$f : \mathbb{R}^n \to \mathbb{R}^m$

is called differentiable

at x_0, with differential

$u : \mathbb{R}^n \to \mathbb{R}^m$, if

$\exists$ a linear map $u : \mathbb{R}^n \to \mathbb{R}^m$

such that

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - u(x - x_0)}{\|x - x_0\|} = 0.$$

ie. The affine linear map

$$g(x) = f(x_0) + u(x - x_0)$$

approximates f well around x_0

in the sense that

$f(x) - g(x) =$ Error goes to 0

faster than $\|x - x_0\| \to 0$.

The differential is denoted by

$d_{x_0} f : \mathbb{R}^n \to \mathbb{R}^m$ or

$(df)(x_0)$

Thm If f is diff at x_0

then ① f is continous at x_0

② $\frac{\partial f_i}{\partial x_j}(x_0)$ exists $\forall \, 1 \le i \le m$, $1 \le j \le n$.

③ The matrix of the lin. map $df(x_0) : \mathbb{R}^n \to \mathbb{R}^m$, $x \to Ax$

is given by the Jacobi Matrix

ie $A = \left(\frac{\partial f_i}{\partial x_j}(x_0) \right)_{1 \le i \le m, 1 \le j \le n}$ //

Thm If $f: \mathbb{R}^n \to \mathbb{R}^m$ has all partial derivatives defined and $\frac{\partial f_i}{\partial x_j}: \mathbb{R}^n \to \mathbb{R}^m$ are continuous. Then f is differentiable.

Thm 1) If $f, g: \mathbb{R}^n \to \mathbb{R}^m$ diff then so is $f \pm g$

and $d_{x_0} f \pm d_{x_0} g = d_{x_0}(f \pm g)$

2) If $f, g: \mathbb{R}^n \to \mathbb{R}$ diff then so is fg and f/g (if $g \neq 0$)

$d_{x_0}(fg) = (d_{x_0} f)\, g(x_0) + f(x_0)\, d_{x_0} g$

The **tangent space** at x_0 is the graph of the affine linear map

$$g(x) = f(x_0) + (d_{x_0} f)(x - x_0).$$

ie. $\{(x, g(x)) \in \mathbb{R}^n \times \mathbb{R}^m \mid g(x) = f(x_0) + d_{x_0} f(x - x_0)\}$

Ex: If $f: \mathbb{R}^2 \to \mathbb{R}$ diff at $\bar{x}_0 = (a, b)$. Then

$$d_{(a,b)} f : \mathbb{R}^2 \to \mathbb{R}$$

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \left(\frac{\partial f}{\partial x}(a,b), \frac{\partial f}{\partial y}(a,b)\right)\begin{pmatrix} x \\ y \end{pmatrix}$$

$$(d_{(a,b)} f)\begin{pmatrix} x \\ y \end{pmatrix} = A_1 x + A_2 y$$

where $A_1 = \frac{\partial f}{\partial x}(a,b)$ $\quad A_2 = \frac{\partial f}{\partial y}(a,b)$

$$\boxed{g(x,y) = f(a,b) + A_1(x-a) + A_2(y-b)}$$

Ex: $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$

$(x,y) \longmapsto \sqrt{x^2+y^2}$

$(x_0,y_0) = (3,4)$

$f(3,4) = 5$

$$\nabla f(3,4) = \begin{pmatrix} \frac{\partial f}{\partial x}(3,4) \\ \\ \frac{\partial f}{\partial y}(3,4) \end{pmatrix} = \begin{pmatrix} 3/5 \\ \\ 4/5 \end{pmatrix}.$$

$$\frac{\partial f}{\partial x} = \frac{1}{2} \cdot \frac{2x}{\sqrt{x^2+y^2}}$$

$$\frac{\partial f}{\partial y} = \frac{1}{2} \cdot \frac{2y}{\sqrt{x^2+y^2}}$$

$$g(x,y) = f(x_0,y_0) + \nabla f \cdot (x-3, y-4).$$

$$g(x,y) = 5 + \frac{3}{5}(x-3) + \frac{4}{5}(y-4).$$

What is the analog of the chain rule

Thm. (Chain rule) let $X \subseteq \mathbb{R}^n$ open, $Y \subseteq \mathbb{R}^m$ open and $f : X \to Y$ and $g : Y \to \mathbb{R}^p$ diff. functions. Then $g \circ f : X \to \mathbb{R}^p$ is differentiable.

If $x_0 \in X$, $f(x_0) = y_0$

$$d_{x_0}(g \circ f) = dg(f(x_0)) \circ df(x_0).$$

$$= d(g \circ f)(x_0).$$

composition of the 2 linear maps corresponding to the differentials

In particular for the Jacobi Matrix of $g \circ f$ satisfies

$$\boxed{J_{g \circ f}(x_0) = J_g(f(x_0)) \cdot J_f(x_0).}$$

$$\begin{tikzcd} \mathbb{R}^n \supseteq X \arrow[r, "f"] \arrow[rr, bend left, "g\circ f"] & Y \subset \mathbb{R}^m \arrow[r, "g"] & \mathbb{R}^p \end{tikzcd}$$

$$x \xrightarrow{A = df} Ax \xrightarrow{B} By.$$

$Ax = y$, $\quad B = dg(y)$

$A = df(x)$

$$d(g\circ f) = BA$$

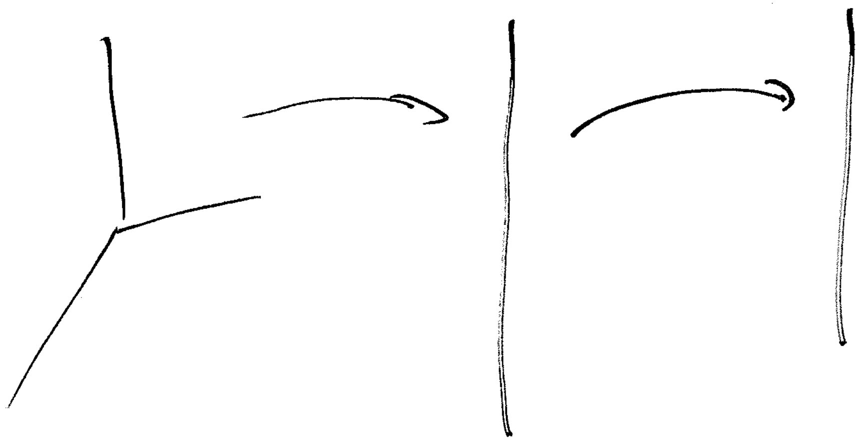
<u>Ex:</u> ① $n = m = p = 1$.

$$\begin{tikzcd} \mathbb{R} \arrow[r, "f"] \arrow[rr, bend right, "g\circ f"'] & \mathbb{R} \arrow[r, "g"] & \mathbb{R} \end{tikzcd}$$

$$(g\circ f)'(x) = g'(f(x)) \cdot f'(x).$$

② $\mathbb{R}^n \xrightarrow{f} \mathbb{R} \xrightarrow{g} \mathbb{R}$

$$g\circ f : \mathbb{R}^n \longrightarrow \mathbb{R}.$$



$$\mathbb{R}^3 \longrightarrow \mathbb{R} \longrightarrow \mathbb{R}.$$

$$d(g\circ f)(x_0) = dg(f(x_0)) \cdot f'(x_0).$$

Ex ① $h : \mathbb{R}^2 \longrightarrow \mathbb{R}$

$$(x,y) \longmapsto e^{xy}$$

$$\boxed{dh(x_0,y_0) = \left(y e^{xy}, x e^{xy}\right)}$$

or we can ~~fo~~ view h as a composition of 2 functions

$f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ $\quad$ $g : \mathbb{R} \longrightarrow \mathbb{R}$

$(x,y) \longrightarrow xy$ $\quad$ $t \longrightarrow e^t$

$$(g \circ f)(x,y) = e^{xy}$$

Using chain rule

$$d(g \circ f) = \underbrace{dg(f(x,y))}\,.\, df(x,y).$$

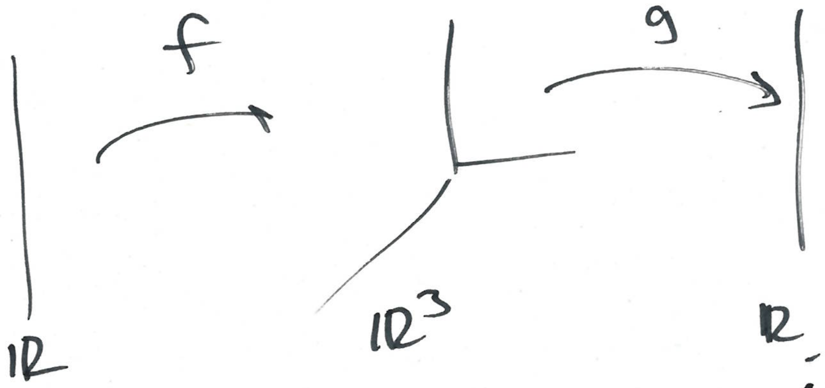
$$e^{xy}\cdot(y, x)$$

$$= \boxed{\left(e^{xy} y, \; x e^{xy}\right)}$$

③. $t \longrightarrow \left(f_1(t), f_2(t) \cdots f_n(t)\right)$

$f : \mathbb{R} \longrightarrow \mathbb{R}^n$.

$g : \mathbb{R}^n \longrightarrow \mathbb{R}$



$$g \circ f : \mathbb{R} \longrightarrow \mathbb{R}$$

$$d(g \circ f) = \underbrace{dg(f(t_0))}\cdot f'(t_0)$$

$$\left(\frac{\partial g}{\partial x_1}(f(t)), \frac{\partial g}{\partial x_2} \cdots \frac{\partial g}{\partial x_n}\right) \cdot \begin{pmatrix} f_1'(t) \\ f_2'(t) \\ \vdots \\ f_n'(t) \end{pmatrix}$$

We can apply this to the following

$$f : \mathbb{R} \to \mathbb{R}^2$$
$$t \to (f_1(t), f_2(t))$$

$$M : \mathbb{R}^2 \longrightarrow \mathbb{R}$$
$$(x,y) \longrightarrow xy$$

$$dM(x,y) = (y, x)$$

$$(M \circ f)(t) = f_1(t) f_2(t)$$

$$d(M \circ f)(t) = dM(f(t)) \cdot f'(t)$$

$$= dM(f_1(t), f_2(t)) \cdot \begin{pmatrix} f_1'(t) \\ f_2'(t) \end{pmatrix}$$

$$= (f_2(t), f_1(t)) \cdot \begin{pmatrix} f_1'(t) \\ f_2'(t) \end{pmatrix}.$$

$$= f_1'(t) \cdot f_2(t) + f_1(t) \cdot f_2'(t)$$

Product rule of f-variable analysis.

$f(x_0) = y_0$

(4)

$$\begin{array}{ccccc} \mathbb{R}^n & \xrightarrow{f} & \mathbb{R}^m & \xrightarrow{g} & \mathbb{R} \\ (x_1, \dots, x_n) & \longrightarrow & y_1, \dots, y_m & \longrightarrow & t \end{array}$$

$g \circ f$

$$J_{g\circ f}(x_0) = \left(\frac{\partial (g\circ f)}{\partial x_1}, \frac{\partial (g\circ f)}{\partial x_2}, \dots \dots \frac{\partial (g\circ f)}{\partial x_n} \right)$$

$$= \underbrace{J_g(f(x_0))}\cdot J_f(x_0).$$

$$\left(\frac{\partial g}{\partial y_1}(y_0), \dots \frac{\partial g}{\partial y_m}(y_0) \right)_{1\times m} \cdot \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

This gives for example

$$\frac{\partial (g\circ f)}{\partial x_1} = \frac{\partial g}{\partial y_1}(y_0)\cdot \frac{\partial f_1}{\partial x_1}$$

$$+ \frac{\partial g}{\partial y_2}(y_0)\cdot \frac{\partial f_2}{\partial x_1}(x_0)$$

$$+ \dots$$

$$+ \frac{\partial g}{\partial y_m}(y_0)\cdot \frac{\partial f_m}{\partial x_1}(x_0).$$

$f: \mathbb{R}^n \to \mathbb{R}$

Rk If f is diff at x_0 $\Rightarrow$ partial derivatives of f exists w.r.t each variable.

Defn

In general given any $v \in \mathbb{R}^n$, we define the directional derivative of f at x_0 in the direction of v as the derivative at $t=0$ of

$$g(t) := f(x_0 + tv)$$

if it exists.

we write

$$d_v f(x_0), \quad df(v; x_0)$$

$$= \lim_{t \to 0} \frac{f(x_0+tv) - f(x_0)}{t}$$

Thm $f : X \to \mathbb{R}^m$, $X \subseteq \mathbb{R}^n$ diff at $x_0 \in X$ and $v \in \mathbb{R}^n$, $v \neq 0$.

Then the directional derivative of f at x_0 in the direction of v exists and

$$\boxed{\left.\frac{d}{dt} f(x_0+tv)\right|_{t=0} = (d_{x_0} f)(v) = J_f(x_0) \cdot v}$$

Ex $f(x,y) = \begin{pmatrix} x^2+y^2 \\ 2x \\ 2y \end{pmatrix}$

$f : \mathbb{R}^2 \to \mathbb{R}^3$

$df(1,2) : \mathbb{R}^2 \to \mathbb{R}^3$

$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto J_f((1,2)) \begin{pmatrix} x \\ y \end{pmatrix}$

$J_f(x,y) = \begin{pmatrix} 2x & 2y \\ 2 & 0 \\ 0 & 2 \end{pmatrix}$

$J_f(1,2) = \begin{pmatrix} 2 & 4 \\ 0 & 2 \\ 2 & 0 \end{pmatrix}$

Hence

$df(1,2)$ is the lin. map which sends

$\begin{pmatrix} x \\ y \end{pmatrix}$ to $\begin{pmatrix} 2 & 4 \\ 0 & 2 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x+4y \\ 2x \\ 2y \end{pmatrix}$

Say $v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

Directional der. of f at $(1, 2)$

in the direction of v is

$$\frac{d}{dt} f\big((1,2) + t(1,1)\big) \Big|_{t=0}.$$

$$= \big(df(1,2)\big)(v) = \begin{pmatrix} 2 & 4 \\ 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \\ 2 \end{pmatrix}$$

Can verify from

$$\frac{d}{dt}\big(f(1+t, 2+t)\big)$$

$$\frac{d}{dt} \begin{pmatrix} (1+t)^2 + (2+t)^2 \\ 2(1+t) \\ 2(2+t) \end{pmatrix}$$

$$= \begin{pmatrix} 2(1+t) + 2(2+t) \\ 2 \\ 2 \end{pmatrix}$$

at $t=0$ this also

gives $\begin{pmatrix} 6 \\ 2 \\ 2 \end{pmatrix}$

The gradient of a function $f: \mathbb{R}^n \to \mathbb{R}$

$(x_1, \dots x_n) \to f(x_1, \dots x_n)$

I $\quad \nabla f(x_0) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x_0) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x_0) \end{pmatrix}$

points in the direction of "greatest increase"!

For any $v \in \mathbb{R}^n$, $v = v_1 \dots v_n$, the directional derivative $(d_{x_0} f)(v) = J_f(x_0) v$ is

$$= \begin{pmatrix} \frac{\partial f}{\partial x_1}(x_0) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x_0) \end{pmatrix} \cdot (v_1 \dots v_n).$$

$$= \langle \nabla_{x_0} f, v \rangle$$

Suppose $\|v\| = 1$

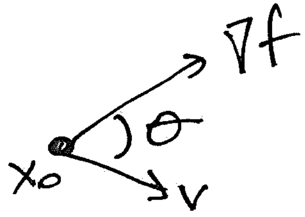
$f: \mathbb{R}^2 \to \mathbb{R}.$

$P = f(x_0, y_0)$ $\quad \nabla f(x_0, y_0) = \begin{pmatrix} \frac{\partial f}{\partial x}(x_0, y_0) \\ \frac{\partial f}{\partial y}(x_0, y_0) \end{pmatrix}$

$$J_f(x_0)(v) = \langle \nabla f, v \rangle$$

$$= \|\nabla f\| \|v\| \cos\theta$$

we used $\|v\| = 1$ $\longrightarrow$ $= \|\nabla f\| \cos\theta$



$\downarrow$

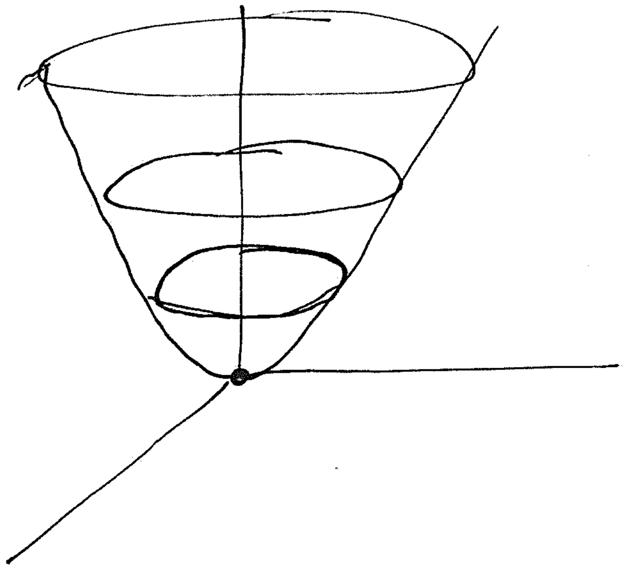
maximized when $\theta = 0$.

=

If we move in the direction of the gradient we maximize the directional derivative.

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}$$

$$(x,y) \longrightarrow x^2 + y^2$$



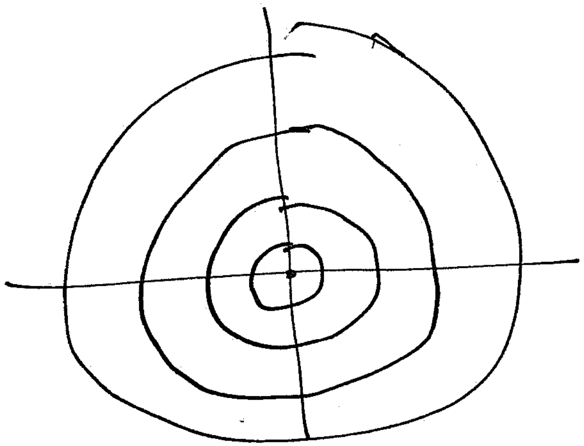
$f(x,y) = \text{constant}$.

$f(x,y) = 0 \Rightarrow (x,y) = (0,0)$

$\{(x,y) : f(x,y) = 1\} = \{(x,y) \mid x^2+y^2 = 1\}$

(II) Consider the level sets of a function $f : \mathbb{R}^n \to \mathbb{R}$.

$$L_c := \{ x \in \mathbb{R}^n \mid f(x) = c \}.$$

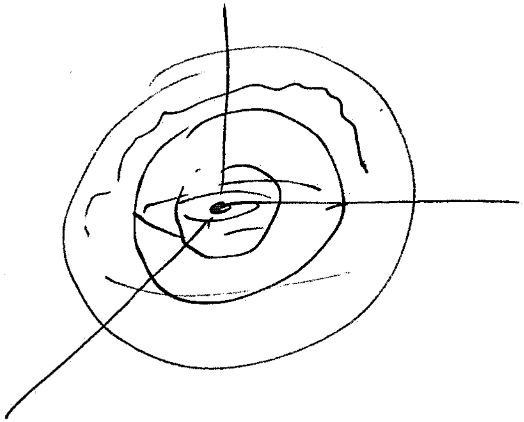


Ex: $f : \mathbb{R}^3 \to \mathbb{R}$

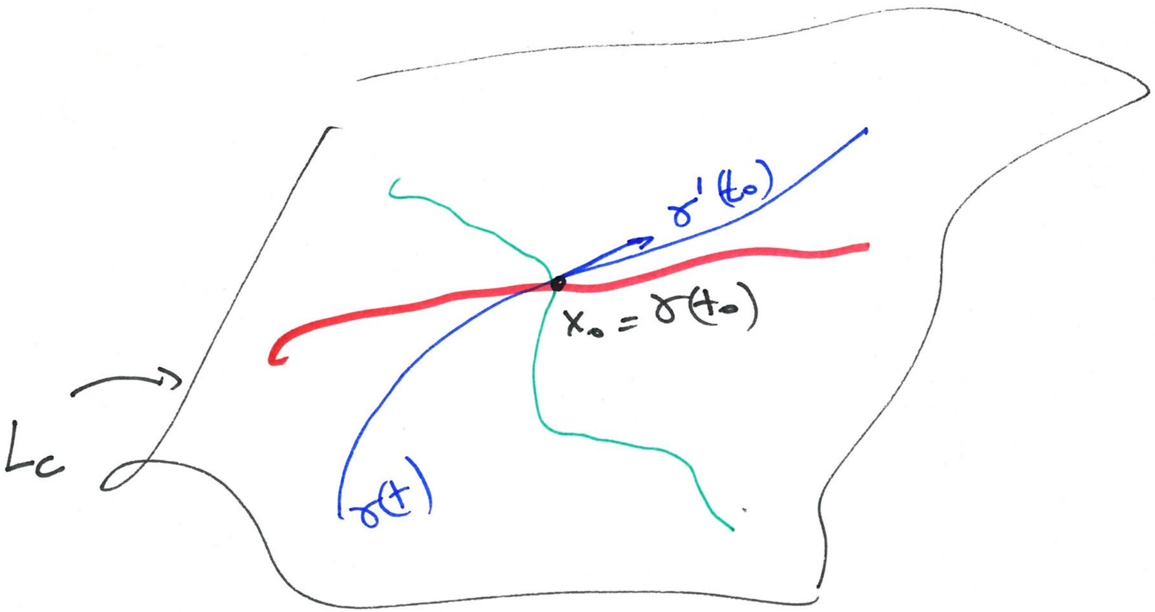
$(x,y,z) \longrightarrow x^2+y^2+z^2$

$L_1 = \{ (x,y,z) \mid x^2+y^2+z^2 = 1 \}$.

sphere of radius 1



let $\gamma : [0,1] \longrightarrow \mathbb{R}^n$ be a diff. curve on the level set of f through x_0 s.t $\gamma(t) \subset L_c$ for some c.



$$L_c = \{ x \in \mathbb{R}^n \mid f(x) = c \} = \text{level set}$$

$f(\gamma(t)) = c$ since $\gamma(t) \subset L_c$

Using the chain rule: we get

$$0 = \left(f(\gamma(t))\right)' = df(\gamma(t)) \cdot \gamma'(t)$$

$$\nabla f(x_0) \cdot \gamma'(t_0) = 0$$

$\Rightarrow$ $\nabla f(x_0)$ is perpendicular to the tangent vektor $\gamma'(t_0)$

This is true for <u>any</u> curve that lies on L_c and goes through x_0

Hence $\nabla f(x_0)$ is perpendicular to the level set

(For $f: \mathbb{R}^2 \to \mathbb{R}$ level sets are curves

For $f: \mathbb{R}^3 \to \mathbb{R}$ level sets are surfaces)

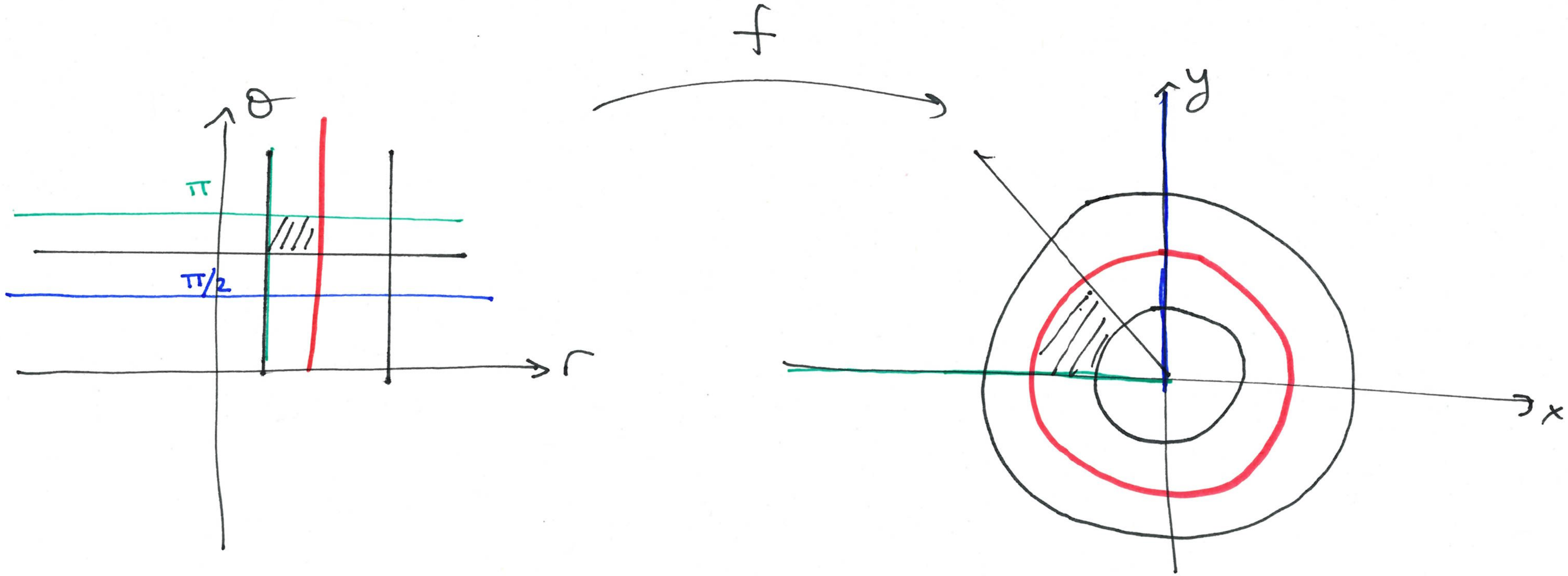
Polar coordinates

$$f : [0,\infty) \times [0, 2\pi) \longrightarrow \mathbb{R}^2$$

$$(r, \theta) \longmapsto \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix} = \begin{pmatrix} f_1(r,\theta) \\ f_2(r,\theta) \end{pmatrix}.$$

$$J_f(r,\theta) = \begin{pmatrix} \frac{\partial f_1}{\partial r} & \frac{\partial f_1}{\partial \theta} \\ \frac{\partial f_2}{\partial r} & \frac{\partial f_2}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}.$$

$$\det J_f(r,\theta) = r(\cos^2\theta + \sin^2\theta) = r$$


θ
π
π/2
r
f
y
x