Basis von $\mathrm{SR}(A)$!

Korollar 2.104. *Wenn man die faule Version in Beispiel 2.103 studiert, kommt man zu den folgenden Erkenntnissen:*

(1) Sei A eine Matrix und $j_1, \dots, j_r$ die Indizes der Spalten mit Pivots einer Matrix, die aus A durch Zeilenoperationen in Stufenform gebracht wurde. Seien $A^{(i)}$ die Spalten von A. Dann ist $\{A^{(j_1)}, \dots, A^{(j_r)}\}$ eine Basis des Spaltenraums von A.

(2) Dies zeigt auch:

$$\dim \mathit{Spaltenraum}(A) = \dim \mathit{Zeilenraum}(A). \tag{2.16}$$

Bemerkung 2.105. Wir werden zwei (vollständige) Beweise von (2.16) geben und bis dann werden wir es nicht verwenden. Es ist aber schon toll sich zu merken, dass man manche solcher Tatsachen beweisen kann. Sie dürfen in Ihren Berechnungen diesen

Trick benutzen. Wenn wir den Orthogonalitäts-Begriff einführen, werden wir noch einen zusätzlichen Trick lernen.

2.3.4 Summen

Wie immer ist V ein Vektorraum über einem Körper K.

Definition 2.106. Seien $U, W \subseteq V$ zwei Untervektorräume. Wir definieren die *Summe* von U und W als

$$U + W = \{u + w \mid u \in U, w \in W\}.$$

Allgemeiner definieren wir für Untervektorräume $U_1, \ldots, U_n$ von V die *Summe* von $U_1, \ldots, U_n$ als

$$U_1 + \ldots + U_n = \{u_1 + \ldots + u_n \mid u_i \in U_i \text{ für } i = 1, \ldots, n\}.$$

Proposition 2.107. *Seien $U, W \subseteq V$ zwei Untervektorräume. Es gilt*

(1) $U + W = \operatorname{Sp}(U \cup W)$ und damit ist $U + W$ der kleinste Untervektorraum, der U und W enthält.

(2) Falls $U + W$ endlich-dimensional ist, dann kann man folgendermassen eine Basis von $U + W$ bilden: Seien $k = \dim(U \cap W)$, $l = \dim(U)$ und $m = \dim W$.

(a) Man wählt eine Basis $p_1, \ldots, p_k$ von $U \cap W$.

(b) Man wählt eine Basis $p_1, \ldots, p_k, u_1, \ldots, u_{l-k}$ für U und eine Basis $p_1, \ldots, p_k, w_1, \ldots, w_{m-k}$ für W.

Dann ist

$$p_1, \ldots, p_k, u_1, \ldots, u_{l-k}, w_1, \ldots, w_{m-k} \tag{2.17}$$

eine Basis für $U + W$.

(3) Insbesondere gilt die sogenannte Dimensionsformel*:*

$$\dim U + W = \dim U + \dim W - \dim U \cap W.$$

Beweis. Teil (1) überlassen wir den Lesern. (Hinweis: Für die zweite Aussage benutzen Sie Lemma 2.36.)

Für (2) müssen wir beweisen, dass die Liste in (2.17) eine Basis von $U + W$ ist. Da die Liste eine Basis für U und W enthält, spannt diese Liste nach (1) $U + W$ auf. Es bleibt also zu zeigen, dass die Liste in (2.17) linear unabhängig ist. Nehmen wir also an, dass

$$a_1 p_1 + \ldots + a_k p_k + b_1 u_1 + \ldots + b_{l-k} u_{l-k} + c_1 w_1 + \ldots + c_{m-k} w_{m-k} = 0 \tag{2.18}$$

beziehungsweise, dass

$$v := c_1 w_1 + \ldots + c_{m-k} w_{m-k} = -(a_1 p_1 + \ldots + a_k p_k + b_1 u_1 + \ldots + b_{l-k} u_{l-k}).$$

für $a_1, \ldots, a_k, b_1, \ldots, b_{l-k}, c_1, \ldots, c_{m-k} \in K$. Einerseits gilt $v \in W$, da die linke Seite in W ist. Andererseits gilt $v \in U$, da die rechte Seite in U ist. Daher ist $v \in U \cap W$. Also existieren $\alpha_1, \ldots, \alpha_p \in K$ mit $v = \alpha_1 p_1 + \ldots + \alpha_k p_k$ und somit

$$\alpha_1 p_1 + \ldots + \alpha_k p_k = c_1 w_1 + \ldots + c_{m-k} w_{m-k}.$$

Da $\{p_1, \ldots, p_k, w_1, \ldots, w_{m-k}\}$ linear unabhängig sind, folgt $c_1 = \ldots = c_{m-k} = \alpha_1 = \ldots = \alpha_p = 0$. Dann vereinfacht sich Gleichung (2.18) zu

$$a_1 p_1 + \ldots + a_k p_k + b_1 u_1 + \ldots + b_{l-k} u_{l-k} = 0,$$

da $c_1 = \ldots = c_{m-k} = 0$. Da $\{p_1, \ldots, p_k, u_1, \ldots, u_{l-k}\}$ auch linear unabhängig ist, gilt

$$a_1 = \ldots = a_k = b_1 = \ldots = b_{l-k} = 0.$$

Dies zeigt, dass die Liste in (2.17) linear unabhängig und daher eine Basis von $U + W$ ist. Es folgt dann auch (3). □

Bemerkung 2.108. Hier lohnt es sich auch den Beweis im Fischer [6, Kap. 1.6.1] zu lesen. Man könnte ausgehend von (2.18) die lineare Unabhängigkeit auch ein bisschen anders zeigen (mittels Proposition 2.68).

Übung 2.109. Seien U_1, U_2, U_3 Untervektorräume eines endlich-dimensionalen Vektorraums. Betrachten Sie die Formel

$$\begin{aligned}\dim(U_1 + U_2 + U_3) = &\dim U_1 + \dim U_2 + \dim U_3 \\ &- \dim(U_1 \cap U_2) - \dim(U_1 \cap U_3) - \dim(U_2 \cap U_3) \\ &+ \dim(U_1 \cap U_2 \cap U_3).\end{aligned}$$

Beweisen Sie die Formel oder geben Sie ein Gegenbeispiel!

Direkte Summen

Proposition 2.110. *Sei V ein endlich-dimensionaler Vektorraum über K und seien $U, W \subseteq V$ zwei Untervektorräume. Dann sind folgende Aussagen äquivalent:*

(1) Es gilt $\dim(U + W) = \dim U + \dim W$.

(2) Es gilt $\dim(U \cap W) = 0$.

(3) Es gilt $U \cap W = \{0\}$.

(4) Jedes $v \in U + W$ *ist eindeutig darstellbar als*

$$v = u + w$$

mit $u \in U$ *und* $w \in W$.

(5) Falls $0 = u + w$ *mit* $u \in U$ *und* $w \in W$, *dann ist* $u = 0$ *und* $w = 0$.

Definition 2.111. Falls eine der äquivalenten Bedingungen in Proposition 2.110 gilt, dann sagen wir, dass $U + W$ die *direkte Summe* von U und W ist und schreiben

$$U + W = U \oplus W.$$

Beweis von Proposition 2.110. Die Äquivalenz von (1) und (2) folgt aus Proposition 2.107 (3). Des Weiteren ist (2) äquivalent zu (3). (Wieso? Es gibt nur einen Untervektorraum mit Dimension 0!) Wir zeigen jetzt, dass (3) die Aussage in (4) impliziert. Seien $u_1 + w_1 = u_2 + w_2$ zwei Darstellungen eines Vektors in $U + W$ mit $u_1, u_2 \in U$ und $w_1, w_2 \in W$. Dann gilt, dass

$$u_1 - u_2 = w_2 - w_1 \in U \cap W.$$

Laut (3) ist $U \cap W = \{0\}$ und daher ist $u_1 - u_2 = w_2 - w_1 = 0$ beziehungsweise $u_1 = u_2$ und $w_2 = w_1$, was (4) zeigt. Des Weiteren gilt „(4) $\Rightarrow$ (5)“. (Wieso? Da $0 = 0 + 0$.) Nun zeigen wir die Implikation „(5) $\Rightarrow$ (3)“: Falls es $0 \neq v \in U \cap W$ gibt, dann ist

$$0 = \underbrace{v}_{\in U} + \underbrace{(-v)}_{\in W}$$

keine triviale Darstellung von 0 als $u + w$ mit $u \in U$ und $w \in W$, im Widerspruch zu (5). □

Proposition 2.112. *Seien* V *ein endlich-dimensionaler Vektorraum und* U *ein Untervektorraum. Dann existiert ein Untervektorraum* W *mit*

$$V = U \oplus W.$$

Definition 2.113. Ein Untervektorraum W wie in Proposition 2.112 wird ein *Komplement* von U in V genannt (oder auch ein *direkter Summand* von V zu U).

Beweis von Proposition 2.112. Sei $u_1, \ldots, u_l$ eine Basis für U und sei $u_1, \ldots, u_l, w_1, \ldots, w_k$ eine Erweiterung der Basis von U zu einer Basis von V. Sei $W := \mathrm{Sp}(w_1, \ldots, w_k)$. Wir behaupten, dass $V = U \oplus W$. Beachten sie dazu zuerst, dass

$$V = U + W = \mathrm{Sp}(U \cup W) = \mathrm{Sp}(u_1, \ldots, u_l, w_1, \ldots, w_k).$$

(Wieso?) Ausserdem gilt $U \cap W = \{0\}$: Falls $v \in U \cap W$, dann existieren $a_1, \ldots, a_l, b_1, \ldots, b_k \in K$ mit

$$v = a_1 u_1 + \ldots + a_l u_l = b_1 w_1 + \ldots b_k w_k.$$

Da $u_1, \ldots, u_l, w_1, \ldots, w_k$ linear unabhängig sind, folgt jedoch

$$a_1 = \ldots = a_l = b_1 = \ldots = b_k = 0$$

und daher ist $v = 0$. Also folgt mit Proposition 2.110 die Aussage. □

Beispiel 2.114. Wir geben einige Beispiele:

(1) Hier ist ein allgemeines Beispiel:

Sei V ein endlich-dimensionaler Vektorraum und $\mathcal{B} = \{v_1, \ldots, v_n\}$ eine Basis von V. Sei $S \subseteq \mathcal{B}$. Dann ist

$$V = \mathrm{Sp}(S) \oplus \mathrm{Sp}(\mathcal{B} \setminus S).$$

Zum Beispiel lässt sich K^n dadurch so schreiben:

$$K^n = \mathrm{Sp}(e_1, \ldots, e_k) \oplus \mathrm{Sp}(e_{k+1}, \ldots, e_n)$$

für alle $0 \leq k \leq n$.

(2) Sei X eine endliche Menge und $P(X)$ die dazugehörige Potenzmenge. Wir betrachten $P(X)$ als Vektorraum über $\mathbb{F}_2$. Für jedes $Y \subseteq X$ gilt:

$$P(X) = P(Y) \oplus P(X \setminus Y).$$

(3) Ein Beispiel in K^3: Für $U = \mathrm{Sp}(e_1, e_2)$ und $V = \mathrm{Sp}(e_2, e_3)$ gilt, dass $U + V = K^3$, aber dies ist keine direkte Summe, weil unter anderem auch

$$U \cap V = \mathrm{Sp}(e_2)$$

gilt.

(4) In der Serie haben Sie das Folgende gezeigt:

Sei $V = \mathbb{R}^{\mathbb{R}} = \mathrm{Abb}(\mathbb{R}, \mathbb{R})$, U die Menge aller geraden Funktionen und V die Menge aller ungeraden Funktionen. Dann gilt

$$V = U \oplus W.$$

Sie werden noch mehr Beispiele in der Übungsstunde und in der Serie sehen.

Zurück zu Korollar 2.98

Antwort auf die Frage „Wieso?" im Beweis von Korollar 2.98. Sei B irgendeine Matrix in Stufenform, die durch Zeilenoperationen aus A entstanden ist. Der Beweis von Lemma 2.93 zeigt, dass

$$\mathrm{ZR}(A) = \mathrm{ZR}(B).$$

Lemma 2.97 besagt, dass

$$\dim \mathrm{ZR}(B) = \text{Anzahl der Pivots in } B.$$

Daher ist diese Anzahl Pivots gleich $\dim \mathrm{ZR}(A)$ und hängt nicht von B ab. □

Changelog: Kapitel 2

- 24.10: In der Definition 2.35 wurden die Begriffe Spann und lineare Hüllen hinzugefügt.
- 26.10: In Beispiel 2.46 wurde $0_{\mathbb{R}^v}$ zu $0_{\mathbb{R}^2}$ geändert.
- 26.10: In Beispiel 2.82 wurde zur Klarstellung eine Referenz hinzugefügt.
- 28.10: Bemerkung 2.54 wurde hinzugefügt.
- 28.10: Im Beweis von Lemma 2.74 wurde $v_j \in \mathrm{Sp}(v_1, \ldots, v_{j-1})$ korrigiert.
- 28.10: In Etappe $1 < j \leq m$ im Beweis von Lemma 2.76 wurde der Satz: „Wir betrachten jetzt die Liste $u_1, \ldots, u_{j-1}, u_j, w_1, \ldots, w_{n-(j-1)}$." hinzugefügt.
- 28.10: Im Beweis von Theorem 2.78 wurde klargestellt, dass die Mengen $\mathcal{A}$ und $\mathcal{B}$ endlich sind.
- 31.10: Im Beweis von Lemma 2.72 Teil (a) wurde der Satz mit $a_1v_1 + \ldots + a_jv_j = a_1v_1 + \ldots + a_nv_n$ (bzw. $a_1v_1 + \ldots + a_jv_j = a_1v_1 + \ldots + a_mv_m$ nach Korrektur des Typos) gelöscht, da er nicht gebraucht wird.
- 31.10: Im Beweis von Lemma 2.76 wurde in Etappe $1 < j \leq m$ die richtige Referenz (Lemma 2.72) eingefügt.
- 31.10: Im Beweis von Lemma 2.74 wurde $0 = a_1v_1 + \ldots a_jv_{j-1} + (-1)v_j$ zu $0 = a_1v_1 + \ldots a_{j-1}v_{j-1} + (-1)v_j$ korrigiert.
- 01.11: In Definition 2.57 wurde hinzugefügt, dass die leere Menge per Konvention linear unabhängig ist und in den Definitionen 2.59 und 2.64 wurde „beliebige Menge" durch „nicht leere Menge" ersetzt.
- 03.11: Bei der Bottom-Up-Methode in Abschnitt 2.3.2 wurde $v_2 \notin \mathrm{Sp}(v_1)$ und $v_3 \notin \mathrm{Sp}(v_1, v_2)$ korrigiert.
- 03.11: In Korollar 2.87 (b) wurde „nicht linear abhängig" mit „linear abhängig" ersetzt.
- 03.11: Bemerkung 2.96 wurde hinzugefügt.
- 03.11: Lemma 2.97 wurde mit einer weiteren Aussage ergänzt. Im Beweis von Lemma 2.97 wurden einige Typos korrigiert.
- 04.11: Der Abschnitt Summen fängt mit Definition 2.106 an.

- 06.11: In Beispiel 2.29 wurden die (möglichen) Vektorräume des Untervektorraums angegeben.
- 07.11: Im Beweis von Lemma 2.79 wurde in Etappe $2 \leq j \leq n$ der Typo $v_1, \ldots, v_n$ zu $v_1, \ldots, v_{j-1}$ korrigiert.
- 07.11: Der Beweis der ersten Aussage in Theorem 2.78 wurde umformuliert.
- 07.11: Die Aussage in Lemma 2.84 wurde genauer formuliert.
- 07:11: Im Beweis von Proposition 2.91 wurde der Typo $n = \dim V = \dim W$ zu $n = \dim U = \dim V$ korrigiert.
- 07:11: Korrektur des Typos $\dim(U + W) = \dim V + \dim W$ in Proposition 2.110 (1) zu $\dim(U + W) = \dim U + \dim W$.
- 07:11: Definition 2.113 wurde hinzugefügt.
- 08.11: In Definition 2.100 wurden bei der Matrix A^T die Indizes angepasst.
- 08.11: In Übung 2.109 wurde $\dim(U_1+U_2+U_3)$ auf der rechten Seite der Gleichung zu $\dim(U_1 \cap U_2 \cap U_3)$ korrigiert.
- 08.11: In Bemerkung 2.16 wurde $s \in L_{A,0}$ zu $s \in L_{A,b}$ korrigiert.
- 01.01: In Satz 2.102 wurde der Typo Zeilenrang(B) zu Zeilenrang(A) korrigiert.

Epilog von Kapitel 2

Basis. Das sollten Sie sich aus Kapitel 2 merken.[18] In jedem Vektorraum gibt es eine Basis und alle Basen haben die gleiche Länge (oder Kardinalität)! Vektorräume sind im Allgemeinen abstrakt und eine Basis ermöglicht uns einen Zugang zu solch einem Vektorraum, auch wenn dieser Vektorraum sehr abstrakt ist. Wie genau? Erinnern Sie sich an Proposition 2.68. Im nächsten Kapitel werden wir Folgendes definieren:

Definition. Sei V ein Vektorraum, $\mathcal{B} = \{v_1, \ldots, v_n\}$ eine Basis von V und $v \in V$ ein beliebiger Vektor. Da $\mathcal{B}$ eine Basis ist, können wir die Koordinaten von v bezüglich $\mathcal{B}$ folgendermassen definieren und mit $[v]_\mathcal{B}$ notieren:

$$[v]_\mathcal{B} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in K^n :\iff v = a_1v_1 + \ldots a_nv_n.$$

Sie sollten darüber so nachdenken: $\mathcal{B}$ macht den abstrakten Raum V für uns „explizit“, indem wir v einfach durch den Vektor $[v]_\mathcal{B} \in K^n$ verstehen können. Etwas anschaulicher stelle ich mir $[v]_\mathcal{B}$ als die Beschreibung von $v \in V$ durch „die Brille $\mathcal{B}$“ vor. Das heisst, dass jede Basis uns eine „Brille“ gibt, mit der wir V „sehen können“. Bevor ich mit dieser Plauderei aufhöre, bemerke ich noch Folgendes: Der Vektor $[v]_\mathcal{B}$ gehört zu K^n. Dies erklärt also auch die folgende Aussage: Ein Vektorraum mit Dimension n hat n „Freiheitsgrade“.

[18] Dies ist keine Zusammenfassung! Wir geben nur Kommentare, woher wir kommen und wohin wir gehen mit dem Stoff.

Kapitel 3

Lineare Abbildungen

Vektorräume zu studieren hat im letzten Kapitel viel Spass gemacht. Aber Vektorräume zu betrachten ohne sie miteinander zu vergleichen wäre schade. Dies wäre wie Mengen zu studieren ohne den Funktionsbegriff. Wir haben schon in der Vorlesung erwähnt, dass man in der Mathematik, und vor allem in der Algebra, gewisse Objekte als *Kategorie* betrachten kann. Grob gesagt, ist eine Kategorie eine Klasse von Objekten und Morphismen (auch Pfeile genannt) zwischen diesen Objekten.[1] Auf jeden Fall sind die Objekte in der linearen Algebra Vektorräume und in diesem Kapitel möchten wir die dazugehörigen Pfeile beschreiben: Lineare Abbildungen.

3.1 Definition einer linearen Abbildung

Wir möchten die Definition einer linearen Abbildung von einem abstrakten Blickwinkel aus motivieren. Seien V und W zwei Vektorräume über einem Körper K. Eine Funktion

$$T : V \to W$$

soll eine „lineare Abbildung" genannt werden, falls sie die Vektorraum-Struktur von V und W „respektiert" beziehungsweise miteinander verbindet. Hier ist die offizielle Definition:

Definition 3.1. Seien V und W Vektorräume über einem Körper K. Eine Funktion

$$T : V \to W$$

heisst *linear* (oder *lineare Abbildung* oder *lineare Transformation*), falls folgende zwei Eigenschaften gelten:

[1]Wer sich dafür interessiert kann hier mehr nachlesen.

$$\begin{aligned} &\forall\, v_1, v_2 \in V : T(v_1 + v_2) = T(v_1) + T(v_2) && \text{(Additivität)} \\ &\forall\, a \in K, v \in V : T(av) = aT(v) && \text{(Homogenität)} \end{aligned} \tag{3.1}$$

Die Menge aller linearen Abbildungen von V nach W bezeichnen wir mit

$$\mathrm{Hom}_K(V, W) \quad \text{oder} \quad \mathrm{Hom}(V, W),$$

falls bei Letzterem K aus dem Zusammenhang klar ist. In diesem Kapitel, falls wir nichts anderes sagen, stehen V und W für Vektorräume über einem Körper K und $T : V \to W$ ist eine lineare Abbildung.

Übung 3.2. Zeigen Sie die folgende Äquivalenz:

$$T \text{ ist linear} \iff \forall a, b \in K,\ v_1, v_2 \in V : T(av_1 + bv_2) = aT(v_1) + bT(v_2).$$

Bemerkung 3.3. Das Wort Abbildung ist ein Synonym für Funktion. Jedoch verwendet man in diesem Zusammenhang hier eher den Ausdruck „lineare Abbildung“ und nicht „lineare Funktion“, obwohl sie beide dasselbe bedeuten.

Bemerkung 3.4. Manchmal schreiben wir der Einfachheit halber Tv statt $T(v)$.

Beispiele

Es ist schwierig zu entscheiden, was das beste erste Beispiel für eine lineare Abbildung ist: Es gibt viele! Einige kommen aus der Geometrie, einige aus der Analysis und es gibt noch mehr. Da wir lineare Abbildungen schlussendlich von einem abstrakten Standpunkt aus betrachten wollen, fangen wir mit abstrakten Beispielen an.

Beispiel 3.5. Sei V ein Vektorraum. Es gibt immer zwei lineare Abbildungen in $\mathrm{Hom}(V, V)$:

- Die Identitäts-Abbildung

$$\begin{aligned} \mathrm{Id}_V : V &\to V, \\ v &\mapsto v, \end{aligned}$$

  die jeden Vektor auf sich selbst abbildet. Manchmal bezeichnen wir die Abbildung als Id oder I, falls der zugrundeliegende Vektorraum aus dem Kontext klar ist.

- Die Null-Abbildung

$$\begin{aligned} 0 : V &\to W, \\ v &\mapsto 0_W. \end{aligned}$$

Beachten Sie, dass wir insbesondere auch $W = V$ wählen können.

(Wieso sind dies lineare Abbildungen? Überprüfen Sie, dass die beiden Abbildungen linear sind!)

Beispiel 3.6. Hier ist ein Beispiel, welches durch die Analysis motiviert ist. Sei K ein Körper. Wir definieren

$$\begin{aligned} D : K[x] &\to K[x], \\ a_0 + a_1 x + \ldots + a_n x^n &\mapsto a_1 + 2a_2 x + 3a_3 x^2 + \ldots + n a_n x^{n-1}. \end{aligned}$$

Hier stehen wie immer die Zahlen $1, 2, 3 \ldots \in K$ für $1, 1+1, 1+1+1, \ldots \in K$ und D steht für Differenziation. Man könnte überprüfen, dass für alle $a, b \in K$ und $p, q \in K[x]$

$$D(ap + bq) = aD(p) + bD(q) \tag{3.2}$$

gilt. Mit der Schreibweise $p' := D(p)$ (bzw. $q' = D(q)$) ist (3.2) genau die Formel

$$(ap + bq)' = ap' + bq',$$

die Sie aus der Mittelschule kennen.

Übung 3.7. (1) Falls $\operatorname{char}(K) = 0$, zeigen Sie, dass

$$\{p \in K[x] \mid D(p) = 0\} = \{\alpha \mid \alpha \in K\} = \{\text{konstante Polynome}\}.$$

(2) Sei p eine Primzahl. Berechnen Sie

$$\{q \in \mathbb{F}_p[x] \mid D(q) = 0\}.$$

Beispiel 3.8. Das Integral ist auch eine lineare Abbildung

$$\begin{aligned} I : I([0,1]) &\to \mathbb{R} \\ f &\mapsto \int_0^1 f(x)\,\mathrm{d}x, \end{aligned}$$

wobei $I([0,1])$ die Menge aller Riemann-integrierbaren Funktionen auf $[0,1]$ ist.

Beispiel 3.9. Viele geometrische Abbildungen der Ebene sind linear. Zum Beispiel Reflexion, Rotation, Streckung und Stauchung und so weiter. Dies besprechen Sie in der Übungsstunde. Um Ihr geometrisches Verständnis zu üben, können Sie sich dieses Video oder auch dieses Video von 3Blue1Brown anschauen.

Beispiel 3.10. Sie besprechen ebenfalls Beispiele mit Matrizenräumen in der Übungsstunde.

Beispiel 3.11. Die Abbildung

$$\begin{aligned} m_{x^{73}} : \mathbb{F}_{37}[x]_{12} &\to \mathbb{F}_{37}[x]_{210} \\ p &\mapsto x^{73} \cdot p. \end{aligned}$$

ist linear.

Beispiel 3.12. Die Abbildung (vgl. Einführung in die Lineare Algebra [1])

$$\begin{aligned} S : K^\infty &\to K^\infty \\ (a_1, a_2, \ldots) &\mapsto (a_2, a_3, \ldots) \end{aligned}$$

ist linear. Diese Abbildung heisst auch *Verschiebungsabbildung* und ist zentral im Gebiet der dynamischen Systeme. Wir haben in unserer Fibonacci-Einführung ([1]) gesehen, dass man S auf den Untervektorraum

$$\mathbf{Fib}(K) := \{(a_1, a_2, \ldots) \in K \mid a_n = a_{n-1} + a_{n-2} \text{ für alle } n \geq 3\} \subseteq K^\infty$$

einschränken und

$$S\mid_{\mathbf{Fib}(K)} : \mathbf{Fib}(K) \to \mathbf{Fib}(K)$$

studieren kann.

Beispiel 3.13 (Mutterbeispiel)**.** Zu guter Letzt geben wir jetzt die wichtigste lineare Abbildung zwischen endlich-dimensionalen Vektorräumen an.

Seien $n, m \in \mathbb{N}$ und sei $A \in M_{m\times n}(K)$. Dann ist die Abbildung

$$\begin{aligned} m_A : K^n &\to K^m \\ x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} &\mapsto Ax = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \end{aligned}$$

linear. Bemerken Sie, Folgendes:

- Das Produkt Ax haben wir in Definition 2.12 definiert.
- Hier sollte man K^n und K^m als K^n_{Spal} und K^m_{Spal} betrachten.
- Wir werden in Kürze zeigen, dass sich jede lineare Abbildung zwischen K^n und K^m als m_A für eine Matrix $A \in M_{m\times n}(K)$ darstellen lässt. Ausserdem werden wir später in diesem Kapitel erklären, dass jede lineare Abbildung zwischen zwei beliebigen endlich-dimensionalen Vektorräumen gewissermassen durch solche Abbildungen darstellbar ist.

- Da diese Abbildung so zentral ist, bezeichnen viele Autoren (und auch wir manchmal) diese Abbildung mit T_A statt m_A.

Proposition 3.14 (Erste Eigenschaften)**.** *Seien V und W zwei Vektorräume über einem Körper K und $T : V \to W$ eine lineare Abbildung. Dann gelten folgende Aussagen:*

(1) Für alle $a_1, \ldots, a_n \in K$ und $v_1, \ldots, v_n \in V$ gilt

$$T\left(\sum_{i=1}^{n} a_i v_i\right) = \sum_{i=1}^{n} a_i T(v_i).$$

Dass heisst, dass das Bild einer Linearkombination der v_i mit Koeffizienten a_i eine Linearkombination der $T(v_i)$ mit Koeffizienten a_i ist.

(2) Es gilt $T(0) = 0$ (oder für die verwirrten Leser: $T(0_V) = 0_W$).

Beweis. (1) Dies zeigt man mittels Induktion mit Hilfe von der Definition in (3.1).

(2) Da $0_V + 0_V = 0_V$ ist, gilt

$$T(0_V) = T(0_V + 0_V) = T(0_V) + T(0_V),$$

wobei wir im letzten Schritt die Linearität der Abbildung T benutzt haben. Wir addieren $-T(0_V) \in W$ auf beiden Seiten und erhalten

$$0_W = T(0_V) - T(0_V) = T(0_V) + T(0_V) - T(0_V) = T(0_V).$$

Dies beweist die Proposition. □

Satz 3.15. *Sei V ein endlich-dimensionaler Vektorraum und $v_1, \ldots, v_n \in V$ eine **Basis** von V. Seien $w_1, \ldots, w_n \in W$ **beliebige** Vektoren. Dann existiert eine eindeutige lineare Abbildung $T : V \to W$ mit*

$$T(v_i) = w_i$$

für $i = 1, \ldots, n$.

Bemerkung 3.16. Denken Sie wie immer bei solchen Aussagen an zwei Tatsachen: Existenz und Eindeutigkeit.

Beweis von Satz 3.15. Wir definieren T durch einen Vorgang, der häufig benutzt wird: „lineare Erweiterung“.

Sei $v \in V$. Wir wollen $T(v)$ definieren. Dafür schreiben wir zuerst v als Linearkombination der Basis $v_1, \ldots, v_n$

$$v = a_1 v_1 + \ldots + a_n v_n = \sum_{i=1}^{n} a_i v_i,$$

wobei $a_1, \ldots, a_n \in K$. Erinnern Sie sich daran, dass diese Schreibweise laut Proposition 2.68 eindeutig ist. Proposition 3.14 sagt uns: Falls es eine lineare Abbildung $T : V \to W$ gäbe, dann würde

$$T(v) = T\left(\sum_{i=1}^{n} a_i v_i\right) = \sum_{i=1}^{n} a_i T(v_i) = \sum_{i=1}^{n} a_i w_i \tag{3.3}$$

gelten. Wir können $T(v)$ genau so definieren, dass (3.3) gilt! Wir definieren also

$$T(v) = \sum_{i=1}^{n} a_i w_i.$$

Dies ist wohl-definiert, da die Schreibweise von v als Linearkombination von $v_1, \ldots, v_n$ eindeutig ist. Wir müssen nun zeigen, dass T linear ist. Seien $v, v' \in V$, so dass

$$v = \sum_{i=1}^{n} a_i v_i \quad \text{und} \quad v' = \sum_{i=1}^{n} b_i v_i$$

gilt für $a_1, \ldots, a_n, b_1, \ldots, b_n \in K$. Dann ist

$$v + v' = \sum_{i=1}^{n} (a_i + b_i) v_i$$

die eindeutige Schreibweise von $v + v'$ bezüglich der Basis $v_1, \ldots, v_n$. Laut unserer Definition von T gilt

$$T(v + v') = \sum_{i=1}^{n} (a_i + b_i) w_i = \sum_{i=1}^{n} (a_i w_i + b_i w_i) = \sum_{i=1}^{n} a_i w_i + \sum_{i=1}^{n} b_i w_i = T(v) + T(v'),$$

was die Additivität von T zeigt. Sei jetzt $\alpha \in K$. Es gilt $\alpha v = \sum_{i=1}^{n} \alpha a_i v_i$ und daher

$$T(\alpha v) = \sum_{i=1}^{n} \alpha a_i w_i = \alpha \sum_{i=1}^{n} a_i w_i = \alpha T(v).$$

Dies zeigt, dass T linear ist und die obige Erklärung zeigt, dass dies die einzig mögliche Definition für eine lineare Abbildung $T : V \to W$ mit $T(v_i) = w_i$ ist. □

Beispiel 3.17. Wieviele lineare Abbildungen $T : \mathbb{F}_3^4 \to \mathbb{F}_3^2$ gibt es oder, in anderen Worten, was ist die Kardinalität von $\operatorname{Hom}(\mathbb{F}_3^4, \mathbb{F}_3^2)$? Wir betrachten die Basis $\{e_1, \ldots, e_4\}$ von $\mathbb{F}_3^4$. Laut Satz 3.15 ergibt jede beliebige Wahl von Bildern $T(e_1), \ldots, T(e_4)$ eine lineare Abbildung $T : \mathbb{F}_3^4 \to \mathbb{F}_3^2$. Da $|\mathbb{F}_3^2| = 9$ ist, gibt es genau 9^4 solche verschiedenen Wahlen von Vektoren. Daher gilt $|\operatorname{Hom}(\mathbb{F}_3^4, \mathbb{F}_3^2)| = 9^4$.

Man könnte dies mit beliebigen Abbildungen (das heisst, Funktionen) von $\mathbb{F}_3^4$ nach $\mathbb{F}_3^2$

vergleichen. Es gibt $|\mathbb{F}_3^2|^{|\mathbb{F}_3^4|} = (3^2)^{(3^4)} = 9^{81}$ solche Abbildungen. (Wieso?) Insbesondere sind $9^{81} - 9^4$ dieser Abbildungen (und somit die überwältigende Mehrheit) nicht linear! Ähnlich gilt

$$|\operatorname{Hom}(V, \mathbb{F}_p^k)| = (p^k)^n$$

wobei p eine Primzahl, k eine natürliche Zahl und V ein n-dimensionaler Vektorraum über $\mathbb{F}_p$ ist.

Wir betrachten jetzt K^n mit der Standard-Basis $e_1, \ldots, e_n$. Seien $w_1, \ldots, w_n \in K^m$ beliebige Vektoren. Laut Satz 3.15 existiert eine eindeutige lineare Abbildung

$$T : K^n \to K^m$$

mit $T(e_i) = w_i$ für $i = 1, \ldots, n$. Sei A die Matrix, deren Spalten $w_1, \ldots, w_n$ sind, also

$$A = \begin{pmatrix} | & | & & | \\ w_1 & w_2 & \cdots & w_n \\ | & | & & | \end{pmatrix}.$$

Dann gilt $m_A(e_i) = Ae_i = w_i$. Das heisst, m_A ist die einzige lineare Abbildung T, die die Eigenschaften von Satz 3.15 erfüllt. Da $w_1 \ldots, w_n$ beliebig sind, folgt das folgende Lemma:

Lemma 3.18. *Für jede lineare Abbildung $T : K^n \to K^m$ existiert ein eindeutiges $A \in M_{m\times n}(K)$, so dass $T = m_A$. Nämlich*

$$A = \begin{pmatrix} | & | & & | \\ T(e_1) & T(e_2) & \cdots & T(e_n) \\ | & | & & | \end{pmatrix} \in M_{m\times n}(K).$$

3.2 Kern und Bild

Auch hier sind V und W zwei Vektorräume über einem beliebigen Körper K. Mit jeder linearen Abbildung sind (zumindest) zwei Untervektorräume verbunden:

Proposition 3.19 (Definition und Proposition)**.** *Sei $T : V \to W$ eine lineare Abbildung.*

(1) Die Menge

$$\operatorname{Ker}(T) = \{v \in V \mid Tv = 0_W\}$$

ist ein Untervektorraum von V und heisst der Kern *von T.*

(2) Das Bild *von T*

$$\mathrm{Im}(T) = \{Tv \mid v \in V\}$$

ist ein Untervektorraum von W.[2]

Erinnerung: Für eine Funktion $f : A \to B$ und $b \in B$ ist

$$f^{-1}(b) = f^{-1}(\{b\}) = \{a \in A : f(a) = b\} \tag{3.4}$$

die Menge aller Urbilder von b. Mit dieser Notation ist $\mathrm{Ker}(T) = T^{-1}(0_W)$.

Beweis. (1) Seien $a, b \in K$ und $v_1, v_2 \in \mathrm{Ker}(T)$. Dann gilt

$$T(av_1 + bv_2) = aTv_1 + bTv_2 = a \cdot 0_W + b \cdot 0_W = 0_W.$$

Ausserdem ist $0_V \in \mathrm{Ker}(T)$ laut Proposition 3.14 und daher ist $\mathrm{Ker}(T)$ ein Untervektorraum von V.

(2) Seien jetzt $w_1, w_2 \in \mathrm{Im}(T)$. Dann existieren $v_1, v_2 \in V$ mit $Tv_1 = w_1$ und $Tv_2 = w_2$ laut der Definition von $\mathrm{Im}(T)$. Es gilt

$$T(av_1 + bv_2) = aTv_1 + bTv_2 = aw_1 + bw_2$$

und daher ist $aw_1 + bw_2 \in \mathrm{Im}(T)$. Auch hier folgt aus Proposition 3.14, dass $0_W \in \mathrm{Im}(T)$ ist, da $T(0_V) = 0_W$. Daher ist $\mathrm{Im}(T)$ ein Untervektorraum von W.

Dies beweist die Proposition. □

Beispiel 3.20 (Kern und homogenene lineare Gleichungssysteme)**.** Sei $A \in M_{m\times n}(K)$. Wir betrachten $m_A : K^n \to K^m$. Dann gilt

$$\mathrm{Ker}(m_A) = \mathrm{Lös}(A, 0_{K^m}).$$

Das heisst, $\mathrm{Ker}(m_A)$ sind die Lösungen des homogenen linearen Gleichungssystems

$$Ax = 0_{K^m}, \quad x \in K^n.$$

Ab jetzt werden wir $\mathrm{Ker}(m_A)$ auch mit $\mathrm{Ker}(A)$ bezeichnen.

Beispiel 3.20 zeigt schon, dass $\mathrm{Ker}(A)$ interessant ist. Hier ist noch ein weiterer interessanter Grund:

Proposition 3.21. *Sei T eine lineare Abbildung. Dann gelten folgende Aussagen:*

(1) T ist injektiv $\iff$ $\mathrm{Ker}(T) = \{0_V\}$.

[2] Viele Autoren verwenden auch $\mathrm{Bild}(T)$ statt $\mathrm{Im}(T)$.

(2) T ist surjektiv $\iff \mathrm{Im}(T) = W$.

Beweis. Bemerken Sie zuerst: In (2) gibt es nichts zu beweisen! Wir beweisen also (1):

„$\Rightarrow$“: Wir haben schon gesehen, dass $\mathrm{Ker}(T)$ die Menge aller Urbilder von 0_W ist. Das heisst, dass $\mathrm{Ker}(T) = T^{-1}(0_W)$. Laut Proposition 3.14 ist 0_V ein Urbild von 0_W. Da T injektiv ist, hat jedes Element von W höchstens ein Urbild und daher ist

$$\mathrm{Ker}(T) = T^{-1}(0_W) = \{0_V\}.$$

„$\Leftarrow$“: Hier brauchen wir die Linearität. Seien $v_1, v_2 \in V$ mit $T(v_1) = T(v_2)$. Da T linear ist, gilt

$$T(v_1 - v_2) = T(v_1) - T(v_2) = 0_W.$$

Daher ist $v_1 - v_2 \in \mathrm{Ker}(T)$ und nach Annahme ist $\mathrm{Ker}(T) = \{0_V\}$. Also folgt, dass $v_1 = v_2$ ist, was die Injektivität zeigt. □

Wir können den Sachverhalt in Proposition 3.21 vage so ausdrücken:

„Injektivität überall ist äquivalent zu Injektivität bei der Null“.

Wir werden sehen, dass es auch andere Eigenschaften ausser der Injektivität gibt, die sich „bei der Null“ überprüfen lassen.

Übung 3.22. Verallgemeinern Sie die Resultate dieses Abschnitts: Sei $T : V \to W$ linear. Zeigen Sie folgende Aussagen:

1. Sei $V' \subseteq V$ ein Untervektorraum. Dann ist $T(V')$ ein Untervektorraum von W.
2. Sei $W' \subseteq W$ ein Untervektorraum. Dann ist $T^{-1}(W')$ ein Untervektorraum von V.[3]

Definition 3.23. Eine lineare Abbildung $T : V \to W$ nennen wir[4]

- einen *Isomorphismus*, falls T bijektiv ist.
- einen *Endomorphismus*, falls $W = V$ ist.
- einen *Automorphismus*, falls $V = W$ und T bijektiv ist.

Falls ein Isomorphismus $T : V \to W$ existiert, sagen wir, dass V und W *isomorph* sind und schreiben $V \cong W$. Die Menge aller Endomorphismen von V bezeichnen wir mit $\mathrm{End}(V)$. Das heisst, $\mathrm{End}(V) := \mathrm{Hom}(V, V)$.

[3] $T^{-1}(W') = \{v \in V : Tv \in W'\}$.

[4] Wir werden folgende Begriffe nicht benutzen, aber andere Autoren definieren auch:

- T heisst *Monomorphismus*, falls T injektiv ist.
- T heisst *Epimorphismus*, falls T surjektiv ist.

Lemma 3.24. *Falls $T : V \to W$ ein Isomorphismus ist, dann existiert $T^{-1} : W \to V$ und ist auch linear.*

Beweis. Seien $a, b \in K$ und $w_1, w_2 \in W$. Seien $v_1, v_2 \in V$ die eindeutigen Urbilder von w_1, w_2. Das heisst, dass $T(v_i) = w_i$ für $i = 1, 2$ oder äquivalent $T^{-1}(w_i) = v_i$ für $i = 1, 2$ gilt. Da T linear ist, gilt

$$T(av_1 + bv_2) = aTv_1 + bTv_2 = aw_1 + bw_2.$$

Das heisst, weil T bijektiv ist, dass

$$T^{-1}(aw_1 + bw_2) = av_1 + bv_2 = aT^{-1}(w_1) + bT^{-1}(w_2).$$

Dies zeigt, dass T^{-1} linear ist. □

Übung 3.25. Zeigen Sie, dass „isomorph“ eine Äquivalenzrelation auf der Menge der Vektorräume ist.

Lemma 3.26. *Sei $T : V \to W$ linear und $v_1, \ldots, v_n$ eine Basis für V. Dann gilt $\mathrm{Im}(T) = \mathrm{Sp}(Tv_1, \ldots, Tv_n)$.*

Beweis. Eine Gleichheit zwischen zwei Mengen zeigt man, indem man zwei Inklusionen zeigt.

„$\supseteq$“: Für $a_1, \ldots, a_n \in K$ gilt

$$\sum_{i=1}^{n} a_i T v_i = T\left(\sum_{i=1}^{n} a_i v_i\right) \tag{3.5}$$

und daher $\sum_{i=1}^{n} a_i T v_i \in \mathrm{Im}(T)$.

„$\subseteq$“: Sei $w \in \mathrm{Im}(T)$ und $v \in V$ mit $Tv = w$. Da $v_1, \ldots, v_n$ eine Basis von V ist, existieren $a_1, \ldots, a_n \in K$ mit $v = \sum_{i=1}^{n} a_i v_i$. Jetzt lesen wir Gleichung (3.5) umgekehrt:

$$w = Tv = T\left(\sum_{i=1}^{n} a_i v_i\right) = \sum_{i=1}^{n} a_i T v_i$$

und daher ist $w \in \mathrm{Sp}(Tv_1, \ldots, Tv_n)$. □

Wir sind jetzt bereit den wichtigsten Satz zu linearen Abbildungen zu zeigen:

Satz 3.27 (Rangsatz, Energieerhaltungs-Satz)**.** *Sei $T : V \to W$ eine lineare Abbildung zwischen Vektorräumen V und W, wobei V endlich-dimensional ist. Dann gilt*

$$\dim \mathrm{Ker}(T) + \dim \mathrm{Im}(T) = \dim V.$$

Beweis. Sei $n = \dim V$ und $u_1, \ldots, u_k$ eine Basis von $\mathrm{Ker}(T)$. Mittels Lemma 2.84 ergänzen wir $u_1, \ldots, u_k$ zu einer Basis von V:

$$u_1, \ldots, u_k, v_1, \ldots, v_{n-k}.$$

Laut Lemma 3.26 gilt

$$\mathrm{Im}(T) = \mathrm{Sp}(Tu_1, \ldots, Tu_k, Tv_1, \ldots, Tv_{n-k}) = \mathrm{Sp}(Tv_1, \ldots, Tv_{n-k}), \tag{3.6}$$

da $Tu_1 = \ldots = Tu_k = 0_W$. Wir behaupten, dass $Tv_1, \ldots, Tv_{n-k}$ eine Basis von $\mathrm{Im}(T)$ ist. Dazu müssen wir nur zeigen, dass $Tv_1, \ldots, Tv_{n-k}$ linear unabhängig sind. (Wieso? Wegen (3.6)!)

Seien also $a_1, \ldots, a_{n-k} \in K$ mit $\sum_{i=1}^{n-k} a_i T v_i = 0_W$. Da T linear ist, gilt

$$T\left(\sum_{i=1}^{n-k} a_i v_i\right) = \sum_{i=1}^{n-k} a_i T v_i = 0_W.$$

Also ist $\sum_{i=1}^{n-k} a_i v_i \in \mathrm{Ker}(T)$. Daher existieren $b_1, \ldots, b_k$ mit

$$\sum_{i=1}^{n-k} a_i v_i = \sum_{i=1}^{k} b_i u_i$$

beziehungsweise

$$\sum_{i=1}^{n-k} a_i v_i + \sum_{i=1}^{k} (-b_i) u_i = 0_W.$$

Weil $u_1, \ldots, u_n, v_1, \ldots, v_{n-k}$ eine Basis von V ist, impliziert das

$$a_1 = \ldots = a_{n-k} = -b_1 = \ldots = -b_k = 0.$$

Also ist insbesondere $a_1 = \ldots = a_{n-k} = 0$ und daraus folgt, dass $Tv_1, \ldots, Tv_{n-k}$ eine Basis von $\mathrm{Im}(T)$ ist. Dies zeigt, dass

$$\dim \mathrm{Im}(T) = n - k.$$

Also gilt

$$\dim V = k + (n - k) = \dim \mathrm{Ker}(T) + \dim \mathrm{Im}(T)$$

per Definition von n und k als $n = \dim V$ und $k = \dim \mathrm{Ker}(T)$. □

Bemerkung 3.28. Ich stelle mir Satz 3.27 immer als „Energieerhaltungs-Satz“ vor: Der Vektorraum V hat $n = \dim V$ Dimensionen/Energie zu geben. Die Abbildung T gibt einige dieser Dimensionen dem Bild $\operatorname{Im}(T)$ und die restlichen Dimensionen müssen dann im Kern $\operatorname{Ker}(T)$ verschwinden.

Korollar 3.29. *Sei $T : V \to W$ eine lineare Abbildung. Dann gelten folgende Aussagen:*

(1) Falls $\dim W < \dim V$, dann ist T nicht injektiv.

(2) Falls $\dim W > \dim V$, dann ist T nicht surjektiv.

(3) Falls $\dim W = \dim V$, dann gilt

$$T \text{ ist bijektiv} \iff T \text{ ist injektiv} \iff T \text{ ist surjektiv.}$$

Beweis. (1) Hier hat V zu viel „Energie“: V kann W nur $\dim W$ Dimensionen „geben“ und daher muss T einige Dimensionen an den Kern abgeben. Also ist der Kern nicht trivial. Laut Proposition 3.21 ist T dann nicht injektiv.

Hier ist eine rigorose Version: Da $\operatorname{Im}(T)$ ein Untervektorraum von W ist, gilt $\dim \operatorname{Im}(T) \leq \dim W < \dim V$. Laut Satz 3.27 folgt dann, dass

$$\dim \operatorname{Ker}(T) = \dim V - \dim \operatorname{Im}(T) > 0.$$

Laut Proposition 3.21 ist T dann nicht injektiv.

(2) Energie-Erhaltungs-Version: $\operatorname{Im}(T)$ braucht $\dim W$ Dimensionen, damit T surjektiv sein kann. Aber $\dim V$ hat nicht genug Energie, da $\dim V < \dim W$.

Rigorose Version: Laut Satz 3.27 gilt

$$\dim \operatorname{Im}(T) = \dim V - \dim \operatorname{Ker}(T) \leq \dim V < \dim W.$$

Also folgt aus Proposition 3.21, dass T nicht surjektiv ist.

(3) Wir müssen nur zeigen, dass T injektiv ist genau dann, wenn T surjektiv ist. (Wieso?) Ohne viele Worte gelten die Äquivalenzen:

$$\begin{aligned} T \text{ ist injektiv} \overset{\text{Proposition 3.21}}{\iff} \operatorname{Ker}(T) = \{0\} &\iff \dim \operatorname{Ker}(T) = 0 \\ &\overset{\text{Satz3.27}}{\iff} \dim V = \dim \operatorname{Im}(T) \\ &\overset{\text{Annahme}}{\iff} \dim W = \dim \operatorname{Im}(T) \\ &\iff W = \operatorname{Im}(T) \\ &\iff T \text{ ist surjektiv.} \end{aligned}$$

Dies zeigt das Korollar. □

Lemma 3.30. *Seien $T : V \to W$ und $S : W \to U$ lineare Abbildungen. Dann ist die Verkettung $S \circ T : V \to U$ auch linear.*

Beweis. Seien $\alpha \in K$ und $v_1, v_2 \in V$. Es gilt

$$S \circ T(v_1 + v_2) \overset{T \text{ linear}}{=} S(Tv_1 + Tv_2) \overset{S \text{ linear}}{=} S \circ T(v_1) + S \circ T(v_2).$$

Ähnlich folgt, dass

$$S \circ T(\alpha v) = S(\alpha T v) = \alpha(S \circ T(v)).$$

Dies beweist das Lemma. □

Übung 3.31. Es gibt auch eine andere Beweis-Strategie für Satz 3.27:

Man fängt mit einer Basis $v_1, \ldots, v_k$ von $\operatorname{Ker}(T)$ und einer Basis $w_1, \ldots, w_l$ von $\operatorname{Im}(T)$ an. Dann wählt man beliebige $u_1, \ldots, u_l \in V$ mit

$$Tu_i = w_i$$

für $i = 1, \ldots, l$. Man zeigt dann, dass $v_1, \ldots, v_k, u_1, \ldots, u_l$ eine Basis von V ist. Der Satz folgt somit.

Führen Sie diese Strategie aus!

Als ob es einfach wäre und nicht so wichtig, folgt jetzt nichtsdestotrotz eines der wichtigsten Resultate der linearen Algebra.

Korollar 3.32. *Zwei endlich-dimensionale Vektorräume V und W sind genau dann isomorph, wenn* $\dim V = \dim W$.

Beweis. „$\Rightarrow$“: Falls $T : V \to W$ ein Isomorphismus ist, dann folgt aus Proposition 3.21, dass $\operatorname{Ker}(T) = \{0\}$ und $W = \operatorname{Im}(T)$. Daher impliziert Satz 3.27, dass

$$\dim W = \dim \operatorname{Im}(T) = \dim V - \dim \operatorname{Ker}(T) = \dim V.$$

„$\Leftarrow$“: Falls $\dim V = \dim W$ ist, seien $(v_1, \ldots, v_n)$ eine Basis von V und $(w_1, \ldots, w_n)$ eine Basis von W. Wir definieren mittels Satz 3.15 die (eindeutige) lineare Abbildung

$$T : V \to W \quad \text{mit} \quad T(v_i) = w_i.$$

Laut Lemma 2.39 ist T surjektiv und laut der Annahme $\dim V = \dim W$ ist T dann auch injektiv (mit Hilfe von Korollar 3.29 (3)). Daher ist T ein Isomorphismus. □

Für später brauchen wir noch eine kleine Definition:

Definition 3.33. Sei $T : V \to W$ eine lineare Abbildung. Der *Rang* von T ist

$$\operatorname{Rang}(T) := \dim \operatorname{Im}(T).$$

Übung 3.34. Seien $T : V \to W$ und $S : W \to U$ zwei lineare Abbildungen zwischen endlich-dimensionalen Vektorräumen U, V und W. Zeigen Sie folgende Aussagen:

(1) Es gilt $\operatorname{Rang}(S \circ T) \leq \min\{\operatorname{Rang}(T), \operatorname{Rang}(S)\}$.

(2) Falls S injektiv ist, dann ist $\operatorname{Rang}(S \circ T) = \operatorname{Rang}(T)$.

(3) Falls T surjektiv ist, dann ist $\operatorname{Rang}(S \circ T) = \operatorname{Rang}(S)$.

(Hinweis: Betrachten Sie $S|_{\operatorname{Im}(T)}$.)

3.3 Koordinaten für Vektorräume und lineare Abbildungen

Alle Vektorräume in diesem Abschnitt sind endlich-dimensional. In diesem Abschnitt werden wir ein bisschen mehr Information bezüglich Korollar 3.32 geben: Ein Vektorraum V über K mit $\dim V = n$ ist isomorph zu K^n.

Für den Beweis und auch für später müssen wir Koordinaten definieren. Dazu müssen wir an Basen als geordnete Mengen denken, das heisst, als eine Liste der Form $v_1, \ldots, v_n$. Dafür schreiben wir $\mathcal{B} = (v_1, \ldots, v_n)$, wenn $\mathcal{B}$ eine *geordnete Basis* ist, welche durch die Liste $v_1, \ldots, v_n$ definiert ist.

Definition 3.35. Sei V ein Vektorraum mit Basis $\mathcal{B} = (v_1, \ldots, v_n)$ und sei $v \in V$. Wir definieren den *Koordinaten-Vektor* von v bezüglich der Basis $\mathcal{B}$ als

$$[v]_{\mathcal{B}} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

wobei $a_1, \ldots, a_n \in K$ sind mit $v = a_1 v_1 + \ldots + a_n v_n$. Da $\mathcal{B}$ eine Basis ist, sind $a_1, \ldots, a_n$ eindeutig definiert, und daher ist $[v]_{\mathcal{B}}$ wohl-definiert.

Die Abbildung

$$\Phi_{\mathcal{B}} : V \to K^n, \quad v \mapsto [v]_{\mathcal{B}}$$

nennen wir die *Koordinaten-Abbildung* bezüglich $\mathcal{B}$.

Proposition 3.36. *Die Abbildung $\Phi_{\mathcal{B}}$ ist linear und bijektiv.*

Beweis. Man kann Satz 3.15 verwenden, um zu zeigen, dass $\Phi_{\mathcal{B}}$ die eindeutige lineare Abbildung mit $T(v_i) = e_i \in K^n$ ist, wobei $e_1, \ldots, e_n$ die Standard-Basis von K^n ist.

Wir geben stattdessen einen direkten Beweis. Seien $a, b \in K$ und $v, v' \in V$. Wir schreiben

$$v = a_1 v_1 + \ldots + a_n v_n,$$
$$v' = b_1 v_1 + \ldots + b_n v_n$$

womit

$$[v]_{\mathcal{B}} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \ [v']_{\mathcal{B}} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Dann gilt

$$av + bv' = (aa_1 + bb_1)v_1 + \ldots + (aa_n + bb_n)v_n$$

und daher

$$\Phi_{\mathcal{B}}(av + bv') = [av + bv']_{\mathcal{B}} = \begin{pmatrix} aa_1 + bb_1 \\ \vdots \\ aa_n + bb_n \end{pmatrix} = a \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + b \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = a\Phi_{\mathcal{B}}(v) + b\Phi_{\mathcal{B}}(v').$$

Dies zeigt die Linearität von $\Phi_{\mathcal{B}}$.

Die Bijektivität folgt, da $\mathcal{B}$ eine Basis ist. (Wieso? Benutzen Sie die Eindeutigkeit – Proposition 2.68.) □

Es ist nützlich zu bemerken, dass die Inverse

$$\Phi_{\mathcal{B}}^{-1} : K^n \to V$$

durch

$$\Phi_{\mathcal{B}}^{-1} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = a_1 v_1 + \ldots + a_n v_n$$

gegeben ist.

Beispiel 3.37. Sei $\mathcal{B} = \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right)$ eine Basis von $\mathbb{Q}^2$ und $\mathcal{E} = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$ die Standardbasis. Dann ist beispielsweise

$$\left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} \right]_{\mathcal{B}} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \quad \text{und} \quad \left[\begin{pmatrix} 0 \\ 1 \end{pmatrix} \right]_{\mathcal{B}} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}.$$

Allgemein ist

$$\left[\begin{pmatrix} x \\ y \end{pmatrix}\right]_{\mathcal{B}} = \begin{pmatrix} \frac{x+y}{2} \\ \frac{x-y}{2} \end{pmatrix},$$

da

$$\frac{x+y}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{x-y}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Bezüglich der Standardbasis gilt

$$\left[\begin{pmatrix} x \\ y \end{pmatrix}\right]_{\mathcal{E}} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

(Wieso?)

Beispiel 3.38. Sei $V = K[X]_3$ und $\mathcal{B} = (1, x, x^2, x^3)$ die Standard-Basis und

$$\mathcal{C} = (1, x+1, x^2+x+1, x^3+x^2+x+1)$$

eine andere Basis. Sei $p = a_0 + a_1x + a_2x^2 + a_3x^3$ ein beliebiger Vektor. Bezüglich der Standard-Basis ist das Leben einfach: Man sieht gleich, dass

$$[p]_{\mathcal{B}} = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} \quad \text{und} \quad [x^3]_{\mathcal{C}} = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}.$$

In diesem Beispiel kann man auch deutlich den Einfluss der Reihenfolge sehen. Zum Beispiel gilt für $\mathcal{B}' = (x^3, x^2, x, 1)$ und $\mathcal{B}'' = (1, x^2, x^3, x)$, dass

$$[p]_{\mathcal{B}'} = \begin{pmatrix} a_3 \\ a_2 \\ a_1 \\ a_0 \end{pmatrix} \quad \text{und} \quad [p]_{\mathcal{B}''} = \begin{pmatrix} a_0 \\ a_2 \\ a_3 \\ a_1 \end{pmatrix}.$$

Für $[p]_{\mathcal{C}}$ müssen wir etwas rechnen. Hier ist die Antwort:

$$[p]_{\mathcal{C}} = \begin{pmatrix} a_0 - a_1 \\ a_1 - a_2 \\ a_2 - a_3 \\ a_3 \end{pmatrix}.$$

Bemerkung. $\Phi_{\mathcal{B}}$ ist die eindeutige lineare Abbildung von V nach K^n, die durch Satz 3.15 mit der Bedingung $\Phi_{\mathcal{B}}(v_i) = e_i \in K^n$ definiert ist, wobei $\mathcal{B} = (v_1, \ldots, v_n)$ und

$(e_1, \ldots, e_n)$ die Standard-Basis von K^n ist.

In Übung 3.25 haben wir gesehen, dass „isomorph sein“ eine Äquivalenzrelation auf der Menge aller Vektorräume ist. Korollar 3.32 besagt, dass die Dimension aller Elemente einer Äquivalenzklasse gleich ist. Jetzt können wir sogar einen Repräsentanten für die Äquivalenzklasse finden:

Satz 3.39. *Sei V ein Vektorraum über K mit* $\dim V = n \in \mathbb{N} \cup \{0\}$. *Dann ist V isomorph zu K^n.*

Beweis. Sei $\mathcal{B}$ eine Basis von V. Dann ist die Abbildung

$$\Phi_{\mathcal{B}} : V \to K^n$$

ein Isomorphismus. □

Bemerkung 3.40 (Nebenbemerkung). Die Abbildung $\Phi_{\mathcal{B}}$ ist nicht-kanonisch. Für ihre Definition müssen wir eine Wahl treffen (die Wahl von $\mathcal{B}$). Vergleichen Sie dies mit der Konstruktion der Quotienten-Abbildung am Ende dieses Kapitels.

Die Koordinaten-Abbildung bringt uns zu unserer nächsten Frage:

Seien V und W zwei Vektorräume mit $\dim V = n$ und $\dim W = m$. Sei $T : V \to W$ eine lineare Abbildung. Wir wählen Basen $\mathcal{B} \subseteq V$ und $\mathcal{C} \subseteq W$ und betrachten das Diagramm:

$$\begin{tikzcd}
V \arrow[r, "T"] & W \arrow[d, "\Phi_{\mathcal{C}}"] \\
K^n \arrow[u, "\Phi_{\mathcal{B}}^{-1}"] \arrow[r, dashed, "\Phi_{\mathcal{C}} \circ T \circ \Phi_{\mathcal{B}}^{-1}"'] & K^m
\end{tikzcd}$$

Laut Lemma 3.18 ist die lineare Abbildung $\Phi_{\mathcal{C}} \circ T \circ \Phi_{\mathcal{B}}^{-1} : K^n \to K^m$ durch eine Matrix $A \in M_{m \times n}(K)$ darstellbar. Das heisst, es existiert eine eindeutige Matrix $A \in M_{m \times n}(K)$, so dass

$$\begin{tikzcd}
V \arrow[r, "T"] & W \arrow[d, "\Phi_{\mathcal{C}}"] \\
K^n \arrow[u, "\Phi_{\mathcal{B}}^{-1}"] \arrow[r, dashed, "m_A"'] & K^m
\end{tikzcd} .$$

kommutiert. Das heisst, dass $\Phi_{\mathcal{C}} \circ T \circ \Phi_{\mathcal{B}}^{-1} = m_A$ gilt.

Toll. Aber ich habe jetzt noch viele Fragen:

$$\begin{array}{l} \text{(1) Wie kann ich } A \text{ berechnen?} \\ \text{(2) Wie sieht } A \text{ bezüglich verschiedenen Basen aus?} \\ \text{(3) Wie sieht die Matrix einer Verkettung aus?} \\ \quad\;\; \text{Gibt uns die Verkettung eine Möglichkeit Matrizen zu multiplizieren?} \end{array} \tag{3.7}$$

Diese Fragen führen uns zu einem neuen Abschnitt!

3.3.1 Matrizen und lineare Abbildungen

Die erste der vorherigen Fragen können wir gleich beantworten. Wir müssen einfach der Definition folgen.

Sei $T : V \to W$ eine lineare Abbildung, $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis des Vektorraums V und $\mathcal{C} = (w_1, \ldots, w_m)$ eine Basis des Vektorraums W. Wir betrachten

$$\begin{tikzcd}
v_i \in V \arrow[r, "T"] \arrow[r, bend left, mapsto] & W \ni Tv_i \arrow[d, "\Phi_{\mathcal{C}}"'] \arrow[d, bend left, mapsto] \\
e_i \in K^n \arrow[u, "\Phi_{\mathcal{B}}^{-1}"'] \arrow[u, bend left, mapsto] \arrow[r, dashed, "m_A"'] & K^m \ni [Tv_i]_{\mathcal{C}}
\end{tikzcd} . \tag{3.8}$$

für $A \in M_{m\times n}(K)$ wie oben.

Was geschieht mit der Standard-Basis $e_1, \ldots, e_n \in K^n$, wenn wir $\Phi_{\mathcal{C}} \circ T \circ \Phi_{\mathcal{B}}^{-1}$ darauf anwenden? Für $i \in \{1, \ldots, n\}$ gilt

$$\Phi_{\mathcal{C}} \circ T \circ \Phi_{\mathcal{B}}^{-1}(e_i) = \Phi_{\mathcal{C}} \circ T(v_i) = \Phi_{\mathcal{C}}(T(v_i)) = [T(v_i)]_{\mathcal{C}}.$$

Dies bedeutet, dass m_A die (eindeutige) lineare Abbildung ist mit $m_A(e_i) = [T(v_i)]_{\mathcal{C}}$. Laut Lemma 3.18 heisst das, dass

$$A = \begin{pmatrix} | & & | \\ [T(v_1)]_{\mathcal{C}} & \cdots & [T(v_n)]_{\mathcal{C}} \\ | & & | \end{pmatrix} \in M_{m\times n}(K). \tag{3.9}$$

Anders gesagt, ist

$$A = (a_{ij})_{\substack{1\le i\le m \\ 1\le j\le n}}$$

mit Einträgen $a_{ij} \in K$, die durch die Gleichung

$$T(v_j) = a_{1j}w_1 + \ldots a_{mj}w_m = \sum_{k=1}^{m} a_{kj}w_k \quad \text{für } j = 1, \cdots, n \tag{3.10}$$

definiert sind.

Dies beantwortet Frage (1). Wir fassen zusammen:

Definition 3.41. Die Matrix A wie in (3.9) heisst die *Darstellungsmatrix* von T bezüglich der Basen $\mathcal{B}$ und $\mathcal{C}$ und wird mit $[T]^{\mathcal{B}}_{\mathcal{C}}$ notiert.

Wir haben bewiesen:

Proposition 3.42. *Sei $T : V \to W$ eine lineare Abbildung zwischen zwei Vektorräumen V und W. Sei $\mathcal{B}$ eine Basis von V und $\mathcal{C}$ eine Basis von W. Für jedes $v \in V$ gilt*

$$[T]^{\mathcal{B}}_{\mathcal{C}}[v]_{\mathcal{B}} = [Tv]_{\mathcal{C}}. \tag{3.11}$$

Beweis. Gleichung (3.11) ist äquivalent dazu zu sagen, dass das folgende Diagramm kommutiert:

$$\begin{tikzcd}
v \arrow[r, phantom, "\in"] \arrow[d, bend right] & V \arrow[r, "T"] \arrow[d, "\Phi_{\mathcal{B}}"] & W \arrow[d, "\Phi_{\mathcal{C}}"'] & w \arrow[l, phantom, "\ni"] \arrow[d, bend left] \\
{[v]_{\mathcal{B}}} \arrow[r, phantom, "\in"] & K^n \arrow[r, "m_{[T]^{\mathcal{B}}_{\mathcal{C}}}"'] & K^m & {[w]_{\mathcal{C}}} \arrow[l, phantom, "\ni"]
\end{tikzcd} . \tag{3.12}$$

Oder anders gesagt, (3.11) ist äquivalent zu $\Phi_{\mathcal{C}} \circ T = m_{[T]^{\mathcal{B}}_{\mathcal{C}}} \circ \Phi_{\mathcal{B}}$. Dies ist jedoch äquivalent zu

$$m_{[T]^{\mathcal{B}}_{\mathcal{C}}} = \Phi_{\mathcal{C}} \circ T \circ \Phi_{\mathcal{B}}^{-1}. \tag{3.13}$$

Wir haben aber $[T]^{\mathcal{B}}_{\mathcal{C}}$ genau so definiert, dass (3.13) gilt (vgl. Diagramm (3.12)). $\square$

Man sollte darüber so nachdenken: Durch die Beschreibung von V durch $\mathcal{B}$ und W durch $\mathcal{C}$ sieht T wie Multiplikation mit $[T]^{\mathcal{B}}_{\mathcal{C}}$ aus (vgl. Epilog Kapitel 2).

Beispiel 3.43. Falls $\mathcal{E}_n$ die Standard-Basis von K^n, $\mathcal{E}_m$ die Standard-Basis von K^m ist und $A \in M_{m\times n}(K)$, dann gilt allgemeiner

$$[m_A]^{\mathcal{E}_n}_{\mathcal{E}_m} = A.$$

Beispiel 3.44. Betrachten wir noch einmal die Ableitungsabbildung

$$D : K[x]_3 \to K[x]_2$$

und seien $\mathcal{E} = (1, x, x^2, x^3)$ die Standard-Basis von $K[x]_3$ und $\mathcal{E}' = (1, x, x^2)$ die Standard-Basis von $K[x]_2$. Dann gilt

$$\begin{aligned}
D(1) &= 0 = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2 \\
D(x) &= 1 = 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 \\
D(x^2) &= 2x = 0 \cdot 1 + 2 \cdot x + 0 \cdot x^2 \\
D(x^3) &= 3x^2 = 0 \cdot 1 + 0 \cdot x + 3 \cdot x^2
\end{aligned}$$

und daher ist

$$[D]_{\mathcal{E}'}^{\mathcal{E}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Falls man D als einen Endomorphismus von $K[x]_3$ betrachtet, das heisst als eine Abbildung $D : K[x]_3 \to K[x]_3$, dann ist

$$[D]_{\mathcal{E}}^{\mathcal{E}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Übung 3.45. Berechnen Sie $[D]_{\mathcal{C}}^{\mathcal{C}}$, $[D]_{\mathcal{B}}^{\mathcal{C}}$ und $[D]_{\mathcal{C}}^{\mathcal{B}}$, wobei $\mathcal{C}$ und $\mathcal{B}$ wie in Beispiel 3.38 sind.

Die Matrizen $[D]_{\mathcal{A}}^{\mathcal{A}'}$ für verschiedene Basen $\mathcal{A}$ und $\mathcal{A}'$ sind im allgemeinen verschieden, rühren aber alle von der Abbildung D her. Sie sind „verschiedene Beschreibungen“ von D bezüglich verschiedenen Basen. Was haben sie gemeinsam? Wie sind sie miteinander verbunden? Diese Fragen hängen beide mit Frage (2) in (3.7) zusammen.

Für Frage (2) sagen wir jetzt nur, dass die Matrix $[T]_{\mathcal{C}}^{\mathcal{B}}$ stark von der Wahl von $\mathcal{B}$ und $\mathcal{C}$ abhängt, mehr Informationen geben wir später. Wir möchten als nächstes Frage (3) beantworten.

Dazu machen wir aber erst eine Pause, um Matrizen zu besprechen und vor allem um die Matrixmultiplikation einzuführen.

3.3.2 Matrizen

Wie wir später sehen werden, wird uns die Definition der Matrixmultiplikation durch die Antwort auf Frage (3) in (3.7) „gegeben“. Ich finde es nichtsdestotrotz verwirrend die Matrixmultiplikation auf diese Art einzuführen. Deshalb werden wir die Geschichte anders erzählen: Wir werden die Matrixmultiplikation einfach definieren und dann zeigen, dass uns dies eine Antwort auf Frage (3) in (3.7) gibt.

Erinnerung: $M_{m\times n}(K)$ bezeichnet die Menge aller Matrizen mit m-Zeilen und n-Spalten und Einträgen in K. Die Menge $M_{m\times n}(K)$ ist ein Vektorraum mit der Addition

$$(a_{ij})_{\substack{1\le i\le m\\1\le j\le n}} + (b_{ij})_{\substack{1\le i\le m\\1\le j\le n}} := (a_{ij}+b_{ij})_{\substack{1\le i\le m\\1\le j\le n}}$$

und mit der Skalarmultiplikation

$$\alpha \cdot (a_{ij})_{\substack{1\le i\le m\\1\le j\le n}} := (\alpha a_{ij})_{\substack{1\le i\le m\\1\le j\le n}}$$

für $\alpha \in K$.

Definition 3.46. Seien $m, n, p \in \mathbb{N}$. Die Matrixmultiplikation ist die Abbildung

$$\begin{aligned} \cdot : M_{m\times n}(K) \times M_{n\times p}(K) &\to M_{m\times p}(K), \\ (A, B) = \left((a_{ij})_{\substack{1\le i\le m \\ 1\le j\le n}}, (b_{ij})_{\substack{1\le i\le n \\ 1\le j\le p}} \right) &\mapsto A \cdot B, \end{aligned}$$

wobei $A \cdot B = (c_{ij})_{\substack{1\le i\le m \\ 1\le j\le p}}$ mit

$$c_{ij} = a_{i1}b_{1j} + \ldots a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}. \tag{3.14}$$

Wir schreiben oft AB statt $A \cdot B$.

Bemerkung 3.47.

- Merken Sie sich: AB ist nur dann definiert, wenn die Anzahl der Spalten von A der Anzahl Zeilen von B entspricht.
- Die Formel in (3.14) ist nicht sehr übersichtlich, aber von einem theoretischen Blickwinkel aus wichtig. Die folgende Darstellung ist übersichtlicher:

$$\begin{array}{cccc} & & & j\text{-te Spalte} \\ & & & \begin{pmatrix} \cdots & \begin{pmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix} & \cdots \end{pmatrix} \\ & & & \downarrow \\ i\text{-te Zeile} & \begin{pmatrix} \vdots \\ (a_{i1} \ldots a_{in}) \\ \vdots \end{pmatrix} & \to & \begin{pmatrix} & \vdots & \\ \cdots & c_{ij} & \cdots \\ & \vdots & \end{pmatrix} \end{array}$$

Zum Beispiel,

$$\begin{array}{ccc} & & \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix} \\ & & \downarrow \\ \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{pmatrix} & \to & \begin{pmatrix} \square & c_{12} & \square \\ \square & \square & \square \\ \square & \square & c_{33} \\ \square & \square & \square \end{pmatrix} \end{array}.$$

Beispiel 3.48. Seien

$$A = \begin{pmatrix} 3 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 4 & 0 \end{pmatrix}.$$

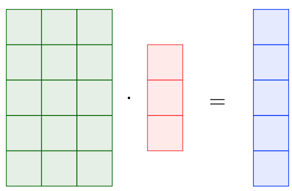
Dann berechnen wir Schritt für Schritt jeden Eintrag:

$$\begin{aligned}
A \cdot B = \begin{pmatrix} 3 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 4 & 0 \end{pmatrix} &= \begin{pmatrix} 3 \cdot 1 + 2 \cdot 0 + 1 \cdot 4 & * \\ * & * \end{pmatrix} = \begin{pmatrix} 7 & * \\ * & * \end{pmatrix} \\
\begin{pmatrix} 3 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 4 & 0 \end{pmatrix} &= \begin{pmatrix} 7 & 3 \cdot 2 + 2 \cdot 1 + 1 \cdot 0 \\ * & * \end{pmatrix} = \begin{pmatrix} 7 & 8 \\ * & * \end{pmatrix} \\
\begin{pmatrix} 3 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 4 & 0 \end{pmatrix} &= \begin{pmatrix} 7 & 8 \\ 1 \cdot 1 + 0 \cdot 1 + 2 \cdot 4 & * \end{pmatrix} = \begin{pmatrix} 7 & 8 \\ 9 & * \end{pmatrix} \\
\begin{pmatrix} 3 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 4 & 0 \end{pmatrix} &= \begin{pmatrix} 7 & 8 \\ 9 & 1 \cdot 2 + 0 \cdot 1 + 2 \cdot 0 \end{pmatrix} = \begin{pmatrix} 7 & 8 \\ 9 & 2 \end{pmatrix}.
\end{aligned}$$

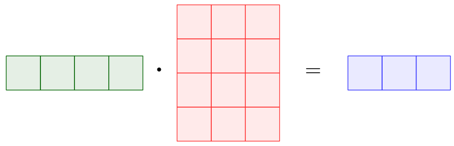
Übung 3.49. In diesem Fall ist auch BA definiert. Berechnen Sie diese Matrix!

Beispiel 3.50 (Spezialfälle und Tricks)**.** Wir betrachten einige Beispiele:

(1) Seien $A \in M_{m\times n}(K)$ und $v \in M_{n\times 1}(K)$. Dann ist $AB \in M_{m\times 1}(K)$. Wenn wir $M_{n\times 1}(K)$ mit $K^n_{\text{Spal}} \cong K^n$ identifizieren, entspricht v einem Spaltenvektor und $Av \in M_{m\times 1}(K) \cong K^m_{\text{Spal}}$. Das Produkt $Av \in K^m_{\text{Spal}}$ ist genau das Produkt, das wir in (2.3) definiert haben.



(2) Ähnlich gilt für $A \in M_{m\times n}(K)$ und $w \in M_{1\times m}(K)$, dass $wA \in M_{1\times n}(K)$. Wenn wir $M_{1\times m}(K)$ mit $K^m_{\text{Zeil}} \cong K^m$ identifizieren, entspricht w einem Zeilenvektor und das Produkt $wA \in K^n_{\text{Zeil}}$ ebenfalls einem Zeilenvektor. Dies definiert ein Produkt zwischen einem Zeilenvektor und einer Matrix.



(3) Wenn man die Spezialfälle (1) und (2) gut beherrscht, dann kann man diese auch für ein allgemeines Produkt AB verwenden: Für

$$A = \begin{pmatrix} - & v_1 & - \\ & \vdots & \\ - & v_m & - \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} | & & | \\ w_1 & \cdots & w_n \\ | & & | \end{pmatrix}$$

gilt

$$AB = A \cdot \begin{pmatrix} | & & | \\ w_1 & \cdots & w_n \\ | & & | \end{pmatrix} = \begin{pmatrix} | & & | \\ Aw_1 & \cdots & Aw_n \\ | & & | \end{pmatrix}$$

beziehungsweise

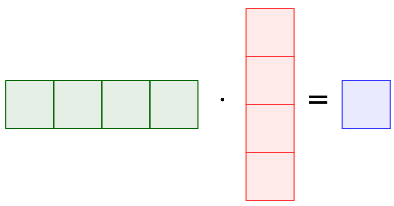
$$AB = \begin{pmatrix} - & v_1 & - \\ & \vdots & \\ - & v_m & - \end{pmatrix} \cdot B = \begin{pmatrix} - & v_1 B & - \\ & \vdots & \\ - & v_m B & - \end{pmatrix}.$$

(4) Seien

$$v = (a_1, \ldots, a_n) \in M_{1\times n}(K) \quad \text{und} \quad w = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in M_{n\times 1}(K).$$

Dann ist $vw \in M_{1\times 1}(K)$. Wenn wir $M_{1\times 1}(K)$ mit K identifizieren, dann heisst dieses Produkt das *Skalarprodukt* von v und w. Dieses Produkt berechnet sich als

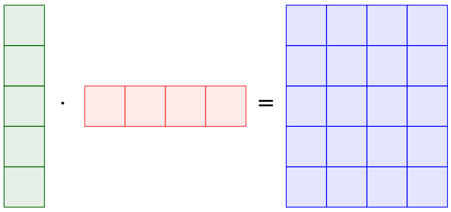
$$vw = (a_1, \ldots, a_n) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + \ldots + a_n b_n.$$



Viel weniger wichtig ist das Produkt $wv \in M_{n\times n}(K)$. Dies berechnet sich als

$$wv = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} (a_1, \ldots, a_n) = \begin{pmatrix} b_1a_1 & \cdots & b_1a_n \\ \vdots & & \vdots \\ b_na_1 & \cdots & b_na_n \end{pmatrix} = (a_ib_j)_{1\leq i,j\leq n}.$$

Beachten Sie, dass wv auch für $w \in M_{m\times 1}(K)$ und $v \in M_{1\times n}(K)$ definiert ist. In diesem Fall ist $wv \in M_{m\times n}(K)$.



Hier ist eine wichtige Matrix und Notation:

Definition 3.51. Wir definieren das *Kronecker-Delta* als

$$\delta_{ij} = \begin{cases} 1, & \text{falls } i = j \\ 0, & \text{falls } i \neq j \end{cases} .$$

Hier gehören i, j normalerweise zu irgendeiner Indexmenge.

Wir definieren die *Einheitsmatrix* I_n für $n \in \mathbb{N}$ als

$$I_n = (\delta_{ij})_{\substack{1\leq i\leq n \\ 1\leq j\leq n}} \in M_{n\times n}(K).$$

Zum Beispiel ist

$$I_1 = (1),\ I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\ I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ usw.}$$

Proposition 3.52. *Es gilt:*

(1) Für $A \in M_{m\times n}(K)$, $B \in M_{n\times p}(K)$ und $C \in M_{p\times q}(K)$ gilt

$$(AB)C = A(BC).$$

Das heisst, dass die Matrixmultiplikation assoziativ ist.

(2) Für $A, B \in M_{m\times n}(K)$ und $C \in M_{n\times p}(K), C' \in M_{q\times m}(K)$ gilt

$$(A+B)C = AC + BC$$
$$C'(A+B) = C'A + C'B.$$

(3) Für alle $A \in M_{m\times n}(K)$ gilt

$$I_m A = A$$
$$AI_n = A.$$

(4) Für alle $\alpha \in K$, $A \in M_{m\times n}(K)$ und $B \in M_{n\times p}(K)$ gilt

$$\alpha(AB) = (\alpha A)B = A(\alpha B).$$

Beweis. All diese Eigenschaften folgen aus einer entsprechenden Berechnung. Dies ist insbesondere richtig für (1), aber die Berechnung ist wirklich nicht angenehm. Nachdem wir eine Verbindung zwischen Matrixmultiplikation und Verkettung herstellen (Lemma 3.61), können wir aus der Tatsache, dass die Verkettung von Funktionen assoziativ ist, herleiten. □

Bemerkung 3.53. Die Menge $M_{n\times n}(K)$ für $n \in \mathbb{N}$ ist laut (1), (2) und (3) in Proposition 3.52 ein Ring mit I_n als Eins. Bemerken Sie, dass dieser Ring für $n \geq 2$ nicht-kommutativ ist, da normalerweise

$$AB \neq BA$$

gilt. Zum Beispiel gilt

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

Unter anderem motiviert Bemerkung 3.53 die folgende Definition:

Definition 3.54. Sei $n \in \mathbb{N}$. Eine Matrix $A \in M_{n\times n}(K)$ heisst *invertierbar*, falls eine Matrix $B \in M_{n\times n}(K)$ existiert, so dass

$$AB = I_n \quad \text{und} \quad BA = I_n$$

gilt. Die Matrix B heisst die *Inverse* von A und wird mit A^{-1} notiert. Die Menge aller invertierbaren Matrizen in $M_{n\times n}(K)$ bezeichnen wir mit

$$\mathrm{GL}_n(K) \quad \text{oder} \quad \mathrm{GL}(n, K)$$

und nennen sie die *allgemeine lineare Gruppe* über K (von Grad n).

Proposition 3.55. *Die allgemeine lineare Gruppe ist eine Gruppe bezüglich Matrixmultiplikation.*

Beweis. Wir müssen zeigen, dass die Matrixmultiplikation eine Verknüpfung auf $\mathrm{GL}_n(K)$ induziert. Das heisst, dass $AB \in \mathrm{GL}_n(K)$ ist für $A, B \in \mathrm{GL}_n(K)$.

Seien also $A, B \in \mathrm{GL}_n(K)$ und $C, D \in M_{n\times n}(K)$ mit

$$\begin{aligned} AC = CA = I_n \\ BD = DB = I_n. \end{aligned}$$

Dann gilt

$$\begin{aligned} (AB)(DC) = A(BD)C = AI_nC = AC = I_n \\ (DC)(AB) = D(CA)B = DI_nB = DB = I_n. \end{aligned}$$

Also folgt, dass $AB \in \mathrm{GL}_n(K)$.

Die Einheitsmatrix $I_n \in \mathrm{GL}_n(K)$, da

$$I_nI_n = I_n$$

und damit ist I_n invertierbar. Des Weiteren ist die Einheitsmatrix das neutrale Element. Laut Proposition 3.52 (1) ist die Matrixmultiplikation assoziativ.

Wenn $A \in \mathrm{GL}_n(K)$ ist, dann ist auch $A^{-1} \in \mathrm{GL}_n(K)$, da A eine Inverse von A^{-1} ist. Daher hat jedes Element $A \in \mathrm{GL}_n(K)$ eine Inverse in $\mathrm{GL}_n(K)$. □

Korollar 3.56. *(1) Die Inverse von A ist eindeutig definiert.*

(2) Es gilt $(A^{-1})^{-1} = A$.

(3) Es gilt $(AB)^{-1} = B^{-1}A^{-1}$.

Beweis. Diese Tatsachen haben wir für alle Gruppen bewiesen. □

Beispiel 3.57 (Inverse und lineare Gleichungssysteme)**.** Bevor wir zu linearen Abbildungen zurückkehren, erklären wir kurz, wieso die Inverse interessant ist.

Seien $A \in M_{n\times n}$ und $x, b \in K^n$, so dass

$$Ax = b. \tag{3.15}$$

Wir nehmen an, dass A invertierbar ist, und dass wir die Inverse A^{-1} kennen. Dann können wir das lineare Gleichungssystem (3.15) in einer Sekunde durch Matrixmultiplikation lösen: Wir multiplizieren (3.15) mit A^{-1} von der linken Seite und erhalten:

$$x = I_n x = (A^{-1}A)x = A^{-1}(Ax) = A^{-1}b.$$

Das heisst, die Lösung von (3.15) ist in diesem Fall $A^{-1}b$!

Korollar 3.58. *Sei $Ax = b$ ein lineares Gleichungssystem mit $A \in M_{n\times n}(K)$. Falls $A \in \mathrm{GL}_n(K)$ ist, dann hat $Ax = b$ für jedes $b \in K^n$ eine eindeutige Lösung x (nämlich $x = A^{-1}b$).*

Bemerkung 3.59. Später werden wir beweisen, dass auch die Umkehrung von Korollar 3.58 gilt.

Für später müssen wir auch noch verstehen wie Transposition und Matrixmultiplikation in Verbindung stehen. Dies ist eine wirklich gute Übung um Matrixmultiplikation zu üben, deshalb lassen wir sie als Übung für die Leser:

Übung 3.60. (1) Sei $A \in M_{m\times n}(K)$ und $B \in M_{n\times p}(K)$. Dann gilt

$$(AB)^T = B^T A^T.$$

(Beachten Sie, dass $A^T B^T$ im Allgemeinen gar nicht definiert ist!)

(2) Sei $A \in \mathrm{GL}_n(K)$. Zeigen Sie, dass

$$A^T \in \mathrm{GL}_n(K)$$

ist und dass

$$(A^T)^{-1} = (A^{-1})^T.$$

3.3.3 Zurück zu Darstellungen linearer Abbildungen

Die Buchstaben V, W und U bezeichnen in diesem Abschnitt immer endlich-dimensionale Vektorräume über K. Das folgende grundlegende Lemma beantwortet Frage (3) in (3.7) und zeigt, dass die Definition der Matrixmultiplikation nicht vom Himmel fällt, sondern einfach von der Darstellung der Verkettung zweier linearer Abbildungen gekommen ist.

Lemma 3.61. *Seien V, W und U drei endlich-dimensionale Vektorräume über K mit jeweiligen Basen $\mathcal{A} = (v_1, \ldots, v_m), \mathcal{B} = (w_1, \ldots, w_n)$ und $\mathcal{C} = (u_1, \ldots, u_p)$. Wir betrachten $T \in \mathrm{Hom}(V, W)$ und $S \in \mathrm{Hom}(W, U)$. Dann gilt*

$$[S \circ T]^{\mathcal{A}}_{\mathcal{C}} = [S]^{\mathcal{B}}_{\mathcal{C}} \cdot [T]^{\mathcal{A}}_{\mathcal{B}},$$

wobei $\cdot$ die Matrixmultiplikation bezeichnet.

Beweis. Der folgende Beweis ist eine sehr gute Indizes-Übung. Seien

$$[T]_{\mathcal{B}}^{\mathcal{A}} = (a_{ij})_{ij}, \quad [S]_{\mathcal{C}}^{\mathcal{B}} = (b_{ij})_{ij} \quad \text{und} \quad [S \circ T]_{\mathcal{C}}^{\mathcal{A}} = (c_{ij})_{ij}$$

(Überlegen Sie sich, welche Werte die Indizes i und j für die jeweiligen Matrizen annehmen.) Diese Matrizen sind durch

$$T(v_j) = a_{1j}w_1 + \ldots a_{nj}w_n = \sum_{k=1}^{n} a_{kj}w_k \quad \text{für } j = 1, \ldots, m, \tag{3.16}$$

$$S(w_j) = b_{1j}u_1 + \ldots + b_{pj}u_p = \sum_{i=1}^{p} b_{ij}u_i \quad \text{für } j = 1, \ldots, n, \tag{3.17}$$

$$(S \circ T)(v_j) = c_{1j}u_1 + \ldots + c_{pj}u_p = \sum_{i=1}^{p} c_{ij}u_i \quad \text{für } j = 1, \ldots, m \tag{3.18}$$

definiert.[5] Aber $(S \circ T)(v_j)$ könnten wir auch durch (3.16) und (3.17) noch anders ausdrücken. Für $j = 1, \ldots, m$ gilt:

$$\begin{aligned}(S \circ T)(v_j) \overset{(3.16)}{=} S\left(\sum_{k=1}^{n} a_{kj}w_k\right) \overset{\text{Linearität}}{=} \sum_{k=1}^{n} a_{kj}S(w_k) \overset{(3.17)}{=} &\sum_{k=1}^{n} a_{kj}\left(\sum_{i=1}^{p} b_{ik}u_i\right)\\ =&\sum_{k=1}^{n}\sum_{i=1}^{p} a_{kj}b_{ik}u_i\\ =&\sum_{i=1}^{p}\sum_{k=1}^{n} b_{ik}a_{kj}u_i.\end{aligned}$$

Daher folgt (Wieso? Vgl. (3.18)), dass

$$c_{ij} = \sum_{k=1}^{n} b_{ik}a_{kj}$$

beziehungsweise $C = BA$, was wir zeigen wollten. □

Dies beantwortet Frage (3) in (3.7), aber auch eine Antwort auf Frage (2) folgt jetzt mittels einem kleinen Trick mit der Identitätsabbildung.

Korollar 3.62. *Sei $T \in \mathrm{Hom}(V, W)$, $\mathcal{B}, \mathcal{B}'$ zwei Basen von V und $\mathcal{C}, \mathcal{C}'$ zwei Basen von W. Dann gilt*

$$[T]_{\mathcal{C}'}^{\mathcal{B}'} = [\mathrm{Id}_W]_{\mathcal{C}'}^{\mathcal{C}}\, [T]_{\mathcal{C}}^{\mathcal{B}}\, [\mathrm{Id}_V]_{\mathcal{B}}^{\mathcal{B}'}. \tag{3.19}$$

[5] Wieso wir die erste Summe mit Indizes k summieren und die restlichen beiden mit i wird später klarer werden.

Beweis. Aus zweifacher Verwendung von Lemma 3.61 folgt, dass die rechte Seite von (3.19)

$$[\mathrm{Id}_W \circ T \circ \mathrm{Id}_V]_{\mathcal{C}'}^{\mathcal{B}'} = [T]_{\mathcal{C}'}^{\mathcal{B}'}.$$

ist. □

Matrizen der Form $[\mathrm{Id}_V]_{\mathcal{B}'}^{\mathcal{B}}$ heissen *Basiswechselmatrizen.* Wir haben noch mehr Information bezüglich solcher Matrizen:

Korollar 3.63. *Seien $\mathcal{B}$ und $\mathcal{B}'$ zwei Basen von V und $n = \dim V$. Dann gilt, dass*

$$[\mathrm{Id}_V]_{\mathcal{B}'}^{\mathcal{B}} \in \mathrm{GL}_n(K).$$

Das heisst, dass $[\mathrm{Id}_V]_{\mathcal{B}'}^{\mathcal{B}}$ invertierbar ist. Ausserdem gilt

$$\left([\mathrm{Id}_V]_{\mathcal{B}'}^{\mathcal{B}}\right)^{-1} = [\mathrm{Id}_V]_{\mathcal{B}}^{\mathcal{B}'}.$$

Beweis. Laut Lemma 3.61 gilt

$$\begin{aligned}
[\mathrm{Id}_V]_{\mathcal{B}'}^{\mathcal{B}} \cdot [\mathrm{Id}_V]_{\mathcal{B}}^{\mathcal{B}'} &= [\mathrm{Id}_V \circ \mathrm{Id}_V]_{\mathcal{B}'}^{\mathcal{B}'} = [\mathrm{Id}_V]_{\mathcal{B}'}^{\mathcal{B}'} \\
[\mathrm{Id}_V]_{\mathcal{B}}^{\mathcal{B}'} \cdot [\mathrm{Id}_V]_{\mathcal{B}'}^{\mathcal{B}} &= [\mathrm{Id}_V \circ \mathrm{Id}_V]_{\mathcal{B}}^{\mathcal{B}} = [\mathrm{Id}_V]_{\mathcal{B}}^{\mathcal{B}}.
\end{aligned}$$

Es folgt aber aus der Definition, dass

$$[\mathrm{Id}_V]_{\mathcal{C}}^{\mathcal{C}} = I_n$$

für jede Basis $\mathcal{C}$ von V (Wieso?). Das Korollar folgt. □

Unter Benutzung der vorherigen Korollare erhalten wir folgendes Resultat für Endomorphismen:

Korollar 3.64. *Sei V ein endlich-dimensionaler Vektorraum mit zwei Basen $\mathcal{B}$ und $\mathcal{B}'$ und $T \in \mathrm{End}(V)$. Dann gilt*

$$[T]_{\mathcal{B}'}^{\mathcal{B}'} = [\mathrm{Id}_V]_{\mathcal{B}'}^{\mathcal{B}} \, [T]_{\mathcal{B}}^{\mathcal{B}} \, [\mathrm{Id}_V]_{\mathcal{B}}^{\mathcal{B}'} = ([\mathrm{Id}_V]_{\mathcal{B}}^{\mathcal{B}'})^{-1} \, [T]_{\mathcal{B}}^{\mathcal{B}} \, [\mathrm{Id}_V]_{\mathcal{B}}^{\mathcal{B}'}. \tag{3.20}$$

Beweis. (Wieso? Benutzen Sie die Korollare 3.62 und 3.63.) □

Bemerkung. Für $T \in \mathrm{End}(V)$ schreibt man manchmal $[T]_{\mathcal{B}} := [T]_{\mathcal{B}}^{\mathcal{B}}$.

Wir können uns nun die Assoziativität der Matrixmultiplikation neu anschauen (vgl. Proposition 3.52 (i)).

Korollar 3.65. *Seien $A \in M_{m\times n}(K)$ und $B \in M_{n\times p}(K)$, dann gilt*

$$m_A \circ m_B = m_{AB}.$$

Beweis. Sei $\mathcal{E}_n$ die Standard-Basis von K^n. Dann gilt laut Beispiel 3.43 und Lemma 3.61, dass
$$[m_A \circ m_B]_{\mathcal{E}_m}^{\mathcal{E}_p} = [m_A]_{\mathcal{E}_m}^{\mathcal{E}_n}[m_B]_{\mathcal{E}_n}^{\mathcal{E}_p} = AB.$$
Aus Proposition 3.42 (und Beispiel 3.43 mit $n = 1$) folgt, dass für jedes $v \in K^p$
$$m_A \circ m_B(v) = [m_A \circ m_B(v)]_{\mathcal{E}_m} = [m_A \circ m_B]_{\mathcal{E}_m}^{\mathcal{E}_p}[v]_{\mathcal{E}_p} = AB[v]_{\mathcal{E}_p} = (AB)v$$
gilt. Dies scheint als eine Art Tautologie, weil Assoziativität, Lemma 3.61 und Proposition 3.42 in der Tat eng miteinander verbunden sind. Also gilt $m_A \circ m_B = m_{AB}$ als Abbildungen von K^p nach K^m. □

Korollar 3.66. *Matrixmultiplikation ist assoziativ.*

Beweis. Seien $A \in M_{m\times n}(K)$, $B \in M_{n\times p}(K)$, $C \in M_{p\times q}(K)$ und sei $\mathcal{E}_l$ die Standard-Basis von K^l. Da Verkettung von Funktionen assoziativ ist, gilt
$$\begin{aligned}(AB)C = [m_{AB}]_{\mathcal{E}_m}^{\mathcal{E}_p}[m_C]_{\mathcal{E}_p}^{\mathcal{E}_q} = [m_A \circ m_B]_{\mathcal{E}_m}^{\mathcal{E}_p}[m_C]_{\mathcal{E}_p}^{\mathcal{E}_q} &= [m_A \circ m_B \circ m_C]_{\mathcal{E}_m}^{\mathcal{E}_q}\\ &= [m_A]_{\mathcal{E}_m}^{\mathcal{E}_n}[m_B \circ m_C]_{\mathcal{E}_n}^{\mathcal{E}_q}\\ &= [m_A]_{\mathcal{E}_m}^{\mathcal{E}_n}[m_{BC}]_{\mathcal{E}_n}^{\mathcal{E}_q} = A(BC).\end{aligned}$$
Das Korollar folgt. □

Gleichungen (3.19) und (3.20) motivieren die folgende Definition:

Definition 3.67.

- Zwei Matrizen $A, B \in M_{m\times n}(K)$ heissen *äquivalent*, falls es $P \in \mathrm{GL}_m(K)$ und $Q \in \mathrm{GL}_n(K)$ gibt, so dass
$$PAQ = B.$$
- Zwei Matrizen $A, B \in M_{n\times n}(K)$ heissen *ähnlich*, falls es $P \in \mathrm{GL}_n(K)$ gibt, so dass
$$P^{-1}AP = B.$$

Übung 3.68. Zeigen Sie, folgende Aussagen:

- Äquivalenz ist eine Äquivalenzrelation auf $M_{m\times n}(K)$.
- Ähnlichkeit ist eine Äquivalenzrelation auf $M_{n\times n}(K)$.
- A ist äquivalent zu B genau dann, wenn es Vektorräume V, W, $T \in \mathrm{Hom}(V, W)$ und Basen $\mathcal{B}, \mathcal{B}' \subseteq V$, $\mathcal{C}, \mathcal{C}' \subseteq W$ gibt, so dass
$$A = [T]_{\mathcal{C}}^{\mathcal{B}} \quad \text{und} \quad B = [T]_{\mathcal{C}'}^{\mathcal{B}'}.$$

- A ist ähnlich zu B genau dann, wenn es einen Vektorraum V, $T \in \mathrm{End}(V)$ und Basen $\mathcal{B}, \mathcal{B}' \subseteq V$ gibt, so dass
$$A = [T]_{\mathcal{B}} \quad \text{und} \quad B = [T]_{\mathcal{B}'}.$$

Um den Namen „Ähnlichkeit“ zu rechtfertigen werden wir sehen, dass ähnliche Matrizen tatsächlich vieles gemeinsam haben. Einen einfachen Repräsentanten für jede Äquivalenzklasse bezüglich Ähnlichkeit auf $M_{n\times n}(K)$ zu finden ist jedoch schwierig und wird uns noch in der Linearen Algebra II beschäftigen. Wir können hingegen problemlos für die Äquivalenzrelation „Äquivalenz“ auf $M_{m\times n}(K)$ einfache Repräsentanten für die Äquivalenzklassen finden.

Sei $D_r = (d_{ij})_{ij} \in M_{m\times n}(K)$ die Matrix mit
$$\begin{cases} d_{ij} = 0, \text{ falls } i \neq j \\ d_{ii} = 1, \text{ falls } 1 \leq i \leq r \\ d_{ii} = 0, \text{ falls } r < i \end{cases}.$$

Das heisst, dass
$$D_r = \begin{pmatrix} I_r & 0_{r,n-r} \\ 0_{m-r,r} & 0_{m-r,n-r} \end{pmatrix} \in M_{m\times n}(K),$$
wobei $0_{l,k} \in M_{l\times k}(K)$ die Nullmatrix ist.

Proposition 3.69. *Sei $T \in \mathrm{Hom}(V,W)$ mit*[6] $\mathrm{Rang}(T) = r$. *Dann existieren Basen $\mathcal{B}$ von V und $\mathcal{C}$ von W, so dass*
$$[T]_{\mathcal{C}}^{\mathcal{B}} = D_r.$$

Beweis. Im Beweis des Rangsatzes 3.27 haben wir eine Basis $\mathcal{B} = (w_1, \ldots, w_r, u_1, \ldots, u_k)$ von V gefunden, so dass
$$Tu_1 = \ldots = Tu_k = 0$$
und $\mathcal{C}' := (Tw_1, \ldots, Tw_r)$ eine Basis von $\mathrm{Im}(T)$ ist. Sei $\mathcal{C}$ eine Ergänzung von $\mathcal{C}'$ zu einer Basis von W. Bezüglich $\mathcal{B}$ und $\mathcal{C}$ gilt
$$[T]_{\mathcal{C}}^{\mathcal{B}} = D_r.$$
(Wieso?) □

Lemma 3.70. *Sei $A \in M_{m\times n}(K)$. Dann gilt*
$$\mathrm{SR}(A) = \mathrm{Im}(m_A).$$

[6] Erinnerung: Für $T \in \mathrm{Hom}(V,W)$ ist $\mathrm{Rang}(T) = \dim \mathrm{Im}(T)$.

Beweis. Dies folgt aus Lemma 3.18 und Lemma 3.26. (Wieso?) □

Proposition 3.71. *Sei* $A \in M_{m\times n}(K)$. *Dann existieren* $P \in \mathrm{GL}_m(K)$ *und* $Q \in \mathrm{GL}_n(K)$ *mit*

$$PAQ = D_r = \begin{pmatrix} I_r & 0_{r,n-r} \\ 0_{m-r,r} & 0_{m-r,n-r} \end{pmatrix},$$

wobei $r =$ *Spaltenrang*(A).

Beweis von Proposition 3.71. Aus Beispiel 3.43 wissen wir, dass

$$[m_A]^{\mathcal{E}_n}_{\mathcal{E}_m} = A,$$

wobei $\mathcal{E}_n, \mathcal{E}_m$ die entsprechenden Standard-Basen sind von K^n und K^m sind. Laut Proposition 3.69 existieren Basen $\mathcal{B}$ von K^n und $\mathcal{C}$ von K^m, mit

$$[\mathrm{Id}_{K^m}]^{\mathcal{E}_m}_{\mathcal{C}} \, [m_A]^{\mathcal{E}_n}_{\mathcal{E}_m} \, [\mathrm{Id}_{K^n}]^{\mathcal{B}}_{\mathcal{E}_n} = D_r.$$

Wir setzen

$$P := [\mathrm{Id}_{K^m}]^{\mathcal{E}_m}_{\mathcal{C}} \quad \text{und} \quad Q := [\mathrm{Id}_{K^n}]^{\mathcal{B}}_{\mathcal{E}_n},$$

was zeigt, dass

$$PAQ = D_r.$$

□

Korollar 3.72. *Es gibt* $\min(m,n)+1$ *Äquivalenzklassen bezüglich der Äquivalenzrelation Äquivalenz auf* $M_{m\times n}(K)$ *mit* D_r *für* $r = 0, \ldots, \min(n,m)$ *als Repräsentanten.*

Einige Autoren (Vgl. Fischer [6, S. 161]) nennen diese Repräsentanten D_r Normalformen. Dies ist nicht so geläufig und man sollte dies nicht mit der sogenannten Jordan-Normalform verwechseln (diese werden wir in der Linearen Algebra II kennenlernen und ist verbunden mit Repräsentanten von Äquivalenzklassen bezüglich Ähnlichkeit).

Bemerkung 3.73. In Kürze werden wir beweisen, dass Spaltenrang(A) = Zeilenrang(A) (vgl. Korollar 2.98). Daher könnte man $r =$ Spaltenrang(A) = Zeilenrang(A) in Proposition 3.71 schreiben. Damit könnte man Korollar 3.72 auch so ausdrücken: Zwei Matrizen sind genau dann äquivalent, wenn sie denselben Rang haben.

Korollar 3.74. *Sei* $T : V \to W$ *ein Isomorphismus und* $\mathcal{B} \subseteq V$, $\mathcal{C} \subseteq W$ *zwei Basen. Dann gilt*

$$[T^{-1}]^{\mathcal{C}}_{\mathcal{B}} = \left([T]^{\mathcal{B}}_{\mathcal{C}}\right)^{-1}.$$

Insbesondere ist $[T]^{\mathcal{B}}_{\mathcal{C}}$ *invertierbar.*

Beweis. Sei $n = \dim V = \dim W$. Laut Lemma 3.61 gilt

$$[T^{-1}]^{\mathcal{C}}_{\mathcal{B}}\,[T]^{\mathcal{B}}_{\mathcal{C}} = [T^{-1} \circ T]^{\mathcal{B}}_{\mathcal{B}} = [\mathrm{Id}_V]^{\mathcal{B}}_{\mathcal{B}} = I_n$$
$$[T]^{\mathcal{B}}_{\mathcal{C}}\,[T^{-1}]^{\mathcal{C}}_{\mathcal{B}} = [T \circ T^{-1}]^{\mathcal{C}}_{\mathcal{C}} = [\mathrm{Id}_W]^{\mathcal{C}}_{\mathcal{C}} = I_n.$$

Das Korollar folgt. □

3.4 Spaltenrang ist gleich Zeilenrang

Satz 3.75. *Sei $A \in M_{m\times n}(K)$. Dann gilt*

$$\textit{Spaltenrang}(A) = \textit{Zeilenrang}(A).$$

Dank diesem Satz können wir definieren, dass

$$\mathrm{Rang}(A) = \mathrm{Spaltenrang}(A) = \mathrm{Zeilenrang}(A).$$

Laut Abschnitt 2.3.3 stimmt diese Definition mit der Definition als die Anzahl der Pivots in einer Stufenform Matrix, die zeilenäquivalent ist zu A überein.

Für den Beweis von Theorem 3.75 brauchen wir einige Lemmas. Erinnern Sie sich an Definition 3.33: $\mathrm{Rang}(T) = \dim \mathrm{Im}(T)$.

Lemma 3.76. *Sei $T \in \mathrm{Hom}(V, W)$ und seien $S : W \to U$ und $L : P \to V$ zwei Isomorphismen. Dann gilt*

- $\mathrm{Rang}(S \circ T) = \mathrm{Rang}(T)$.
- $\mathrm{Rang}(T \circ L) = \mathrm{Rang}(T)$.

Insbesondere gilt auch, dass $\mathrm{Rang}(S \circ T \circ L) = \mathrm{Rang}(T)$ ist.

Beweis. Dies ist ein Spezialfall von Übung 3.34, die Sie in der Serie haben. □

Lemma 3.77. *Sei $A \in M_{n\times n}(K)$. Dann gilt, dass $A \in \mathrm{GL}_n(K)$ genau dann, wenn $m_A : K^n \to K^n$ ein Isomorphismus ist.*

Beweis. „$\Rightarrow$": Die lineare Abbildung $m_{A^{-1}} : K^n \to K^n$ ist die Inverse von m_A.

„$\Leftarrow$": Aus Korollar 3.74 folgt, dass

$$A = [m_A]^{\mathcal{E}_n}_{\mathcal{E}_n} \tag{3.21}$$

invertierbar ist, wobei $\mathcal{E}_n$ die Standard-Basis ist. □

Lemma 3.78. *Seien $A, B \in M_{m\times n}(K)$ mit*

$$B = PAQ, \tag{3.22}$$

wobei $P \in \mathrm{GL}_m(K)$ und $Q \in \mathrm{GL}_n(K)$. Dann gilt

(1) Spaltenrang(A) = Spaltenrang(B).

(2) Zeilenrang(A) = Zeilenrang(B).

Beweis. Betrachten wir alle Matrizen im Lemma als lineare Abbildungen, ist (3.22) äquivalent zu

$$m_B = m_P \circ m_A \circ m_Q$$

und laut Lemma 3.76 gilt

$$\mathrm{Rang}(m_A) = \mathrm{Rang}(m_B),$$

da m_Q und m_P laut Lemma 3.77 Isomorphismen sind. Wir haben aber gesehen (Lemma 3.70), dass

$$\mathrm{Rang}(m_C) = \mathrm{Spaltenrang}(C)$$

für jede Matrix C ist. Daher gilt

$$\mathrm{Spaltenrang}(A) = \mathrm{Spaltenrang}(B).$$

Dies beweist (1). Für (2) transponieren wir (3.22) mit Hilfe von Übung 3.60:

$$B^T = Q^T A^T P^T.$$

Aus Übung 3.60 wissen wir auch, dass $Q^T \in \mathrm{GL}_n(K)$ und $P^T \in \mathrm{GL}_m(K)$ und daher folgt aus (1), dass

$$\mathrm{Spaltenrang}(A^T) = \mathrm{Spaltenrang}(B^T),$$

was äquivalent zu

$$\mathrm{Zeilenrang}(A) = \mathrm{Zeilenrang}(B)$$

ist. □

Jetzt folgt Satz 3.75 aus Proposition 3.71:

Beweis von Satz 3.75. Wegen Proposition 3.71 existieren $P \in \mathrm{GL}_n(K)$ und $Q \in \mathrm{GL}_n(K)$, so dass

$$PAQ = D_r = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} =: B.$$

Für B gilt Spaltenrang(B) = Zeilenrang$(B) = r$. Laut Lemma 3.78 gelten

$$\text{Spaltenrang}(A) = \text{Spaltenrang}(B) \quad \text{und} \quad \text{Zeilenrang}(A) = \text{Zeilenrang}(B).$$

Daher folgt, dass Spaltenrang(A) = Zeilenrang(A). □

Bemerkung 3.79. Ich habe diesen Beweis aus dem Buch von Fischer [6, Kap. 2.6.6] gelernt. Beachten Sie, dass wir in 2.102 (noch) einen anderen Beweis skizziert haben. Falls Sie dies wünschen, kann ich in der Nachbesprechungszeit einen Beweis mittels Gauss und dem Rangsatz ebenfalls skizzieren.

3.5 Hom(V, W) als Vektorraum

Wir werden jetzt sehen, dass Hom(V, W) ein Vektorraum ist, und dass wir dies eigentlich schon wissen (zumindest im Fall, wo V und W endlich-dimensional sind).

Seien V und W zwei Vektorräume über einem Körper K. Jetzt betrachten wir Hom(V, W) als eine Menge. Betrachten Sie die folgenden Operationen auf Hom(V, W):

Seien $T, S \in \text{Hom}(V, W)$. Wir definieren Addition auf Hom(V, W):

$$T + S : V \to W$$

durch

$$(T + S)(v) := T(v) +_W S(v).$$

Sei $\alpha \in K$ und $T \in \text{Hom}(V, W)$. Wir definieren Skalarmultiplikation auf Hom(V, W):

$$\alpha T : V \to W$$

durch

$$(\alpha T)(v) := \alpha \cdot_W T(v).$$

Lemma 3.80. *Für $T, S \in \text{Hom}(V, W)$ und $\alpha \in K$ gilt, dass $T + S \in \text{Hom}(V, W)$ und $\alpha T \in \text{Hom}(V, W)$.*

Beweis. Seien $a, b \in K$ und $v_1, v_2 \in V$. Dann ist

$$\begin{aligned}(T + S)(av_1 + bv_2) &= T(av_1 + bv_2) + S(av_1 + bv_2)\\ &= (aTv_1 + bTv_2) + (aSv_1 + bSv_2)\\ &= a(Tv_1 + Sv_1) + b(Tv_2 + Sv_2)\\ &= a((T + S)(v_1)) + b((T + S)(v_2)).\end{aligned}$$

Wir überlassen es den Lesern die analoge Überprüfung für αT zu machen. □

Proposition 3.81. Hom(V, W) *mit der obigen Addition und Skalarmultiplikation ist ein Vektorraum.*

Beweis. Wir überlassen diesen Beweis den Lesern und bemerken nur, dass der Nullvektor in Hom(V, W) die lineare Abbildung

$$\begin{aligned} 0 : V &\to W, \\ v &\mapsto 0_W \end{aligned}$$

ist. □

Für endlich-dimensionale Vektorräume V und W „kennen" wir eigentlich diesen Vektorraum schon:

Satz 3.82. *Seien V und W zwei Vektorräume mit* $\dim V = n$ *und* $\dim W = m$ *mit jeweiligen Basen $\mathcal{B} = (v_1, \dots, v_n) \subseteq V$ und $\mathcal{C} = (w_1, \dots, w_m) \subseteq W$. Die Abbildung*

$$\begin{aligned} \psi_{\mathcal{C}}^{\mathcal{B}} : \mathrm{Hom}(V, W) &\to M_{m\times n}(K) \\ T &\mapsto [T]_{\mathcal{C}}^{\mathcal{B}} \end{aligned}$$

ist ein Isomorphismus. Das heisst, sie ist linear und bijektiv.

Beweis. Wir zeigen zuerst, dass $\psi_{\mathcal{C}}^{\mathcal{B}}$ linear ist: Per Definition gilt

$$[T+S]_{\mathcal{C}}^{\mathcal{B}} = \begin{bmatrix} | & & | \\ [T+S(v_1)]_{\mathcal{C}} & \cdots & [T+S(v_n)]_{\mathcal{C}} \\ | & & | \end{bmatrix},$$

wobei $\mathcal{B} = (v_1, \dots, v_n)$. Es gilt aber für jedes $i = 1, \dots, n$, dass

$$[T+S(v_i)]_{\mathcal{C}} = [Tv_i + Sv_i]_{\mathcal{C}} = [Tv_i]_{\mathcal{C}} + [Sv_i]_{\mathcal{C}},$$

da $v \mapsto [v]_{\mathcal{C}}$ linear ist. Daher gilt

$$\begin{aligned} &\begin{bmatrix} | & & | \\ [T+S(v_1)]_{\mathcal{C}} & \cdots & [T+S(v_n)]_{\mathcal{C}} \\ | & & | \end{bmatrix} \\ &\quad = \begin{bmatrix} | & & | \\ [Tv_1]_{\mathcal{C}} & \cdots & [Tv_n]_{\mathcal{C}} \\ | & & | \end{bmatrix} + \begin{bmatrix} | & & | \\ [Sv_1]_{\mathcal{C}} & \cdots & [Sv_n]_{\mathcal{C}} \\ | & & | \end{bmatrix} = [T]_{\mathcal{C}}^{\mathcal{B}} + [S]_{\mathcal{C}}^{\mathcal{B}}. \end{aligned}$$

Die Überprüfung der Homogenität ist sehr ähnlich und daher überlassen wir dies den Lesern.

Des Weiteren ist $\psi_{\mathcal{C}}^{\mathcal{B}}$ injektiv: Falls $[T]_{\mathcal{C}}^{\mathcal{B}} = 0$ ist, dann ist $T(v_i) = 0_W$ für alle $i = 1, \dots n$ und daher ist $T = 0$.

Wir müssen noch zeigen, dass $\psi_{\mathcal{C}}^{\mathcal{B}}$ surjektiv ist: Sei $A = (a_{ij}) \in M_{m\times n}(K)$. Mittels Satz 3.15 definieren wir T durch

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i.$$

Dann gilt $[T]_{\mathcal{C}}^{\mathcal{B}} = A$. (Wieso? Vgl. (3.10).) □

Korollar 3.83. *Falls V und W endlich-dimensionale Vektorräume sind, dann gilt*

$$\dim \operatorname{Hom}(V, W) = \dim V \cdot \dim W.$$

Bemerkung 3.84 (Nebenbemerkung)*.* Wie die Notation schon andeutet, hängt der Isomorphismus $\psi_{\mathcal{C}}^{\mathcal{B}}$ von der Wahl der Basen $\mathcal{B}$ und $\mathcal{C}$ ab. In diesem Sinne ist $\psi_{\mathcal{C}}^{\mathcal{B}}$ „nicht kanonisch“.

Korollar 3.85. *Der Vektorraum* $\operatorname{End}(V) = \operatorname{Hom}(V, V)$ *ist ein Ring mit Verkettung als Multiplikation und* Id_V *als Einselement. Falls V endlich-dimensional ist (mit Dimension n) und $\mathcal{B} \subseteq V$ eine Basis ist, dann ist*

$$\begin{aligned} \psi_{\mathcal{B}}^{\mathcal{B}} : \operatorname{End}(V) &\to M_{n\times n}(K) \\ T &\mapsto [T]_{\mathcal{B}} := [T]_{\mathcal{B}}^{\mathcal{B}}. \end{aligned}$$

ein Ring-Isomorphismus.

Beweis. Man überprüft die Ring-Axiome direkt, zum Beispiel:

Seien $v \in V$ und $T_1, T_2, T_3 \in \operatorname{End}(V)$. Dann gilt

$$T_1 \circ (T_2 + T_3)(v) = T_1(T_2 v + T_3 v) \overset{T_1 \text{linear}}{=} T_1 T_2 v + T_1 T_3 v = (T_1 \circ T_2 + T_1 \circ T_3)(v).$$

Daher gilt $T_1(T_2 + T_3) = T_1 T_2 + T_1 T_3$.

Sei jetzt $n = \dim V < \infty$. Die Tatsache, dass $\psi_{\mathcal{B}}^{\mathcal{B}}$ ein Ring-Isomorphismus ist, folgt aus Satz 3.82 und Lemma 3.61. Beachten Sie, dass $[\mathrm{Id}_V]_{\mathcal{B}} = I_n$ und für $0 \in \operatorname{End}(V)$ ist $[0]_{\mathcal{B}} = 0 \in M_{n\times n}(K)$, unabhängig davon wie $\mathcal{B}$ gewählt wurde.[7] □

[7] Man könnte daher sagen, dass die Koordinaten der Identitätsabbildung und der Nullabbildung nicht von der Wahl einer Basis abhängen, also kanonisch sind.

3.6 Die Inverse, Elementarmatrizen und lineare Gleichungssysteme

Dieser Abschnitt ist eine Art Zusammenfassung von verschiedenen Ergebnissen, die wir bisher gesehen haben.

Erinnern Sie sich an die Abbildung

$$\Phi : K^k \to \text{Lös}(A, 0),$$

wobei $A \in M_{m\times n}(K)$ ist und $k = n - \text{Rang}(A)$, die wir im Fischer [6, Kap. 0.4] besprochen haben. Die Abbildung Φ hat die Form (hier ist $\lambda = (\lambda_1, \dots, \lambda_k) \in K^k$)

$$\Phi(\lambda) = \Phi(\lambda_1, \dots, \lambda_k) = (x_1(\lambda), \dots, x_r(\lambda), \lambda_1, \dots, \lambda_k)$$

bis auf Umordnung der Variablen, wobei $r = \text{Rang}(A)$. Wenn wir die Struktur der Abbildung betrachten, sehen wir, dass $x_i(\lambda)$ linear ist in $\lambda = (\lambda_1, \dots, \lambda_k)$ und wir können daraus schliessen, dass Φ eine bijektive lineare Abbildung ist. Dies zeigt:

Korollar 3.86. *Sei $A \in M_{m\times n}(K)$. Dann ist*

$$U = \textit{Lös}(A, 0) = \text{Ker}(A) = \text{Ker}(m_A)$$

ein Untervektorraum von K^n von Dimension $n - \text{Rang}(A)$. Die Abbildung Φ ist ein expliziter Isomorphismus von $K^{n-\text{Rang}(A)}$ nach U.

Korollar 3.87. *Sei $b \in K^m$ und $A \in M_{m\times n}(K)$. Es gilt*

(1) $\textit{Lös}(A, b) \neq \emptyset$ genau dann, wenn $b \in \text{SR}(A)$.

(2) Falls $\textit{Lös}(A, b) \neq \emptyset$ und $y \in \textit{Lös}(A, b)$ ist, dann ist

$$\textit{Lös}(A, b) = y + \textit{Lös}(A, 0) := \{y + x \mid x \in \textit{Lös}(A, 0)\}. \tag{3.23}$$

Beweis. (1) Dies folgt aus dem „Trick“ aus Lemma 2.50:

$$Ax = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 \cdot \begin{pmatrix} | \\ v_1 \\ | \end{pmatrix} + x_2 \cdot \begin{pmatrix} | \\ v_2 \\ | \end{pmatrix} + \dots + x_n \cdot \begin{pmatrix} | \\ v_n \\ | \end{pmatrix}.$$

(2) Dies folgt aus der Linearität von m_A: $A(x + y) = Ax + Ay$. (Wieso?)

□

Mengen der Form (3.23) heissen *affine Unterräume*. Wir werden mehr dazu in Abschnitt 3.7 sagen. Wir besprechen jetzt zwei klassische Ergebnisse „in der Sprache der linearen Gleichungssysteme“ und beweisen sie mit unserer neu entwickelten Theorie.

Proposition 3.88. *Sei $A \in M_{m\times n}(K)$ und $b \in K^m (= K^m_{Spal})$. Dann sind folgende Aussagen äquivalent:*

(1) Es gilt $\mathrm{Rang}(A) = \mathrm{Rang}(A|b)$.

(2) Die Gleichung $Ax = b$ hat eine Lösung.

Beweis. (1) ist äquivalent zu $\mathrm{Spaltenrang}(A) = \mathrm{Spaltenrang}(A|b)$, was äquivalent ist zu $\mathrm{Sp}(v_1, \ldots, v_n) = \mathrm{Sp}(v_1, \ldots, v_n, b)$, wobei $v_1, \ldots, v_n$ die Spalten von A sind. Dies wiederum ist äquivalent zu $b \in \mathrm{Sp}(v_1, \ldots, v_n) = \mathrm{SR}(A)$ (Wieso?). Wegen Korollar 3.87 (1) ist dies äquivalent zu (2). □

Beispiel 3.89. Die Matrix $\begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$ hat Rang 2. Die Matrix

$$\begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix} = \left(\begin{array}{cc|c} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{array}\right)$$

hat auch Rang 2 und daher hat

$$x + 4y = 7 \tag{3.24}$$
$$2x + 5y = 8 \tag{3.25}$$
$$3x + 6y = 9 \tag{3.26}$$

eine Lösung. In der Tat ist $x = -1$, $y = 2$ eine Lösung.

Beispiel 3.90. Wenn A eine Matrix in Stufenform ist mit Rang r, dann sieht $(A \mid b)$ so aus:

$$\left(\begin{array}{cccc|c} 1 & * & \cdots & * & b_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & & \ddots & * & \\ & & & 1 & b_r \\ \hline 0 & & \cdots & & b_{r+1} \\ \vdots & & & & \vdots \\ 0 & & \cdots & & b_m \end{array}\right),$$

wobei $*$ nicht Nulleinträge sind. Hier sieht man, dass $\mathrm{Rang}(A \mid b) = \mathrm{Rang}(A)$ genau dann, wenn $b_{r+1} = \ldots = b_m = 0$.

Ein ähnlicher Beweis zeigt:

Proposition 3.91. *Sei $A \in M_{m\times n}(K)$ und $b \in K^m (= K^m_{Spal})$. Dann sind folgende Aussagen äquivalent:*

(1) Es gilt $\mathrm{Rang}(A) = \mathrm{Rang}(A|b) = n$.

(2) Die Gleichung $Ax = b$ hat eine eindeutige Lösung.

Beweis. Sei

$$A = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix}.$$

Da $\mathrm{Rang}(A) = \mathrm{Spaltenrang}(A)$ ist, gilt:

$$\text{Die Vektoren } v_1, \dots, v_n \text{ bilden eine Basis von } \mathrm{SR}(A) \iff \mathrm{Rang}(A) = n. \tag{3.27}$$

Des Weiteren ist wie zuvor

$$\mathrm{Rang}(A|b) = \mathrm{Rang}(A) \iff b \in \mathrm{SR}(A). \tag{3.28}$$

Daher ist (1) äquivalent zu $b \in \mathrm{SR}(A)$ und $\{v_1, \dots, v_n\}$ ist eine Basis von $\mathrm{SR}(A)$, was genau äquivalent ist zu (2) (laut Proposition 2.68). □

Beispiel 3.92. Zurück zu Beispiel 3.89: Da $A := \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \in M_{3\times 2}(\mathbb{Q})$ und $\mathrm{Rang}(A) = 2$ ist die Lösung (3.24) eindeutig.

Proposition 3.93. *Sei $A \in M_{n\times n}(K)$. Dann sind folgende Aussagen äquivalent:*

(1) A ist invertierbar.

(2) $m_A : K^n \to K^n$ ist ein Isomorphismus.

(3) $m_A : K^n \to K^n$ ist injektiv.

(4) $m_A : K^n \to K^n$ ist surjektiv.

(5) $\mathrm{Rang}(A) = n$.

(6) Spaltenrang$(A) = n$.

(7) Zeilenrang$(A) = n$.

(8) Die Spalten von A sind eine Basis von K^n.

(9) Die Spalten von A erzeugen K^n.

(10) Die Spalten von A sind linear unabhängig.

(11) $Ax = 0$ hat nur die triviale Lösung $x = 0 \in K^n$.

(12) $Ax = b$ hat für alle $b \in K^n$ eine eindeutige Lösung.

(13) A^T ist invertierbar.

(14) $\mathrm{Rang}(A^T) = n$.

(15) Die Zeilen von A sind eine Basis.

(16) Die Zeilen von A erzeugen K^n.

(17) Die Zeilen von A sind linear unabhängig.

(18) A^{-1} ist invertierbar.

(19) A ist zeilenäquivalent zu I_n. Das heisst, dass man A mittels Zeilenoperationen zu I_n überführen kann.

(20) A ist spaltenäquivalent zu I_n. Das heisst, dass man A mittels Spaltenoperationen zu I_n überführen kann.

Beweis. (1) $\iff$ (2): Dies entspricht Lemma 3.77.

(2), (3) und (4) sind äquivalent wegen Korollar 3.29 (Korollar zum Rangsatz).

(4) $\iff$ (6): Dies gilt, weil $\mathrm{SR}(A) = \mathrm{Sp}(m_A(e_1), \dots, m_A(e_n))$.

(5), (6) und (7) sind äquivalent wegen Satz 3.75.

(6) $\iff$ (8): Dies gilt wegen Satz 2.85 (Gleichgewicht).

(8), (9) und (10) sind wiederum wegen Satz 2.85 (Gleichgewicht) äquivalent.

(10) $\iff$ (11): Dies folgt aus dem Multiplikationstrick in Lemma 2.50.

(12) $\iff$ (8): Dies gilt wiederum wegen dem Multiplikationstrick und Proposition 2.68.

(1) $\iff$ (13): Dies gilt wegen Übung 3.60 (für „$\Rightarrow$“), und Übung 3.60 zusammen mit $(A^T)^T = A$ (für „$\Leftarrow$“).

Die Äquivalenzen von (13), (14), (15), (16) und (17) folgen aus den Äquivalenzen von (1), (5), (8), (9) und (10) angewandt auf A^T.

(1) $\iff$ (18): Dies folgt aus Korollar 3.56 (2).

Dies zeigt die Äquivalenzen (1)-(18).

Wir zeigen die Implikation (5) $\Rightarrow$ (19) sorgfältig: Mit Gauss'scher Elimination haben wir gezeigt, dass A zeilenäquivalent ist zu einer Matrix der Form

$$B = \begin{pmatrix} 1 & * & \cdots & * \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & * \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Beachten Sie, dass es keine Nullen auf der Diagonalen gibt, da $\text{Rang}(A) = n$ laut (5).

Bemerken Sie für später auch, dass wir bis jetzt immer Zeilenoperationen der Form $L_i \to L_i + \alpha L_j$ nur mit $j < i$ benutzt haben, was wir „reinigen nach unten“ genannt haben. Wir können aber auch „nach oben reinigen“: Mit Zeilenoperationen der Form $L_i + \alpha L_j \to L_i$, wobei $j > i$, können wir $\mathcal{B}$ nach I_n überführen, was (19) zeigt.

(14) $\Rightarrow$ (20): Dies folgt, wenn wir die Implikation (5) $\Rightarrow$ (19) für A^T benutzen und aus der Tatsachen, dass sich Zeilenoperationen auf A^T wie Spaltenoperationen auf A verhalten.

Wir zeigen jetzt (19) $\Rightarrow$ (7): Sei A eine Matrix, die zeilenäquivalent zu I_n ist. Laut Lemma 2.93 ist $\text{ZR}(A) = \text{ZR}(I_n) = K^n$, was insbesondere (7) zeigt. Mit Transposition zeigt dies auch, dass (20) $\Rightarrow$ (6), was den Beweis dieser Proposition beendet! $\square$

Später werden wir einen anderen Beweis von (19) $\Rightarrow$ (1) geben, der sogar einen Algorithmus für die Berechnung der Inverse gibt. Dafür führen wir Elementarmatrizen im nächsten Abschnitt ein.

Bemerkung 3.94. Ein Hauptziel des nächsten Kapitels wird sein die Matrix A mit einem Skalar $\det A$ zu versehen, so dass die Bedingungen in Proposition 3.93 äquivalent sind zur Aussage $\det A \neq 0$. Der Skalar $\det A$ wird *Determinante* von A genannt.

Übung 3.95. Falls Sie nicht warten können, dann können Sie versuchen zu zeigen, dass

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc \tag{3.29}$$

genau die gewünschte Bedingung für 2×2 Matrizen in Bemerkung 3.94 erfüllt.

Ein anderes Ziel des nächsten Kapitels wird es sein ein Formel für die Inverse einer Matrix zu geben.

Übung 3.96. Wiederum, wenn Sie nicht warten können, dann können Sie zeigen: Falls $\det A \neq 0$, dann ist

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

wobei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ und $\det A$ wie in (3.29) gegeben ist.

3.6.1 Elementarmatrizen und Zeilen- und Spaltenoperationen

Wir geben jetzt einen anderen Beweis der Tatsache:

$$A \text{ ist zeilenäquivalent zu } I_n \Rightarrow A \text{ ist invertierbar}$$

(Siehe Proposition 3.93).

Dafür definieren wir eine spezielle Famile von Matrizen, sogenannte Elementarmatrizen, welche uns erlauben werden Zeilen- und Spaltenoperationen als Matrixmultiplikationen aufzufassen.

Erinnern Sie sich daran, dass wir mit E_{ij} die Matrix bezeichnet, die überall Nulleinträge hat ausser beim ij-ten Eintrag, welcher 1 ist.

Definition 3.97. Sei $n \in \mathbb{N}$. Wir definieren[8] die folgenden quadratischen Matrizen:

- Für $1 \leq i \neq j \leq n$ und $\alpha \in K$ definieren wir

$$Q_{i,j}(\alpha) := I_n + \alpha E_{ij} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} + \alpha \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$
$$= \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \alpha & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

- Für $1 \leq i, j \leq n$ definieren wir

$$P_{i,j} = I_n - E_{ii} - E_{jj} + E_{ij} + E_{ji}.$$

Wir nennen P_{ij} eine *Permutationsmatrix.*

[8] Leider gibt es keine Standard-Notation dieser Matrizen

- Für $1 \leq i \leq n$ und $\alpha \in K^\times$ definieren wir

$$S_i(\alpha) = \begin{pmatrix} 1 & 0 & \cdots & & & \cdots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & 1 & \ddots & & & \\ & & 0 & \alpha & 0 & & \\ & & & \ddots & 1 & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \cdots & & & \cdots & 0 & 1 \end{pmatrix},$$

  wobei α im ii-ten Eintrag steht. Matrizen dieser Form heissen *Elementarmatrizen.*

Einige Autoren nennen $Q_{i,j}(\alpha)$ Elementarmatrizen von *Typ I*, $P_{i,j}$ von *Typ II* und $S_i(\alpha)$ von *Typ III.*

Beispiel 3.98. In $M_{4\times 4}(K)$ ist

$$Q_{1,3}(\alpha) = \begin{pmatrix} 1 & 0 & \alpha & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad Q_{3,1}(\beta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \beta & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$P_{1,2} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad P_{2,4} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$S_3(\alpha) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Lemma 3.99. *Sei $A \in M_{n\times p}(K)$. Es gilt*

- *Linksmultiplikation mit $Q_{i,j}(\alpha)$ entspricht der Zeilenoperation $L_i + \alpha L_j \to L_i$. Das heisst*
$$A \xrightarrow{L_i+\alpha L_j \to L_i} Q_{i,j}(\alpha)A.$$

- *Linksmultiplikation mit $P_{i,j}$ entspricht der Zeilenoperation $L_i \leftrightarrow L_j$. Das heisst*
$$A \xrightarrow{L_i \leftrightarrow L_j} P_{i,j}A.$$

- *Linksmultiplikation mit $S_i(\alpha)$ entspricht der Zeilenoperation $\alpha L_i \to L_i$. Das heisst*

$$A \xrightarrow{\alpha L_i \to L_i} S_i(\alpha)A.$$

Ähnlich gilt für Spaltenoperationen[9]*: Sei $A \in M_{m\times n}(K)$, dann gilt:*

- *Rechtsmultiplikation mit $Q_{i,j}(\alpha)$ entspricht der Spaltenoperation $C_j + \alpha C_i \to C_j$. Das heisst*
$$A \xrightarrow{C_j+\alpha C_i \to C_j} AQ_{i,j}(\alpha).$$

- *Rechtsmultiplikation mit $P_{i,j}$ entspricht der Spaltenoperation $C_i \leftrightarrow C_j$. Das heisst*

$$A \xrightarrow{C_i \leftrightarrow C_j} AP_{i,j}.$$

- *Rechtsmultiplikation mit $S_i(\alpha)$ entspricht der Spaltenoperation $\alpha C_i \to C_i$. Das heisst*
$$A \xrightarrow{\alpha C_i \to C_i} AS_i(\alpha).$$

Beweis. All diese Aussagen sind lediglich eine Rechnung. □

Lemma 3.100. *Es gilt:*

- *Für alle $\alpha \in K$ ist*
$$(Q_{i,j}(\alpha))^{-1} = Q_{i,j}(-\alpha).$$

- *Die Inverse von $P_{i,j}$ ist gegeben durch $(P_{i,j})^{-1} = P_{i,j}$.*

- *Für alle $\alpha \in K^\times$ ist*
$$(S_i(\alpha))^{-1} = S_i\left(\frac{1}{\alpha}\right).$$

Insbesondere ist jede Elementarmatrix invertierbar und ihre Inverse ist auch elementar.

Beweis. Dies ist wieder eine Berechnung. Diesmal können wir dies jedoch mit Lemma 3.99 sehen. Zum Beispiel entspricht

$$Q_{i,j}(\alpha)^{-1} = Q_{i,j}(-\alpha)$$

der Tatsache, dass $L_i - \alpha L_j \to L_i$ die Inverse Operation zu $L_i + \alpha L_j \to L_i$ ist. Ähnlich entspricht
$$(P_{i,j})^{-1} = P_{j,i}$$
der Tatsache, dass $L_i \leftrightarrow L_j$ die Inverse Operation zu $L_i \leftrightarrow L_j$ ist und so weiter. □

[9] Streng genommen haben wir nur Zeilenoperationen definiert, aber die Definition der Spaltenoperationen ist ganz analog.

Satz 3.101. *Jede Matrix $A \in \mathrm{GL}_n(K)$ ist ein endliches Produkt von Elementarmatrizen.*

Beweis. Im Beweis von Proposition 3.93 haben wir gesehen, dass jede invertierbare Matrix zeilenäquivalent ist zu I_n. Das heisst, dass man A mit endlich vielen Zeilenoperationen nach I_n überführen kann.

Wenn wir das mit Lemma 3.99 betrachten, entsprechen diese Zeilenoperationen Linksmultiplikation mit Elementarmatrizen. Daher existieren Elementarmatrizen $T_1, \ldots, T_k$ mit

$$T_k \cdots T_2 T_1 A = I_n. \tag{3.30}$$

Wir multiplizieren (3.30) mit

$$(T_k \cdots T_1)^{-1} = T_1^{-1} \cdots T_{k-1}^{-1} T_k^{-1}$$

und sehen, dass

$$A = T_1^{-1} \cdots T_k^{-1}.$$

Laut Lemma 3.100 sind alle T_i^{-1} auch elementar und das Lemma folgt. □

Bemerkung 3.102. In der Sprache der Gruppentheorie besagt Lemma 3.100, dass $\mathrm{GL}_n(K)$ durch Elementarmatrizen erzeugt ist.

Korollar 3.103. *Mit der Notation des Beweises von Satz 3.101 ist $A^{-1} = T_k \cdots T_2 T_1$.*

Diese Tatsache ermöglicht uns ein einfaches Verfahren zum Bestimmen der Inversen einer Matrix. Sei $A \in M_{n\times n}(K)$. Wir möchten wissen, ob A invertierbar ist und falls ja, möchten wir auch A^{-1} berechnen können.

<u>Hier ist der Trick:</u> Für $B, C \in M_{n\times n}(K)$ schreiben wir $(B \mid C) \in M_{n\times 2n}(K)$ für die zusammengefügte Matrix.

Lemma 3.104. *Seien $B, C, D \in M_{n\times n}(K)$. Es gilt*

$$B(C \mid D) = (BC \mid BD).$$

Beweis. Noch einmal eine direkte Berechnung! □

Dies können wir benutzen, um Folgendes zu zeigen:

Satz 3.105. *Seien $n \in \mathbb{N}$ und $A \in M_{n\times n}(\mathbb{R})$. Man versucht die $n\times 2n$ zusammengefügte Matrix $(A \mid I_n)$ mittels Zeilenoperationen zur Form $(I_n \mid B)$ zu bringen. Wenn dies möglich ist, dann ist A invertierbar und es gilt $A^{-1} = B$. Falls dies nicht möglich ist, dann ist A nicht invertierbar.*

Beweis. Beachten Sie, dass $(A \mid I_n)$ zeilenäquivalent zu $(I_n \mid B)$ ist genau dann, wenn A zeilenäquivalent zu I_n ist. Dies ist laut Proposition 3.93 äquivalent zur Invertierbarkeit von A. Dies zeigt die zweite Aussage. Nehmen wir also an, dass A zeilenäquivalent ist zu I_n. Laut Lemma 3.100 existieren Elementarmatrizen $T_1, \ldots, T_k$ mit

$$T_k \cdots T_2 T_1 A = I_n$$

(vgl. Gleichung (3.30)). Erinnern Sie sich daran, dass wir die Zeilenoperationen auf $(A \mid I_n)$ anwenden. Laut Lemma 3.104 entspricht das dem Produkt

$$T_k \cdots T_1 (A \mid I_n) \overset{3.104}{=} (T_k \cdots T_1 A \mid T_k \cdots T_1 I_n) = (I_n \mid T_k \cdots T_1) \overset{3.103}{=} (I_n \mid A^{-1}),$$

was genau der ersten Aussage entspricht. □

Einige Beispiele

Hier geben wir einige Beispiele für verschiedene Teile dieses Kapitels.

Beispiel 3.106 (Injektivität und Surjektivität)**.** Beachten Sie, dass wir Korollar 3.29 nicht benutzen können, wenn die Vektorräume nicht endlich-dimensional sind: Die Abbildung

$$\begin{aligned} m_{x^{73}} : K[x] &\to K[x] \\ p &\mapsto x^{73} p \end{aligned}$$

ist injektiv aber nicht surjektiv. Ähnlich ist

$$\begin{aligned} D : K[x] &\to K[x] \\ p &\mapsto p' \end{aligned}$$

surjektiv aber nicht injektiv.

Beispiel 3.107. Angenommen Sie betrachten $T \in \operatorname{End}(\mathbb{Q}[x]_2)$ und wollen T bezüglich der Basis

$$\mathcal{C} = (1 + 2x + 2x^2, 2 + 4x + x^2, x)$$

berechnen (vielleicht wegen einer bestimmten Anwendung). Normalerweise wird T bezüglich der Standard-Basis $\mathcal{E} = (1, x, x^2)$ gegeben sein. Also ist es eine direkte Rechnung $[T]_{\mathcal{E}}$ zu berechnen. Statt nun $[T]_{\mathcal{C}}$ direkt zu berechnen (wodurch wir viele lineare Gleichungssysteme lösen müssten), benutzen wir die Transformationsformel

$$[T]_{\mathcal{C}} = [\mathrm{Id}]^{\mathcal{E}}_{\mathcal{C}} [T]^{\mathcal{E}}_{\mathcal{E}} [\mathrm{Id}]^{\mathcal{C}}_{\mathcal{E}},$$

wobei $\mathrm{Id} = \mathrm{Id}_{\mathbb{Q}[x]_2}$. Beachten Sie, dass es sehr einfach ist $[\mathrm{Id}]_{\mathcal{E}}^{\mathcal{C}}$ zu berechnen. Es gilt

$$[\mathrm{Id}]_{\mathcal{E}}^{\mathcal{C}} = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 4 & 1 \\ 2 & 1 & 0 \end{pmatrix}.$$

Die Matrix $[\mathrm{Id}]_{\mathcal{C}}^{\mathcal{E}}$ zu berechnen ist hingegen nicht sehr schön. Es ist viel einfach sich daran zu erinnern, dass

$$[\mathrm{Id}]_{\mathcal{C}}^{\mathcal{E}} = ([\mathrm{Id}]_{\mathcal{E}}^{\mathcal{C}})^{-1}$$

und Gauss-Jordan zu benutzen. Wir suchen also die Inverse von

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 4 & 1 \\ 2 & 1 & 0 \end{pmatrix}.$$

Wir betrachten $(A \mid I) \in M_{3\times 6}(\mathbb{R})$ und führen die Gauss'sche Elimination durch:

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 2 & 4 & 1 & 0 & 1 & 0 \\ 2 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow[L_3 - 2L_1 \to L_3]{L_2 - 2L_1 \to L_2} \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & -3 & 0 & -2 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{L_3 \to L_2} \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & -3 & 0 & -2 & 0 & 1 \\ 0 & 0 & 1 & -2 & 1 & 0 \end{array}\right)$$

$$\xrightarrow[-\frac{1}{3}L_2 \to L_2]{L_1 + \frac{2}{3}L_2 \to L_1} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{3} & 0 & \frac{2}{3} \\ 0 & 1 & 0 & \frac{2}{3} & 0 & -\frac{1}{3} \\ 0 & 0 & 1 & -2 & 1 & 0 \end{array}\right)$$

Somit ist die Inverse von A

$$A^{-1} = \begin{pmatrix} -\frac{1}{3} & 0 & \frac{2}{3} \\ \frac{2}{3} & 0 & -\frac{1}{3} \\ -2 & 1 & 0 \end{pmatrix}.$$

Es gilt also

$$[T]_{\mathcal{C}} = [\mathrm{Id}]_{\mathcal{C}}^{\mathcal{E}}[T]_{\mathcal{E}}^{\mathcal{E}}[\mathrm{Id}]_{\mathcal{E}}^{\mathcal{C}} = \begin{pmatrix} -\frac{1}{3} & 0 & \frac{2}{3} \\ \frac{2}{3} & 0 & -\frac{1}{3} \\ -2 & 1 & 0 \end{pmatrix} [T]_{\mathcal{E}}^{\mathcal{E}} \begin{pmatrix} 1 & 2 & 0 \\ 2 & 4 & 1 \\ 2 & 1 & 0 \end{pmatrix}.$$

Beispiel 3.108. Wir betrachten den Vektorraum $\mathbf{Fib} \subseteq K^\infty$. Die Folgen

$$F_{1,0} = (1, 0, 1, 1, 2, 3, \ldots)$$
$$F_{0,1} = (0, 1, 1, 2, 3, \ldots)$$

bilden eine Basis $\mathcal{B} = (F_{1,0}, F_{0,1})$ von $\mathbf{Fib}$. Wir haben ein beliebiges Element von $\mathbf{Fib}$ als

$$F_{a,b} = (a, b, a + b, \ldots)$$

notiert. Es gilt

$$[F_{a,b}]_\mathcal{B} = \begin{pmatrix} a \\ b \end{pmatrix},$$

da $F_{a,b} = aF_{1,0} + bF_{0,1}$. Wir haben die Basis $\mathcal{C} = (F_{1,\varphi}, F_{1,\psi})$ gefunden, wobei

$$\varphi = \frac{1+\sqrt{5}}{2} \quad \text{und} \quad \psi = \frac{1-\sqrt{5}}{2}.$$

Wie wir in der Einleitung erklärt haben, kann man diese Basis finden, wenn man die „Symmetrie“ der Verschiebungsabbildung

$$S : \mathbf{Fib} \to \mathbf{Fib}$$
$$(a, b, a + b, a + 2b, \ldots) \mapsto (b, a + b, a + 2b, \ldots)$$

betrachtet. Wir werden dies noch rigoros erklären, wenn wir Eigenwerte und Eigenvektoren besprechen. Bezüglich $\mathcal{B}$ sieht S so aus

$$[S]_\mathcal{B} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

Proposition 3.42 besagt, dass

$$[S]_\mathcal{B}[F_{a,b}]_\mathcal{B} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} b \\ a+b \end{pmatrix} = [S(F_{a,b})]_\mathcal{B}.$$

Daraus folgt[10]: Falls $F = (a_0, a_1, a_2, \ldots) \in \mathbf{Fib}$, dann ist

$$([S]_\mathcal{B})^n[F]_\mathcal{B} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix}.$$

Wenn man die Potenzen $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n$ berechnen kann, dann findet man eine Formel für

[10] man könnte es auch direkt sehen.

a_n, die nur von a_0 und a_1 abhängt. Aber die Potenzen von $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ zu berechnen ist schwierig. Man könnte sich fragen, ob es eine ähnliche Matrix zu $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ gibt, deren Potenzen leicht zu berechnen sind. Ähnliche Matrizen zu $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ entsprechen $[S]_{\mathcal{D}}$ für andere Basen $\mathcal{D}$. Daher könnte man sich fragen: Gibt es eine Basis, auf deren Elemente die Abbildung S besonders einfach operiert? In der Linearen Algebra II werden wir lernen wie man solche Basen findet. Für S haben wir eine solche Basis jedoch schon gefunden: Bezüglich $\mathcal{C}$ sieht S so aus:

$$[S]_{\mathcal{C}} = \begin{pmatrix} \varphi & 0 \\ 0 & \psi \end{pmatrix},$$

da $S(F_{1,\varphi}) = \varphi F_{1,\varphi}$ und $S(F_{1,\psi}) = \psi F_{1,\psi}$. Laut der Transformationsformel 3.62 sind $[S]_{\mathcal{B}}$ und $[S]_{\mathcal{C}}$ ähnlich:

$$[S]_{\mathcal{B}} = [\mathrm{Id}_{\mathbf{Fib}}]^{\mathcal{C}}_{\mathcal{B}}[S]_{\mathcal{C}}[\mathrm{Id}_{\mathbf{Fib}}]^{\mathcal{B}}_{\mathcal{C}}.$$

Wir setzen $P := [\mathrm{Id}_{\mathbf{Fib}}]^{\mathcal{B}}_{\mathcal{C}}$. Dann gilt

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = P^{-1} \begin{pmatrix} \varphi & 0 \\ 0 & \psi \end{pmatrix} P. \tag{3.31}$$

Beachten Sie: Potenzen von $[S]_{\mathcal{C}} = \begin{pmatrix} \varphi & 0 \\ 0 & \psi \end{pmatrix}$ sind einfach zu berechnen:

$$\begin{pmatrix} \varphi & 0 \\ 0 & \psi \end{pmatrix}^n = \begin{pmatrix} \varphi^n & 0 \\ 0 & \psi^n \end{pmatrix}.$$

Aus (3.31) folgt

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n = P^{-1} \begin{pmatrix} \varphi & 0 \\ 0 & \psi \end{pmatrix}^n P = P^{-1} \begin{pmatrix} \varphi^n & 0 \\ 0 & \psi^n \end{pmatrix} P$$

(Wieso?). Beachten Sie, dass P und P^{-1} ohne Potenzen erscheinen. Dies gibt uns eine Motivation, um P und P^{-1} zu berechnen. Eine davon ist leicht zu berechnen und die andere bereitet ein bisschen Mühe. In der Tat ist

$$P^{-1} = [\mathrm{Id}_{\mathbf{Fib}}]^{\mathcal{C}}_{\mathcal{B}}$$

einfacher zu berechnen. Da $F_{1,\varphi} = 1 \cdot F_{1,0} + \varphi \cdot F_{0,1}$ und $F_{1,\psi} = 1 \cdot F_{1,0} + \psi \cdot F_{0,1}$ ist

$$P^{-1} = \begin{pmatrix} 1 & 1 \\ \varphi & \psi \end{pmatrix}.$$

Da $P = (P^{-1})^{-1} = \begin{pmatrix} 1 & 1 \\ \varphi & \psi \end{pmatrix}^{-1}$ ist, könnte man unsere obige Methode benutzen, um

$$P = \frac{1}{\psi - \varphi} \begin{pmatrix} \psi & -1 \\ -\varphi & 1 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} -\psi & 1 \\ \varphi & -1 \end{pmatrix}.$$

zu berechnen. Daher gilt

$$\begin{aligned} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n &= \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 1 \\ \varphi & \psi \end{pmatrix} \begin{pmatrix} \varphi^n & 0 \\ 0 & \psi^n \end{pmatrix} \begin{pmatrix} -\psi & 1 \\ \varphi & -1 \end{pmatrix} \\ &= \frac{1}{\sqrt{5}} \begin{pmatrix} \varphi^n & \psi^n \\ \varphi^{n+1} & \psi^{n+1} \end{pmatrix} \begin{pmatrix} -\psi & 1 \\ \varphi & -1 \end{pmatrix} \\ &= \frac{1}{\sqrt{5}} \begin{pmatrix} \varphi\psi^n - \psi\varphi^n & \varphi^n - \psi^n \\ \varphi\psi^{n+1} - \psi\varphi^{n+1} & \varphi^{n+1} - \psi^{n+1} \end{pmatrix}. \end{aligned}$$

Daraus folgt beispielsweise das Folgende:

$$F_{0,1} = (b_0, b_1, b_2, \ldots)$$

Dann ist

$$\begin{pmatrix} b_n \\ b_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} \varphi^n - \psi^n \\ \varphi^{n+1} - \psi^{n+1} \end{pmatrix}$$

oder

$$b_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1 + \sqrt{5}}{2} \right)^n - \left(\frac{1 - \sqrt{5}}{2} \right)^n \right).$$

Übung 3.109. Finden Sie eine Formel für a_n, wobei $F_{a,b} = (a_0, a_1, a_2, \ldots)$, indem Sie

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n \begin{pmatrix} a \\ b \end{pmatrix}$$

berechnen.

3.7 Fasern und Quotientenräume

3.7.1 Fasern

Wir fangen mit einer (neuen) Notation an. Für zwei Teilmengen $A, B \subseteq V$ eines Vektorraums V definieren wir

$$A + B := \{a + b \mid a \in A,\, b \in B\}.$$

Wir schreiben auch für $a \in V$ und $B \subseteq V$

$$a + B := \{a\} + B.$$

Proposition 3.110. *Sei $T : V \to W$ eine lineare Abbildung zwischen zwei Vektorräumen V und W. Dann gilt:*

(1) Falls $w \in \mathrm{Im}(T)$, dann ist $T^{-1}(w) = v + \mathrm{Ker}(T)$, wobei v ein beliebiger Vektor ist mit $Tv = w$.

(2) Falls $w \notin \mathrm{Im}(T)$, dann gilt $T^{-1}(w) = \emptyset$.

Beweis. (1) Sei $v \in V$ mit $Tv = w$. Wir zeigen, dass $T^{-1}(w) = v + \mathrm{Ker}(T)$ ist, indem wir zwei Inklusionen zeigen.

„$\supseteq$": Sei $u \in \mathrm{Ker}(T)$. Dann gilt

$$T(v + u) \overset{T \text{ linear}}{=} Tv + Tu = w + 0 = w.$$

Also gilt, dass $T^{-1}(w) \supseteq v + \mathrm{Ker}(T)$.

„$\subseteq$": Sei $v' \in T^{-1}(w)$. Es ist genug zu zeigen, dass $v' - v \in \mathrm{Ker}(T)$. (Wieso?) Da T linear ist, gilt

$$T(v - v') = Tv - Tv' = w - w = 0.$$

Daher folt $T^{-1}(w) \subseteq v + \mathrm{Ker}(T)$.

(2) Hier gibt es nichts zu beweisen! Dies ist einfach die Definition von $\mathrm{Im}(T)$.

□

Die Menge $T^{-1}(w)$ heisst die *Faser* von w bezüglich T.

Beispiel 3.111. Wir betrachten die lineare Abbildung

$$\begin{aligned} T : \mathbb{R}^2 &\to \mathbb{R} \\ (x, y) &\mapsto x - y. \end{aligned}$$

Die Abbildung T ist surjektiv und für $z \in \mathbb{R}$ gilt beispielsweise, dass

$$T(z,0) = z$$

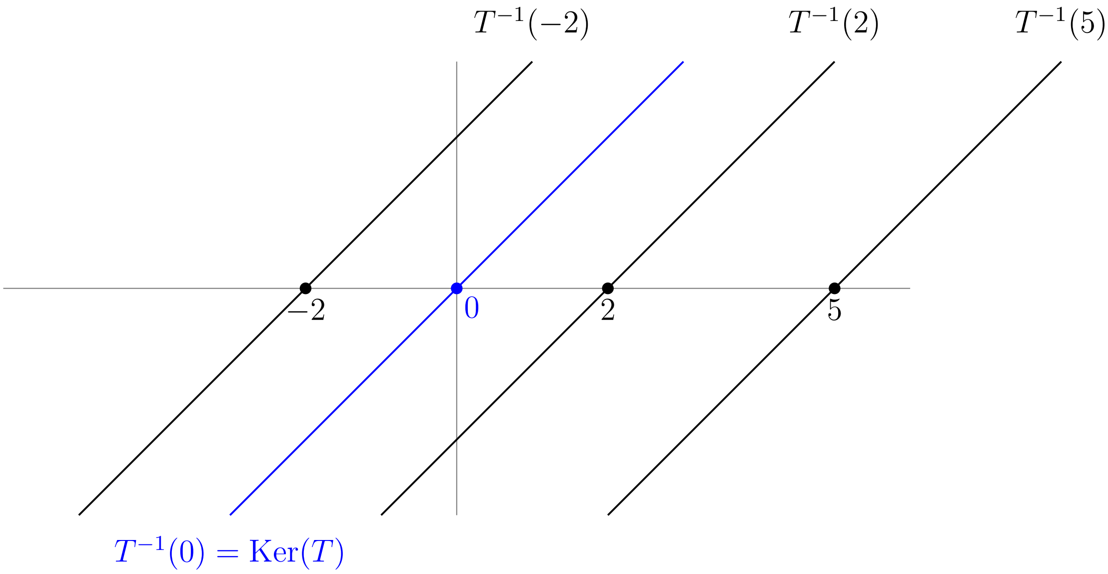
und daher ist

$$T^{-1}(z) = (z,0) + \mathrm{Ker}(T).$$

Der Kern $\mathrm{Ker}(T)$ berechnet sich als

$$\{(x,y) \mid T(x,y) = 0\} = \{(x,y) \mid x - y = 0\} = \{(x,y) \mid x = y\}.$$

Also sehen die Fasern von T so aus:



Wir möchten einen neuen Vektorraum aus den Fasern bilden. In anderen Worten möchten wir Fasern miteinander addieren. Wir betrachten noch einmal eine beliebige lineare Abbildung $T : V \to W$ zwischen zwei Vektorräumen V und W. Wegen Proposition 3.110 wissen wir, dass alle nicht leeren Fasern die Form $v + \mathrm{Ker}(T)$ haben. Da $\mathrm{Ker}(T)$ ein Untervektorraum von V ist, erlaubt uns die nachfolgende Konstruktion, Fasern zu addieren, und ist auch etwas allgemeiner.

3.7.2 Quotientenraum

Sei V ein Vektorraum über K und $U \subseteq V$ ein Untervektorraum[11]. Wir definieren eine Relation $\sim_U$ auf V durch

$$v_1 \sim_U v_2 \;:\Longleftrightarrow\; v_1 - v_2 \in U.$$

Ohne viele Worte können wir zwei einfache Lemmata beweisen.

[11]Die Leser können an $U = \mathrm{Ker}(T)$ denken.

Lemma 3.112. *Die Relation $\sim_U$ ist eine Äquivalenzrelation.*

Beweis. Seien $v, v_1, v_2, v_3 \in V$.

Reflexion: Es gilt $v \sim_U v$, da $v - v = 0 \in U$.

Symmetrie: Falls $v_1 \sim_U v_2$, dann ist $v_1 - v_2 \in U$. Dann folgt auch

$$(v_2 - v_1) = -(v_1 - v_2) \in U.$$

Also ist $v_2 \sim_U v_1$.

Transitivität: Seien $v_1 \sim_U v_2$ und $v_2 \sim_U v_3$. Dann ist $v_1 - v_2 \in U$ und $v_2 - v_3 \in U$. Daher ist $v_1 - v_3 = (v_1 - v_2) + (v_2 - v_3) \in U$. Dies zeigt $v_1 \sim_U v_3$. □

Lemma 3.113 (Äquivalenzklasse)**.** *Für jedes $v \in V$ gilt*

$$[v]_{\sim_U} = v + U.$$

Beweis. Es gelten folgende Äquivalenzen:

$$\begin{aligned} v' \in [v]_{\sim_U} &\iff v' - v \in U \\ &\iff \exists\, u \in U : v' = v + u \\ &\iff v' \in v + U. \end{aligned}$$

□

Bemerkung 3.114. Im Fall $U = \operatorname{Ker}(T)$ sind die Äquivalenzklassen genau die nicht leeren Fasern. Die entsprechende Partition

$$V = \bigsqcup_{v \in V} [v]_{\sim_{\operatorname{Ker}(T)}} = \bigsqcup_{v \in V} v + \operatorname{Ker}(T)$$

von V heisst in diesem Fall eine *Faserung* von V durch T.

Beispiel 3.115. Für $U = \operatorname{Ker}(T)$ mit T wie in Beispiel 3.111 sind alle Fasern/Äquivalenzklassen Geraden mit Steigung 1.

Definition 3.116. Mengen der Form $A = v + U$ heissen *affine Unterräume* von V. Die *Dimension* $\dim A$ von A ist definiert als $\dim U$ und man sagt, dass A *parallel* zu U ist. Der Untervektorraum U nennen wir den Untervektorraum *assoziiert* zum affinen Unterraum A.

In Beispiel 3.115 sind alle Geraden affine Unterräume, die zu $U = \operatorname{Ker}(T)$ assoziiert sind.

Wir möchten zeigen, dass die Menge aller affinen Unterräume, welche zu einem gegebenen Unterraum U assoziiert sind, einen Vektorraum bilden.

Definition 3.117. Mit der vorherigen Konstruktion, nennen wir die Menge

$$V/U := V/\sim_U = \{v + U \mid v \in V\}$$

den *Quotientenraum* oder den *Faktorraum* von V nach U.

Wie der Name des Quotientenraums (oder des Faktorraums) schon andeutet, wollen wir eine Vektorraum-Struktur auf V/U definieren.

Definition der Addition auf V/U:

Seien $v_1 + U, v_2 + U \in V/U$. Wir definieren die Addition $+ : V/U \times V/U \to V/U$ durch

$$(v_1 + U) + (v_2 + U) := v_1 + v_2 + U.$$

Lemma 3.118. *Die zuvor definierte Addition ist wohl-definiert.*

Beweis. Seien $v_1, v_1', v_2, v_2' \in V$ mit $v_1 + U = v_1' + U$ und $v_2 + U = v_2' + U$. Um zu zeigen, dass die Addition wohl-definiert ist, müssen wir zeigen, dass

$$v_1 + v_2 + U = v_1' + v_2' + U.$$

Dies ist äquivalent zu

$$v_1 + v_2 - (v_1' + v_2') \in U \quad \text{bzw.} \quad (v_1 - v_1') + (v_2 - v_2') \in U.$$

Letzteres ist jedoch wahr, da nach Annahme $v_1 - v_1'$, $v_2 - v_2' \in U$, und U ein Untervektorraum ist. □

Definition der Skalarmultiplikation auf V/U:

Wir definieren die Skalarmultiplikation auf V/U durch

$$\begin{aligned} \cdot : K \times V/U &\to V/U \\ (\alpha, v + U) &\mapsto \alpha v + U. \end{aligned}$$

Übung 3.119. Zeigen Sie, dass die Skalarmultiplikation wohl-definiert ist.

Man könnte jetzt überprüfen, dass die Axiome eines Vektorraums für V/U gelten. (Dies ist ziemlich leicht, da alles von den entsprechenden Axiomen für V und den Untervektorraum-Axiomen für U folgt.) Zum Protokoll:

Proposition 3.120. *Mit der obigen Addition und Skalarmultiplikation ist V/U ein Vektorraum über K.*

Beispiel 3.121. Hier sind zwei Beispiele in $\mathbb{R}^3$:

- Sei U eine Gerade, die durch den Ursprung geht. Dann ist $\mathbb{R}^3/U$ die Menge aller Geraden, die zu U parallel sind.
- Sei W eine Ebene in $\mathbb{R}^3$, die durch den Ursprung geht. Dann ist $\mathbb{R}^3/W$ die Menge aller Ebenen, die parallel sind zu W.

Proposition 3.122. *Sei $\pi_U : V \to V/U$ die Abbildung*

$$\pi_U(v) = v + U.$$

Dann ist π_U linear, $\mathrm{Ker}(\pi_U) = U$ *und* $\mathrm{Im}(\pi_U) = V/U$.

Bemerkung 3.123. Die Abbildung π_U heisst die *kanonische Quotientenabbildung.*

Beweis von Proposition 3.122. Wir müssen drei Dinge überprüfen.
Linearität: Seien $v_1, v_2 \in V$, dann gilt

$$\pi_U(v_1 + v_2) = v_1 + v_2 + U = (v_1 + U) + (v_2 + U) = \pi_U(v_1) + \pi_U(v_2),$$

wobei die zweite Gleichheit wegen der Definition der Addition gilt. Ähnlich gilt

$$\pi_U(\alpha v_1) = \alpha v_1 + U = \alpha(v_1 + U) = \alpha \pi_U(v_1).$$

$\mathrm{Ker}(\pi_U) = U$: Der Nullvektor in V/U ist $U = 0 + U$. Wir berechnen

$$\mathrm{Ker}(\pi_U) = \{v \in V \mid v + U = 0 + U\} = \{v \in V \mid v - 0 \in U\} = \{v \in V \mid v \in U\} = U.$$

$\mathrm{Im}(\pi_U) = V/U$: (Wieso? Weil $\pi_U(v) = v + U$.) □

Durch den Rangsatz 3.27 erhalten wir folgendes Korollar:

Korollar 3.124. *Sei V ein endlich-dimensionaler Vektorraum. Für die Dimension des Quotientenraums V/U gilt*

$$\dim V/U = \dim V - \dim U.$$

Beweis. Laut Proposition 3.122 gilt

$$\dim V/U = \dim \mathrm{Im}(\pi_U) = \dim V - \dim \mathrm{Ker}(\pi_U) = \dim V - \dim U,$$

wobei die zweite Gleichheit aus dem Rangsatz 3.27 folgt. □

Dies gibt den Fasern einer linearen Abbildung $T : V \to W$ die Struktur eines Vektorraums und dieser ist (kanonisch) isomorph zu $\mathrm{Im}(T)$:

Proposition 3.125. *Sei $T : V \to W$ eine lineare Abbildung. Es gilt*

$$V/\operatorname{Ker}(T) \cong \operatorname{Im}(T).$$

Beweis. Wir definieren

$$\begin{aligned} \varphi : V/\operatorname{Ker}(T) &\to \operatorname{Im}(T) \\ v + \operatorname{Ker}(T) &\mapsto T(v). \end{aligned}$$

Die einzige Schwierigkeit besteht darin zu zeigen, dass die Definition von φ Sinn ergibt. Das heisst, dass φ wohl-definiert ist. Seien dazu $v, v' \in V$ mit $v + \operatorname{Ker}(T) = v' + \operatorname{Ker}(T)$. Dies ist äquivalent zu $v - v' \in \operatorname{Ker}(T)$. Um zu zeigen, dass φ wohl-definiert ist, müssen wir überprüfen, dass

$$\varphi(v' + \operatorname{Ker}(T)) = \varphi(v + \operatorname{Ker}(T))$$

gilt. Wir berechnen

$$\begin{aligned} \varphi(v' + \operatorname{Ker}(T)) = T(v') &= T(v - (v - v')) \\ &= T(v) - T(v - v') \\ &= T(v) - 0 \\ &= T(v) \\ &= \varphi(v + \operatorname{Ker}(T)). \end{aligned}$$

Ausserdem ist φ linear (Wieso?[12]) Die Abbildung φ ist sicher surjektiv (Wieso?). Da $\dim \operatorname{Im}(T)$ und $\dim (V/\operatorname{Ker}(T))$ gleich sind (wegen dem Rangsatz 3.27 und Korollar 3.124) folgt, dass φ ein Isomorphismus ist. □

Bemerkung 3.126. Vielleicht haben Sie bemerkt, dass ich hier schon zweimal das magische Wort „kanonisch“ verwendet habe. Ich möchte darauf nicht tief eingehen und nur Folgendes sagen: In der Definition von π_U oder auch in der Definition von φ haben wir keine Wahl getroffen. Zum Beispiel haben wir diese Abbildungen ohne eine gewisse Wahl einer Basis definiert. Für die Definition der Abbildung π_U verwendet man nur die Informationen, die man über V und U hat und nicht mehr. In diesem Sinne ist π_U kanonisch definiert!

[12] Für alle $a, b \in K$ und $v_1, v_2 \in V$ ist

$$\begin{aligned} \varphi(a(v_1 + \operatorname{Ker}(T)) + b(v_2 + \operatorname{Ker}(T))) &= \varphi(av_1 + bv_2 + \operatorname{Ker}(T)) \\ &= T(av_1 + bv_2) \\ &= aT(v_1) + bT(v_2) \\ &= a\varphi(v_1 + \operatorname{Ker}(T)) + b\varphi(v_2 + \operatorname{Ker}(T)). \end{aligned}$$

Beispiel 3.127. Für $U = \mathrm{Ker}(T)$ wie in Beispiel 3.111 ist V/U die Menge aller Geraden mit Steigung 1. Jede dieser Geraden kann als

$$(z, 0) + \mathrm{Ker}(T)$$

dargestellt werden und wegen des Beweises von Proposition 3.125 wird jede dieser Geraden durch $\varphi : V/U \to \mathrm{Im}(T)$ auf

$$T(z, 0) = z - 0 = z \in \mathbb{R}$$

abgebildet.

Hier ist eine Verbindung zwischen direkten Summen, Komplement und Quotientenräume:

Proposition 3.128. *Sei $U \subseteq V$ ein Untervektorraum und W ein Komplement von U. Das heisst, dass $W \subseteq V$ ein Untervektorraum ist, so dass $V = U \oplus W$. Dann ist*

$$\pi_U|_W : W \to V/U$$

ein Isomorphismus.

Beweis. Jede Einschränkung einer linearen Abbildung ist wieder linear. Sei nun $v \in V$. Laut Proposition 2.110 (4) existieren eindeutige $u \in U$, $w \in W$ mit $v = u + w$. Daher ist w das eindeutige Urbild von $v + U \in V/U$. Dies zeigt, dass $\pi_U|_W$ sowohl injektiv als auch surjektiv ist (Wieso?). $\square$

Zu guter Letzt:

Proposition 3.129 (Universelle Eigenschaft von V/U)**.** *Sei $T : V \to W$ eine lineare Abbildung zwischen zwei Vektorräumen V und W und sei $U \subseteq V$ ein Untervektorraum mit $U \subseteq \mathrm{Ker}(T)$. Dann existiert eine eindeutige lineare Abbildung $\widetilde{T} : V/U \to W$, so dass folgendes Diagramm kommutiert:*

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "\pi_U"'] & W \\ V/U \arrow[ur, "\widetilde{T}"'] & \end{tikzcd} .$$

Oder in anderen Worten, so dass

$$T = \widetilde{T} \circ \pi_U \tag{3.32}$$

Beweis-Skizze. Damit (3.32) gilt, muss man $\widetilde{T}$ folgendermassen definieren:

$$\widetilde{T}(v + U) = T(v). \tag{3.33}$$

Man zeigt nun, dass $\widetilde{T}$ wohl-definiert ist (hier benutzt man $U \subseteq \mathrm{Ker}(T)$), und dass $\widetilde{T}$ linear ist. Dies zeigt die Existenz. Für die Eindeutigkeit benutzen wir, dass wir keine andere Wahl für $\widetilde{T}$ haben ausser (3.33), so dass (3.32) gilt. □

Bemerkung 3.130. Im Bereich der abstrakten Algebra, versucht man Objekte mittels solcher universellen Eigenschaften zu definieren. Dies kann ein gutes Thema für die „Nachbesprechungszeit“ sein.

Changelog: Kapitel 3

- 11.11: Lemma 3.26 ist jetzt vor dem Rangsatz.
- 11.11: Beispiel 3.8 ist jetzt separat aufgeführt; vorher war es Teil der vorhergehenden Übung.
- 11.11: In Beispiel 3.17 wurden einige Typos korrigiert.
- 11.11: Im Beweis von Satz 3.15 wurde der Typo $T(v + v') = \ldots = T(v) + T(w)$ zu $T(v + v') = T(v) + T(v')$ korrigiert.
- 11.11: In Beispiel 3.5 wurde die zweite Abbildung (die Nullabbildung) etwas allgemeiner aufgeschrieben, nämlich von $V \to W$, wobei V und W zwei beliebige Vektorräume sind.
- 11.11: In Beispiel 3.9 wurden zwei Videos, um das geometrische Verständnis zu üben, verlinkt.
- 16.11: Lemma 3.30 wurde allgemeiner von $S : W \to U$ statt $S : W \to V$ formuliert.
- 16.11: Im Beweis von Satz 3.27 wurde die Notation angepasst: Statt w_i steht im Beweis jetzt v_i.
- 23.11: Im Beweis von Proposition 3.69 wurde $\mathcal{B} = (u_1, \ldots, u_k, w_1, \ldots, w_r)$ zu $\mathcal{B} = (w_1, \ldots, w_r, u_1, \ldots, u_k)$ geändert.
- 23.11: In Abschnitt 3.5 wurde der Typo $(\alpha T) := \alpha \cdot_W T(v)$ zu $(\alpha T)(v) := \alpha \cdot_W T(v)$ korrigiert.
- 24.11: Im Beweis von Korollar 3.66 wurden einige Typos korrigiert.
- 24.11: Zu Beginn von Abschnitt 3.4 wurde der Typo Spaltenrang(A) = Zeilenrang(B) zu Spaltenrang(A) = Zeilenrang(A) korrigiert.
- 24.11: In Satz 3.27 wurde die Annahme V endlich-dimensional hinzugefügt.
- 25.11: Im Beweis von Proposition 3.93 wurde in der dritten Zeile $(4) \iff (5)$ zu $(4) \iff (6)$ korrigiert.
- 26.11: In Beispiel 3.90 wurden die Einträge oberhalb der Diagonalen in der linken oberen Teilmatrix zu $*$ (nicht Nulleinträge) korrigiert.
- 26.11: Im Beweis von Proposition 3.93 wurden einige Typos korrigiert.

- 1.12: In den Beweisen von Satz 3.101, Satz 3.105 und in Korollar 3.103 wurden alle T_n zu T_k geändert.
- 2.12: Im Beweis von Korollar 3.85 wurde $T_1T_2v+T_1T_2v$ zu $T_1T_2v+T_1T_3v$ geändert.
- 2.12: In der Definition der Matrix $Q_{i,j}(\alpha)$ (Definition 3.97) wurde λ zu α geändert.
- 15.12: Übung 3.45 wurde angepasst.
- 24.12: In Korollar 3.86 wurde K^n zu $K^{n-\mathrm{Rang}(A)}$ korrigiert.
- 01.01: In Lemma 3.99 wurde $A \xrightarrow{\alpha L_i \to L_i} S_i(\alpha)$ zu $A \xrightarrow{\alpha L_i \to L_i} S_i(\alpha)A$ korrigiert.

Kapitel 4

Determinanten

Wie bereits erwähnt ist eines der Hauptziele dieses Kapitels jede quadratische Matrix $A \in M_{n\times n}(K)$ mit einem Skalar $\det A \in K$ zu versehen, so dass die Bedingung

$$\det A \neq 0$$

äquivalent zu allen anderen Bedingungen in Proposition 3.93 ist. Wir werden aber sehen, dass der Skalar $\det A$ viele andere Anwendungen hat. Um diese einzuführen machen wir zuerst ein bisschen Magie mit 2×2-Matrizen.

4.1 Magie in K^2 mit 2×2-Matrizen

Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_{2\times 2}(K)$. Wir fragen uns, ob A invertierbar ist. Wir könnten das Gauss-Jordan Verfahren[1] benutzen (Satz 3.105), um das Folgende zu sehen:

Proposition 4.1. *Die Matrix A ist invertierbar genau dann, wenn $ad - bc \neq 0$. Falls sie invertierbar ist, dann gilt*

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Wir schreiben $\det A = ad - bc$*, und dann haben wir unser Hauptziel für dieses Kapitel für $n = 2$ schon erreicht.*

Wir könnten uns fragen, was für Eigenschaften die Zahl $ad - bc$ hat. Die Tatsache, dass sie Invertierbarkeit einer Matrix A charakterisiert, impliziert „starke“ Eigenschaften dieser Zahl. Wenn wir uns das geometrisch überlegen (zum Beispiel für $K = \mathbb{R}$), sehen wir, dass $\det A = 0$ genau dann, wenn die Abbildung m_A nicht surjektiv ist,

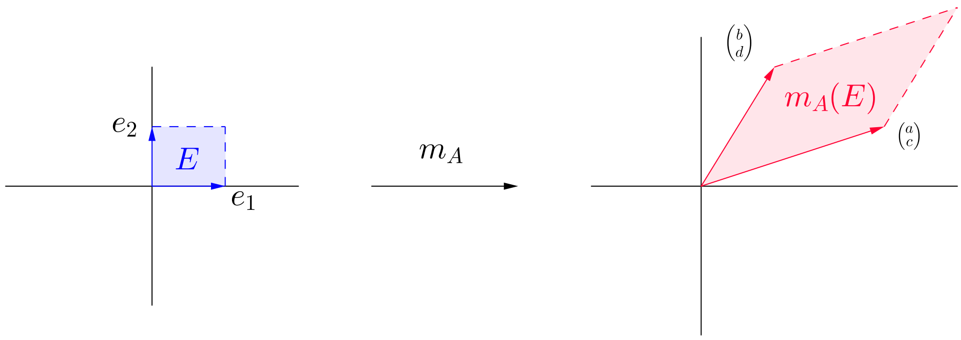
[1]Sie werden genau das in der Übungsstunde besprechen.

was genau dann gilt, wenn das Bild von m_A eine Gerade in $\mathbb{R}^2$ ist (oder sogar nur der Ursprung, falls $A = 0$ ist). Falls wir den „geometrischen Effekt“ der Abbildung m_A für Diagonalmatrizen wie $B = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ betrachten, sehen wir andererseits, dass m_B den Flächeninhalt von Bereichen $\mathbf{B} \subseteq \mathbb{R}^2$ mit einem Faktor $\det B = 2 \cdot 3$ skaliert. Dies stimmt immer noch für allgemeine Diagonalmatrizen $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ mit $a, b > 0$ oder $\begin{pmatrix} -a & 0 \\ 0 & -b \end{pmatrix}$ mit $a, b > 0$. Für Matrizen der Form

$$\begin{pmatrix} -2 & 0 \\ 0 & 3 \end{pmatrix}$$

ist der Skalierungsfaktor $|\det A| = |-2 \cdot 3| = 6$. Wir könnten uns auch fragen, wie das Vorzeichen der Determinante die Geometrie beeinflusst. Bevor wir das machen, kann man mit diesem Applet spielen, um zu sehen, dass die Skalierungseigenschaft nicht nur für Diagonalmatrizen stimmt (oder für Matrizen mit Determinante 0)[2], sondern auch für allgemeine Matrizen. Das ruft nach einem Beweis!

Proposition 4.2. *Sei E das Quadrat aufgespannt von e_1 und e_2 in $\mathbb{R}^2$. Das Bild $m_A(E)$ von E unter A ist das Parallelogramm aufgespannt von $m_A(e_1) = \begin{pmatrix} a \\ c \end{pmatrix}$ und $m_A(e_2) = \begin{pmatrix} b \\ d \end{pmatrix}$.*

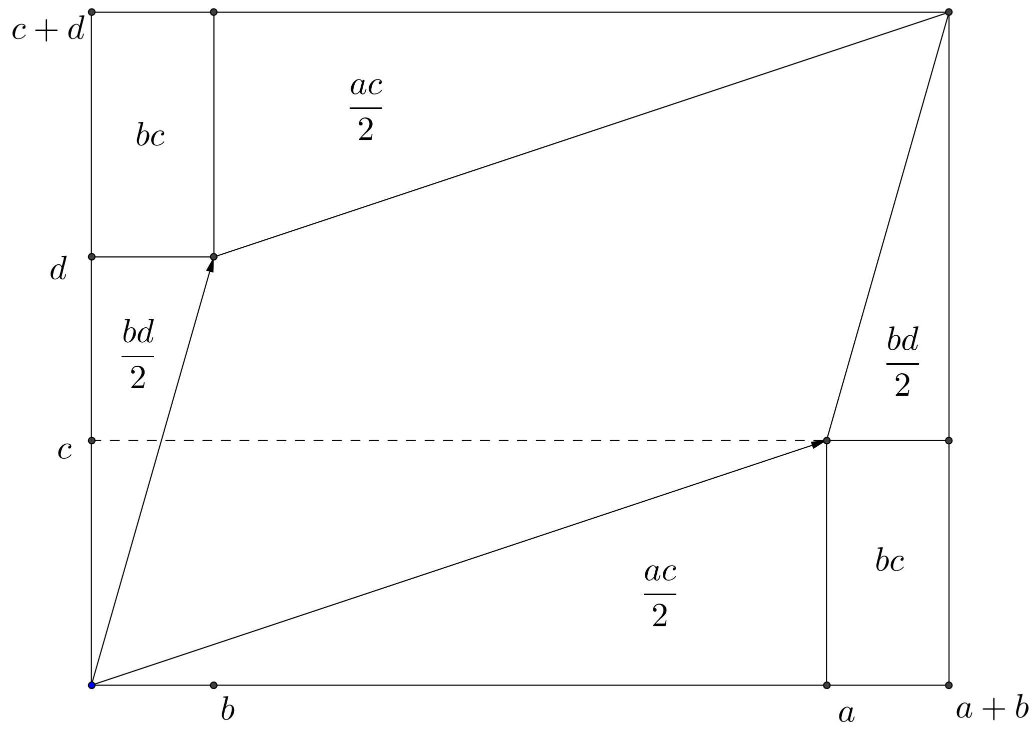


Dann ist der Flächeninhalt von $m_A(E)$ gegeben durch

$$\textit{Flächeninhalt}(m_A(E)) = |ad - bc| = |\det A| \,.$$

Beweis. Wir betrachten

[2] Wie wir vorher gesehen haben, „quetscht“ eine Matrix mit Determinante 0 den ganz Raum $\mathbb{R}^2$ zu einem Unterraum kleinerer Dimension zusammen. Also ist der Skalierungsfaktor 0, wie die Determinante der Matrix.



Figur 4.1: Berechnung des Flächeninhaltes.

und berechnen

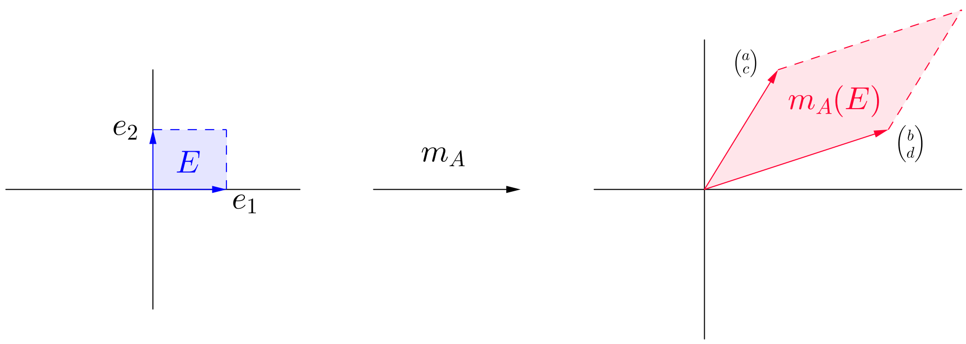
$$\begin{aligned}(a+b)(c+d)-\left(bc+bc+\frac{bd}{2}+\frac{bd}{2}+\frac{ac}{2}+\frac{ac}{2}\right) &= ac+ad+bc+bd-2bc-bd-ac\\ &= ad-bc\\ &= \det A.\end{aligned}$$

□

Bemerkung 4.3. Sie könnten sich beschweren, dass wir eigentlich gezeigt haben, dass der Skalierungsfaktor $\det A$ und nicht $|\det A|$ ist und Sie würden sich zu Recht beschweren. Der Grund dafür ist, dass es eine versteckte Annahme in unserem Bild 4.1 gibt. Die Annahme ist, dass das Paar von Vektoren

$$m_A(e_1) = \begin{pmatrix} a \\ c \end{pmatrix} \quad \text{und} \quad m_A(e_2) = \begin{pmatrix} b \\ d \end{pmatrix}$$

die gleiche „Orientierung“ wie e_1 und e_2 haben. Was meinen wir damit? Beachten Sie, dass e_2 „links“ von e_1 liegt und wir $m_A(e_2)$ auch links von $m_A(e_1)$ gezeichnet haben. Wenn das der Fall ist, dann sagen wir, dass A *orientierungserhaltend* ist. Falls wir das Bild so gezeichnet hätten:



Dann hätten wir

$$\text{Flächeninhalt}(m_A(E)) = bc - ad = |\det A|$$

erhalten, was bedeuten würde, dass die Determinante in diesem Fall negativ ist. In diesem Fall sagt man, dass m_A *orientierungsumkehrend* ist.

Fazit:

- Die Abbildung m_A ist orientierungserhaltend genau dann, wenn $\det A > 0$.
- Die Abbildung m_A ist orientierungsumkehrend genau dann, wenn $\det A < 0$.
- In beiden Fällen skaliert m_A den Flächeninhalt mit dem Faktor $|\det A|$.

Bemerkung 4.4. Wiederum könnten Sie argumentieren, dass wir Proposition 4.2 nicht für einen beliebigen Bereich in $\mathbb{R}^2$ kennen, da wir nur die Proposition nur für das von e_1 und e_2 aufgespannten Quadrat gezeigt haben. Und auch Letzteres haben wir durch ein Bild gezeigt, welches viele bestimmte Wahlen beinhaltete. Sie hätten recht, aber was wir gezeigt haben ist nicht weit entfernt von einem kompletten Beweis: Man muss einen allgemeinen Bereich in kleine Quadrate teilen (was Sie wahrscheinlich schon viele Male in der Analysis gemacht haben) und dann den Beweis, den wir gegeben haben, so modifizieren, dass er für beliebige (kleine) Quadrate gilt.

Die obige geometrische Interpretation der Determinante wird in allen Dimensionen wahr sein.

4.2 Volumen-Funktionen und die abstrakte Definition der Determinante

Wir werden die folgende Notation während dieses Abschnitts verwenden:

Sei $A \in M_{n\times n}(K)$ eine Matrix. Wir schreiben $A_{(1)}, \ldots, A_{(n)}$ für die Zeilen von A und $A^{(1)}, \ldots, A^{(n)}$ für die Spalten von A. Wir schreiben

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in M_{n\times n}(K)$$

für die Matrix, deren Zeilen $v_1, \ldots, v_n$ sind.

Man kann die Determinante auf zumindest zwei Arten explizit definieren:

(1) Mittels Permutationen (nach Leibniz, welcher die Determinante zuerst für 3×3-Matrizen im Jahre 1683 in einem Brief an L'Hôpital definierte).

(2) Induktiv unter der Benutzung des Entwicklungssatz von Laplace, welchen wir später sehen werden.

Die eleganteste Art (jedoch mit dem Nachteil, dass sie abstrakt ist[3]) die Determinante einzuführen, ist es der Charakterisierung von Weierstrass (aus dem Jahre 1903) zu folgen und drei einfache Axiome zu benutzen:

Definition 4.5 (Definition und Satz)**.** Es existiert eine eindeutige Funktion

$$\det : M_{n\times n}(K) \to K,$$

die *Determinante* heisst und die folgenden Eigenschaften[4] erfüllt:[5]

(D1) Die Abbildung det ist *multilinear* bezüglich Zeilen[6] Dies bedeutet:

(a) Für alle $v_1, \ldots, v_n, w \in K^n_{\text{Zeil}}$ und alle $1 \leq i \leq n$ gilt

$$\det \begin{pmatrix} v_1 \\ \vdots \\ v_{i-1} \\ v_i + w \\ v_{i+1} \\ \vdots \\ v_n \end{pmatrix} = \det \begin{pmatrix} v_1 \\ \vdots \\ v_{i-1} \\ v_i \\ v_{i+1} \\ \vdots \\ v_n \end{pmatrix} + \det \begin{pmatrix} v_1 \\ \vdots \\ v_{i-1} \\ w \\ v_{i+1} \\ \vdots \\ v_n \end{pmatrix}.$$

[3]Wenn man darauf besteht eine explizite (nicht abstrakte) Definition der Determinanten anzugeben mittels Permutationen, dann muss man ein Meister im Umgang mit Indizes werden. (Siehe den wunderbaren Beweis von $\det AB = \det A \det B$ in [8, S. 229].)

[4]Diese Eigenschaften nennt man die *Axiome der Determinante*.

[5]Wir wählen die Notation (D1)-(D3), so dass Sie problemlos den entsprechenden Stoff im Fischer [6] nachlesen können.

[6]Genau so gut könnten wir dies bezüglich Spalten definieren und würden dieselbe Determinante erhalten.

(b) Für alle $a \in K$, alle $v_1, \ldots, v_n \in K^n_{\text{Zeil}}$ und alle $1 \leq i \leq n$ gilt

$$\det \begin{pmatrix} v_1 \\ \vdots \\ v_{i-1} \\ av_i \\ v_{i+1} \\ \vdots \\ v_n \end{pmatrix} = a \det \begin{pmatrix} v_1 \\ \vdots \\ v_{i-1} \\ v_i \\ v_{i+1} \\ \vdots \\ v_n \end{pmatrix}.$$

(D2) Die Abbildung det ist *alternierend*. Das heisst, falls A zwei gleiche Zeilen hat, dann ist $\det A = 0$.

(D3) Die Abbildung $\det A$ ist *normiert*. Das heisst, dass $\det I_n = \det \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = 1$.

Für

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \in M_{n \times n}(K)$$

schreibt man häufig $|A| := \det A$ oder

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

Wir werden den Existenz-Teil dieses Satzes später beweisen. Jetzt beweisen wir zunächst weitere Eigenschaften der Determinante, die aus den drei Axiomen (D1)-(D3) folgen, und zum Schluss folgt auch die Eindeutigkeit der Abbildung det.

Dies ist ziemlich cool: Wir wissen noch nicht, ob solche Objekte überhaupt existieren, werden jetzt aber weitere Eigenschaften von diesem hypothetischen Objekt finden bis zu dem Punkt, wo wir fast schon wissen wie man zeigen kann, dass dieses Objekt überhaupt existiert.

In diesem ganzen Abschnitt bezeichnet det eine Funktion, die (D1)-(D3) in Definition 4.5 erfüllt. Das heisst, man könnte zu jeder Aussage in diesem Abschnitt den Satz „Sei det eine Funktion, die die Definition 4.5 erfüllt“ hinzufügen.

Proposition 4.6. *Es gilt:*

(D4) $\det(\lambda A) = \lambda^n \det A$.

(D5) Falls A eine Nullzeile hat, dann ist $\det A = 0$.

Beweis. Beide folgen aus (D1) (b) (Wieso?). Wir geben jedoch noch einen ausführlichen Beweis:

(D4) Mit mehrfacher Anwendung von (D1) (b) folgt, dass

$$\det \lambda A = \det \begin{pmatrix} \lambda A_{(1)} \\ \vdots \\ \lambda A_{(n)} \end{pmatrix} \overset{(D1)(b)}{=} \lambda^n \det \begin{pmatrix} A_{(1)} \\ \vdots \\ A_{(n)} \end{pmatrix} = \lambda^n \det A.$$

(D5) Falls die Nullzeile $A_{(i_0)}$ ist, dann gilt $0 \cdot A_{(i_0)} = A_{(i_0)}$ und daher folgt

$$\det \begin{pmatrix} A_{(1)} \\ \vdots \\ A_{(i_0)} \\ \vdots \\ A_{(n)} \end{pmatrix} = \det \begin{pmatrix} A_{(1)} \\ \vdots \\ 0 \cdot A_{(i_0)} \\ \vdots \\ A_{(n)} \end{pmatrix} \overset{(D1)(b)}{=} 0 \det \begin{pmatrix} A_{(1)} \\ \vdots \\ A_{(i_0)} \\ \vdots \\ A_{(n)} \end{pmatrix} = 0.$$

□

Wir würden gerne den Effekt der Gauss'schen Elimination, also Zeilenumformungen, auf die Determinante verstehen. Da wir bereits den Zusammenhang von Zeilenumformungen und Elementarmatrizen verstehen, werden wir diesen auch in der Formulierung folgender Proposition verwenden:

Proposition 4.7 (Gauss und die Determinante[7]). *Es gelten folgende weitere Eigenschaften der Determinante:*

(D1b) Sei $\lambda \in K$ und $A \xrightarrow{\lambda L_i \to L_i} B$. Dann ist $\det B = \lambda \det A$. *Mit Elementarmatrizen[8] ausgedrückt heisst das, dass*

$$\det (S_i(\lambda) A) = \lambda \det A.$$

(D6) Falls $A \xrightarrow{L_i \leftrightarrow L_j} B$, dann ist $\det A = -\det B$. *Mit Elementarmatrizen ausgedrückt heisst das, dass*

$$\det (P_{i,j} A) = -\det A.$$

[7]Hinter dieser Überschrift steckt eigentlich ein Wortwitz: Gauss gab ursprünglich der Determinante ihren Namen, aber in einem anderen Kontext. Er studierte quadratische Formen in zwei Variablen, welche er mit einer Zahl versehen hat, die ihre arithmetischen Eigenschaften determiniert (bestimmt). Also hat er diese Zahl Determinante genannt.

[8]Obwohl $S_i(\lambda)$ nur dann eine Elementarmatrix ist, wenn $\lambda \neq 0$.

(D7) Sei $\lambda \in K$ *und* $A \xrightarrow{L_i+\lambda L_j \to L_i} B$*, wobei* $i \neq j$*. Dann ist* $\det A = \det B$*. Mit Elementarmatrizen ausgedrückt heisst das, dass*

$$\det\left(Q_{i,j}(\lambda)A\right) = \det A.$$

Beweis. Wie die Nummerierung andeutet, ist die erste Aussage äquivalent zu (D1)(b).

Für (D6) nehmen wir ohne Beschränkung der Allgmeinheit an, dass $i < j$ ist und betrachten die Matrix

$$B = \begin{pmatrix} A_{(1)} \\ \vdots \\ A_{(i)} + A_{(j)} \\ \vdots \\ A_{(i)} + A_{(j)} \\ \vdots \\ A_{(n)} \end{pmatrix},$$

wobei der Zeilenvektor $A_{(i)} + A_{(j)}$ in der i-ten beziehungsweise j-ten Zeile steht. Dann gilt

$$
\begin{aligned}
0 &\overset{(D2)}{=} \det B \overset{(D1)(a)}{=} \det \begin{pmatrix} A_{(1)} \\ \vdots \\ A_{(i)} \\ \vdots \\ A_{(i)} + A_{(j)} \\ \vdots \\ A_{(n)} \end{pmatrix} + \det \begin{pmatrix} A_{(1)} \\ \vdots \\ A_{(j)} \\ \vdots \\ A_{(i)} + A_{(j)} \\ \vdots \\ A_{(n)} \end{pmatrix} \\
&\overset{(D1)(a)}{=} \det \begin{pmatrix} A_{(1)} \\ \vdots \\ A_{(i)} \\ \vdots \\ A_{(i)} \\ \vdots \\ A_{(n)} \end{pmatrix} + \det \begin{pmatrix} A_{(1)} \\ \vdots \\ A_{(i)} \\ \vdots \\ A_{(j)} \\ \vdots \\ A_{(n)} \end{pmatrix} + \det \begin{pmatrix} A_{(1)} \\ \vdots \\ A_{(j)} \\ \vdots \\ A_{(i)} \\ \vdots \\ A_{(n)} \end{pmatrix} + \det \begin{pmatrix} A_{(1)} \\ \vdots \\ A_{(j)} \\ \vdots \\ A_{(j)} \\ \vdots \\ A_{(n)} \end{pmatrix} \\
&\overset{(D2)}{=} \det \begin{pmatrix} A_{(1)} \\ \vdots \\ A_{(i)} \\ \vdots \\ A_{(j)} \\ \vdots \\ A_{(n)} \end{pmatrix} + \det \begin{pmatrix} A_{(1)} \\ \vdots \\ A_{(j)} \\ \vdots \\ A_{(i)} \\ \vdots \\ A_{(n)} \end{pmatrix}
\end{aligned}
$$

Dies zeigt (D6).

Für (D7) benutzen wir (D1):

$$
\det B = \det \begin{pmatrix} A_{(1)} \\ \vdots \\ A_{(i-1)} \\ A_{(i)} + \lambda A_{(j)} \\ A_{(i+1)} \\ \vdots \\ A_{(n)} \end{pmatrix} = \det \begin{pmatrix} A_{(1)} \\ \vdots \\ A_{(i-1)} \\ A_{(i)} \\ A_{(i+1)} \\ \vdots \\ A_{(n)} \end{pmatrix} + \lambda \det \underbrace{\begin{pmatrix} A_{(1)} \\ \vdots \\ A_{(i-1)} \\ A_{(j)} \\ A_{(i+1)} \\ \vdots \\ A_{(n)} \end{pmatrix}}_{=:C} = \det A + \lambda \cdot 0 = \det A,
$$

wobei wir benutzt haben, dass C zwei gleiche Zeilen ($i \neq j$) hat. Also ist $\det C = 0$ nach (D2).

□

Bemerkung 4.8. Bemerken Sie, dass (D6) eigentlich äquivalent ist zu (D2) (im Fall $\mathrm{char}(K) \neq 2$). Das erklärt den Begriff „alternierend".

Nach dieser Proposition sollte der Leser sagen: „Cool!" Was ich im „Endgame" machen muss, ist alles, was ich jetzt wissen muss. Also was die Determinante einer Matrix in Stufenform ist. Ein Fall ist bereits klar: Falls die Stufenform eine Nullzeile hat, dann ist die Determinante 0 nach (D5).

Proposition 4.9. *Für eine obere Dreiecksmatrix*

$$A = \begin{pmatrix} a_1 & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & a_n \end{pmatrix}.$$

gilt $\det A = a_1 \cdots a_n$. *Dieselbe Aussage gilt für untere Dreiecksmatrizen*

$$\det \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ * & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ * & \cdots & * & a_n \end{pmatrix} = a_1 \cdots a_n.$$

Beweis. Fall $a_i \neq 0$ ist für alle $i \in \{1, \cdots, n\}$, können wir „nach oben reinigen" und verändern damit laut (D7) die Determinante nicht. Daher gilt

$$\det A \overset{(D7)}{=} \det \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_n \end{pmatrix} \overset{(D1b)}{=} a_1 \cdots a_n \det(I_n) \overset{(D3)}{=} a_1 \cdots a_n.$$

Falls ein $i \in \{1, \cdots, n\}$ mit $a_i = 0$ existiert, dann sei $i_0 = \max\{i \mid a_i = 0\}$. Wir können dann „nach oben reinigen" mit $a_{i_0+1}, \ldots, a_n$, die alle verschieden von Null sind, ohne die Determinante zu ändern. Dabei ist dann die i_0-Zeile eine Nullzeile, und aus (D5) folgt, dass

$$\det A = 0 = a_1 \cdots a_n.$$

Derselbe Beweis, aber mit „nach unten reinigen" und mit min statt max zeigt die zweite Aussage bezüglich unterer Dreiecksmatrizen. □

Korollar 4.10. *Es gilt*

$$\det A \neq 0 \iff \operatorname{Rang}(A) = n \iff A \textit{ ist invertierbar.}$$

Beweis. Die zweite Äquivalenz haben wir bereits in Proposition 3.93 gezeigt. Für die erste sei B eine Matrix in Stufenform, die zeilenäquivalent zu A ist. Laut Proposition 4.7 gilt

$$\det A \neq 0 \iff \det B \neq 0.$$

Falls $\operatorname{Rang}(B) < n$ ist, dann hat B eine Nullzeile[9]. Nach (D5) gilt dann $\det B = 0$. Falls $\operatorname{Rang}(B) = n$ ist, so ist B eine obere Dreiecksmatrix mit von Null verschiedenen Einträgen auf der Hauptdiagonale. Laut Proposition 4.9 gilt also $\det B \neq 0$. Dies zeigt $\det B \neq 0 \iff \operatorname{Rang}(B) = n$. Da $\operatorname{Rang}(A) = \operatorname{Rang}(B)$ ist, folgt das Korollar. □

Erinnern Sie sich, dass diese Eigenschaft der Determinante unser Hauptziel war. Das einzige Problem: Wir wissen noch nicht, ob die Determinante überhaupt existiert! Diese Erkenntnis gibt uns aber die Motivation, die Existenz zu zeigen, und ausserdem Methoden zur Berechnung der Determinante zu entwickeln und vielleicht auch Methoden zur Berechnung der Inverse mittels Determinanten zu finden (Beispielsweise Verallgemeinerungen von Proposition 4.1 für $n \in \mathbb{N}$ zu finden).

Bevor wir weitere Eigenschaften dieses tollen hypothetischen Objekts zeigen, bemerken wir, dass wir die Determinante mittels Proposition 4.7 berechnen können.

Beispiel 4.11. Um die Determinante $|A|$ von $A = \begin{pmatrix} 0 & 1 & 2 \\ 3 & -2 & 0 \\ 4 & 1 & 4 \end{pmatrix}$ zu bestimmen, verwenden wir Zeilen-Operationen, um eine Dreiecksmatrix zu bekommen.

$$\overbrace{\begin{pmatrix} 0 & 1 & 2 \\ 3 & -2 & 0 \\ 4 & 1 & 4 \end{pmatrix}}^{=A} \xrightarrow{L_1 \leftrightarrow L_2} \overbrace{\begin{pmatrix} 3 & -2 & 0 \\ 0 & 1 & 2 \\ 4 & 1 & 4 \end{pmatrix}}^{:=A_1} \xrightarrow{L_3 - \frac{4}{3}L_1 \to L_3} \overbrace{\begin{pmatrix} 3 & -2 & 0 \\ 0 & 1 & 2 \\ 0 & \frac{11}{3} & 4 \end{pmatrix}}^{:=A_2}$$

$$\xrightarrow{3L_3 \to L_3} \overbrace{\begin{pmatrix} 3 & -2 & 0 \\ 0 & 1 & 2 \\ 0 & 11 & 12 \end{pmatrix}}^{:=A_3} \xrightarrow{L_3 - 11L_2 \to L_3} \overbrace{\begin{pmatrix} 3 & -2 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & -10 \end{pmatrix}}^{:=A_4}$$

Laut Proposition 4.7 gilt

- $|A| = -|A_1|$,

[9] Alternativ ist B eine obere Dreiecksmatrix mit 0 auf der Hauptdiagonale und hat daher Determinante 0.

- $|A_1| = |A_2|$,
- $3\,|A_2| = |A_3|$,
- $|A_3| = |A_4|$,

also $-3\,|A| = |A_4|$. Laut Proposition 4.9 ist $|A_4| = 3 \cdot 1 \cdot (-10) = -30$. Daher gilt $|A| = \frac{-30}{-3} = 10$.

Hier ist eine verwirrende Frage: Wie kann es sein, dass wir etwas berechnen können, dessen Existenz wir noch nicht sicherstellen können? Wenn wir es berechnen können, zeigt dies nicht seine Existenz? (Hinweis: Wenn man versucht mittels Proposition 4.7 die Determinante zu definieren und dabei die Existenz zu etablieren, gibt es ein kleines Problem... die Wohldefiniertheit!)

Für den Beweis der Eindeutigkeit und um zu zeigen, dass die Determinante multiplikativ ist, wäre es hilfreich die Determinante der Elementarmatrizen zu kennen.

Lemma 4.12. *Für jedes $\alpha \in K$ gilt*

(1) $\det(S_i(\alpha)) = \alpha$,

(2) $\det(P_{i,j}) = -1$.

(3) $\det(Q_{i,j}(\alpha)) = 1$

Ausserdem gilt für jede Elementarmatrix T und jede beliebige Matrix B, dass

$$\det(TB) = \det T \det B.$$

Beweis. Wir geben unten einen Beweis, bemerken aber kurz, wie man auch durch zwei andere Argumente argumentieren könnte: (1) und (3) folgen direkt aus Proposition 4.9. Ausserdem kann man alle drei Aussagen auch mit Zeilenumformungen zeigen, wie die Leser überprüfen können.

Wir zeigen das Lemma nun aber mit der Elemementarmatrizen-Formulierung von Proposition 4.7: Teil (1) folgt aus

$$1 = \det I_n \overset{\text{Lemma 3.100}}{=} \det\left(S_i\left(\frac{1}{\alpha}\right) S_i(\alpha)\right) \overset{\text{Prop. 4.7}}{=} \frac{1}{\alpha} \det S_i(\alpha).$$

Teil (2) folgt aus

$$1 = \det I_n \overset{\text{Lemma 3.100}}{=} \det(P_{i,j} P_{i,j}) \overset{\text{Prop. 4.7}}{=} -\det P_{i,j}.$$

Teil (3) folgt aus

$$1 = \det I_n \overset{\text{Lemma 3.100}}{=} \det(Q_{i,j}(-\alpha) Q_{i,j}(\alpha)) \overset{\text{Prop. 4.7}}{=} \det Q_{i,j}(\alpha).$$

Die letzte Aussage folgt jetzt, da sie mit (1), (2) und (3) genau äquivalent zu der Elementarmatrizen-Formulierung von Proposition 4.7 ist.

□

Um weiter zu zeigen, wie toll die Elementarmatrizen sind, benutzen wir sie jetzt, um die folgende grundlegende Eigenschaft der Determinante zu beweisen:

Proposition 4.13. *Seien* $A, B \in M_{n\times n}(K)$. *Dann gilt*

$$\det(AB) = \det A \det B.$$

Beweis. Falls $\operatorname{Rang}(A) < n$, dann ist auch $\operatorname{Rang}(AB) < n$ (Wieso? Denken Sie an Matrizen als lineare Abbildungen). Laut Korollar 4.10 sind dann beide Seiten der Gleichung Null.

Nehmen wir also an, dass $\operatorname{Rang}(A) = n$ und daher, dass $A \in \mathrm{GL}_n(K)$. Laut Satz 3.101 ist A ein endliches Produkt von Elementarmatrizen, also

$$A = T_1 \cdots T_k$$

mit Elementarmatrizen T_i, $1 \leq i \leq k$. Wir beweisen jetzt (4.13) mit Induktion über k. Für $k = 1$ haben wir genau die letzte Aussage in Lemma 4.12. Wir nehmen an, dass die Aussage für $k - 1$ gilt und betrachten

$$\begin{aligned}\det(AB) = \det(T_1T_2\cdots T_kB) &\overset{\text{Fall } k=1}{=} \det T_1 \det(T_2\cdots T_kB)\\ &\overset{\text{Ind.-Annahme}}{=} \det T_1 \det(T_2\cdots T_k)\det B\\ &\overset{\text{Fall } k=1}{=} \det(T_1T_2\cdots T_k)\det B = \det A \det B.\end{aligned}$$

Also folgt die Proposition. □

Als eine letzte Anwendung von Elementarmatrizen beweisen wir die Eindeutigkeit der Determinante.

Korollar 4.14. *Seien* $\det$ *und* $\widetilde{\det}$ *zwei Funktionen, die* $(D1) - (D3)$ *erfüllen. Dann ist* $\det = \widetilde{\det}$.

Beweis. Für A mit $\operatorname{Rang}(A) < n$ muss laut Korollar 4.10

$$\det(A) = 0 = \widetilde{\det}(A)$$

gelten. Für A mit $\operatorname{Rang}(A) = n$ existieren Elemenatarmatrizen $T_1, \cdots, T_k$ mit

$$A = T_1 \cdots T_k.$$

Laut Proposition 4.13 und Lemma 4.12 gilt

$$\begin{aligned} \det A &= \det T_1 \cdots \det T_k \\ &= \widetilde{\det} T_1 \cdots \widetilde{\det} T_k = \widetilde{\det}(A), \end{aligned} \tag{4.1}$$

daher gilt $\det = \widetilde{\det}$. □

Bemerkung 4.15. Wir haben willkürlich gewählt, die Theorie der Determinante über Zeilenoperationen zu entwickeln [10]. Im nächsten Kapitel werden wir zeigen, dass $\det(A) = \det(A^T)$ ist, was auch zeigt, dass die analogen Aussagen aller obigen Aussagen auch für Spalten gelten. Die Leser sollten sich daher merken, dass es äussert nützlich ist, sowohl Zeilen- als auch Spaltenumformungen in der Berechnung einer Determinante zu benutzen.

4.3 Existenz der Determinante

Es gibt zumindest zwei Wege die Existenz der Determinante zu zeigen. Einer davon ist es, eine induktive Definition zu benutzen mittels dem Entwicklungssatz von Laplace, welchen wir später sehen werden (vgl. Prop. 4.5 in [10]). Der zweite Weg benutzt Permutationen. Da induktive Definitionen nie erleuchtend sind, gehen wir den zweiten Weg.

Um diesen Text nicht noch länger zu machen, erklären wir jetzt nicht, wieso uns Permutationen eigentlich aufgezwungen werden durch die vorherigen abstrakten Definitionen (der Determinante), wenn wir die Existenz der Determinante zeigen wollen. Die interessierten Leser können den Dozenten in der Nachbesprechungszeit fragen und versuchen folgende Übung zu lösen:

Übung 4.16. Zeigen Sie nur unter Verwendung von ((D1)-(D3) und vielleicht (D6)), dass

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

die einzige Funktion ist, die (D1)-(D3) erfüllt.

<u>Hinweis:</u> Für eine Funktion Φ, die (D1)-(D3) erfüllt, betrachten Sie $\Phi\left(\begin{pmatrix} ae_1 + be_2 \\ ce_1 + de_2 \end{pmatrix}\right)$.

4.3.1 Permutationen

Sei $S_n = \mathrm{Abb}(\{1, \ldots, n\})$. Das heisst, S_n ist die Gruppe aller bijektiven Funktionen von $\{1, \ldots, n\}$ auf sich selbst mit der Verkettung als Verknüpfung. Die Elemente von

[10] Eigentlich nicht ganz willkürlich: so kann man diesen Stoff auch in Fischer nachlesen.

S_n heissen *Permutationen* (von $\{1, \ldots, n\}$). Wir benutzen die folgende Schreibweise: Für $\sigma \in S_n$ schreiben wir

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}. \tag{4.2}$$

Per Definition der Multiplikation als Verkettung gilt für alle $\sigma, \tau \in S_n$

$$\begin{aligned} \tau \cdot \sigma &= \begin{pmatrix} 1 & 2 & \cdots & n \\ \tau(1) & \tau(2) & \cdots & \tau(n) \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & \cdots & n \\ \tau\sigma(1) & \tau\sigma(2) & \cdots & \tau\sigma(n) \end{pmatrix}. \end{aligned}$$

Erinnern Sie sich daran, dass die Permutation

$$\mathrm{Id} := \begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix}$$

das neutrale Element von S_n ist, und dass die Inverse σ^{-1} die Umkehrfunktion von σ ist. Einige Elemente spielen eine spezielle Rolle:

Definition 4.17. Sei $n \in \mathbb{N}$. Wir definieren das Element $\sigma_{i,j} \in S_n$ als die Permutation, die i und j vertauscht, und sonst nichts ändert. Das heisst

$$\sigma_{i,j}(k) := \begin{cases} j, & \text{falls } k = i \\ i, & \text{falls } k = j \\ k, & \text{sonst} \end{cases}.$$

Ein Element von S_n, das von der Form $\sigma_{i,j}$ ist, nennen wir eine *Transposition*.

Lemma 4.18. *Sei $\tau \in S_n$ mit $\tau(1) = i$ und $\tau(2) = j$. Dann gilt $\tau\sigma_{1,2}\tau^{-1} = \sigma_{i,j}$.*

Beweis. Wir berechnen: Für $k \notin \{i, j\}$ ist $\tau^{-1}(k) \notin \{1, 2\}$ und daher ist

$$\sigma_{1,2}(\tau^{-1}(k)) = \tau^{-1}(k)$$

beziehungsweise $\tau(\sigma_{1,2}(\tau^{-1}(k))) = \tau(\tau^{-1}(k)) = k$. Für i und j gilt

$$\begin{aligned} \tau\sigma_{1,2}\tau^{-1}(i) &= \tau(\sigma_{1,2}(1)) = \tau(2) = j \\ \tau\sigma_{1,2}\tau^{-1}(j) &= \tau(\sigma_{1,2}(2)) = \tau(1) = i \end{aligned}$$

und das Lemma folgt. □

Beispiel 4.19. Die Gruppe S_1 ist ziemlich langweilig. Sie hat nur ein Element $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Beispiel 4.20. Die Gruppe S_2 ist nicht viel interessanter, hat aber zumindest eine Transposition

$$S_2 = \left\{ \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \sigma_{1,2} \right\}$$

Beispiel 4.21. Jede Transposition hat Ordnung zwei. Das heisst, für jedes $\sigma_{i,j} \in S_n$ gilt

$$(\sigma_{i,j})^2 := \sigma_{i,j} \cdot \sigma_{i,j} = \sigma_{i,j} \circ \sigma_{i,j} = \begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix}.$$

Dies bedeutet auch, dass $(\sigma_{i,j})^{-1} = \sigma_{i,j}$. Bemerken Sie ausserdem, dass $\sigma_{i,j} = \sigma_{j,i}$.

Beispiel 4.22. Die Gruppe S_3 ist schon interessanter. Wir schreiben alle ihre Elemente auf:

$$\begin{aligned}
\mathrm{Id} &:= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \\
\sigma &:= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \sigma_{1,3} \cdot \sigma_{1,2} = \sigma_{1,2} \cdot \sigma_{2,3} && (\star) \\
\sigma \cdot \sigma &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = \sigma_{1,2} \cdot \sigma_{1,3} \\
\tau &:= \sigma_{1,2} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \sigma_{2,1} \\
\tau \cdot \sigma &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \sigma_{2,3} = \sigma_{3,2} \\
\sigma \cdot \tau &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \sigma_{3,1} = \sigma_{1,3}.
\end{aligned}$$

Es gibt keine anderen Elemente in S_3 (Wieso?). Ausserdem kann man sich als Beispiel merken, dass

$$\sigma\tau\sigma^{-1} = \sigma\sigma_{1,2}\sigma^{-1} = \sigma_{2,3}$$

aus Lemma 4.18 folgt, was man auch direkt berechnen kann.

Hier sind zwei Aussagen, die aus dem Obigen folgen:

- S_3 ist nicht-abelsch, da $\tau\sigma \neq \sigma\tau$ ist.
- Jedes Element von S_3 ist entweder eine Transposition oder ein Produkt von Transpositionen (aber nicht auf eine eindeutige Weise wie zum Beispiel $(\star)$ zeigt).

Die beiden letzten Aussagen werden wir verallgemeinern. Die erste zu verallgemeinern ist sehr einfach:

Übung 4.23. Benutzen Sie, dass S_3 nicht-abelsch ist, um zu zeigen, dass S_n für $n \geq 3$ nicht-abelsch ist.

4.3.2 Signum einer Permutation

Definition 4.24. Sei $\sigma \in S_n$. Ein *Fehlstand*[11] von σ ist ein Paar $i < j$ für $i, j \in \{1, \ldots, n\}$, so dass
$$\sigma(i) > \sigma(j).$$

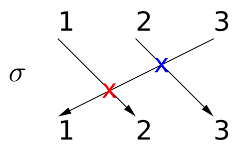
Beispiel 4.25. Falls $n = 3$ ist, gibt es drei mögliche Paare mit $i < j$ in $\{1, 2, 3\}$, nämlich $\{(1, 2), (1, 3), (2, 3)\}$. Zwei davon sind Fehlstände für die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} :$$

- Für $1 < 2$ ist $\sigma(1) = 2 < 3 = \sigma(2)$. Also ist $(1, 2)$ kein Fehlstand.
- Für $1 < 3$ ist $\sigma(1) = 2 > 1 = \sigma(3)$. Also ist $(1, 3)$ ein Fehlstand.
- Für $2 < 3$ ist $\sigma(2) = 3 > \sigma(3) = 1$. Also ist $(2, 3)$ ein Fehlstand.

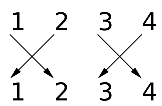
Man sagt: $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ hat zwei Fehlstände.

Dies könnte man versuchen zu veranschaulichen:



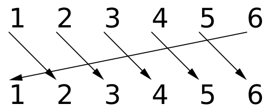
Das blaue Kreuz entspricht dem Fehlstand $(2, 3)$ und das rote Kreuz entspricht dem Fehlstand $(1, 3)$. Man überprüft, wie viele Überkreuzungen es gibt. Jede Überkreuzung entspricht einem Fehlstand.

Beispiel 4.26. (1) Die Permutation $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$ hat zwei Fehlstände:



[11] Inversion auf Englisch.

(2) Die Permutation $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 4 & 5 & 6 & 1 \end{pmatrix}$ hat fünf Fehlstände:



(3) Ähnlich hat die Permutation $\begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ 2 & 3 & \cdots & n & 1 \end{pmatrix}$ genau $n-1$ Fehlstände.

Beispiel 4.27. Die Transposition $\sigma_{1,2} \in S_n$ hat nur $(1,2)$ als Fehlstand. (Wieso?)

Definition 4.28. Wir definieren das *Signum* einer Permutation $\sigma \in S_n$ als[12]

$$\begin{aligned} \operatorname{sign}(\sigma) &= \begin{cases} +1, & \text{falls } \sigma \text{ eine gerade Anzahl von Fehlständen hat} \\ -1, & \text{falls } \sigma \text{ eine ungerade Anzahl von Fehlständen hat} \end{cases} \\ &= (-1)^k, \end{aligned}$$

wobei k die Anzahl der Fehlstände von σ ist. Eine Permutation σ heisst *gerade*, falls $\operatorname{sign}(\sigma) = +1$ und *ungerade*, falls $\operatorname{sign}(\sigma) = -1$ ist.

Beispiel 4.29.

- $\operatorname{sign}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = (-1)^2 = 1.$
- $\operatorname{sign}(\sigma_{1,2}) = (-1)^1 = -1.$
- $\operatorname{sign}\begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ 2 & 3 & \cdots & n & 1 \end{pmatrix} = (-1)^{n-1}.$

Wir werden jetzt die folgenden zwei Lemmata zeigen:

Lemma 4.30. *Für $\sigma \in S_n$ gilt*

$$\operatorname{sign}(\sigma) = \prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j - i} = \frac{\prod_{i<j}(\sigma(j) - \sigma(i))}{\prod_{i<j}(j-i)}. \tag{4.3}$$

Lemma 4.31. *Für alle $\sigma, \tau \in S_n$ gilt* $\operatorname{sign}(\tau\sigma) = \operatorname{sign}(\tau)\operatorname{sign}(\sigma)$. *In anderen Worten ist*

$$\begin{aligned} S_n &\to \mathbb{Z}/2\mathbb{Z} \cong \{1, -1\} \\ \sigma &\mapsto \operatorname{sign}(\sigma), \end{aligned}$$

ein Homomorphismus von S_n nach $\mathbb{Z}/2\mathbb{Z} \cong \{1, -1\}$.

[12] Viele Autoren schreiben sgn statt sign (inklusive dem Dozenten).

Diese Lemmata und ihre zugehörigen Beweise mögen schwierig erscheinen, sind aber eigentlich elementar. Hier sind einige elementare Beobachtungen, die uns im Beweis von Lemma 4.30 helfen:

- Sowohl der Nenner als auch der Zähler in (4.3) enthalten alle möglichen Paare. Im Nenner kommen sie immer in der Form $j-i$ vor, was immer eine positive Zahl ist. Im Zähler hingegen könnte σ das Vorzeichen kehren. Insbesondere ist

$$\frac{\prod_{i<j} |\sigma(j)-\sigma(i)|}{\prod_{i<j}(j-i)} = 1, \tag{4.4}$$

da alle Faktoren sich miteinander kürzen.

- Genauer gesagt, ändert σ das Vorzeichen im Zähler genau dann, wenn (i,j) ein Fehlstand ist:

$$\begin{aligned} \frac{\sigma(j)-\sigma(i)}{j-i} > 0 &\iff (i,j) \text{ ist kein Fehlstand von } \sigma \\ \frac{\sigma(j)-\sigma(i)}{j-i} < 0 &\iff (i,j) \text{ ist ein Fehlstand von } \sigma. \end{aligned}$$

Jetzt sind wir bereit für den Beweis von Lemma 4.30. Wir erklären zur Vorbereitung zuerst alle Schritte mit Hilfe der Permutation

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

Beispiel 4.32. Wir überprüfen Lemma 4.30 für $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$. Es gilt

$$\{i<j \mid i,j \in \{1,2,3\}\} = \{(1,2),(1,3),(2,3)\}.$$

Daher besagt das Lemma, dass

$$\begin{aligned} \operatorname{sign}(\sigma) &= \frac{\sigma(2)-\sigma(1)}{2-1} \cdot \frac{\sigma(3)-\sigma(1)}{3-1} \cdot \frac{\sigma(3)-\sigma(2)}{3-2} \\ &= \underbrace{\frac{3-2}{2-1}}_{\text{(1,2) ist kein Fehlstand}} \cdot \underbrace{\frac{1-2}{3-1}}_{\text{(1,3) ist ein Fehlstand}} \cdot \underbrace{\frac{1-3}{3-2}}_{\text{(2,3) ist ein Fehlstand}} \\ &\overset{\text{kürzen}}{=} \underbrace{(-1)^2}_{\text{2 ist die Anzahl der Fehlstände}} = (-1)^{\text{Anzahl der Fehlstände}}. \end{aligned}$$

Der allgemeine Beweis ist ganz ähnlich.

Beweis von Lemma 4.30. Sei $\sigma \in S_n$. Für $i < j$ gilt:

$$(i,j) \text{ ist ein Fehlstand } \iff \sigma(j) - \sigma(i) < 0.$$

Sei s die Anzahl der Fehlstände von σ. Dann gilt

$$\begin{aligned}\prod_{i<j} \frac{\sigma(j)-\sigma(i)}{j-i} &= \prod_{\substack{i<j,\\ \sigma(i)<\sigma(j)}} \frac{\sigma(j)-\sigma(i)}{j-i} \cdot \left((-1)^s \prod_{\substack{i<j,\\ \sigma(i)>\sigma(j)}} \frac{|\sigma(j)-\sigma(i)|}{j-i} \right) \\ &= (-1)^s \prod_{i<j} \frac{|\sigma(j)-\sigma(i)|}{j-i} = (-1)^s,\end{aligned}$$

wobei die letzte Gleichheit aus (4.4) folgt. □

Für den Beweis von Lemma 4.31 benutzen wir Lemma 4.30:

Beweis von Lemma 4.31. Es gilt

$$\begin{aligned}\operatorname{sign}(\tau\sigma) &\overset{\text{Lemma 4.30}}{=} \prod_{i<j} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{j-i} \\ &= \prod_{i<j} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j)-\sigma(i)} \prod_{i<j} \frac{\sigma(j)-\sigma(i)}{j-i} \\ &\overset{\text{Lemma 4.30}}{=} \underbrace{\prod_{i<j} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j)-\sigma(i)}}_{(*)} \operatorname{sign}(\sigma).\end{aligned}$$

Es genügt also zu zeigen, dass $(*) = \operatorname{sign}(\tau)$. Machen Sie an diesem Punkt eine Pause und überzeugen Sie sich davon, dass es tatsächlich genügt, $(*) = \operatorname{sign}(\tau)$ zu zeigen.

Es gilt

$$\begin{aligned}(*) &= \prod_{i<j} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j)-\sigma(i)} \\ &= \prod_{\substack{i<j,\\ \sigma(i)<\sigma(j)}} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j)-\sigma(i)} \cdot \prod_{\substack{i<j,\\ \sigma(i)>\sigma(j)}} \underbrace{\frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j)-\sigma(i)}}_{(\star)}\end{aligned} \tag{4.5}$$

Betrachten Sie nun genau den Term $(\star)$. Was passiert, wenn wir i mit j vertauschen und j mit i? Nichts passiert, da wir sowohl den Nenner und als auch den Zähler einfach mit -1 multiplizieren. Das sagt uns, dass wir die Rollen von i und j einfach vertauschen können im zweiten Produkt, ohne dass sich etwas ändert. Also ist

$$\prod_{\substack{i<j,\\ \sigma(i)>\sigma(j)}} (\star) = \prod_{\substack{j<i,\\ \sigma(j)>\sigma(i)}} (\star).$$

Wenn wir das in (4.5) einsetzen, erhalten wir

$$\begin{aligned}
(*) &= \prod_{\substack{i<j,\\ \sigma(i)<\sigma(j)}} \frac{\tau(\sigma(j))-\tau(\sigma(i))}{\sigma(j)-\sigma(i)} \cdot \prod_{\substack{j<i,\\ \sigma(i)<\sigma(j)}} \frac{\tau(\sigma(j))-\tau(\sigma(i))}{\sigma(j)-\sigma(i)} \\
&= \prod_{\sigma(i)<\sigma(j)} \frac{\tau(\sigma(j))-\tau(\sigma(i))}{\sigma(j)-\sigma(i)} \\
&\overset{\widetilde{i}:=\sigma(i),\widetilde{j}:=\sigma(j)}{=} \prod_{\widetilde{i}<\widetilde{j}} \frac{\tau(\widetilde{j})-\tau(\widetilde{i})}{\widetilde{j}-\widetilde{i}} \\
&\overset{\text{Lemma 4.30}}{=} \operatorname{sign}(\tau).
\end{aligned}$$

Dies beendet den Beweis des Lemmas. □

Korollar 4.33. *Für alle* $\sigma \in S_n$ *gilt* $\operatorname{sign}(\sigma^{-1}) = \operatorname{sign}(\sigma)$.

Beweis. Nach Lemma 4.31 gilt $1 = \operatorname{sign}(\mathrm{Id}) = \operatorname{sign}(\sigma)\operatorname{sign}(\sigma^{-1})$, also ist

$$\operatorname{sign}(\sigma^{-1}) = (\operatorname{sign}(\sigma))^{-1} = \operatorname{sign}(\sigma).$$

□

Korollar 4.34. *Für jedes* $i \neq j$ *gilt* $\operatorname{sign}(\sigma_{i,j}) = -1$.

Beweis. Aus Beispiel 4.29 wissen wir, dass $\operatorname{sign}(\sigma_{1,2}) = -1$. Sei nun $\tau \in S_n$ mit $\tau(1) = i$ und $\tau(2) = j$. Laut Lemma 4.18 wissen wir, dass

$$\tau\sigma_{1,2}\tau^{-1} = \sigma_{i,j}.$$

Daher ist

$$\operatorname{sign}(\sigma_{i,j}) = \operatorname{sign}(\tau)\operatorname{sign}(\sigma_{1,2})\operatorname{sign}(\tau^{-1}) \overset{\text{Korollar 4.33}}{=} \operatorname{sign}(\sigma_{1,2}) = -1.$$

Das Korollar folgt. □

Korollar 4.35. *Es gilt:*

(1) Die Abbildung $\operatorname{sign}\colon S_n \to \{\pm 1\}$ *ist surjektiv, falls* $n \geq 2$.

(2) Die Menge $A_n := \{\sigma \in S_n \mid \operatorname{sign}(\sigma) = 1\} = \{$*alle geraden Permutationen*$\}$ *ist eine Untergruppe von* S_n.

(3) Für jede Transposition[13] $\tau \in S_n$ *gilt*

$$\begin{aligned} A_n\tau &:= \{\sigma\tau \mid \sigma \in A_n\} \\ &= \{\sigma \in S_n \mid \operatorname{sign}(\sigma) = -1\} \\ &= \{\textit{alle ungeraden Permutationen}\} \\ &= S_n \setminus A_n. \end{aligned}$$

(4) Ausserdem gilt $|A_n| = |A_n\tau| = \frac{1}{2}|S_n| = \frac{n!}{2}$.

Beweis. (1) Es gilt zum Beispiel, dass $\operatorname{sign}(\mathrm{Id}) = 1$ und $\operatorname{sign}(\sigma_{1,2}) = -1$. Also ist die Abbildung sign surjektiv.

(2) Da wir diese Aussage nicht brauchen werden, überlassen wir diesen Beweis dem Leser als Übung. Erinnern Sie sich daran, dass A_n eine Untergruppe ist, wenn $\mathrm{Id} \in A_n$ ist, und wenn A_n abgeschlossen ist bezüglich Multiplikation und Inverse.

(3) Sei τ eine Transposition. Bemerken Sie zuerst, dass

$$\begin{aligned} A_n &\to A_n\tau \\ \sigma &\mapsto \sigma\tau \end{aligned} \tag{4.6}$$

eine Bijektion ist, da die Funktion $\eta \mapsto \eta\tau^{-1}$ ihre Umkehrfunktion ist. Wir behaupten des Weiteren, dass $A_n\tau$ die Menge aller ungeraden Permutationen ist:

- Jede Permutation der Form $\sigma\tau$ für $\sigma \in A_n$ erfüllt

$$\operatorname{sign}(\sigma\tau) = \operatorname{sign}(\sigma)\operatorname{sign}(\tau) = +1 \cdot (-1) = -1.$$

- Falls η Signum $\operatorname{sign}(\eta) = -1$ hat, dann ist

$$\operatorname{sign}(\eta\tau^{-1}) = \operatorname{sign}(\eta)\operatorname{sign}(\tau^{-1}) = -1 \cdot (-1) = 1,$$

wobei wir benutzt haben, dass $\operatorname{sign}(\tau^{-1}) = \operatorname{sign}(\tau) = -1$. Also ist $\eta\tau^{-1} \in A_n$ und daher ist

$$\eta = (\eta\tau^{-1})\tau \in A_n\tau.$$

(4) Da (4.6) eine Bijektion ist, folgt $|A_n| = |A_n\tau|$. Des Weiteren folgt, dass

$$n! = |S_n| = |A_n \sqcup A_n\tau| = |A_n| + |A_n\tau|,$$

was (4) impliziert (Wieso die erste Gleichheit gilt, überlassen wir den Lesern).

□

[13] Eigentlich gilt es sogar für jede ungerade Permutation.

Wir sind jetzt bereit, die Existenz der Determinante zu zeigen.

4.3.3 Existenz

Wir haben bereits die Eindeutigkeit der Determinante gezeigt. Wenn man die Eindeutigkeit direkt zeigen möchte (siehe auch Übung 4.16 für den Fall $n = 2$), dann führt dies zur Betrachtung des folgenden Ausdrucks:

$$\begin{aligned}\det A = \det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} &:= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdots a_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}.\end{aligned} \tag{4.7}$$

Das mag auf den ersten Blick vielleicht furchteinflössend wirken. Schauen wir uns also konkrete Formeln für $n = 2, 3$ an. Danach werden wir zeigen, dass Definition die (4.7) tatsächlich (D1)-(D3) erfüllt, was dann auch die Existenz der Determinante zeigt.

Beispiel 4.36 ($n = 2$)**.** Es gilt, dass $S_2 = \left\{ \mathrm{Id}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \right\}$. Daher ist

$$\begin{aligned}\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} &= \sum_{\sigma \in S_2} \operatorname{sign}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \\ &= \operatorname{sign}(\mathrm{Id}) a_{11} a_{22} + \operatorname{sign}\left(\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \right) a_{12} a_{21} \\ &= a_{11} a_{22} - a_{12} a_{21}.\end{aligned}$$

Beispiel 4.37. Unter Benutzung von Beispiel 4.22 erinnern wir uns an alle Elemente von S_3 und ihre jeweiligen Signums:

$$\begin{aligned}A_3 &= \left\{ \mathrm{Id}, \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \sigma^2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \right\} \\ S_3 \setminus A_3 &= \left\{ \tau = \mathrm{Id}\, \tau = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \sigma\tau = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \sigma^2\tau = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \right\}.\end{aligned}$$

Also ist

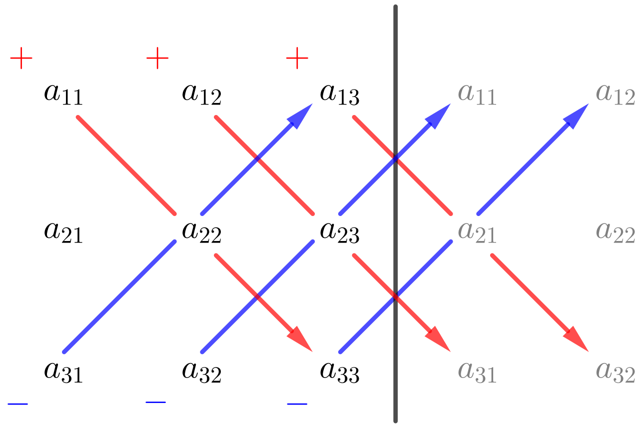
$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \sum_{\eta \in S_3} \operatorname{sign}(\eta) a_{1\eta(1)} a_{2\eta(2)} a_{3\eta(3)}$$

$$= \sum_{\eta \in A_3} a_{1\eta(1)} a_{2\eta(2)} a_{3\eta(3)} - \sum_{\eta \in S_3 \setminus A_3} a_{1\eta(1)} a_{2\eta(2)} a_{3\eta(3)}$$

$$= \underbrace{a_{11}a_{22}a_{33}}_{\eta=\mathrm{Id}} + \underbrace{a_{12}a_{23}a_{31}}_{\eta=\sigma} + \underbrace{a_{13}a_{21}a_{32}}_{\eta=\sigma^2}$$

$$- (\underbrace{a_{12}a_{21}a_{33}}_{\eta=\tau} + \underbrace{a_{13}a_{22}a_{31}}_{\eta=\sigma\tau} + \underbrace{a_{11}a_{23}a_{32}}_{\eta=\sigma^2\tau}).$$

Es gibt eine coole Art und Weise, wie man sich diese Formel merken kann, und zwar mit der *Regel von Sarrus*:



Bemerkung 4.38. Nachdem sie diese Regel gesehen haben, googeln viele Studenten nach einer ähnlichen Formel für 4×4-Matrizen. Leider existiert eine solche Formel nicht. Aber Sie müssen keine Angst haben: Wir werden im nächsten Abschnitt den Laplace-Entwicklungssatz sehen, welcher die beste Methode ist, um Determinanten für grosse Matrizen zu berechnen (zumindest wenn man es von Hand macht).

Jetzt wo wir keine Angst mehr haben vor Ausdrücken wie in (4.7), sind wir bereit die Existenz der Determinante zu zeigen.

Satz 4.39. *Sei* $\det$ *so definiert wie in* (4.7)*. Dann erfüllt* $\det$ *die Bedingungen (D1)-(D3) aus Satz 4.5.*

Beweis. (D1): Um (D1) (a) zu zeigen, benutzen wir folgende nützliche Notation aus dem Buch von Fischer [6]:

$$\begin{aligned}\det\begin{pmatrix}\vdots\\ a_i'+a_i''\\ \vdots\end{pmatrix} &= \sum_{\sigma\in S_n}\operatorname{sign}(\sigma)\cdot a_{1\sigma(1)}\cdot\ldots\cdot(a'_{i\sigma(i)}+a''_{i\sigma(i)})\cdot\ldots\cdot a_{n\sigma(n)}\\ &= \sum_{\sigma\in S_n}\operatorname{sign}(\sigma)\cdot a_{1\sigma(1)}\cdot\ldots\cdot a'_{i\sigma(i)}\cdot\ldots\cdot a_{n\sigma(n)}\\ &+ \sum_{\sigma\in S_n}\operatorname{sign}(\sigma)\cdot a_{1\sigma(1)}\cdot\ldots\cdot a''_{i\sigma(i)}\cdot\ldots\cdot a_{n\sigma(n)}\\ &= \det\begin{pmatrix}\vdots\\ a_i'\\ \vdots\end{pmatrix}+\det\begin{pmatrix}\vdots\\ a_i''\\ \vdots\end{pmatrix}\end{aligned}$$

Also gilt (D1) (a).

Wir zeigen nun, dass auch (D1) (b) gilt: Für jedes $1\leq i\leq n$ und $\lambda\in K$ gilt

$$\begin{aligned}\det\begin{pmatrix}a_{11}&\cdots&a_{1n}\\ \vdots&&\vdots\\ \lambda a_{i1}&\cdots&\lambda a_{in}\\ \vdots&&\vdots\\ a_{n1}&\cdots&a_{nn}\end{pmatrix} &= \sum_{\sigma\in S_n}\operatorname{sign}(\sigma)a_{1\sigma(1)}\cdot\cdots\ldots\cdot\lambda a_{i\sigma(i)}\cdot\ldots\cdot a_{n\sigma(n)}\\ &= \lambda\cdot\sum_{\sigma\in S_n}\operatorname{sign}(\sigma)a_{1\sigma(1)}\cdot\ldots\cdot a_{n\sigma(n)}\\ &= \lambda\det A.\end{aligned}$$

(D2): Fast alles, was wir bis jetzt besprochen haben über S_n werden wir für den Beweis von (D2) verwenden: Angenommen die Matrix A hat zwei gleiche Zeilen $A_{(i)}$ und $A_{(j)}$ mit $i\neq j$. Sei $\tau=\sigma_{i,j}$. Wir müssen zeigen, dass $\det A=0$. Erinnern Sie sich daran, dass $S_n=A_n\sqcup A_n\tau=\{\text{alle geraden Permutationen}\}\sqcup\{\text{alle ungeraden Permutationen}\}$ (Korollar 4.35), wobei $A_n\tau=\{\sigma\tau\mid\sigma\in A_n\}$. Daher gilt

$$\begin{aligned}\det A &= \underbrace{\sum_{\sigma\in A_n}a_{1\sigma(1)}\cdot\ldots\cdot a_{n\sigma(n)}}_{\text{Summe über }A_n}-\underbrace{\sum_{\sigma\in A_n}a_{1\sigma(\tau(1))}\cdot\ldots\cdot a_{n\sigma(\tau(n))}}_{\text{Summe über }S_n\setminus A_n=A_n\tau}\\ &= \underbrace{\sum_{\sigma\in A_n}a_{1\sigma(1)}\cdot\ldots\cdot a_{n\sigma(n)}-a_{1\sigma(\tau(1))}\cdot\ldots\cdot a_{n\sigma(\tau(n))}}_{=:\diamondsuit}.\end{aligned}$$

Beachten Sie, dass die linke und rechte Seite (links bzw. rechts des Minus) jedes Summanden von $\diamondsuit$ fast die gleichen Terme haben bis auf gewisse Änderungen, welche von τ verursacht werden.

Erinnern Sie sich jetzt daran, dass $A_{(i)} = A_{(j)}$ gilt, was impliziert, dass

$$a_{ik} = a_{jk} \tag{4.8}$$

für alle $k = 1, \ldots, n$ gilt. Dies zeigt, dass das Produkt der Terme, die von τ geändert werden, dasselbe ist:

$$a_{i\sigma(\tau(i))} a_{j\sigma(\tau(j))} \overset{\text{Definition von } \tau}{=} a_{i\sigma(j)} a_{j\sigma(i)} \overset{(4.8)}{=} a_{j\sigma(j)} a_{i(\sigma(i)}. \tag{4.9}$$

Dies zeigt, dass jeder Summand von $\diamondsuit$ gleich 0 ist. Also ist $\det A = 0$.

(D3): Um ein gutes Ende zu haben: Der Beweis von (D3) ist einfach: Für

$$\mathrm{I_n} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

haben wir, dass $a_{ij} = 0$, falls $i \neq j$. Der einzige Summand im Ausdruck (4.7), welcher keinen Faktor der Form a_{ij} für $i \neq j$ enthält, ist der Summand, welcher der Permutation $\mathrm{Id} = \begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix}$ entspricht. Also ist

$$\det I_n = \operatorname{sign}(\mathrm{Id}) a_{11} \cdot \ldots \cdot a_{nn} = 1 \cdot 1 \cdots 1 = 1.$$

Also erfüllt det wie in (4.7) die Bedingungen (D1)-(D3). Dies zeigt die Existenz der Determinante. □

Nun da wir eine explizite Beschreibung der Determinante haben, können wir einiges mehr über sie beweisen. Insbesondere ist es nützlich den folgenden Satz zu kennen:

Satz 4.40. *Für jede Matrix $A \in M_{n\times n}(K)$ gilt*

$$\det A^T = \det A.$$

Dies zeigt, dass alle Resultate, welche Zeilenumformungen beinhalten auch für Spaltenumformungen gelten.

Beweis. Sei $A = (a_{ij})$. Die Determinante der Transponierten von A berechnet sich dann als

$$\det A^T = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \underbrace{a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n}}_{(\star)}.$$

Um $\det A^T$ mit $\det A$ in Verbindung zu bringen, betrachten wir das Produkt ($\star$) und versuchen die Reihenfolge der Indizes zu ändern. Für jedes $\sigma \in S_n$ und jedes $1 \leq i \leq n$ gilt

$$a_{\sigma(i)i} = a_{\sigma(i)\sigma^{-1}(\sigma(i))}.$$

Daher gilt

$$(\star) = a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n} = a_{1\sigma^{-1}(1)} a_{2\sigma^{-1}(2)} \cdot \ldots \cdot a_{n\sigma^{-1}(n)}.$$

Ausserdem gilt $\operatorname{sign}(\sigma) = \operatorname{sign}(\sigma^{-1})$ und daher ist

$$(\star) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma^{-1}) a_{1\sigma^{-1}(1)} \cdot \ldots \cdot a_{n\sigma^{-1}(n)} = \sum_{\sigma^{-1} \in S_n} \operatorname{sign}(\sigma^{-1}) a_{1\sigma^{-1}(1)} \cdot \ldots \cdot a_{n\sigma^{-1}(n)} = \det A.$$

Die letzte Gleichheit folgt, da die Abbildung $\sigma \mapsto \sigma^{-1}$ eine Bijektion von S_n auf sich selbst ist und deshalb ändern wir bei Summation über σ^{-1} nur die Summationsreihenfolge. Dies beendet den Beweis des Satzes. $\square$

Korollar 4.41. *Alle Resultate, welche wir für Zeilen(-Umformungen) hatten, gelten auch für Spalten(-Umformungen). Insbesondere ist das Analogon von Proposition 4.7 sehr nützlich.*

Hier ist ein klassisches Resultat, das zeigt, wie man Spalten- und Zeilenoperationen benutzen kann.

Beispiel 4.42 (Vandermonde-Determinante)**.** Seien $x_1, x_2, \ldots, x_n \in K$. Wir möchten die *Vandermonde-Determinante*

$$V_{x_1,\ldots,x_n} = \det \begin{pmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ \vdots & & & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{pmatrix}$$

berechnen. Wieso diese Determinante und ihre Berechnung interessant sind, erfahren Sie in Serie 13. Wir lösen den Fall $n = 2, 3$:

$$\begin{vmatrix} 1 & x_1 \\ 1 & x_2 \end{vmatrix} = x_2 - x_1$$

Seien C_1, C_2, C_3 die Spalten von V_{x_1,x_2,x_3}. Dann ist

$$\begin{aligned}
\begin{vmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{vmatrix} &\overset{C_3 - x_1 C_2 \to C_3}{=} \begin{vmatrix} 1 & x_1 & 0 \\ 1 & x_2 & x_2^2 - x_1 x_2 \\ 1 & x_3 & x_3^2 - x_1 x_3 \end{vmatrix} \\
&\overset{C_2 - x_1 C_1 \to C_2}{=} \begin{vmatrix} 1 & 0 & 0 \\ 1 & x_2 - x_1 & x_2(x_2 - x_1) \\ 1 & x_3 - x_1 & x_3(x_3 - x_1) \end{vmatrix} \\
&\overset{\text{det Blockmatrix}}{=} \begin{vmatrix} x_2 - x_1 & x_2(x_2 - x_1) \\ x_3 - x_1 & x_3(x_3 - x_1) \end{vmatrix} \\
&\overset{(D1)(b)\text{Zeilen}}{=} (x_2 - x_1)(x_3 - x_1) \begin{vmatrix} 1 & x_2 \\ 1 & x_3 \end{vmatrix} = \prod_{1 \le i < j \le 3} (x_j - x_i).
\end{aligned}$$

Im Allgemeinen gilt

$$V_{x_1,\dots,x_n} = \prod_{1 \le i < j \le n} (x_j - x_i),$$

wie Sie in der Serie beweisen werden. Die Vandermonde-Determinante hat viele Anwendungen. Eine davon werden Sie in der Serie sehen.

In Beispiel 4.42 haben wir etwas benutzt, das Sie in der Übungsserie gesehen haben:

Übung 4.43. Seien $k \le n \in \mathbb{N}$ und $A \in M_{n\times n}(K)$ eine Blockmatrix der Form

$$A = \left(\begin{array}{c|c} B_1 & * \\ \hline 0 & B_2 \end{array}\right)$$

mit $B_1 \in M_{k\times k}(K)$ und $B_2 \in M_{(n-k)\times(n-k)}(K)$. Dann gilt

$$\det A = \det B_1 \cdot \det B_2.$$

Mit Induktion gilt eine ähnliche Aussage für Matrizen mit mehreren Blöcken. Des Weiteren gilt mit Transposition eine ähnliche Aussage für Matrizen der Form

$$\left(\begin{array}{c|c} B_1 & 0 \\ \hline * & B_2 \end{array}\right).$$

4.3.4 Determinante über einem Ring

Mittels der expliziten Formel für die Determinante sehen wir, dass für $A \in M_{n\times n}(\mathbb{Z})$ (das heisst, für eine Matrix mit ganzzahligen Einträgen) die Determinante $\det A$ ganzzahlig ist, also

$$\det A \in \mathbb{Z}.$$

Allgemeiner sei R ein Ring (zum Beispiel $R = M_{l\times l}(K)$). Sei

$$\begin{aligned} M_{n\times n}(R) &= \{\text{die Menge aller } n\times n\text{-Matrizen mit Einträgen in } R\} \\ &= \{(a_{ij})_{1\leq i,j\leq n} \mid a_{ij} \in R\}. \end{aligned}$$

(Falls $R = M_{l\times l}(K)$, dann sind die Elemente von $M_{n\times n}(K)$ Matrizen, deren Einträge $l \times l$-Matrizen über K sind!)

Für $A \in M_{n\times n}(R)$ definieren wir $\det A$ genau wie in 4.7:

$$\det A = \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}.$$

Dann gilt $\det A \in R$. Wir werden diese Tatsache für $R = M_{l\times l}(K)$ in der Linearen Algebra II benutzen.

4.4 Magie mit Minoren

In diesem Abschnitt wollen wir die folgende Beobachtung verallgemeinern: Für $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ definieren wir

$$\operatorname{adj}(A) := \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Dann folgt mit einer kurzen Rechnung, dass

$$A \cdot \operatorname{adj}(A) = \operatorname{adj}(A) \cdot A = \begin{pmatrix} \det A & 0 \\ 0 & \det A \end{pmatrix}.$$

Daraus folgt eine Formel für die Inverse, falls sie existiert: Falls $\det A \neq 0$, dann ist

$$A^{-1} = \frac{1}{\det A} \operatorname{adj}(A).$$

Eigentlich folgt noch viel mehr aus dieser Verallgemeinerung, wie der Titel dieses Abschnitts schon andeutet. Wir besprechen in diesem Abschnitt die folgenden Resultate für eine Matrix $A \in M_{n\times n}(K)$:

(1) Die Konstruktion einer Matrix $\operatorname{adj}(A)$ mit der Eigenschaft, dass

$$\operatorname{adj}(A) \cdot A = A \cdot \operatorname{adj}(A) = \det A \cdot I_n.$$

Dies gibt uns insbesondere eine Formel für die Inverse (falls $\det A \neq 0$).

(2) Den Laplace-Entwicklungssatz, ein nützliches Resultat, um die Determinante zu berechnen (zumindest, wenn man die Berechnung von Hand macht).

(3) Die Cramersche Regel, eine Formel für die Lösung eines linearen Gleichungssystems $Ax = b$ mit einer Matrix A der Grösse $n \times n$.

Die nachfolgenden Resultate sind alle eng miteinander verwandt, so dass man sie voneinander herleiten kann.

Wir beginnen mit dem Laplace-Entwicklungssatz, da dieser am natürlichsten aus unseren vorherigen Resultaten folgt (nach Meinung des Dozenten).

Für eine Matrix $A \in M_{n\times n}(K)$ sei $A_{ij} \in M_{(n-1)\times(n-1)}(K)$ die Matrix, welche durch das Streichen der i-ten Zeile und j-ten Spalte entstanden ist.

Satz 4.44 (Laplace-Entwicklungssatz)**.** *Sei $A \in M_{n\times n}(K)$. Für jedes $1 \le i \le n$ kann man die Determinante von A bezüglich der i-ten Zeile folgendermassen entwickeln:*

$$\begin{aligned} |A| &= (-1)^{i+1} a_{i1} |A_{i1}| + (-1)^{i+2} a_{i2} |A_{i2}| + \ldots + (-1)^{i+n} a_{in} |A_{in}| && (4.10)\\ &= \sum_{k=1}^{n} (-1)^{i+k} a_{ik} |A_{ik}| . && (4.11) \end{aligned}$$

Ähnlich kann man die Determinante bezüglich der j-ten Spalte folgendermassen entwickeln:

$$|A| = \sum_{k=1}^{n} (-1)^{k+j} a_{kj} |A_{kj}|$$

Beispiel 4.45. Sei

$$A = \begin{pmatrix} -2 & 2 & -3 \\ -1 & 1 & 3 \\ 2 & 0 & -1 \end{pmatrix}.$$

Wir entwickeln nach der zweiten Spalte ($j = 2$ in Satz 4.44). Es ist nützlich, das Vorzeichen in Abhängigkeit der Indizes zu zeichnen:

$$\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}.$$

Damit berechnen wir:

$$\begin{aligned}\det A &= (-1)^{1+2}\cdot 2\cdot \left|\begin{pmatrix}-1 & 3\\ 2 & -1\end{pmatrix}\right| + (-1)^{2+2}\cdot 1\cdot \left|\begin{pmatrix}-2 & -3\\ 2 & -1\end{pmatrix}\right| + (-1)^{3+2}\cdot 0\cdot \left|\begin{pmatrix}-2 & -3\\ -1 & 3\end{pmatrix}\right|\\
&= (-2)\cdot((-1)\cdot(-1) - 2\cdot 3) + 1\cdot((-2)\cdot(-1) - 2\cdot(-3)) + 0\\
&= (-2)\cdot(-5) + 8\\
&= 18.\end{aligned}$$

Beweis von Satz 4.44. Bemerken Sie zunächst, dass die Entwicklung bezüglich Spalten durch Transponieren und durch die Identität $|A| = \left|A^T\right|$ aus der entsprechenden Entwicklung bezüglich Zeilen folgt. Daher genügt es den ersten Teil des Satzes zu beweisen.

Wir beweisen zuerst den Satz für die erste Zeile, das heisst, $i = 1$. Seien

$$B_k = \begin{pmatrix} 0 & \cdots & 0 & a_{1k} & 0 & \cdots & 0\\ a_{21} & \cdots & & \cdots & & \cdots & a_{2n}\\ \vdots & & & & & & \vdots\\ a_{n1} & \cdots & & \cdots & & \cdots & a_{nn}\end{pmatrix}$$

für $k = 1, \dots, n$. Dann folgt aus der Linearität der Determinante bezüglich der ersten Zeile, dass

$$|A| = \sum_{k=1}^{n} |B_k|.$$

Daher genügt es zu zeigen, dass

$$|B_k| = (-1)^{1+k} a_{1k} |A_{1k}|. \tag{4.12}$$

Wir fangen mit B_1 an:

$$|B_1| = \left|\left(\begin{array}{c|ccc} a_{11} & 0 & \cdots & 0\\ \hline a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{array}\right)\right| \overset{\text{Blockmatrix}}{=} a_{11} |A_{11}| = (-1)^{1+1} a_{11} |A_{11}|.$$

Für B_2 vertauschen wir zuerst die ersten beiden Spalten:

$$
\begin{aligned}
|B_2| = \left| \begin{pmatrix} 0 & a_{12} & 0 & \cdots & 0 \\ a_{21} & \cdots & & \cdots & a_{2n} \\ \vdots & & & & \vdots \\ a_{n1} & \cdots & & \cdots & a_{nn} \end{pmatrix} \right| &\overset{C_1 \leftrightarrow C_2}{=} (-1) \left| \left(\begin{array}{c|ccc} a_{12} & 0 & \cdots & 0 \\ \hline a_{22} & a_{21} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n2} & a_{n1} & \cdots & a_{nn} \end{array}\right) \right| \\
&\overset{\text{Fall } i=1}{=} (-1) a_{12} \left| A_{12} \right| \\
&= (-1)^{1+2} a_{12} \left| A_{12} \right| .
\end{aligned}
$$

Beachten Sie, dass die Blockmatrix unten rechts genau A_{12} entspricht, da wir die ersten beiden Spalten vertauscht haben. Für B_3 vertauschen wir die Spalten C_2 und C_3 und benutzen den vorherigen Fall, so dass wir

$$
|B_3| = (-1)(-1)^{1+2} a_{13} \left| A_{13} \right| = (-1)^{1+3} a_{13} \left| A_{13} \right| .
$$

Ähnlich (mit Induktion über k und Spaltenvertauschen $C_{k-1} \leftrightarrow C_k$) folgt (4.12). Dies zeigt die Entwicklung bezüglich $i = 1$. Die Entwicklung bezüglich $i = 2$ folgt jetzt mit der Zeilenvertauschung $L_1 \leftrightarrow L_2$: Sei B, so dass $A \xrightarrow{L_1 \leftrightarrow L_2} B$. Dann gilt

$$
|A| = - |B| \overset{\text{Entwicklung } i=1}{=} -1 \cdot (-1)^{1+k} b_{1k} \left| B_{1k} \right| \overset{\text{Definition } B}{=} (-1)^{2+k} a_{2k} \left| A_{2k} \right| ,
$$

was die Entwicklung bezüglich der zweiten Zeile zeigt. Die Entwicklung bezüglich der p-ten Zeile folgt genau wie oben aus der Entwicklung der $(p-1)$-Zeile und dem Zeilenaustausch $L_{p-1} \leftrightarrow L_p$, was den Beweis beendet. □

Man kann den Sachverhalt in Satz 4.44 schön in einer Matrix darstellen:

Definition 4.46. Die *komplementäre Matrix* oder die *Adjunkte*[14] einer Matrix $A \in M_{n \times n}(K)$ ist

$$
\operatorname{adj}(A) := \left((-1)^{i+j} \left| A_{ji} \right| \right)_{1 \leq i,j \leq n} .
$$

Bemerkung 4.47.

- Die Vertauschung der Indizes in $|A_{ji}|$ ist kein Tippfehler!
- Wir definieren in der Linearen Algebra II die Adjungierte einer Matrix, welche nicht mit der Adjunkten verwechselt werden sollte.
- Aus Abschnitt 4.3.4 folgt: Falls $A \in M_{n \times n}(R)$ für einen Ring R, dann ist $\operatorname{adj}(A)$ ebenfalls in $M_{n \times n}(R)$.

[14] „Adjunct“ oder „adjugate“ auf Englisch.

Beispiel 4.48. Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Dann ist

$$\operatorname{adj}(A) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

(Wieso?)

Der folgende Satz folgt aus dem Entwicklungssatz von Laplace 4.44 (und ist eigentlich äquivalent zu ihm):

Satz 4.49. *Sei $A \in M_{n\times n}(K)$. Dann ist*

$$A \cdot \operatorname{adj}(A) = \operatorname{adj}(A) \cdot A = \det A \cdot I_n = \begin{pmatrix} \det A & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \det A \end{pmatrix}.$$

Beweis. Berechnen wir zuerst die Einträge von $A \cdot \operatorname{adj}(A)$ auf der Hauptdiagonalen:

$$(A \operatorname{adj}(A))_{ii} = \sum_{k=1}^{n} a_{ik} \underbrace{(-1)^{k+i} |A_{ik}|}_{(\operatorname{adj}(A))_{ki}} \overset{\text{Laplace } i\text{-te Zeile}}{=} \det A.$$

Für die Einträge ausserhalb der Hauptdiagonalen sei nun $i \neq j$: Dann ist

$$(A \operatorname{adj}(A))_{ij} = \sum_{k=1}^{n} a_{ik} \underbrace{(-1)^{k+j} |A_{jk}|}_{(\operatorname{adj}(A))_{kj}} =: (\star)$$

Bemerken Sie nun, dass $(\star)$ die Entwicklung bezüglich der j-ten Zeile der Matrix B ist, wobei B die Matrix ist, welche durch das Ersetzen der j-ten Zeile von A mit der i-ten Zeile von A entsteht. (Ersetzen! Nicht vertauschen!) Daher ist $(\star) = \det B$. Da B zwei gleiche Zeilen hat, folgt $(\star) = 0$.

Die Gleichung $\operatorname{adj}(A) \cdot A = \det A \cdot I_n$ zeigt man ähnlich (mit Spaltenentwicklung statt Zeilenentwicklung). $\square$

Korollar 4.50. *Falls $A \in \mathrm{GL}_n(K)$, dann gilt*

$$A^{-1} = \frac{1}{\det A} \operatorname{adj}(A) \tag{4.13}$$

Korollar 4.51 (Cramersche Regel)**.** *Sei $A \in \mathrm{GL}_n(K)$. Wir betrachten das lineare Gleichungssystem*

$$Ax = b \tag{4.14}$$

für $b \in K^n = K^n_{Spal}$*. Sei* A_i *die Matrix, welche* b *als* i*-te Spalte hat und die restlichen Spalten stimmen mit den Spalten von* A *überein. Das heisst*

$$A_i = \left(A^{(1)} \cdots A^{(i-1)}\ b\ A^{(i+1)} \cdots A^{(n)}\right) \tag{4.15}$$

für $1 \le i \le n$*. Dann berechnet sich die eindeutige Lösung* $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ *von (4.14) als*

$$x_i = \frac{|A_i|}{|A|}.$$

Beweis. Dies folgt unmittelbar aus (4.14):

$$x = A^{-1}b = \frac{1}{|A|}\operatorname{adj}(A)b$$

beziehungsweise

$$x_i = \frac{1}{|A|}\sum_{k=1}^{n} \operatorname{adj}(A)_{ik} b_k = \frac{1}{|A|}\sum_{k=1}^{n} (-1)^{i+k}\left|A_{ki}\right| b_k \overset{\text{Laplace}}{=} \frac{1}{|A|}\left|A_i\right|.$$

Vielleicht fällt es Ihnen schwer die letzte Gleichheit zu sehen. Falls dies der Fall ist, dann betrachten Sie

$$A_i = \begin{pmatrix} | & & | & b_1 & | & & | \\ A^{(1)} & \dots & A^{(i-1)} & \vdots & A^{(i+1)} & \dots & A^{(n)} \\ | & & | & b_k & | & & | \end{pmatrix}$$

und machen Sie die Laplace-Entwicklung auf der i-ten Spalte! □

Hier sind konkrete Formeln für $n = 2, 3$ (Wir überlassen den Fall $n = 1$ dem Leser.)):

Explizite Formel für 2×2-Matrizen: Wir betrachten folgendes lineares Gleichungssystem

$$\begin{cases} a_1 x + b_1 y = c_1 \\ a_2 x + b_2 y = c_2 \end{cases},$$

welches in Matrixschreibweise von der Form

$$\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

ist. Wir nehmen an, dass $a_1b_2 - b_1a_2 \neq 0$. Dann können wir x und y mit Hilfe der Cramerschen Regel folgendermassen finden:

$$x = \frac{\left|\begin{pmatrix} c_1 & b_1 \\ c_2 & b_2 \end{pmatrix}\right|}{\left|\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}\right|} = \frac{c_1b_2 - b_1c_2}{a_1b_2 - b_1a_2} \quad \text{und} \quad y = \frac{\left|\begin{pmatrix} a_1 & c_1 \\ a_2 & c_2 \end{pmatrix}\right|}{\left|\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}\right|} = \frac{a_1c_2 - c_1a_2}{a_1b_2 - b_1a_2}.$$

Explizite Formel für 3×3-Matrizen: Die Regel für 3×3-Matrizen sind ähnlich. Sei

$$\begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \\ a_3x + b_2y + c_3z = d_3 \end{cases},$$

welches in Matrixschreibweise von der Form

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}.$$

Dann sind

$$x = \frac{\begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}}, \; y = \frac{\begin{vmatrix} a_1 & d_1 & c_1 \\ a_2 & d_2 & c_2 \\ a_3 & d_3 & c_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}} \quad \text{und} \quad z = \frac{\begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}}.$$

In der Übungsstunde werden Sie vielleicht folgendes Beispiel betrachten:

$$\begin{aligned} 82x_1 + 45x_2 + 9x_3 &= 1 \\ 27x_1 + 16x_2 + 3x_3 &= 1 \\ 9x_1 + 5x_2 + 1x_3 &= 0 \end{aligned}$$

Dann ist die erweiterte Koeffizienten-Matrix gegeben durch:

$$(A|b) = \left(\begin{array}{ccc|c} 82 & 45 & 9 & 1 \\ 27 & 16 & 3 & 1 \\ 9 & 5 & 1 & 0 \end{array}\right).$$

Nach der Cramerschen Regel berechnet sich die Lösung des linearen Gleichungssystems als:

$$x_1 = \frac{\det A_1}{\det A} = \frac{\det \begin{pmatrix} 1 & 45 & 9 \\ 1 & 16 & 3 \\ 0 & 5 & 1 \end{pmatrix}}{\det \begin{pmatrix} 82 & 45 & 9 \\ 27 & 16 & 3 \\ 9 & 5 & 1 \end{pmatrix}} = \frac{1}{1} = 1, \; x_2 = \frac{\det A_2}{\det A} = \frac{\det \begin{pmatrix} 82 & 1 & 9 \\ 27 & 1 & 3 \\ 9 & 0 & 1 \end{pmatrix}}{\det \begin{pmatrix} 82 & 45 & 9 \\ 27 & 16 & 3 \\ 9 & 5 & 1 \end{pmatrix}} = \frac{1}{1} = 1$$

$$\text{und} \quad x_3 = \frac{\det A_3}{\det A} = \frac{\det \begin{pmatrix} 82 & 45 & 1 \\ 27 & 16 & 1 \\ 9 & 5 & 0 \end{pmatrix}}{\det \begin{pmatrix} 82 & 45 & 9 \\ 27 & 16 & 3 \\ 9 & 5 & 1 \end{pmatrix}} = \frac{-14}{1} = -14.$$

4.5 Verschiedene Wege in der Welt der Determinante

In dieser Vorlesung ist unser Ziel, verschiedene Ideen zu präsentieren. Das bedeutet manchmal auch, dass wir gewisse Resultate nicht auf die einfachste oder kürzeste Weise erhalten wollen. Es ist aber interessant zu wissen, wie man anders vorgehen könnte. Hier sind ein paar Beispiele:

- Man könnte die Determinante induktiv definieren mittel Laplace-Entwicklung: Für $\begin{pmatrix} a \end{pmatrix} \in M_{1\times 1}(K)$ definiert man

  $$\det \begin{pmatrix} a \end{pmatrix} = a$$

  und für $A \in M_{n\times n}(K)$ definiert man $\det A$ induktiv unter Benutzung von (4.10). Man beweist dann, dass diese Definition (D1)-(D3) erfüllt und geniesst dann die Tatsache, dass Satz 4.44 per Definition wahr ist. Beachten Sie, dass man dabei nicht einmal wissen muss, was eine Permutation ist und man braucht auch nicht den Gruppenbegriff. Der Nachteil so einer Definition ist, dass induktive Definitionen selten einleuchtend sind und Sie würden nicht viel von so einem Vorgehen lernen. Der interessierte Leser kann dazu gerne [10] konsultieren.

- Man könnte die Charakterisierung der Determinante durch (D1)-(D3) komplett weglassen und alles direkt mit der Gruppe S_n machen (siehe [8]).

- Falls das, was wir hier gemacht haben, nicht abstrakt genug ist, könnte man die Determinante $\det T$ direkt für Endomorphismen $T \in \mathrm{End}(V)$ über einem endlichdimensionalen Vektorraum V definieren (siehe [9]). Wir werden $\det T$ im nächsten Abschnitt definieren.

- Fischer beweist, mit ähnlichen Argumenten wie wir sie zu Beginn benutzt haben, zuerst Satz 4.49 und dann Satz 4.44 als ein Korollar davon. Wir wählen einen anderen Weg, da die Ideen im Beweis von Satz 4.44 ein schönes Beispiel sind wieso Reduktionen nützlich sind.

Unter Verwendung von verschiedenen Vorgehensweisen werden einige Resultate einfacher, andere hingegen werden schwieriger. Einige unserer späteren Resultate folgen aus unseren vorherigen Resultaten. Hier ist ein cooles Beispiel wie man die Cramersche Regel nur unter Benutzung von (D1) und (D2) (für Spalten) beweisen kann:

Dieser Beweis benutzt auch ein grundlegendes Resultat aus Kapitel 2:

Alternativer Beweis von Korollar 4.51. Sei $A \in \mathrm{GL}_n(K)$. Wir betrachten die Abbildung

$$\begin{aligned} \Phi : K^n &\to K^n \\ b &\mapsto \frac{1}{|A|} \cdot (|A_1|, \ldots, |A_n|), \end{aligned}$$

wobei A_i genau wie in (4.15) definiert ist. Wegen (D1) (angewandt auf die Spalten) ist diese Abbildung linear. Des Weiteren ist das Bild der Spalte $A^{(i)}$ genau

$$(0, \ldots, 1, \ldots, 0) = e_i$$

wegen (D2). Also stimmt Φ mit der Abbildung $m_{A^{-1}}$ über der Basis $(A^{(1)}, \ldots, A^{(n)})$ überein. Also ist nach Satz 3.15

$$\Phi = m_{A^{-1}},$$

was äquivalent ist zur Aussage der Cramerschen Regel. □

Frage: Hier eine „Meta-Frage“: Wieso ergibt es Sinn, dass die Regel von Cramer unabhängig ist von (D3)?

4.6 Determinante und Rang

Wir haben gesehen, dass die Determinante einer Matrix $A \in M_{n\times n}(K)$ weiss, ob $\mathrm{Rang}(A) = n$ ist oder nicht. Weiss die Determinante auch, ob $\mathrm{Rang}(A) = r$ ist für $0 \leq r < n$?

Die Antwort ist „Ja“. Um dies zu sehen brauchen wir die folgende Definition:

Definition 4.52 (Minoren)**.** Seien $1 \leq k \leq n, m \in \mathbb{N}$. Ein *$k$-Minor* von $A \in M_{m\times n}(K)$ ist die Determinante einer $k \times k$-Matrix, die durch Streichen von $(n-k)$-Spalten und $(m-k)$-Zeilen von A entstanden ist. Eine solche Matrix heisst *$k \times k$-Untermatrix* von A.

Satz 4.53. *Sei $A \in M_{m\times n}(K)$. Dann gilt*[15]

$$\operatorname{Rang}(A) = \max\left\{k \mid \textit{es existiert ein } k\textit{-Minor, der verschieden von Null ist}\right\}.$$

Beweis. Sei $r = \operatorname{Rang}(A)$. Wir müssen zeigen, dass jeder k-Minor mit $k > r$ Null ist, und dass ein r-Minor existiert, der nicht Null ist.

Sei also $k > r$. Da $\operatorname{Rang}(A) = \operatorname{Spaltenrang}(A)$ ist, sind je k Spalten von A linear abhängig. Also sind die k Spalten einer beliebigen $k \times k$-Untermatrix von A auch linear abhängig. Also ist jeder k-Minor gleich Null.

Jetzt müssen wir einen r-Minor finden, welcher nicht Null ist. Nehmen wir an, dass $A \neq 0$ (siehe Fussnote). Da $\operatorname{Spaltenrang}(A) = r > 0$ ist, existieren

$$j_1 < \ldots < j_r,$$

so dass $A^{(j_1)}, \ldots, A^{(j_r)}$ linear unabhängig sind. Also hat

$$B := \left(A^{(j_1)} \cdots A^{(j_r)}\right)$$

Rang r. Da $\operatorname{Zeilenrang}(B) = \operatorname{Rang}(B) = r$ ist, existieren

$$i_1 < \ldots < i_r$$

mit

$$C := \begin{pmatrix} B_{(i_1)} \\ \vdots \\ B_{(i_r)} \end{pmatrix} \in M_{r\times r}(K)$$

und $\operatorname{Rang}(C) = r$. Daher ist $|C| \neq 0$. Die Matrix C ist eine $r \times r$-Untermatrix von A, was den Beweis beendet. $\square$

[15] Falls hier die leere Menge steht (d.h. falls $A = 0$), dann verstehen wir dieses Maximum als 0, damit der Satz gilt.

4.7 Einige Korollare und die Determinante eines Endomorphismus

In diesem Abschnitt geben wir einige einfache Korollare zur Multiplikativität der Determinante, welche die Definition der Determinante eines Endomorphismus ermöglicht:

Korollar 4.54. *Für jedes* $A \in \mathrm{GL}_n(K)$ *gilt*

$$\det(A^{-1}) = \det(A)^{-1}.$$

Beweis. Es gilt

$$1 = \det(I_n) = \det(AA^{-1}) = \det(A)\det(A^{-1})$$

und das Korollar folgt. □

Korollar 4.55. *Seien* $A, B \in M_{n\times n}(K)$ *ähnliche Matrizen. Dann gilt*

$$\det(A) = \det(B).$$

Beweis. Sei $P \in \mathrm{GL}_n(K)$ mit $A = P^{-1}BP$. Dann ist

$$\det(A) = \det(P^{-1})\det(B)\det(P) = \det(P)^{-1}\det(B)\det(P) = \det(B)$$

nach Korollar 4.54. □

Das letzte Korollar ist für das nächste Kapitel wichtig und ermöglicht folgende Definition:

Definition 4.56. Sei V ein Vektorraum über K mit $\dim(V) = n < \infty$ und sei $T \in \mathrm{End}(V)$. Wir definieren die Determinante von T als

$$\det(T) := \det([T]_{\mathcal{B}}),$$

wobei $\mathcal{B}$ eine beliebige (geordnete) Basis von V ist.

Lemma 4.57. *Diese Definition ergibt Sinn. Das heisst,* $\det(T)$ *ist wohl-definiert.*

Beweis. Seien $\mathcal{B}, \mathcal{C}$ zwei geordnete Basen von V. Laut Korollar 3.64 sind die Matrizen $[T]_{\mathcal{B}}$ und $[T]_{\mathcal{C}}$ ähnlich zueinander.

Daher folgt aus Korollar 4.55, dass

$$\det([T]_{\mathcal{B}}) = \det([T]_{\mathcal{C}}).$$

Das bedeutet genau, dass der Wert $\det(T)$ nicht von der Wahl der Basis abhängt, was zu zeigen war. □

Wie bei Matrizen misst die Determinante, ob ein Endomorphismus invertierbar ist oder nicht. Hier ist die grosse Schwester von Proposition 3.93 (A.K.A. ugly Lemma):

Korollar 4.58. *Seien $T \in \mathrm{End}(V)$ wobei $\dim V < \infty$ und $\mathcal{B}$ eine geordnete Basis von V. Die folgenden Aussagen sind äquivalent:*

(1) T ist ein Isomorphismus.

(2) T ist injektiv.

(3) T ist surjektiv.

(4) $[T]_{\mathcal{B}}$ ist invertierbar.

(5) $[T]_{\mathcal{B}}$ erfüllt eine der äquivalenten Bedingungen in Proposition 3.93.

(6) $\det(T) \neq 0$.

Beweis. Die Äquivalenz der ersten drei Aussagen folgt aus dem Rangsatz. Die Abbildung T ist genau dann ein Isomorphismus, wenn $[T]_{\mathcal{B}}$ invertierbar ist. (Dies folgt zum Beispiel aus Korollar 3.85.) Dies ist äquivalent zur Aussage, dass

$$\det([T]_{\mathcal{B}}) \neq 0.$$

Da (per Definition) $\det(T) = \det([T]_{\mathcal{B}})$, folgen alle andere Äquivalenzen. □

Obwohl wir dies im Moment nicht brauchen, möchten wir noch eine andere Grösse erwähnen, die unter Ähnlichkeit invariant ist:

Definition 4.59. Sei $A \in M_{n\times n}(K)$. Die *Spur* von A ist die Summe aller Einträge auf der Hauptdiagonalen. Das heisst,

$$\mathrm{spur}(A) := a_{11} + \ldots + a_{nn}.$$

Lemma 4.60. *Für $A, B \in M_{n\times n}(K)$ gilt*

$$\mathrm{spur}(AB) = \mathrm{spur}(BA).$$

Beweis. Laut der Definition der Matrixmultiplikation gilt

$$\mathrm{spur}(AB) = \sum_{i=1}^{n}(AB)_{ii} = \sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik}b_{ki} = \sum_{k=1}^{n}\sum_{i=1}^{n} b_{ki}a_{ik} = \sum_{k=1}^{n}(BA)_{kk} = \mathrm{spur}(BA).$$

Also folgt das Lemma. □

Korollar 4.61. *Falls* $A, B \in M_{n\times n}(K)$ *ähnlich sind, dann gilt*

$$\mathrm{spur}(A) = \mathrm{spur}(B).$$

Beweis. Sei $P \in \mathrm{GL}_n$ mit $A = P^{-1}BP$. Dann ist

$$\mathrm{spur}(A) = \mathrm{spur}(P^{-1}(BP)) = \mathrm{spur}((BP)P^{-1}) = \mathrm{spur}(B)$$

und das Korollar folgt. □

4.8 Volumen und Orientierung

Wir haben in Abschnitt 4.1 für $A \in M_{2\times 2}(\mathbb{R})$ gesehen:

- Die Abbildung m_A ist orientierungserhaltend genau dann, wenn $\det(A) > 0$.
- Die Abbildung m_A ist orientierungsumkehrend genau dann, wenn $\det(A) < 0$.

Für einige Sätze, die Sie dieses Jahr in der Analysis sehen werden, und auch für später im Studium ist es wichtig, den Begriff der Orientierung für $\mathbb{R}^n$ und sogar für allgemeine endlich-dimensionale reelle Vektorräume zu verstehen.

Orientierung

Sei V ein Vektorraum über $\mathbb{R}$ mit $\dim(V) = n < \infty$. Sei X_V die Menge aller geordneten Basen auf V. Wir definieren folgendermassen eine Äquivalenzrelation auf X_V: Für $\mathcal{B} = (v_1, \ldots, v_n)$ und $\mathcal{C} = (w_1, \ldots, w_n)$ sei $T \in \mathrm{End}(V)$ die eindeutige Abbildung mit $T(v_i) = w_i$ für $i \in \{1, \ldots, n\}$. Wir sagen, dass $\mathcal{B}$ und $\mathcal{C}$ die gleiche Orientierung haben, und schreiben $\mathcal{B} \sim \mathcal{C}$, falls $\det(T) > 0$.

Übung 4.62. Es gilt

$$\mathcal{B} \sim \mathcal{C} \iff \det([\mathrm{Id}_V]^{\mathcal{B}}_{\mathcal{C}}) > 0 \iff \det([\mathrm{Id}_V]^{\mathcal{C}}_{\mathcal{B}}) > 0.$$

(Hinweis: Berechnen Sie $[T]^{\mathcal{B}}_{\mathcal{C}}$ für T wie oben und benutzen Sie $[T]^{\mathcal{C}}_{\mathcal{C}} = [T]^{\mathcal{B}}_{\mathcal{C}}[\mathrm{Id}_V]^{\mathcal{C}}_{\mathcal{B}}$.)

Die obige Relation $\sim$ definiert in der Tat eine Äquivalenzrelation auf X_V und $X_V/\!\sim$ enthält genau zwei Äquivalenzklassen falls $\dim(V) > 0$.

Eine Wahl einer dieser beiden Äquivalenzklassen auf V heisst eine *Orientierung* auf V.

Beispiel 4.63. In $\mathbb{R}^n$ wählt man normalerweise die Äquivalenzklasse der Basis $(e_1, \ldots, e_n)$ und nennt die entsprechende Äquivalenzklasse die *positive Orientierung*. Basen in dieser

Äquivalenzklasse nennt man *positiv orientiert* und Basen in der anderen Äquivalenzklasse nennt man *negativ orientiert.*

In $\mathbb{R}^2$ ist dann (e_1, e_2) positiv orientiert und (e_2, e_1) negativ orientiert. In $\mathbb{R}^3$ ist eine „rechtshändige Basis“ positiv orientiert und eine „linkshändige Basis“ negativ orientiert.

Frage: Angenommen Sie sind eine Ameise auf einem Möbiusband. Wie könnten Sie herausfinden, dass Sie in der Tat auf einem Möbiusband leben, unter Verwendung des Inhalts dieses Abschnittes?

Intermezzo über das Volumen

In Analysis II werden Sie beweisen, dass für gewisse Teilmengen $X \subseteq \mathbb{R}^n$ und $m_A : \mathbb{R}^n \to \mathbb{R}^n$ gilt

$$\mathrm{Vol}_{\mathbb{R}^n}(m_A(X)) = |\det(A)|\,\mathrm{Vol}_{\mathbb{R}^n}(X). \tag{4.16}$$

Hier ist eine ungefähre Skizze, wieso (4.16) stimmen sollte.

Jede vernünftige Volumenfunktion auf $\mathbb{R}^n$ sollte Eigenschaften analog zu (D1)–(D3) erfüllen, aber mit Absolutbeträgen. Dies ist äquivalent dazu, dass (4.16) gilt, wenn A eine Elementarmatrix ist. Zum Beispiel ändert Multiplikation mit der Elementarmatrix $P_{i,j}$ das Volumen nicht oder Multiplikation mit $S_i(\alpha)$ multipliziert das Volumen mit $|\alpha|$. Unter Verwendung dieser Tatsachen kann man (4.16) „beweisen" für alle $A \in \mathrm{GL}_n(\mathbb{R})$:

Schreibe $A = T_1 \cdots T_k$ für $T_1, \ldots, T_k$ Elementarmatrizen. Dann gilt

$$\begin{aligned} \mathrm{Vol}_{\mathbb{R}^n}(m_A(X)) = \mathrm{Vol}_{\mathbb{R}^n}(T_1 \cdots T_k X) &\overset{\textit{Induktion}}{=} |\det(T_1)| \cdots |\det(T_k)| \,\mathrm{Vol}_{\mathbb{R}^n}(X) \\ &= |\det(T_1 \cdots T_k)| \,\mathrm{Vol}_{\mathbb{R}^n}(X) = |\det(A)| \,\mathrm{Vol}_{\mathbb{R}^n}(X). \end{aligned}$$

Für $A \notin \mathrm{GL}_n(\mathbb{R})$ könnte man (4.16) auch „beweisen": In diesem Fall gilt

$$\mathrm{Im}(m_A) = m_A(\mathbb{R}^n) \subsetneq V$$

und daher ist $m_A(\mathbb{R}^n)$ eine Hyperebene mit Dimension $< n$. Solche Hyperebenen sollten Volumen Null haben für jede vernünftige Volumenfunktion auf $\mathbb{R}^n$. Da $\det A = 0$ ist für $A \notin \mathrm{GL}_n(\mathbb{R})$, gilt (4.16) auch in diesem Fall.

Changelog: Kapitel 4

- 2.12: Im Beweis von Proposition 4.6 wurde (D4) (b) zu (D1) (b) geändert.
- 7.12: Im Beweis von Lemma 4.18 wurden einige Typos korrigiert.
- 7.12: In Beispiel 4.25 wurde ganz zu Beginn die Permutation σ angepasst, so dass sie mit dem Rest des Beispiels übereinstimmt.
- 7.12: In Beispiel 4.29 wurde die erste Permutation angepasst.
- 9.12: In Gleichung (4.4) steht jetzt $j-i$ im Nenner statt $j-1$.
- 9.12: In Übung 4.43 wurde $\det A = \det B_1 + \det B_2$ zu $\det A = \det B_1 \cdot \det B_2$ korrigiert.
- 9.12: In Satz 4.44 wurde im zweiten Teil j-te Zeile zu j-te Spalte korrigiert.
- 10.12: Im ersten Teil von Satz 4.44 wurde der Koeffizient a_{ik} bei der letzten Gleichheit in der Summe hinzugefügt.
- 10.12: Gleichung (4.9) wurde angepasst.
- 10.12: In Definition 4.46 stand fälschlicherweise Adjungierte statt Adjunkte.
- 11.12: In Definition 4.17 wurde der Begriff Transposition hinzugefügt.
- 15.12: In Gleichung (4.12) wurde der Term a_{1k} hinzugefügt.
- 15.12: Im Beweis von Satz 4.49 wurde die Definition der Matrix B geändert (i-te Zeile wurde zu j-ter Zeile und j-te Spalte wurde zu i-ter Spalte).
- 17.12: Definition 4.52 und Satz 4.53 wurden allgemeiner für $m \times n$-Matrizen formuliert.
- 24.12: Im Teil (D1) (b) im Beweis von Satz 4.39 wurde $\operatorname{sign}(\sigma)$ in der Summe nach der zweiten Gleichheit hinzugefügt.

Kapitel 5

Eigenvektoren und Eigenwerte

5.1 Einführung

Sei V ein Vektorraum über K und $T \in \mathrm{End}(V)$. Wie wir in Kapitel 0 erwähnt haben, kann man bei T an eine Symmetrie von V denken. Man sucht einen Vektor $v \in V$, auf welchem T besonders einfach agiert. In Kapitel 0 konnten wir die Basis $\mathcal{B} = (\mathcal{F}_{1,\varphi}, \mathcal{F}_{1,\psi})$ für **Fib** finden, auf deren Elemente die Verschiebungsabbildung

$$\begin{aligned} S : \mathbf{Fib} &\to \mathbf{Fib} \\ (a_0, a_1, \ldots) &\mapsto (a_1, a_2, \ldots) \end{aligned}$$

besonders einfach agiert.:

$$S(\mathcal{F}_{1,\varphi}) = \varphi\mathcal{F}_{1,\varphi} \quad \text{und} \quad S(\mathcal{F}_{1,\psi}) = \psi\mathcal{F}_{1,\psi}.$$

Wir mussten aber $\mathcal{F}_{1,\varphi}$ und $\mathcal{F}_{1,\psi}$ „erraten“ und hatten leider keinen Algorithmus zur Verfügung, um diese besonderen Vektoren zu finden. Später haben wir in Beispiel 3.108 gesehen, dass S bezüglich der Basis $\mathcal{B}$ eine besonders einfache Darstellungsmatrix hat, nämlich

$$[S]_{\mathcal{B}} = \begin{pmatrix} \varphi & 0 \\ 0 & \psi \end{pmatrix}. \tag{5.1}$$

Damit kann man die Matrizen S^n für $n \in \mathbb{N}$ problemlos berechnen und eine Formel für Fibonacci-Folgen finden. Dies bezieht sich einfach auf die Tatsache, dass man leicht Potenzen von Diagonalmatrizen berechnen kann. In der Terminologie, die wir in diesem Kapitel entwickeln, bedeutet (5.1), dass S *diagonalisierbar* ist. Diagonalisierung von Endomorphismen hat viele Anwendungen sowohl in der Mathematik als auch in allen möglichen Gebieten der Wissenschaft. Eine grundlegende Anwendung in der Physik ist beispielsweise die Analyse von Schwingungen mit kleinen Auslenkungen. Um zu verstehen, wann ein Endomorphismus diagonalisiert werden kann und wie man dies explizit machen kann, braucht man die Begriffe von Eigenwerten und Eigenvektoren.

5.2 Definitionen

Sei V ein Vektorraum über K.

Definition 5.1. Ein Endomorphismus $T \in \mathrm{End}(V)$, wobei $\dim V = n < \infty$, heisst *diagonalisierbar*, falls es eine Basis $\mathcal{B} = (v_1, \ldots, v_n)$ gibt, so dass

$$[T]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} \tag{5.2}$$

eine Diagonalmatrix ist.

Bemerken Sie, dass (5.2) genau dann gilt, wenn

$$Tv_i = \lambda_i v_i$$

für $i = 1, \ldots, n$. Vektoren mit dieser Eigenschaft sind die Helden dieses Kapitels.

Definition 5.2. Ein Vektor $v \in V$ heisst ein *Eigenvektor* von T, falls $v \neq 0$ und es $\lambda \in K$ gibt, so dass

$$Tv = \lambda v.$$

Der Skalar λ heisst der zum Eigenvektor v zugehörige *Eigenwert*.

Bemerkung 5.3. (1) Man könnte genau so gut Folgendes definieren: Ein Skalar $\lambda \in K$ heisst ein *Eigenwert* von T, falls es $v \in V$ gibt mit $v \neq 0$ und

$$Tv = \lambda v. \tag{5.3}$$

Jeder Vektor v mit (5.3), der ungleich Null ist, heisst ein *Eigenvektor* von T zum Eigenwert λ.

(2) Eigenvektoren sind immer verschieden vom Nullvektor! Eigenwerte können jedoch 0 sein, wie das folgende Beispiel zeigt.

Beispiel 5.4. Sei $T : V \to V$ eine lineare Abbildung mit $\mathrm{Ker}(T) \neq \{0\}$. Dann ist 0 ein Eigenwert von T und jeder Vektor $v \in \mathrm{Ker}(T)$, der ungleich Null ist, ist ein Eigenvektor zum Eigenwert 0. Anders gesagt, gilt

$$\mathrm{Ker}(T) = \{0_V\} \sqcup \{\text{alle Eigenvektoren mit Eigenwert } 0\}.$$

Beispiel 5.5. Man könnte jetzt zu Korollar 4.58 Folgendes hinzufügen:

$$T \text{ ist ein Isomorphismus} \iff 0 \text{ ist kein Eigenwert von } T$$

Beispiel 5.6. Die Eigenwerte von $S\colon \mathbf{Fib} \to \mathbf{Fib}$ sind φ und ψ und $\mathcal{F}_{1,\varphi}$, $\mathcal{F}_{1,\psi}$ sind die entsprechenden Eigenvektoren.

Das folgende Lemma gilt per Definition, aber es ist sehr wichtig dessen Inhalt im Fall einer Matrix zu verstehen.

Lemma 5.7. *Ein Endomorphismus $T \in \mathrm{End}(V)$ ist diagonalisierbar genau dann, wenn eine Basis existiert, die nur aus Eigenvektoren von T besteht.*

Beweis. Die Aussage des Lemmas gilt in der Tat per Definition von Diagonalisierbarkeit, von Eigenvektoren und der Darstellungsmatrix. □

Definition 5.8. (1) Ein Skalar $\lambda \in K$ heisst ein *Eigenwert* einer Matrix $A \in M_{n\times n}(K)$, falls λ ein Eigenwert von $m_A\colon K^n \to K^n$ ist.

(2) Ein Vektor $v \in K^n$ heisst ein *Eigenvektor* einer Matrix $A \in M_{n\times n}(K)$, falls v ein Eigenvektor von $m_A\colon K^n \to K^n$ ist.

(3) A heisst *diagonalisierbar*, falls m_A diagonalisierbar ist.

5.2.1 Definition 5.8 explizit für Matrizen

Der Spezialfall $T = m_A$ ist so wichtig und teilweise auch verwirrend für einige Studenten, dass wir Definition 5.8 noch einmal explizit für Matrizen angeben. Es ist aber wichtig zu wissen, dass dieser Abschnitt streng genommen nicht nötig ist.

Teile (1) und (2) von Definition 5.8 können so geschrieben werden:

Definition 5.9. Sei $A \in M_{n\times n}(K)$. Ein Skalar $\lambda \in K$ heisst ein *Eigenwert* von A, falls es einen Vektor $v \in K^n$ gibt, so dass $v \neq 0$ und

$$Av = \lambda v.$$

Ein Vektor $v \in K^n$ mit dieser Eigenschaft heisst ein *Eigenvektor* zum Eigenwert λ.

Beispiel 5.10. Sei

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

Dann sind $\{\lambda_1, \ldots, \lambda_n\}$ die Eigenwerte von A. Es gilt

$$\begin{aligned}\{v \in V \mid v \text{ ist ein Eigenvektor von } A \text{mit Eigenwert } \lambda_i\} &= \{ae_i \mid a \neq 0\} \\ &= \mathrm{Sp}(e_i) \setminus \{0\}.\end{aligned}$$

Beispiel 5.11. Für $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ gilt

$$A \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Daher ist 1 ein Eigenwert von A und $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ ist ein Eigenvektor von A mit Eigenwert 1. Ausserdem ist $a \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} a \\ -a \end{pmatrix}$ ebenfalls ein Eigenvektor von A mit Eigenwert 1. Ähnlich gilt, dass

$$A \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = 3 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

und daher ist 3 ein Eigenwert von A und

$$\left\{ \begin{pmatrix} a \\ a \end{pmatrix} \mid a \neq 0 \right\}$$

ist die Menge aller Eigenvektoren von A mit Eigenwert 3.

Geometrisch ist es einfach zu verstehen, wie eine Matrix $A \in M_{n\times n}(K)$ auf einem Eigenvektor $v \neq 0$ mit Eigenwert λ agiert: A skaliert v mit einem Faktor λ.

Übung 5.12. Versuchen Sie zu zeigen, dass $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ nur 1 als Eigenwert hat und alle Eigenvektoren mit Eigenwert 1 sind durch

$$\{ae_1 \mid a \neq 0\}$$

gegeben.

Die folgende Definition ist äquivalent zu Definition 5.8 (3):

Definition 5.13. Ein Matrix $A \in M_{n\times n}(K)$ heisst *diagonalisierbar*, falls m_A diagonalisierbar ist. Äquivalent ist A diagonalisierbar, falls $P \in \mathrm{GL}_n(K)$ existiert, so dass

$$P^{-1}AP = D$$

ist, wobei D eine Diagonalmatrix ist.

Dies bietet auch eine hilfreiche Art und Weise, um zu sehen, was Lemma 5.7 im Fall $T = m_A$ aussagt, wie die Leser überprüfen können:

Lemma 5.14 (Lemma 5.7 für $T = m_A$). *Eine Matrix A ist diagonalisierbar genau dann, wenn es eine Basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ von K^n gibt, die nur Eigenvektoren von A enthält.“*

In diesem Fall ist

$$P = [\mathrm{Id}_{K^n}]^{\mathcal{B}}_{\mathcal{E}} = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix},$$

wobei $\mathcal{E}$ die Standard-Basis von K^n ist und

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

wobei λ_i der zugehörige Eigenwert zum Eigenvektor v_i ist.

Hier ist die Verbindung zwischen allgemeinen Endomorphismen und Matrizen:

Übung 5.15. Sei $T \in \mathrm{End}(V)$ mit $\dim V < \infty$ und $\mathcal{B}$ eine Basis von V. Dann gilt

$$\begin{gathered} v \text{ ist ein Eigenvektor von } T \text{ mit Eigenwert } \lambda \\ \Longleftrightarrow \\ [v]_{\mathcal{B}} \text{ ist ein Eigenvektor von } [T]_{\mathcal{B}} \text{ mit Eigenwert } \lambda. \end{gathered}$$

5.2.2 Eigenvektoren mit verschiedenen Eigenwerten sind linear unabhängig

Wir wechseln nun wieder zu dem allgemeinen Fall $T \in \mathrm{End}(V)$. Lemma 5.7 und 5.14 zeigen, dass es eine zentrale Frage ist, ob wir eine Basis aus Eigenvektoren finden können. Die folgende Proposition hilft uns in dieser Sache:

Proposition 5.16. *Sei $T \in \mathrm{End}(V)$ und $\{v_1, \ldots, v_n\}$ Eigenvektoren zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_n$. Dann ist $\{v_1, \ldots, v_n\}$ linear unabhängig.*

Korollar 5.17. *Falls $\dim V = n$ ist und $T \in \mathrm{End}(V)$ n paarweise verschiedene Eigenwerte hat, dann ist T diagonalisierbar.*

Beweis von Proposition 5.16. Wir beweisen die Aussage mit Induktion über n. Für $n = 1$ erinnern Sie sich daran, dass jeder Eigenvektor per Definition verschieden von Null ist. Daher ist $\{v_1\}$ linear unabhängig.

Nehmen wir an, dass die Aussage wahr ist für $n-1$ Vektoren und seien $\{v_1, \ldots, v_n\}$ Eigenvektoren mit paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_n$. Wir müssen zeigen, dass

$$\underbrace{a_1 v_1 + \ldots + a_n v_n = 0}_{(\diamondsuit)} \implies a_1 = \ldots = a_n = 0.$$

Wir wenden T auf $(\diamondsuit)$ an:

$$\begin{aligned} T(a_1 v_1 + \ldots + a_n v_n) &= a_1 T v_1 + \ldots + a_n T v_n \\ &= \underbrace{a_1 \lambda_1 v_1 + \ldots + a_n \lambda_n v_n = T(0) = 0}_{(\star)}. \end{aligned}$$

Wir betrachten jetzt $(\star) - \lambda_n(\diamondsuit)$:

$$a_1 \lambda_1 v_1 + \ldots + \cancel{a_n \lambda_n v_n} - (\lambda_n a_1 v_1 + \ldots + \cancel{\lambda_n a_n v_n}) = 0 - \lambda_n 0 = 0$$

beziehungsweise

$$a_1(\lambda_1 - \lambda_n) v_1 + \ldots + a_{n-1}(\lambda_{n-1} - \lambda_n) v_{n-1} = 0.$$

Laut der Induktionsannahme sind $v_1, \ldots, v_{n-1}$ linear unabhängig und daher ist

$$a_i(\lambda_i - \lambda_n) = 0$$

für alle $i = 1, \ldots, n-1$. Da $\lambda_1, \ldots, \lambda_n$ paarweise verschieden sind, folgt $\lambda_i - \lambda_n \neq 0$ für alle $i = 1, \ldots, n-1$. Also folgt, dass

$$a_1 = \ldots = a_{n-1} = 0.$$

Wenn wir dies in $(\diamondsuit)$ einsetzen, dann folgt auch $a_n = 0$, was den Beweis beendet. $\square$

Hier ist eine nette Anwendung von Proposition 5.16:

Beispiel 5.18. Die Menge

$$A = \{F_t = (1, t, t^2, t^3, \ldots) \mid t \in \mathbb{R}\} \subseteq \mathbb{R}^\infty$$

ist eine linear unabhängige Menge der Kardinalität $\mathbb{R}$. Wieso? Es gibt verschiedene Wege dies zu zeigen. Wir präsentieren hier jedoch dieses tolle Argument:

Wir betrachten die Verschiebungsabbildung

$$\begin{aligned} S : \mathbb{R}^\infty &\to \mathbb{R}^\infty \\ (a_1, a_2, a_3, \ldots) &\mapsto (a_2, a_3, \ldots). \end{aligned}$$

Übung 5.19. Zeigen Sie, dass A keine Basis von $\mathbb{R}^\infty$ ist. (Hinweis: Was ist die Wachstumsrate von Linearkombinationen von Elementen von A?)

Zurück zum Beispiel: Der Vektor F_t ist ein Eigenvektor von S mit Eigenwert t. Laut Proposition 5.16 ist also jede endliche Teilmenge von A linear unabhängig, was per Definition bedeutet, dass A linear unabhängig ist. Es folgt $\dim \mathbb{R}^\infty \geq |\mathbb{R}|$. Da $\mathbb{R}^\infty$ und $\mathbb{R}$ dieselbe Kardinalität haben, folgt $\dim \mathbb{R}^\infty = |\mathbb{R}|$.

Zwei grundlegende Fragen bleiben offen:

(1) Wie finden wir alle möglichen Eigenwerte zu einem gegebenen Endomorphismus $T \in \operatorname{End}(V)$ (bzw. zu einer gegebenen Matrix $A \in M_{n\times n}(K)$)?

(2) Nachdem wir wissen, dass $\lambda \in K$ ein Eigenwert von T ist (bzw. von A). Wie finden wir alle Eigenvektoren mit Eigenwert λ?

Beide Fragen werden wir im nächsten Abschnitt beantworten.

5.3 Das charakteristische Polynom

Die Antwort auf die zweite Frage (2) ist eigentlich sehr einfach.

Definition 5.20. Sei $T \in \operatorname{End}(V)$ und λ ein Eigenwert von T. Der *Eigenraum* von T bezüglich λ ist

$$\begin{aligned}\operatorname{Eig}_T(\lambda) &:= \operatorname{Ker}(T - \lambda \operatorname{Id}_V)\\ &= \{v \in V \mid (T - \lambda \operatorname{Id}_V)v = 0\}\\ &= \{v \in V \mid Tv = \lambda v\}.\end{aligned}$$

Falls $A \in M_{n\times n}(K)$ ist, dann ist $\operatorname{Eig}_A(\lambda) := \operatorname{Eig}_{m_A}(\lambda)$.

Übung 5.21. Machen Sie die Definition von $\operatorname{Eig}_A(\lambda)$ explizit! (Die Lösung ist in der Fussnote[1])

Lemma 5.22. *Es gilt*

(1) Der Eigenraum $\operatorname{Eig}_T(\lambda)$ *ist ein Untervektorraum.*

(2) Es gilt $\operatorname{Eig}_T(\lambda) \neq \{0_V\}$ *genau dann, wenn* λ *ein Eigenwert von* T *ist. Allgemeiner gilt*

$$\operatorname{Eig}_T(\lambda) = \{0_V\} \sqcup \{\textit{alle Eigenvektoren mit Eigenwert } \lambda\}. \tag{5.4}$$

[1] Es gilt $\operatorname{Eig}_A(\lambda) = \operatorname{Ker}(A - \lambda I_n) = \{v \in K^n \mid Av = \lambda v\}$.

(3) Falls $\lambda_1 \neq \lambda_2$ ist, dann gilt

$$\mathrm{Eig}_T(\lambda_1) \cap \mathrm{Eig}_T(\lambda_2) = \{0_V\}.$$

Beweis. (1) Der Kern einer linearen Abbildung ist immer ein Untervektorraum.

(2) Jeder von Null verschiedene Vektor in $\mathrm{Eig}_T(\lambda)$ erfüllt

$$Tv = \lambda v$$

und daher ist so ein Vektor per Definition ein Eigenvektor von T mit Eigenwert λ. Also folgt Gleichung (5.4).

(3) Dies folgt aus Proposition 5.16: Falls $v \in \mathrm{Eig}_T(\lambda_1) \cap \mathrm{Eig}_T(\lambda_2)$, dann ist $\{v, v\}$ linear unabhängig laut Proposition 5.16, was ein Widerspruch ist.

Hier ist ein anderer direkter Beweis: Für $v \in \mathrm{Eig}_T(\lambda_1) \cap \mathrm{Eig}_T(\lambda_2)$ gilt

$$\lambda_1 v = Tv = \lambda_2 v$$

und daher ist $(\lambda_1 - \lambda_2)v = 0$, was $v = 0$ impliziert, da $\lambda_1 - \lambda_2 \neq 0$.

□

Korollar 5.23. *Sei $T \in \mathrm{End}(V)$. Es gilt*

$$\lambda \textit{ ist ein Eigenwert von } T \iff \mathrm{Ker}(T - \lambda \, \mathrm{Id}_V) \neq 0.$$

Dieses Korollar ist unser Schlüssel für die Antwort auf Frage (2) zumindest, wenn $\dim V < \infty$.

Definition 5.24 (Charakteristisches Polynom)**.** Seien V ein Vektorraum über K mit $\dim V < \infty$, $T \in \mathrm{End}(V)$ und x eine Variable. Wir definieren das *charakteristische Polynom* von T als[2]

$$p_T(x) = \det(T - x \, \mathrm{Id}_V).$$

Übung 5.25. Es gilt natürlich $p_A(x) = p_{m_A}(x)$. Überzeugen Sie sich davon.

[2] Erinnern Sie sich an Definition 4.56.

Beispiel 5.26. Sei $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ dieselbe Matrix wie in Beispiel 5.11. Dann ist

$$\begin{aligned} p_A(x) = \det\left(\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} - x\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) &= \det\begin{pmatrix} 2-x & 1 \\ 1 & 2-x \end{pmatrix} \\ &= (2-x)^2 - 1 \qquad\qquad = x^2 - 4x + 3 \\ &= (x-1)(x-3). \end{aligned}$$

Die Nullstellen von $p_A(x)$ sind genau die Eigenwerte von A, die wir in Beispiel 5.11 gefunden haben.

Das war kein Zufall, wie die nächste Proposition zeigt.

Proposition 5.27. *Sei V ein Vektorraum über K und $T \in \mathrm{End}(V)$ mit $\dim V < \infty$. Die Nullstellen von $p_T(x)$ in K sind genau die Eigenwerte von T.*

Beweis. Laut Korollar 5.23 ist $\lambda \in K$ ein Eigenwert von T genau dann, wenn

$$\mathrm{Ker}(T - \lambda\,\mathrm{Id}_V) \neq \{0\}.$$

Laut Korollar 4.58 ist dies äquivalent zu $\det(T - \lambda\,\mathrm{Id}_V) = 0$. Das heisst zu $p_T(\lambda) = 0$. □

Bemerkung 5.28. Für eine Matrix $A \in M_{n\times n}(K)$ lautet diese Proposition:

$$\lambda \in K \text{ ist ein Eigenwert von } A \iff p_A(\lambda) = 0.$$

Beispiel 5.29. Jetzt können wir Übung 5.12 „algorithmisch". Wir suchen alle möglichen Eigenvektoren von $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ um herauszufinden, ob $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ diagonalisierbar ist oder nicht. Laut Proposition 5.27 sind alle möglichen Eigenwerte die Nullstellen von

$$\begin{aligned} p_A(x) = \left|\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} - x\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right| &= \begin{vmatrix} 1-x & 1 \\ 0 & 1-x \end{vmatrix} \\ &= (1-x)^2 - 0 \\ &= (1-x)^2. \end{aligned}$$

Daher ist 1 der einzig mögliche Eigenwert von A. Es folgt, dass die einzig möglichen Eigenvektoren von A, diejenigen Vektoren sind die nicht Null sind und in

$$\mathrm{Ker}(A - 1\cdot I_2) = \mathrm{Ker}\left(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right) = \left\{\begin{pmatrix} a \\ 0 \end{pmatrix} \,\middle|\, a \in K\right\}$$

liegen. Daher sind

$$\left\{ \begin{pmatrix} a \\ 0 \end{pmatrix} \middle| \; a \neq 0 \right\}$$

die möglichen Eigenvektoren von A. Insbesondere existiert keine Basis von K^2, die nur aus Eigenvektoren von A besteht. Also ist A nicht diagonalisierbar.

Bemerkung 5.30. Das letzte Beispiel mag zwar einfach sein, man kann jedoch viel von ihm lernen. Stellen Sie sicher, dass sie alle Details in diesem Beispiel verstehen.

Wir können also ein Rezept herleiten: Sei $T \in \operatorname{End}(V)$ mit $\dim V = n < \infty$. Wenn man alle möglichen Eigenvektoren von T finden will (beispielsweise, um zu wissen ob T diagonalisierbar ist oder nicht), macht man Folgendes:

1. Man berechnet alle Nullstellen des charakteristischen Polynoms. Bezeichnen wir diese mit $\lambda_1, \dots, \lambda_k$ (wir werden zeigen $0 \leq k \leq n$). Dann sind das alle Eigenwerte von T.

2. Für jedes $i = 1, \dots, k$ berechnet man $\operatorname{Eig}_T(\lambda_i) = \operatorname{Ker}(T - \lambda_i \operatorname{Id}_V)$. Alle Vektoren in $\operatorname{Eig}_T(\lambda_i)$, die verschieden von Null sind, sind Eigenvektoren von T mit λ_i als Eigenwert.

Sei $\mathcal{B}_i$ eine Basis von $\operatorname{Eig}_T(\lambda_i)$ für alle $i = 1, \dots, k$. Dann werden wir (mittels Proposition 5.16) sehen, dass die Vereinigung von $\mathcal{B}_i$ für alle $i = 1, \dots, k$ eine linear unabhänige Menge ist. Dann folgern wir, dass T genau dann diagonalisierbar ist, wenn diese Vereinigung eine Basis von V ist.

All dies besprechen wir im ersten Teil der Linearen Algebra II!

Changelog: Kapitel 5

- 17.12: Die Aussage in Lemma 5.14 wurde zu einer „genau dann, wenn“ Aussage verallgemeinert.
- 17.12: Im Beweis von Proposition 5.16 wurde $(\diamondsuit)-\lambda_n(\star)$ zu $(\star)-\lambda_n(\diamondsuit)$ korrigiert.
- 17.12: In Definition 5.20 wurde I_n zu Id_V korrigiert.

Literaturverzeichnis

[1] Menny Aka, *Einführung in die Lineare Algebra*, (2020), Vorlesungsnotizen ETH Zürich.

[2] ———, *Ergänzungen zu Kapitel 0 im Fischer*, (2020), Vorlesungsnotizen ETH Zürich.

[3] Menny Aka, Manfred Einsiedler, and Thomas Ward, *A Journey Through The Realm of Numbers From Quadratic Equations to Quadratic Reciprocity*, (2020), https://www.springer.com/gp/book/9783030552329.

[4] Vladimir Igorevič Arnold, *Lectures and problems: A gift to young mathematicians*, vol. 17, American Mathematical Soc., 2015.

[5] S. Axler, *Linear Algebra Done Right*, Springer, 2015, https://www.springer.com/gp/book/9783319110790.

[6] G. Fischer, *Lineare Algebra: Eine Einführung für Studienanfänger*, Grundkurs Mathematik, Springer, 2008.

[7] P.R. Halmos, *Finite-Dimensional Vector Spaces*, Courier Dover Publications, 2017, https://www.springer.com/gp/book/9780387900933.

[8] Christian Karpfinger and Hellmuth Stachel, *Lineare algebra*.

[9] Emmanuel Kowalski, *Linear Algebra*, (2016), Vorlesungsnotizen ETH Zürich.

[10] S. H. Friedberg, A. J. Insel, L. E. Spence, *Linear Algebra: Pearson New International Edition*, Pearson Higher Ed., 2013.

[11] J. P. Serre, *A course in arithmetic*, vol. 7, Springer Service & Business Media, 2012, https://link.springer.com/book/10.1007%2F978-1-4684-9884-4.