2.1 – Functions of Several Variables; Graphing Surfaces

<u>Def:</u> A <u>function</u> is a trio:

1. domain – where inputs come from;
2. codomain (target) – where outputs go;
3. rule of assignment – defines exactly one output for each input.

<u>Ex:</u> $\ln : (0,\infty) \to \mathbb{R}$

<u>Ex:</u> $f : \{\text{students}\} \to \mathbb{Z}$, defined by Student ID #.

Similar, but <u>different</u>: $g : \{\text{students}\} \to \mathbb{R}$

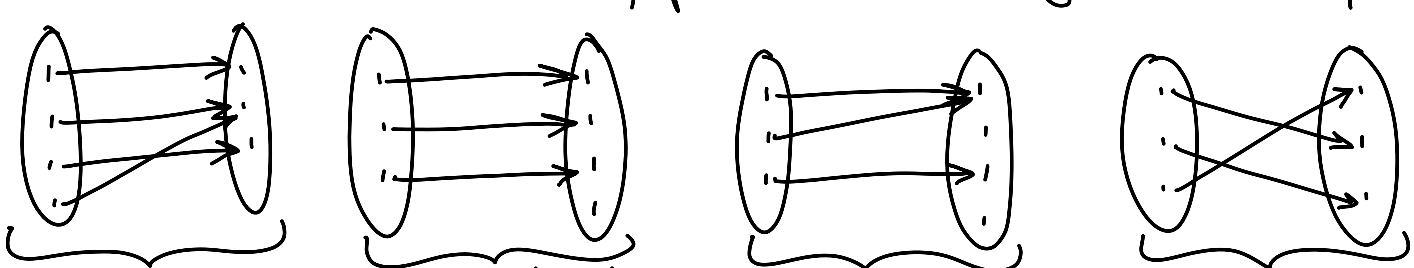
$h : \{\text{students in this course}\} \to \mathbb{Z}$

The image (or range) is the set of actual outputs.

<u>Ex:</u> $f : \mathbb{R} \to \mathbb{R}$ with $f(x) = x^2$ has image $[0,\infty)$.

If $f(x) = y$, then ① y is the image of x;
② x is a pre-image of y.

"onto" = "surjective": Every point in the target has ≥ 1 pre-image.
"1-1" = "injective": Every point in the target has ≤ 1 pre-image.
"invertible" = "bijective": Every point in the target has 1 pre-image.



onto, not 1-1 — 1-1, not onto — neither! — both = invertible

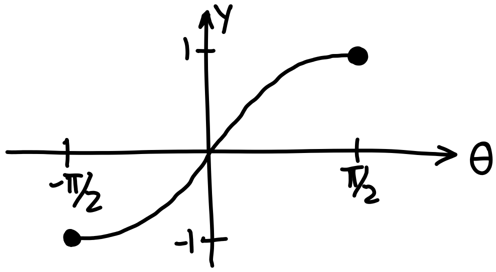
Ex: $f:\mathbb{R}\to\mathbb{R}$, $f(x)=x^2$ is not invertible.

$g:[0,\infty)\to[0,\infty)$, $g(x)=x^2$ is invertible!

$s:[0,\infty)\to[0,\infty)$, $s(x)=\sqrt{x}$ is g^{-1}.

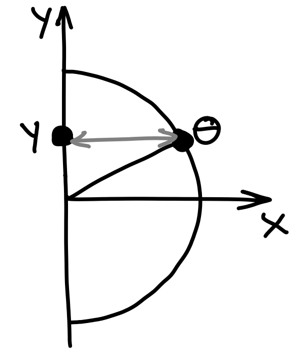
Ex: $f:[-\pi/2,\pi/2]\to[-1,1]$, $f(\theta)=\sin\theta$ is invertible

"arcsin" is the inverse of this function.



$y=\sin\theta$

$\theta=\arcsin y$



For $f:\mathbb{R}^n\to\mathbb{R}^m$, we can write either

$$f(x_1,\dots,x_n)=(y_1,\dots,y_m) \quad \underline{\text{or}} \quad f(\vec{x})=\vec{y}$$

$m=1$: "real-valued" or "scalar-valued"

$m>1$: "vector-valued"

$n=1$: "single variable"

$n>1$: "multivariable"

We can also write the "component functions":

$$y_1:\mathbb{R}^n\to\mathbb{R},\ \dots,\ y_m:\mathbb{R}^n\to\mathbb{R}$$

Ex: Vector-valued multivariable functions are everywhere!

Weather:

$$f\begin{pmatrix}\eta\\ \varepsilon\\ a\\ t\end{pmatrix}=\begin{pmatrix}T\\ P\\ W\\ h\end{pmatrix}$$

lat. long alt. time

temperature pressure windspeed humidity

Visualizing Functions

Graphs

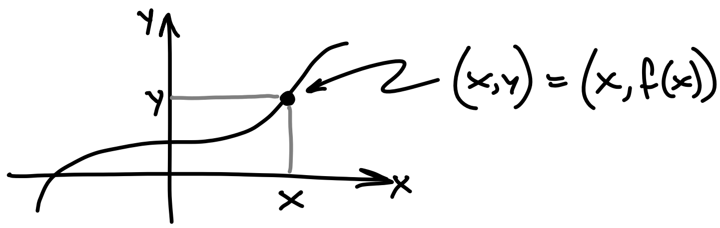
The *graph* of a function is a set of points whose coordinates represent the input and output values.

Def:) The graph of $f:(D\subset\mathbb{R}^n)\to\mathbb{R}^m$, a subset of $\mathbb{R}^{n+m}$, is

$$\left\{ (\vec{x},\vec{y}) = (x_1,\dots,x_n,y_1,\dots,y_m) \;\middle|\; \vec{x}\in D,\ \vec{y}=f(\vec{x}) \right\}$$

Ex:) $f:\mathbb{R}^1\to\mathbb{R}^1$

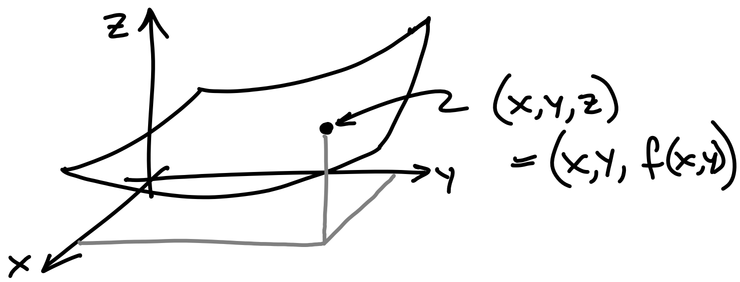
graph $= \{(x,y) \mid y=f(x)\}$



Ex:) $f:\mathbb{R}^2\to\mathbb{R}^1$

Say we view $z=f(x,y)$.

graph $= \{(x,y,z) \mid z=f(x,y)\}$



Ex:) $f:\mathbb{R}^3\to\mathbb{R}^1$, say $w=f(x,y,z)$.

graph $= \{(x,y,z,w) \mid w=f(x,y,z)\}\subset\mathbb{R}^4$ (Of course we can't draw or visualize!)

$$\boxed{\text{Don't conflate a graph with the function itself}}$$

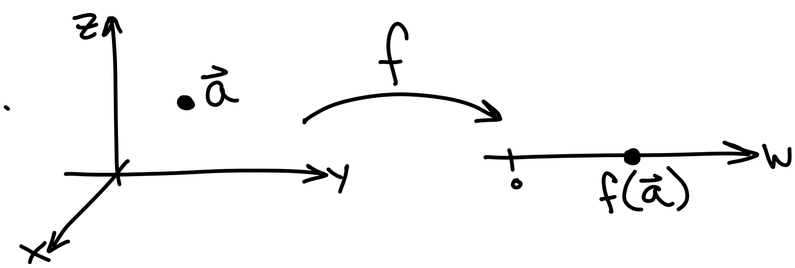
There are other important ways to think of functions geometrically!

"Literal" picture

A function "takes" points from the domain to the target.

Ex:) $f: \mathbb{R}^3 \to \mathbb{R}^1$

Doesn't show the "whole" function.
But does show something!
Better than the graph!



Level sets

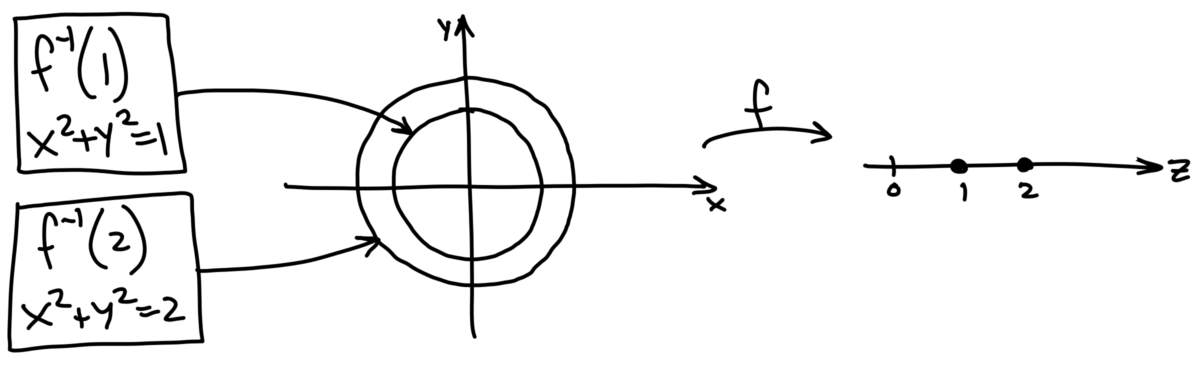
A level set is a pre-image.

Def:) With $f: \mathbb{R}^n \to \mathbb{R}^m$ and $\vec{b} \in \mathbb{R}^m$, the <u>level set</u> $f^{-1}(\vec{b}) \subset \mathbb{R}^n$ is

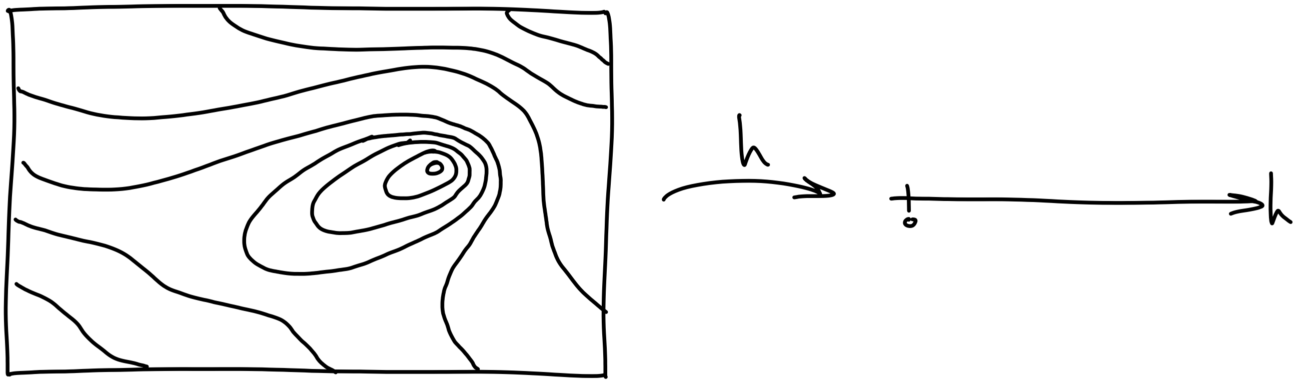
$$f^{-1}(\vec{b}) = \{\vec{x} \in \mathbb{R}^n \mid f(\vec{x}) = \vec{b}\}$$

This is most useful for $f: \mathbb{R}^n \to \mathbb{R}^1$.

Ex:) $f: \mathbb{R}^2 \to \mathbb{R}^1$ with $f(x,y) = x^2 + y^2$.



Ex:) $h: (M \subset \mathbb{R}^2) \to \mathbb{R}^1$, where M is a map and h is the altitude
The level sets are known as "contour lines".



Suppose $f: \mathbb{R}^2 \to \mathbb{R}^1$...

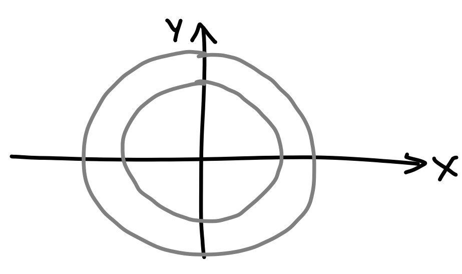
Graph: $f(x,y) = z$ Level curves: $f(x,y) = c$

Horiz. cross section: $\left.\begin{aligned} f(x,y) &= z \\ z &= c \end{aligned}\right\}$ Same!

So, level curves are horizontal cross sections of graphs.

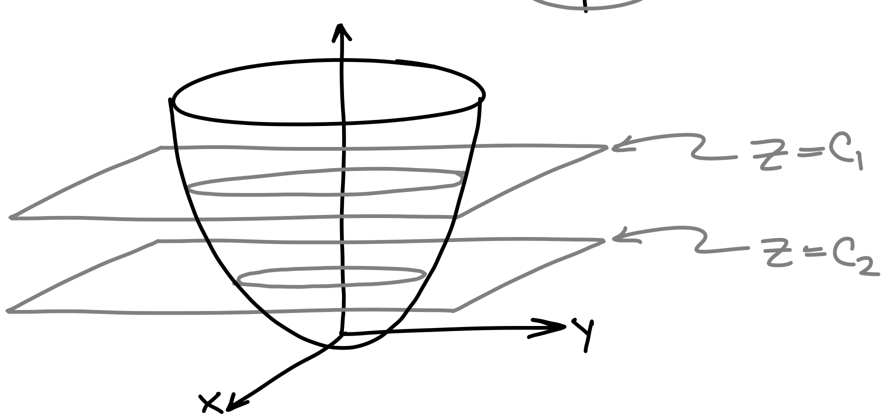
Ex:) $f(x,y) = x^2 + y^2$

Level sets: $x^2 + y^2 = c$
are circles



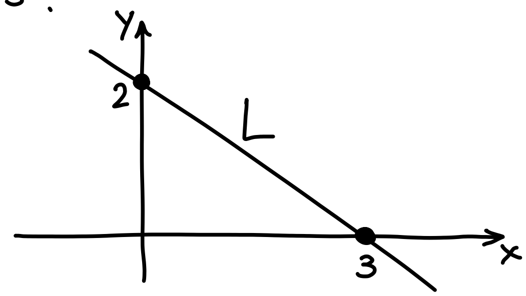
Graph: $x^2 + y^2 = z$

Cross sections $(z = c)$
are
$x^2 + y^2 = c$
same circles!
(projecting to the xy-plane)



We have now several different connections between functions and pictures. Don't confuse them!

Ex:) What functions relate to the line L with equation $2x+3y=6$? In what ways?



① Graph of $f:\mathbb{R}^1 \to \mathbb{R}^1$, $f(x) = \frac{6-2x}{3}$

② A level set of $g:\mathbb{R}^2 \to \mathbb{R}^1$, $g(x,y) = 2x+3y$

③ Parametrized by $\vec{x}:\mathbb{R}^1 \to \mathbb{R}^2$, $\vec{x}(t) = (3t, 2-2t)$

Calculus will look different in these different pictures!

Ex:) From above:

Deriv. of f gives slope...

Derivs. of $g, \vec{x}$ do not — we don't even know yet how to take those derivs!

These derivatives will have important geometric interpretations, different from slope!

Obs.) Every graph is a level set.

Consider $f: \mathbb{R}^n \to \mathbb{R}^1$

Graph of f → $y = f(x_1, \dots, x_n)$

$$\Updownarrow$$

$y - f(x_1, \dots, x_n) = 0$ ← a level set of $g: \mathbb{R}^{n+1} \to \mathbb{R}^1$
$g(x_1, \dots, x_n, y) = y - f(x_1, \dots, x_n)$

Very handy if you have a graph, and need to use the calculus of level sets.

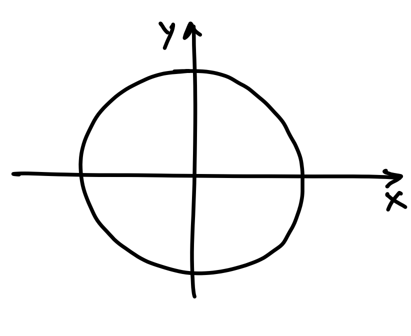
Ex:) $f(x,y) = x^2 + y^2$ has graph $z = x^2 + y^2$

$$\Updownarrow$$

This is the $g^{-1}(0)$ level set of $g(x,y,z) = z - x^2 - y^2$. $\quad z - x^2 - y^2 = 0$

Obs:) Not every level set is a graph.

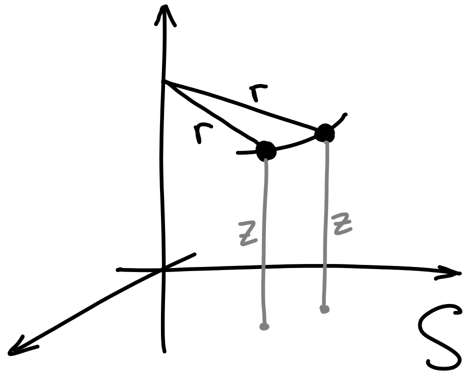
CtrEx:) $x^2 + y^2 = 1$ is $g^{-1}(1)$ of $g(x,y) = x^2 + y^2$



But you can't solve for x or y!
(Fails the vertical line test.)

Rotations

If an equation involves x's and y's <u>only</u> as part of $\sqrt{x^2+y^2}$, then it is <u>rotationally symmetric</u> around the z-axis.



These two points have...

- different x, y, but
- <u>same</u> $\sqrt{x^2+y^2}$, <u>same</u> z!

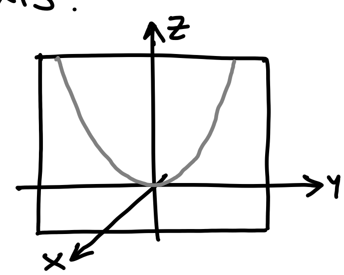
So either they both work, or neither!

(Equiv.: the cylindrical equation has no θ's.)

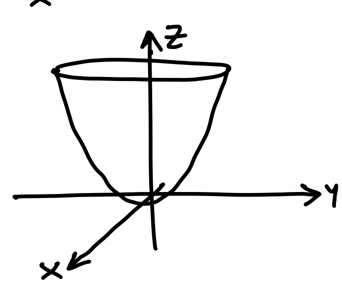
And if a surface has this symmetry, then you can make it by rotating its cross-section in any plane containing the z-axis!

Ex:) What is $z = x^2 + y^2$

1. Rotationally symmetric around z-axis.



2. Cross-section in $x=0$ is $z = y^2$ (parabola).
3. So we can just rotate this parabola around the z-axis.

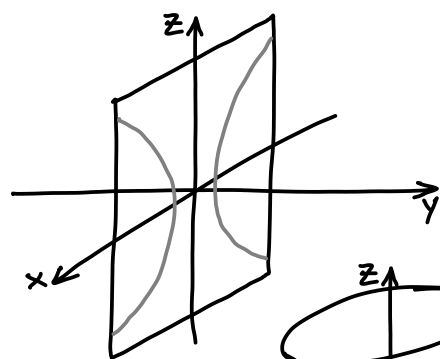


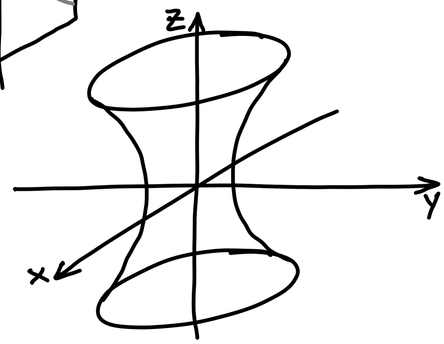
This is a "paraboloid".

Ex: What is $x^2+y^2-z^2=1$?

① Rot. symm. around z-axis.

② In $y=0$ we have $x^2-z^2=1$





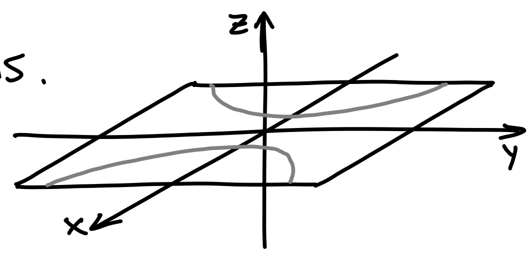
③ Rotating around z-axis gives a "hyperboloid of 1 sheet"

A similar argument can be made for symmetries/rotations around the other axes.

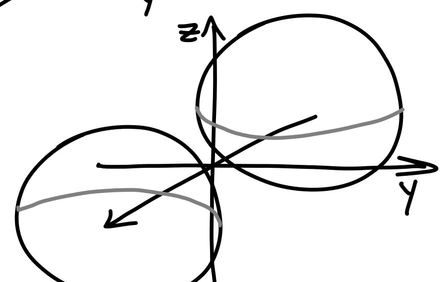
Ex: $x^2-y^2-z^2=1$ $\left(\Longleftrightarrow x^2-(y^2+z^2)=1\right)$

① Rot. symm around x-axis.

② In $z=0$ we have
$$x^2-y^2=1$$



③ Rotating around x-axis:
This is a "hyperboloid of 2 sheets".



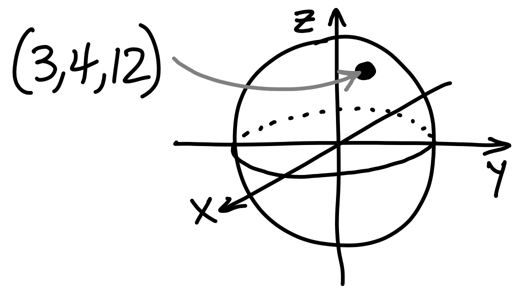
Deformations

What happens to a surface if in its equation you replace all "z"'s with "2z"'s?

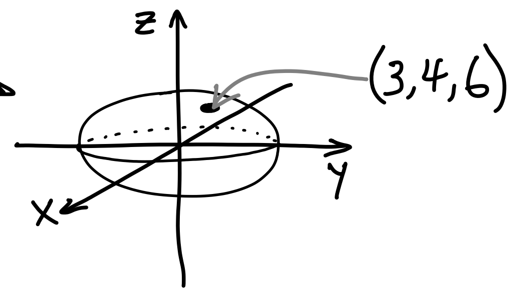
Ex: $x^2+y^2+z^2=13^2$ $\qquad$ $x^2+y^2+(2z)^2=13^2$

$\begin{cases} (3,4,12) \text{ doesn't work} \\ (3,4,6) \text{ does!} \end{cases}$

All z-coordinates are halved!

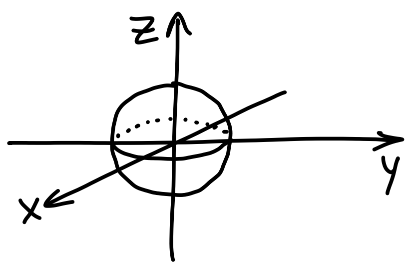


"z" $\mapsto$ "2z"



What you "do" to a variable in an equation, the "opposite" happens to the surface.

Ex: $x^2+y^2+z^2=1$ $\qquad$ $\left(\frac{x}{a}\right)^2+\left(\frac{y}{b}\right)^2+\left(\frac{z}{c}\right)^2=1$

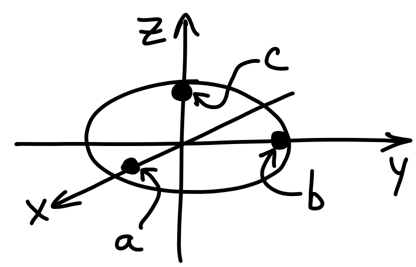


unit sphere

"x" $\mapsto$ "$\frac{x}{a}$"

"y" $\mapsto$ "$\frac{y}{b}$"

"z" $\mapsto$ "$\frac{z}{c}$"



ellipsoid

Ex:) $z = x^2 + y^2$ $\longrightarrow$ $z = \frac{x^2}{a^2} + \frac{y^2}{b^2}$

paraboloid $\qquad$ elliptical paraboloid

Ex:) $x^2 - y^2 - z^2 = 1$ $\longrightarrow$ $\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$

hyperboloid of 2 sheets $\qquad$ elliptical hyperboloid of 2 sheets

See other examples in the book.

2.3 – The Derivative

There are lots of kinds of derivatives of multivariable functions.

Unifying feature:
Relate input changes to output changes.

Partial derivatives

We start with these because:

1. easy to compute
2. easy to relate to single variable derivatives
3. the book does.

Warning – these are not really the best as a primary point of view on derivatives!!

We compute as follows:

1. Hold all variables constant except for one (x_i), view as a single variable function.
2. Take the usual derivative of that function.

Notation: $\frac{\partial f}{\partial x_i} = \frac{\partial}{\partial x_i} f = f_{x_i}$

N.B: "∂" $\neq$ "d" $\neq$ "δ" !   Important!

Ex:) Compute $\frac{\partial}{\partial y}\left(e^x y^2 \cot(x^2 e^z)\right)$.

Importantly x and z are constant, so

$$\underbrace{e^x}_{\text{constant}} y^2 \underbrace{\cot(x^2 e^z)}_{\text{constant}}$$

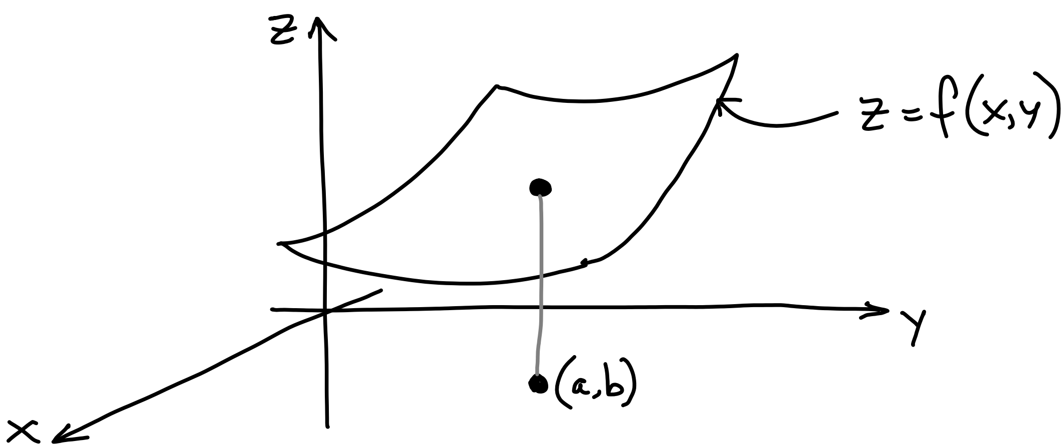
$$\frac{\partial}{\partial y}\left(e^x y^2 \cot(x^2 e^z)\right) = e^x \, 2y \, \cot(x^2 e^z) = 2e^x y \cot(x^2 e^z)$$

Ex:) $\frac{\partial}{\partial x}(x^2+1)^{y^2} = y^2 (x^2+1)^{y^2-1}(2x)$

Ex:) $\frac{\partial}{\partial z}(x^2+y^2)^{z} = (x^2+y^2)^z \ln(x^2+y^2)$

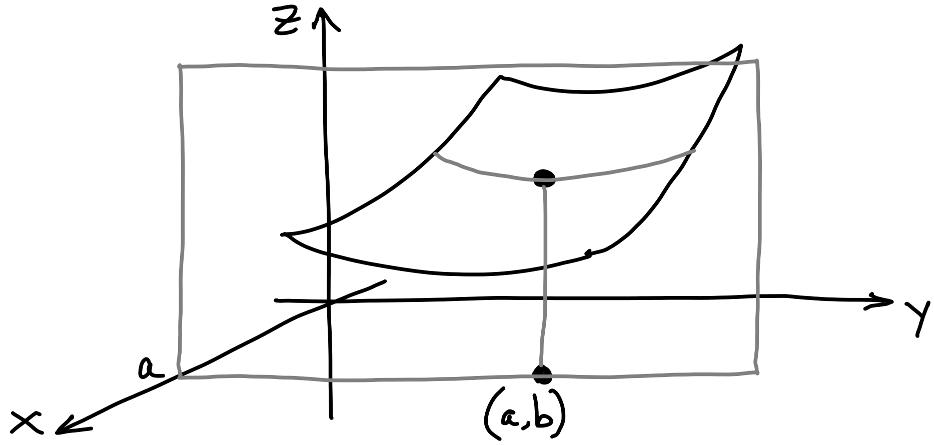
What are the geometric interpretations?
Depends on what kind of picture you are looking at!

Graphs: Say $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = z$. Consider $\frac{\partial f}{\partial y}(a,b)$.

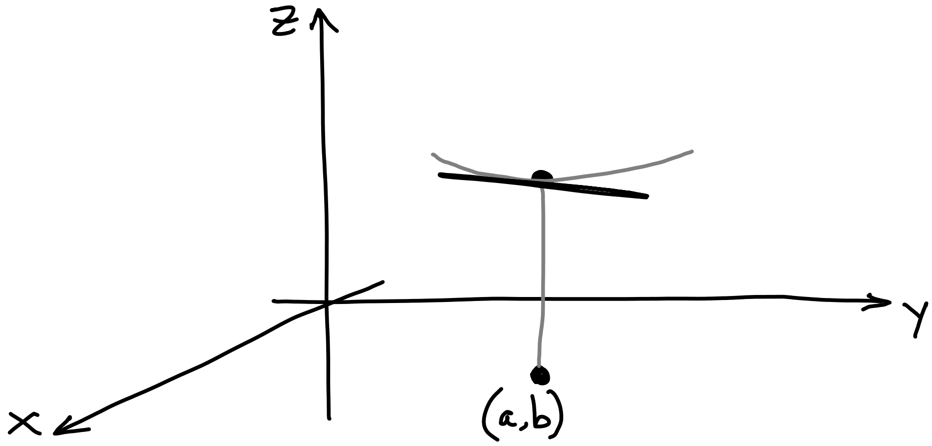


Each step of $\frac{\partial}{\partial y}$ has an existing interpretation!

① Fix all other variables (just x!). Set $x=a$.
This is a vertical cross section in the y-direction.



② Take derivative w.r.t. y. This means we are taking the slope of the tangent line to that curve.



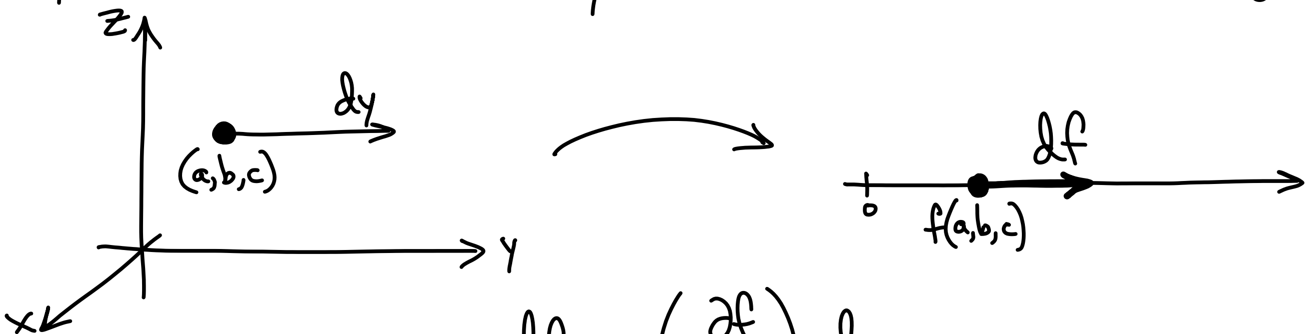
So: $\frac{\partial f}{\partial y}$ is the <u>slope</u> of the <u>tangent line</u> to the <u>graph</u> of f in <u>the y-direction</u>.

Also: $\frac{\partial f}{\partial x}$ is the <u>slope</u> of the <u>tangent line</u> to the <u>graph</u> of f in <u>the x-direction</u>.

(Think through this picture!)

"Literal" picture

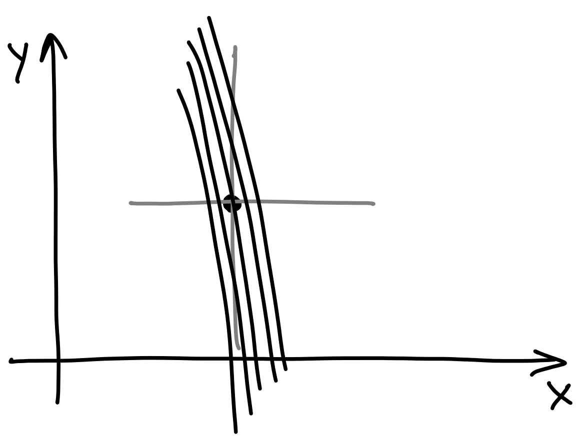
Say $f: \mathbb{R}^3 \to \mathbb{R}$. $\quad \frac{\partial f}{\partial y}(a,b,c)$ relates these changes.



$$df = \left(\frac{\partial f}{\partial y} \right) dy$$

Level sets

Partial derivs are inversely related to how close level sets are in that direction.



$\frac{\partial f}{\partial x}$ is large

$\frac{\partial f}{\partial y}$ is small

Linear approximations ("affine", not "linear"!)

$\boxed{\text{If it exists (?!)}}$, then the linear approximation is the unique degree 1 polynomial with the same value and first derivative(s) as the given function at a given point.

Ex:) $f:\mathbb{R}^1\to\mathbb{R}^1$. Note that

$$L(x)=f(a)+f'(a)(x-a)$$

has

$$\text{① } L(a)=f(a)$$
$$\text{② } L'(a)=f'(a)$$

Ex:) $f:\mathbb{R}^n\to\mathbb{R}^1$. Check that

$$L(\vec{x})=f(\vec{a})+\left.\frac{\partial f}{\partial x_1}\right|_{\vec{a}}(x_1-a_1)+\dots+\left.\frac{\partial f}{\partial x_n}\right|_{\vec{a}}(x_n-a_n)$$

has

$$\text{① } L(\vec{a})=f(\vec{a})$$
$$\text{② } \left.\frac{\partial L}{\partial x_1}\right|_{\vec{a}}=\left.\frac{\partial f}{\partial x_1}\right|_{\vec{a}},\ \dots,\ \left.\frac{\partial L}{\partial x_n}\right|_{\vec{a}}=\left.\frac{\partial f}{\partial x_n}\right|_{\vec{a}}$$

As a notational convenience, we define the gradient as

$$\nabla g=\left(\frac{\partial g}{\partial x_1},\dots,\frac{\partial g}{\partial x_n}\right)$$

and then write

$$L(\vec{x})=f(\vec{a})+\nabla f(\vec{a})\cdot(\vec{x}-\vec{a})$$

Ex:) $f(x,y)=x^2+y^2$, at $\vec{a}=(1,2)$.

$$\left.\begin{matrix}\frac{\partial f}{\partial x}=2x, & \frac{\partial f}{\partial x}(\vec{a})=2\\ \frac{\partial f}{\partial y}=2y, & \frac{\partial f}{\partial y}(\vec{a})=4\end{matrix}\right\}\Rightarrow \begin{aligned}&\nabla f(\vec{a})=(2,4)\\ &L(x,y)=5+\begin{pmatrix}2\\4\end{pmatrix}\cdot\begin{pmatrix}x-1\\y-2\end{pmatrix}\\ &\quad=2x+4y-5\end{aligned}$$

Ex: $f:\mathbb{R}^n \to \mathbb{R}^m$. As a notational convenience we let

$$Df = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & & \frac{\partial f_m}{\partial x_n} \end{pmatrix} \begin{array}{l} \leftarrow \text{partials of } f_1 \\ \vdots \\ \leftarrow \text{partials of } f_m \end{array}$$

w.r.t. x_1 ... w.r.t. x_n

Check that

$$L(\vec{x}) = f(\vec{a}) + Df(\vec{a})\,(\vec{x}-\vec{a})$$

has

1. $L(\vec{a}) = f(\vec{a})$
2. $\frac{\partial L_i}{\partial x_j}(\vec{a}) = \frac{\partial f_i}{\partial x_j}(\vec{a})$ for all i,j

Ex: $f(x,y) = (x^2+y^2,\ xy,\ 3x^2y^3)$, at $\vec{a}=(1,2)$.

$$Df = \begin{pmatrix} 2x & 2y \\ y & x \\ 6xy^3 & 9x^2y^2 \end{pmatrix} \qquad Df(1,2) = \begin{pmatrix} 2 & 4 \\ 2 & 1 \\ 48 & 36 \end{pmatrix}$$

So

$$L(\vec{x}) = f(\vec{a}) + Df(\vec{a})\,(\vec{x}-\vec{a}) = \begin{pmatrix} 5 \\ 2 \\ 24 \end{pmatrix} + \begin{pmatrix} 2 & 4 \\ 2 & 1 \\ 48 & 36 \end{pmatrix}\begin{pmatrix} x-1 \\ y-2 \end{pmatrix}$$

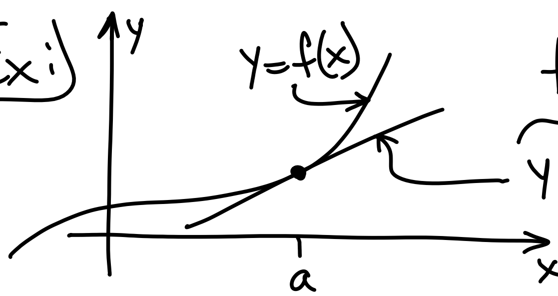
Obs: $\begin{pmatrix} \vert & & \vert \\ \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \\ \vert & & \vert \end{pmatrix} = Df = \begin{pmatrix} \text{---} \nabla f_1 \text{---} \\ \vdots \\ \text{---} \nabla f_m \text{---} \end{pmatrix}$

Obs: L is merely affine — but Df <u>is</u> linear!

Tangent planes

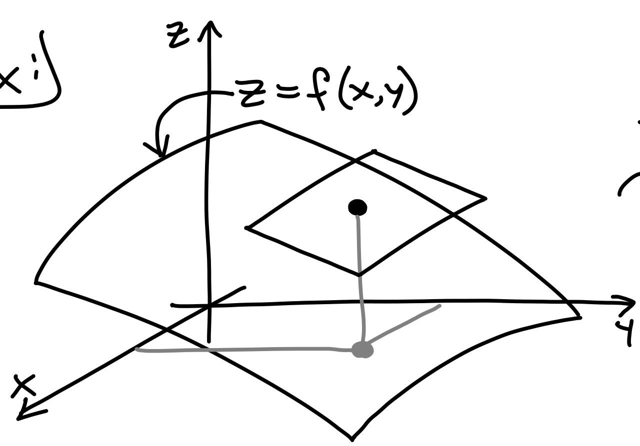
Like the tangent line, the tangent plane is the <u>graph</u> of the linear approximation.

Ex:



$$\overbrace{y}^{\text{tangent line}} = \underbrace{L(x) = f(a) + f'(a)(x-a)}_{\text{linear approximation}}$$

Ex:



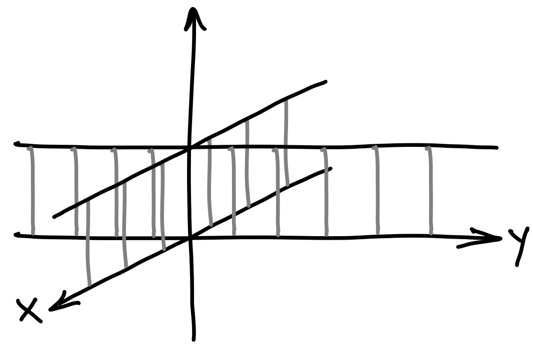
$$\overbrace{z = L(x,y)}^{\text{tangent plane}} = \underbrace{f(\vec{a}) + \nabla f(\vec{a}) \cdot (\vec{x}-\vec{a})}_{\text{linear approximation}}$$

Differentiabitity

$\boxed{\text{Is } L \text{ \underline{really} a linear approximation?}}$ <u>Maybe not!</u>

<u>Ex:</u> Let

$$f(x,y) = \begin{cases} 1 & \text{if } x=0 \text{ or } y=0 \\ 0 & \text{else} \end{cases}, \text{ at } \vec{a}=\vec{0}.$$



$f(\vec{a})=1$, $\nabla f(\vec{a})=(0,0)$

$\Rightarrow L(\vec{x})=1$. <u>Not good</u>,

even though all partials exist!

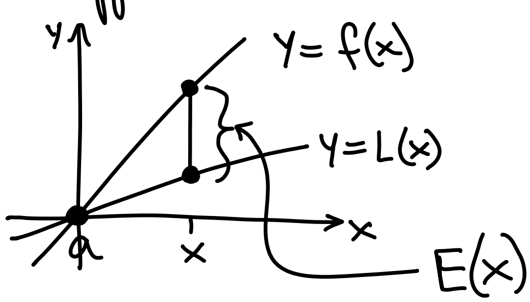
<u>Def:</u> The "error" in an approximation L of f is

$$E(\vec{x}) = f(\vec{x}) - L(\vec{x})$$

Can we just require $E(\vec{x}) \to 0$? <u>Nope!</u>

<u>Ex:</u> $f(x)=5x$ at $a=0$; is $L(x)=3x$ a good linear approximation?

No:



And yet: $\lim\limits_{x\to 0} E(x)$

$$= \lim(5x-3x) = \lim(2x) = 0!$$

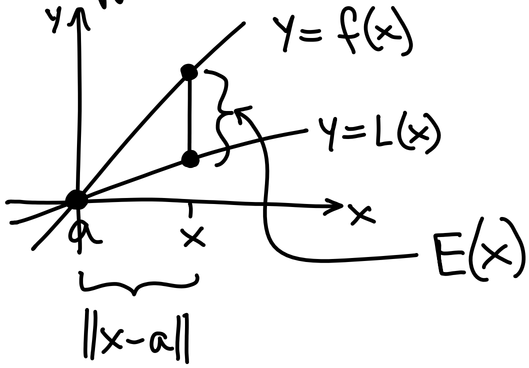
So it is not good enough to require $\lim E(\vec{x}) = 0$.

Def:) The "relative error" in an approximation L of f is

$$R(\vec{x}) = \frac{\|E(\vec{x})\|}{\|\vec{x}-\vec{a}\|} \quad \begin{array}{l} \leftarrow \text{(how far off is the output)} \\ \leftarrow \text{(how far away is input)} \end{array}$$

Ex:) $f(x) = 5x$ at $a=0$; is $L(x) = 3x$ a good linear approximation?

No:



And $\displaystyle\lim_{x\to 0} \frac{\|E(x)\|}{\|x-a\|} = \lim \left|\frac{5x-3x}{x}\right|$

$= 2 \neq 0$

"Vanishing relative error" is the stronger test we need!

Def:) $f : (X \subset \mathbb{R}^n) \to \mathbb{R}^m$ is <u>differentiable</u> at $\vec{a} \in X$ and $L(\vec{x}) = f(\vec{a}) + Df(\vec{a})(\vec{x}-\vec{a})$ is the <u>linear approximation</u> iff

$$R(\vec{x}) = \frac{\|E(x)\|}{\|\vec{x}-\vec{a}\|} = \frac{\|f(\vec{x}) - L(x)\|}{\|\vec{x}-\vec{a}\|} = \frac{\|f(\vec{x}) - f(\vec{a}) - Df(\vec{a})(\vec{x}-\vec{a})\|}{\|\vec{x}-\vec{a}\|}$$

approaches zero as $\vec{x} \to \vec{a}$.

Alt:) $E(\vec{x}) = f(\vec{x}) - L(\vec{x}) = \underbrace{\left(f(\vec{x}) - f(\vec{a})\right)}_{\text{increment}} - \underbrace{\left(Df(\vec{a})(\vec{x}-\vec{a})\right)}_{\text{"differential"} = df}$

Continuous differentiability

In practice, how do we know when f is differentiable?

Def:) f is continuously differentiable if all partials ① exist, and ② are continuous.

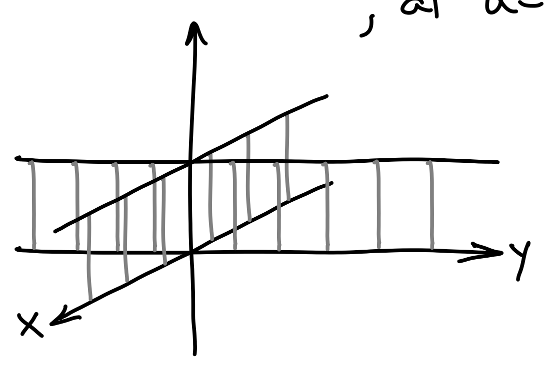
Thm:) $\underbrace{f \text{ is continuously differentiable}}_{\text{easy to check!}} \Rightarrow \underbrace{f \text{ is differentiable}}_{\text{otherwise hard to check!}}$

Ex:) Is $f(x,y,z) = x^2z - yz^3 + xy^3z$ differentiable?
All partials are polynomials, all of which are continuous, so f is cont. diff. and therefore differentiable.

Also useful:

Thm:) f is differentiable $\Rightarrow$ f is continuous
a.k.a. f is not continuous $\Rightarrow$ f is not differentiable

Ex:) Let
$$f(x,y) = \begin{cases} 1 & \text{if } x=0 \text{ or } y=0 \\ 0 & \text{else} \end{cases}, \text{ at } \vec{a} = \vec{0}.$$



f not continuous, so not differentiable.

The Derivative

For a differentiable function $f: \mathbb{R}^n \to \mathbb{R}^m$, we rewrite our linear approximation as

$$L(\vec{x}) = f(\vec{a}) + Df(\vec{a})\,(\vec{x}-\vec{a})$$

$$L(\vec{x}) = L(\vec{a}) + Df(\vec{a})\,(\vec{x}-\vec{a})$$

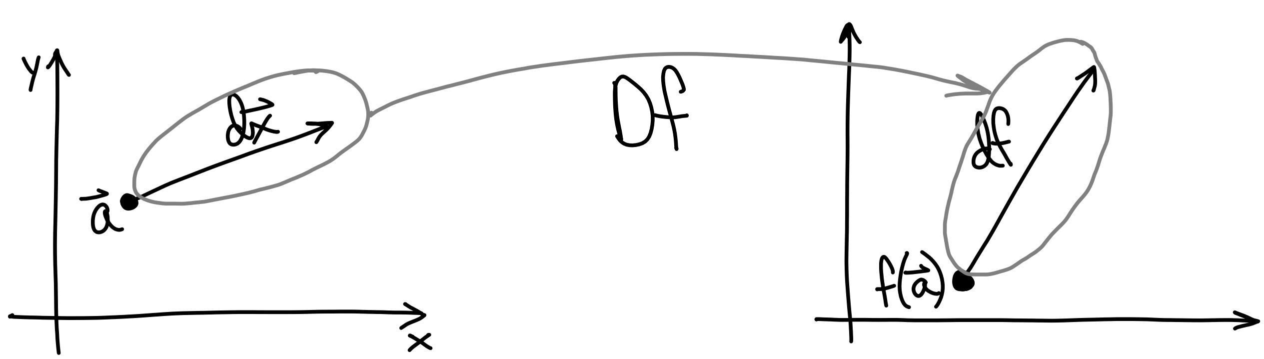
$$\underbrace{L(\vec{x}) - L(\vec{a})}_{\substack{\text{output differential} \\ = df}} = Df(\vec{a})\,\underbrace{(\vec{x}-\vec{a})}_{\substack{\text{input differential} \\ = d\vec{x}}}$$

$$df = Df(\vec{a})\, d\vec{x}$$

So Df is the thing that <u>turns input changes into output changes!</u>

We call Df "the derivative", "the derivative matrix", or "the derivative transformation" of f.

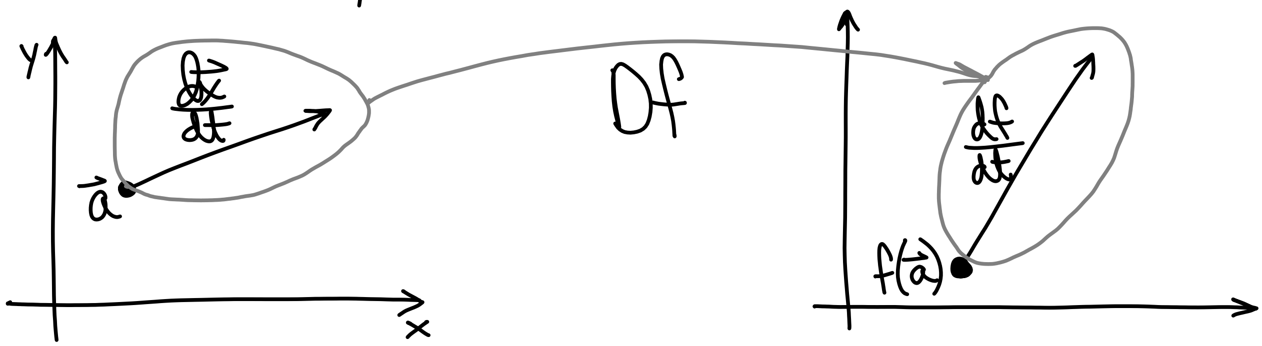
(Also sometimes written as Jf.)



Similarly if the "changes" are velocities; divide by dt.

$$df = Df(\vec{a})\, d\vec{x}$$

$$\underbrace{\frac{df}{dt}}_{\text{output velocity}} = Df(\vec{a})\, \underbrace{\frac{d\vec{x}}{dt}}_{\text{input velocity}}$$

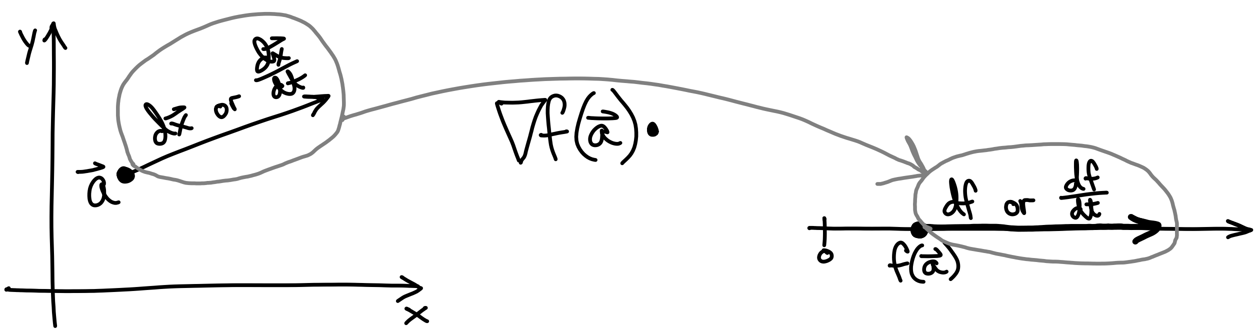


For a real-valued function, similarly the gradient is the derivative.

$$\underbrace{L(\vec{x}) - L(\vec{a})}_{\substack{\text{output differential} \\ = df}} \overset{!}{=} \nabla f(\vec{a}) \cdot \underbrace{(\vec{x} - \vec{a})}_{\substack{\text{input differential} \\ = d\vec{x}}}$$

$$df = \nabla f(\vec{a}) \cdot d\vec{x}$$

(Note also that in this case $Df = \left(\frac{\partial f}{\partial x_1} \cdots \frac{\partial f}{\partial x_n}\right)$, so $Df(\vec{a})\, d\vec{x} = \nabla f(\vec{a}) \cdot d\vec{x}$, so this is really the same!)



2.4 – Properties; Higher Order Partial Derivatives

We already saw that $Df: \mathbb{R}^n \to \mathbb{R}^m$ is linear.
But "D" is linear too:

$$D : \{ f : \mathbb{R}^n \to \mathbb{R}^m \mid f \text{ diff'bl} \} \longrightarrow \{ m \times n \text{ matrices} \}$$

That is, $D(af + bg) = a\,Df + b\,Dg$

<u>Ex:</u>) Consider $f(x,y) = (x, 2y)$, $g(x,y) = (3y, 4x)$.

$$Df = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \qquad Dg = \begin{pmatrix} 0 & 3 \\ 4 & 0 \end{pmatrix}$$

Consider now $h = 5f + 6g = (5x + 18y, 24x + 10y)$.

$$Dh = \begin{pmatrix} 5 & 18 \\ 24 & 10 \end{pmatrix}$$

Then

$$5Df + 6Dg = 5\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} + 6\begin{pmatrix} 0 & 3 \\ 4 & 0 \end{pmatrix} = \begin{pmatrix} 5 & 18 \\ 24 & 10 \end{pmatrix} = Dh$$
$$= D(5f + 6g)$$

as expected by this linearity.

D also satisfies a product rule and quotient rule in some cases (in the book). Check by direct computation!

Higher order partials

If f is a function of $x_1, \dots, x_n$, then so is $\frac{\partial f}{\partial x_i}$!
So you can take partials of partials.

Ex:)

$$
\begin{tikzcd}[column sep=small]
 & & f = x^5 y^3 \arrow[dl, "\frac{\partial}{\partial x}"'] \arrow[dr, "\frac{\partial}{\partial y}"] & & \\
 & f_x = \frac{\partial f}{\partial x} = 5x^4y^3 \arrow[dl, "\frac{\partial}{\partial x}"'] \arrow[d, "\frac{\partial}{\partial y}"] & & 3x^5y^2 = \frac{\partial f}{\partial y} = f_y \arrow[d, "\frac{\partial}{\partial x}"'] \arrow[dr, "\frac{\partial}{\partial y}"] & \\
f_{xx} = \frac{\partial^2 f}{\partial x^2} & \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x} & & \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} = f_{yy}
\end{tikzcd}
$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = (f_x)_y = f_{xy} = 15x^4y^2 \qquad \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = (f_y)_x = f_{yx} = 15x^4y^2$$

Notice that $f_{xy} = f_{yx}$ here... Coincidence? Nope!

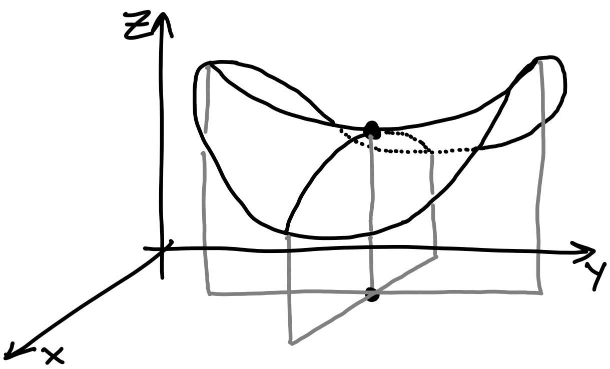
Def:) $C^k = \{ f \mid \text{all } k\text{th order partials exist and are continuous} \}$

Thm:) If $f \in C^k$, and $(i_1, \dots, i_k)$ and $(j_1, \dots, j_k)$ are rearrangements of each other, then

$$\frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}} = \frac{\partial^k f}{\partial x_{j_1} \cdots \partial x_{j_k}}$$

What do second partials look like (on graphs)?

f_{xx}, f_{yy} relate to concavity – because they are effectively single variable second derivatives.



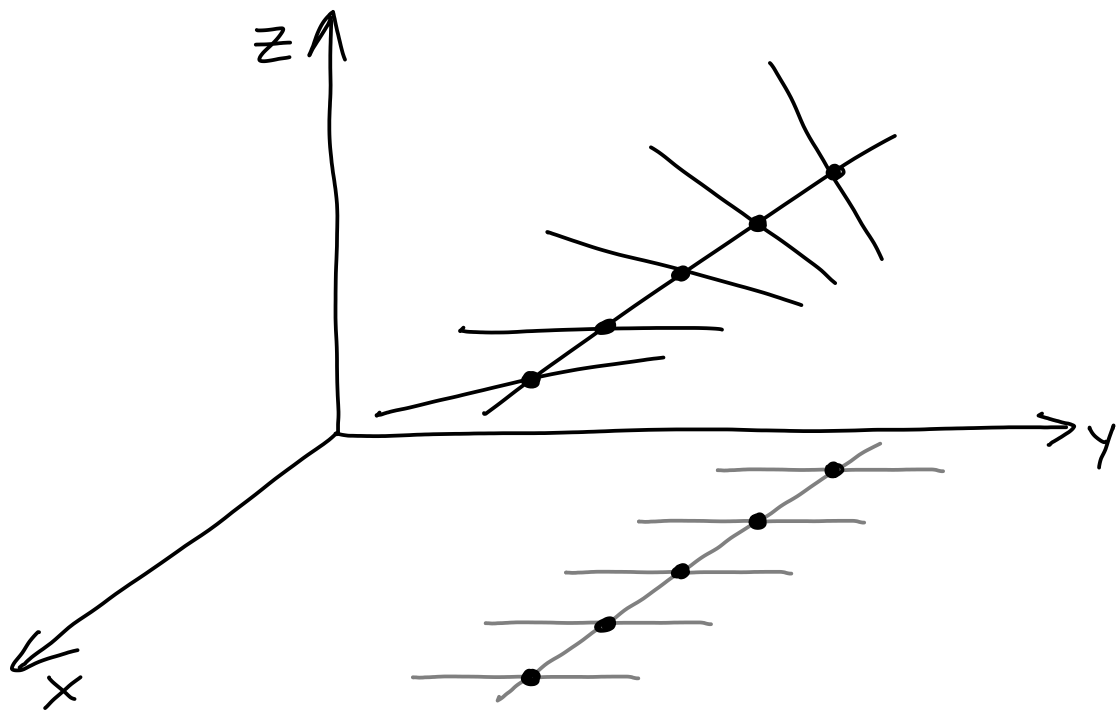
$$f_{xx} < 0$$

$$f_{yy} > 0$$

What about $\dfrac{\partial^2 f}{\partial x \partial y} = \underbrace{\dfrac{\partial}{\partial x}}\underbrace{\left(\dfrac{\partial f}{\partial y}\right)}$?

rate of change moving in x-direction

slope of tangent line in y-direction



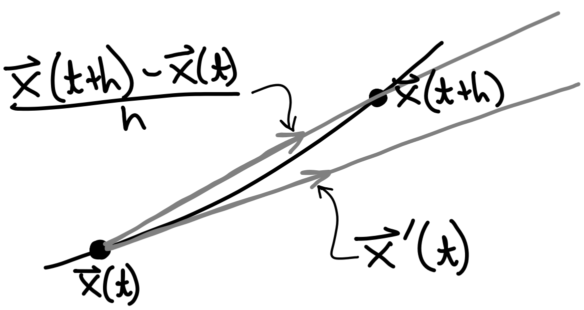
This indicates a "twisting" of the graph, called "torsion".

3.1 – Parametrized Curves

We have previously discussed parametric curves, $\vec{x} : \mathbb{R} \to \mathbb{R}^n$, $\vec{x}(t) = \big(x_1(t), \ldots, x_n(t)\big)$.

The derivative matrix is

$$D\vec{x} = \begin{pmatrix} \frac{dx_1}{dt} \\ \vdots \\ \frac{dx_n}{dt} \end{pmatrix} = \frac{d\vec{x}}{dt} = \lim_{h \to 0} \frac{\vec{x}(t+h) - \vec{x}(t)}{h}$$



① $\vec{x}'$ points in the direction the curve is going;

② $\|\vec{x}'\| = \lim \frac{\|d\vec{x}\|}{|t|} = \text{speed}$

So $\vec{x}'(t) = \text{velocity} = \vec{v}(t)$.

Similarly, $\vec{x}''(t) = \vec{v}'(t) = \text{acceleration} = \vec{a}(t)$.

Ex: Say $\vec{v}(t) = (2-3\sin t,\ e^t)$ and $\vec{x}(0) = (5,6)$.

① What is $\vec{a}(t)$?

$$\vec{a} = \begin{pmatrix} 2-3\sin t \\ e^t \end{pmatrix}' = \begin{pmatrix} -3\cos t \\ e^t \end{pmatrix}$$

② What is $\vec{x}(t)$?

$$\vec{x} = \int \vec{v}\,dt = \int \begin{pmatrix} 2-3\sin t \\ e^t \end{pmatrix} dt = \begin{pmatrix} \int 2-3\sin t\,dt \\ \int e^t\,dt \end{pmatrix} = \begin{pmatrix} 2t+3\cos t \\ e^t \end{pmatrix} + \vec{c}$$

$$@\ t=0:\quad \begin{pmatrix} 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} + \vec{c} \implies \vec{c} = \begin{pmatrix} 2 \\ 5 \end{pmatrix}$$

NB – $\vec{c}$ is <u>not</u> always equal to $\vec{x}_0$! When that happens it is a <u>peculiarity</u> of the specific functions involved.

NB – these are vector-valued antiderivatives — <u>not</u> area under a curve!

(Skip Kepler stuff — but <u>do</u> see:

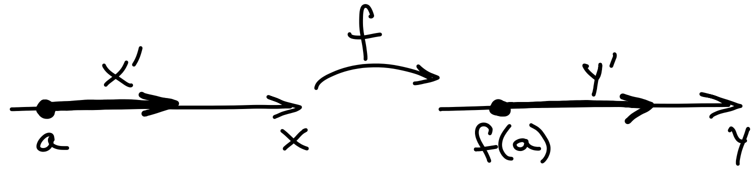
<u>Prop. 1.4</u>: If $\vec{x}(t), \vec{y}(t)$ are differentiable paths, then

$$\frac{d}{dt}(\vec{x}\cdot\vec{y}) = \frac{d\vec{x}}{dt}\cdot\vec{y} + \vec{x}\cdot\frac{d\vec{y}}{dt}$$

$$\frac{d}{dt}(\vec{x}\times\vec{y}) = \frac{d\vec{x}}{dt}\times\vec{y} + \vec{x}\times\frac{d\vec{y}}{dt}$$

2.5 – The Chain Rule

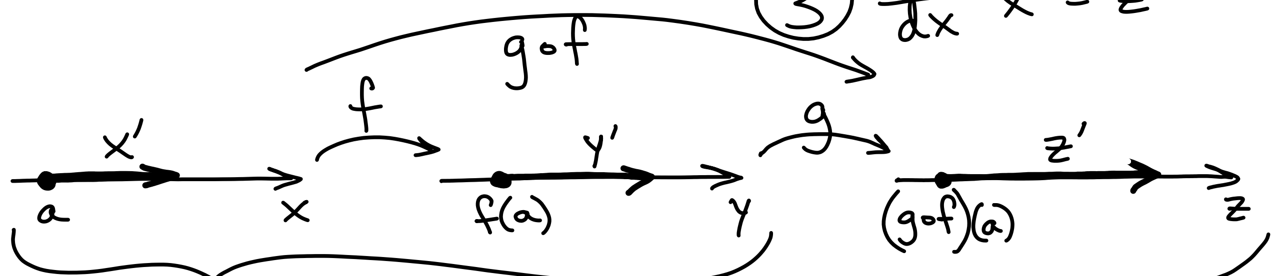
Recall from single variable, a derivative relates velocities.



$$\frac{dy}{dx} x' = y'$$

Take this point of view on a sequence of two functions...
... and their composition:

③ $\frac{dz}{dx} x' = z'$



① $\frac{dy}{dx} x' = y'$

② $\frac{dz}{dy} y' = z'$

Putting ① into ② and comparing with ③ gives

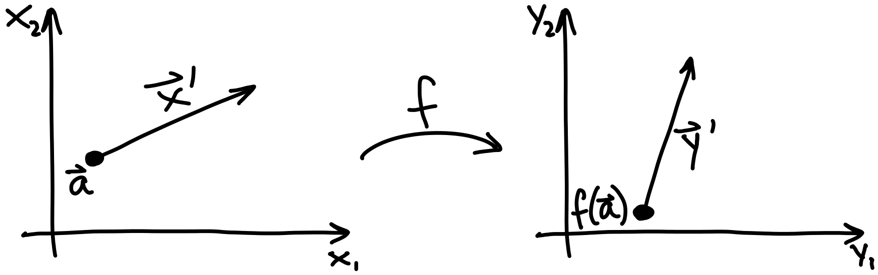
$$\frac{dz}{dy}\left(\frac{dy}{dx} x'\right) = z'$$

$$\frac{dz}{dy}\frac{dy}{dx}(x') = z'$$

$$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx}$$

So, differentiation turns a composition into a product.

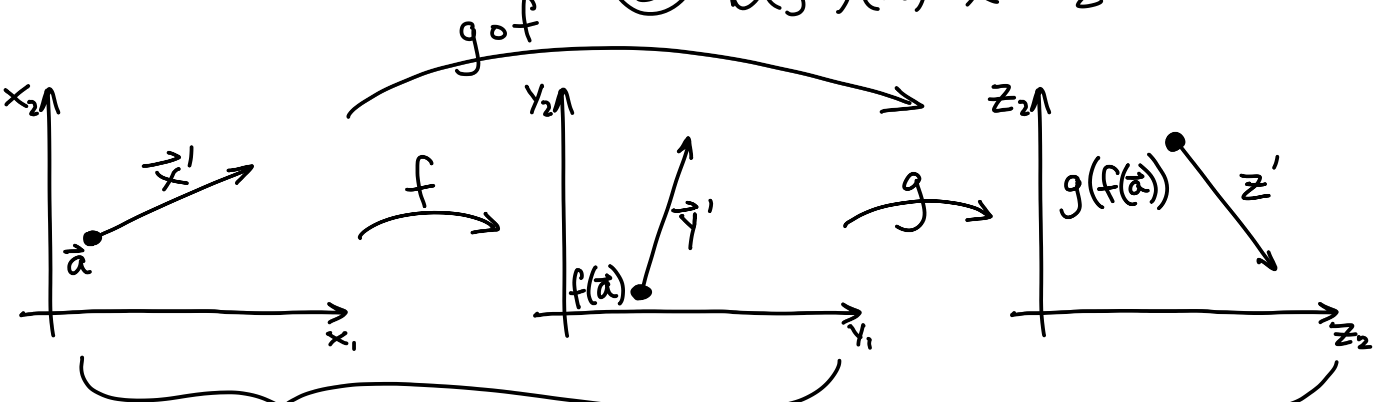
Similarly for multivariable, a derivative relates velocities.



$$Df(\vec{a})\,\vec{x}' = \vec{y}'$$

Composing a sequence gives:

③ $D(g\circ f)(\vec{a})\,\vec{x}' = \vec{z}'$



① $Df(\vec{a})\,\vec{x}' = \vec{y}'$

② $Dg(f(\vec{a}))\,\vec{y}' = \vec{z}'$

Putting ① into ② and comparing with ③ gives

$$Dg(f(\vec{a}))\left(Df(\vec{a})\,\vec{x}'\right) = \vec{z}'$$

$$\left(Dg(f(\vec{a}))\,Df(\vec{a})\right)\vec{x}' = \vec{z}'$$

$$D(g\circ f)(\vec{a}) = Dg(f(\vec{a}))\,Df(\vec{a})$$

Again, differentiation turns a composition into a product, this time a product of matrices!

Likewise for any combinations of dimensions!

Ex:) Say $f(x_1,x_2)=(y_1,y_2,y_3)$, $g(y_1,y_2,y_3)=(z_1,z_2,z_3)$ (both differentiable) and

$$f(1,2)=(3,1,0),\quad Df(1,2)=\begin{pmatrix}1&0\\3&4\\5&-1\end{pmatrix},\quad Dg(3,1,0)=\begin{pmatrix}0&1&0\\1&0&1\\0&1&0\end{pmatrix}$$

Compute $D(g\circ f)(1,2)$, and $\frac{\partial z_3}{\partial x_1}$.

Well,

$$\begin{aligned}D(g\circ f)(1,2) &= Dg(f(1,2))\,Df(1,2)\\ &= Dg(3,1,0)\,Df(1,2)\\ &= \begin{pmatrix}0&1&0\\1&0&1\\0&1&0\end{pmatrix}\begin{pmatrix}1&0\\3&4\\5&-1\end{pmatrix}=\begin{pmatrix}3&-1\\6&-1\\3&-1\end{pmatrix}\end{aligned}$$

And $\frac{\partial z_3}{\partial x_1}$ is the element in the 1st column, 3rd row, so $\frac{\partial z_3}{\partial x_1}=3$.

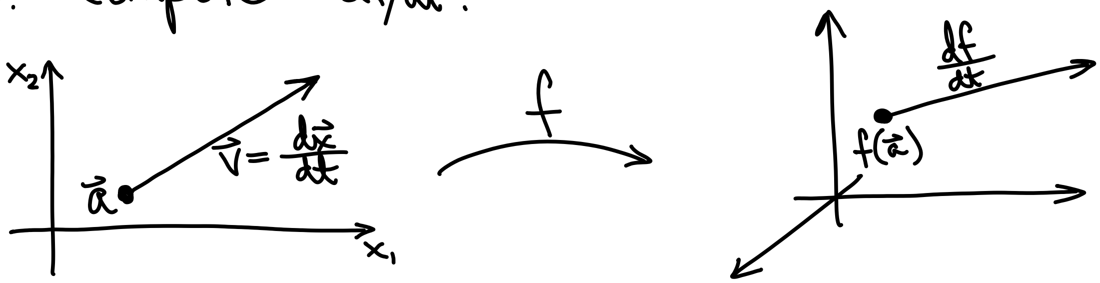
Special cases

Ex:) Say $\underset{(t)}{\mathbb{R}^1}\xrightarrow{\vec{x}}\underset{(\vec{x})}{\mathbb{R}^n}\xrightarrow{f}\underset{(\vec{y})}{\mathbb{R}^m}$. Then the chain rule says

$$\begin{pmatrix}\frac{dy_1}{dt}\\ \vdots\\ \frac{dy_m}{dt}\end{pmatrix}=\begin{pmatrix} & & \\ & Df & \\ & & \end{pmatrix}\begin{pmatrix}\frac{dx_1}{dt}\\ \vdots\\ \frac{dx_n}{dt}\end{pmatrix}$$

This effectively restates that the derivative relates velocities.

Ex: Suppose we have $f(x_1, x_2) = (x_1 x_2,\ x_1 - x_2,\ 3x_1 + 4x_2)$, and that a particle is at $\vec{a} = (2,1)$ and moving with velocity $\vec{v} = (4,3)$. Compute df/dt.



$$Df = \begin{pmatrix} x_2 & x_1 \\ 1 & -1 \\ 3 & 4 \end{pmatrix}, \quad Df(\vec{a}) = \begin{pmatrix} 1 & 2 \\ 1 & -1 \\ 3 & 4 \end{pmatrix}$$

$$\frac{df}{dt} = Df(\vec{a}) \frac{d\vec{x}}{dt} = \begin{pmatrix} 1 & 2 \\ 1 & -1 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 4 \\ 3 \end{pmatrix} = \begin{pmatrix} 10 \\ 1 \\ 24 \end{pmatrix}$$

Ex: Say $\underset{(t)}{\mathbb{R}^1} \xrightarrow{\vec{x}} \underset{(\vec{x})}{\mathbb{R}^n} \xrightarrow{f} \underset{(y)}{\mathbb{R}}$. Then the chain rule says

$$\frac{dy}{dt} = \begin{pmatrix} \frac{\partial y}{\partial x_1} & \cdots & \frac{\partial y}{\partial x_n} \end{pmatrix} \begin{pmatrix} \frac{dx_1}{dt} \\ \vdots \\ \frac{dx_n}{dt} \end{pmatrix}$$

$$\frac{dy}{dt} = \nabla f \cdot \frac{d\vec{x}}{dt}$$

Again, for real-valued functions, the derivative is a dot with the gradient.

Ex: A bird is at $\vec{p} = (3,2,1)$ with velocity $\vec{v} = (2,0,1)$. Temperature is given by $T(x,y,z) = e^{x-y} - z$. Find $\frac{dT}{dt}$.

$$\nabla T = \left(e^{x-y}, -e^{x-y}, -1\right), \quad \nabla T(\vec{p}) = (e, -e, -1).$$

$$\frac{dT}{dt} = \nabla T(\vec{p}) \cdot \frac{d\vec{x}}{dt} = (e, -e, -1) \cdot (2,0,1) = 2e - 1$$

Intermediate variables

Suppose $f(x_1, \dots, x_n) = (y_1, \dots, y_m)$, $g(y_1, \dots, y_m) = (z_1, \dots, z_p)$.

The chain rule tells us

$$\begin{pmatrix} & \vdots & \\ \cdots\cdots & \frac{\partial z_i}{\partial x_j} & \cdots \\ & \vdots & \end{pmatrix} = \begin{pmatrix} \frac{\partial z_i}{\partial y_1} & \cdots & \frac{\partial z_i}{\partial y_m} \\ & & \end{pmatrix} \begin{pmatrix} & \frac{\partial y_1}{\partial x_j} & \\ & \vdots & \\ & \frac{\partial y_m}{\partial x_j} & \end{pmatrix}$$

$$\frac{\partial z_i}{\partial x_j} = \frac{\partial z_i}{\partial y_1}\frac{\partial y_1}{\partial x_j} + \cdots + \frac{\partial z_i}{\partial y_m}\frac{\partial y_m}{\partial x_j}$$

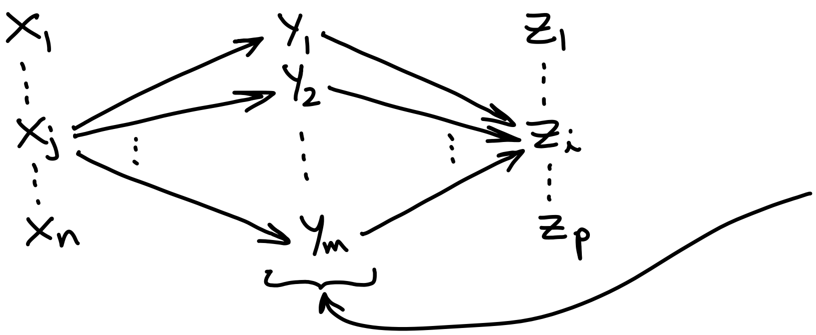
Warning: You can't "cancel" the y's! That's not how this notation works!

$$\underbrace{\frac{\partial z_i}{\partial y_1}}_{\text{here, } y_2, \dots, y_m \text{ are required to be constant!}} \quad \underbrace{\frac{\partial y_1}{\partial x_j}}_{\text{here, } y_2, \dots, y_m \text{ are not constant!}}$$

We can interpret this equation

$$\frac{\partial z_i}{\partial x_j} = \frac{\partial z_i}{\partial y_1}\frac{\partial y_1}{\partial x_j} + \cdots + \frac{\partial z_i}{\partial y_m}\frac{\partial y_m}{\partial x_j}$$

by how changes "propogate" along "paths" between the variables.



Each such path goes through one of these "intermediate variables".

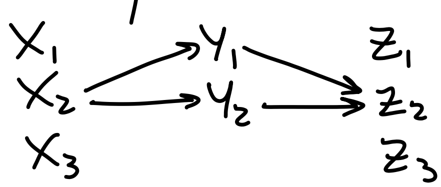
This point of view can be more convenient than matrices in some cases.

<u>Ex:</u>) Say $f(x_1, x_2, x_3) = (x_1 - x_2,\ x_1^2 x_2) = (y_1, y_2)$

$$g(y_1, y_2) = (y_1^2 - y_2^2,\ y_1^3,\ y_2^3) = (z_1, z_2, z_3)$$

Compute $\frac{\partial z_2}{\partial x_2}$.

Note that the only intermediate variables are y_1, y_2



So

$$\frac{\partial z_2}{\partial x_2} = \frac{\partial z_2}{\partial y_1}\frac{\partial y_1}{\partial x_2} + \frac{\partial z_2}{\partial y_2}\frac{\partial y_2}{\partial x_2}$$

$$= (3y_1^2)(-1) + (0)(x_1^2) = -3(x_1 - x_2)^2$$

Warning – don't confuse variables!

<u>Ex:</u> $f(x) = x^2$, $g(x) = x^3$. What is $(g \circ f)'$?

<u>Wrong:</u> $(g \circ f)' = g' f' = (3x^2)(2x) = 6x^3$

<u>Valid:</u> $(g \circ f)' = g'(f(x))\, f'(x) = \left(3(x^2)^2\right)(2x) = 6x^5$

<u>Better!:</u> Rewrite to avoid confusion of variables!

$$\underbrace{x \xrightarrow{f} y}_{f(x) = x^2 = y} \underbrace{\xrightarrow{g} z}_{g(y) = y^3 = z}$$

Then $(g \circ f)' = g' f' = (3y^2)(2x) = 6y^2 x = 6x^5$

<u>Ex:</u> $f(x,y) = (x^2 + y^2, x^2 - y^2)$, $g(x,y) = (2y^2, 3x^2 + 1)$

Compute $D(g \circ f)$.

$$\underbrace{\begin{matrix} x \\ y \end{matrix} \xrightarrow{f} \begin{matrix} u \\ v \end{matrix}} \underbrace{\xrightarrow{g}}$$

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x^2 + y^2 \\ x^2 - y^2 \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix} \qquad g\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 2v^2 \\ 3u^2 + 1 \end{pmatrix}$$

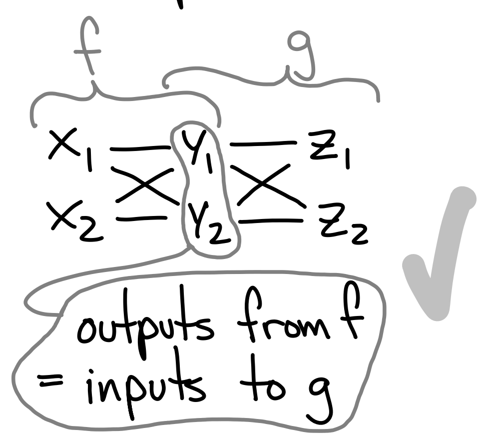
$$Df = \begin{pmatrix} 2x & 2y \\ 2x & -2y \end{pmatrix} \qquad Dg = \begin{pmatrix} 0 & 4v \\ 6u & 0 \end{pmatrix}$$

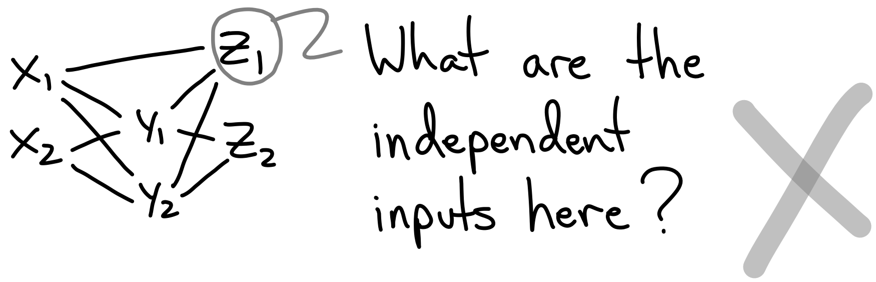
$$D(g \circ f) = Dg\, Df = \begin{pmatrix} 0 & 4v \\ 6u & 0 \end{pmatrix} \begin{pmatrix} 2x & 2y \\ 2x & -2y \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 4(x^2 - y^2) \\ 6(x^2 + y^2) & 0 \end{pmatrix} \begin{pmatrix} 2x & 2y \\ 2x & -2y \end{pmatrix}$$

Warning (related) — diagrams <u>must</u> represent <u>compositions</u> of <u>functions</u>!

- unambiguous & uniquely named inputs, outputs.
- inputs are independent.
- output from one = inputs from next.





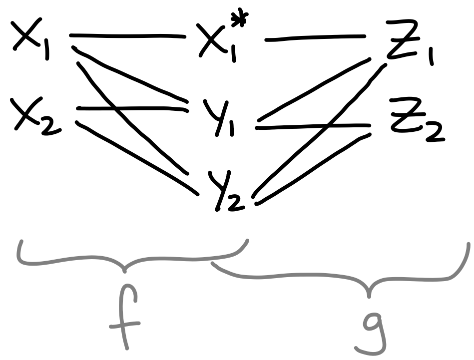
What are the independent inputs here?

What does $\frac{\partial z_1}{\partial x_1}$ even mean?

You might be able to fix problems (such as above) by introducing "new" variables.

<u>Ex:</u> $y_1 = x_1 + x_2$, $y_2 = x_1 - x_2$, $z_1 = x_1 + y_1 + y_2$, $z_2 = y_1 - y_2$.

Let $x_1^* = x_1$. Then we rewrite $z_1 = x_1^* + y_1 + y_2$ and have



Now, what is $\frac{\partial z_1}{\partial x_1}$?

$$\frac{\partial z_1}{\partial x_1} = \frac{\partial z_1}{\partial x_1^*}\frac{\partial x_1^*}{\partial x_1} + \frac{\partial z_1}{\partial y_1}\frac{\partial y_1}{\partial x_1} + \frac{\partial z_1}{\partial y_2}\frac{\partial y_2}{\partial x_1}$$

$$\frac{\partial z_1}{\partial x_1} = \frac{\partial z_1}{\partial x_1^*} + \frac{\partial z_1}{\partial y_1}\frac{\partial y_1}{\partial x_1} + \frac{\partial z_1}{\partial y_2}\frac{\partial y_2}{\partial x_1}$$

$\vdots$

Second derivatives

NB – if z is a function of x, y ... then so are its partials!

Ex:) Say $z = \overset{C^2}{f}(x,y)$, r, θ are the usual polar coordinates. Compute $\frac{\partial^2 z}{\partial r^2}$ in terms of r, θ, and partials of f.

①
$$\begin{tikzcd} r \arrow[r] \arrow[dr] & x \arrow[r] & z \\ \theta & y \arrow[ur] & \end{tikzcd}$$

$$\frac{\partial z}{\partial r} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial r}$$

$$= z_x \cos\theta + z_y \sin\theta$$

②
$$\begin{tikzcd} r \arrow[r] \arrow[dr] & x \arrow[r] \arrow[dr] \arrow[ddr] & z \\ \theta & y \arrow[ur] \arrow[r] \arrow[dr] & z_x \\ & & z_y \end{tikzcd}$$

$$\frac{\partial^2 z}{\partial r^2} = \frac{\partial}{\partial r}\left(\frac{\partial z}{\partial r}\right)$$

$$= \frac{\partial}{\partial r}\left(z_x \cos\theta + z_y \sin\theta\right)$$

$$= \left(\frac{\partial}{\partial r} z_x\right)\cos\theta + \left(\frac{\partial}{\partial r} z_y\right)\sin\theta$$ ← (b/c $\cos\theta, \sin\theta$ are constants. Lucky!)

$$= \left(\frac{\partial z_x}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial z_x}{\partial y}\frac{\partial y}{\partial r}\right)\cos\theta + \left(\frac{\partial z_y}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial z_y}{\partial y}\frac{\partial y}{\partial r}\right)\sin\theta$$

$$= \left(z_{xx}\cos\theta + z_{xy}\sin\theta\right)\cos\theta + \left(z_{xy}\cos\theta + z_{yy}\sin\theta\right)\sin\theta$$

$$= z_{xx}\cos^2\theta + 2 z_{xy}\sin\theta\cos\theta + z_{yy}\sin^2\theta$$

Application

The "Laplacian" operator is needed to understand how waves flow, how heat moves, and how electrons behave.

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

But very often, f is much easier to describe in spherical coordinates!

Using the method on the previous page, we can rewrite

$$\Delta f = \frac{1}{\rho^2}\frac{\partial}{\partial \rho}\left(\rho^2 \frac{\partial f}{\partial \rho}\right) + \frac{1}{\rho^2 \sin\phi}\frac{\partial}{\partial \phi}\left(\sin\phi \frac{\partial f}{\partial \phi}\right) + \frac{1}{\rho^2 \sin^2\phi}\frac{\partial^2 f}{\partial \theta^2}$$

from the chain rule applied to: $\begin{matrix}\rho \\ \phi \\ \theta\end{matrix} \to \begin{matrix}x \\ y \\ z\end{matrix} \to f$

<u>Ex:</u>) Gravitational and electric potentials f must have $\Delta f = 0$.

Consider the candidate $f = \frac{k}{\rho} = \frac{k}{\sqrt{x^2+y^2+z^2}}$

<u>Option 1:</u> $\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\left(\frac{k}{\sqrt{x^2+y^2+z^2}}\right)$ is hard to compute.

<u>Option 2:</u>

$$\left[\frac{1}{\rho^2}\frac{\partial}{\partial \rho}\left(\rho^2 \frac{\partial}{\partial \rho}\right) + \frac{1}{\rho^2 \sin\phi}\frac{\partial}{\partial \phi}\left(\sin\phi \frac{\partial}{\partial \phi}\right) + \frac{1}{\rho^2 \sin^2\phi}\frac{\partial^2}{\partial \theta^2}\right]\left(\frac{k}{\rho}\right)$$

is easy to compute!

2.6 - Directional Derivatives and the Gradient

Recall, for differentiable $f : \mathbb{R}^n \to \mathbb{R}^1$, the partial $\frac{\partial f}{\partial x_i}$ is:

① the <u>slope</u> of the graph of f in the $\vec{e}_i$-direction

② the factor you multiply by dx_i (or dx_i/dt) to get df (or df/dt) as you move in the $\vec{e}_i$-direction.

We can rewrite the definition of $\frac{\partial f}{\partial x_i}$ as

$$\frac{\partial f}{\partial x_i}(\vec{a}) = \frac{d}{dt}\Big|_{t=0} f(\vec{a} + t\vec{e}_i)$$

...and compute it with

$$\frac{\partial f}{\partial x_i}(\vec{a}) = \nabla f(\vec{a}) \cdot \vec{e}_i$$

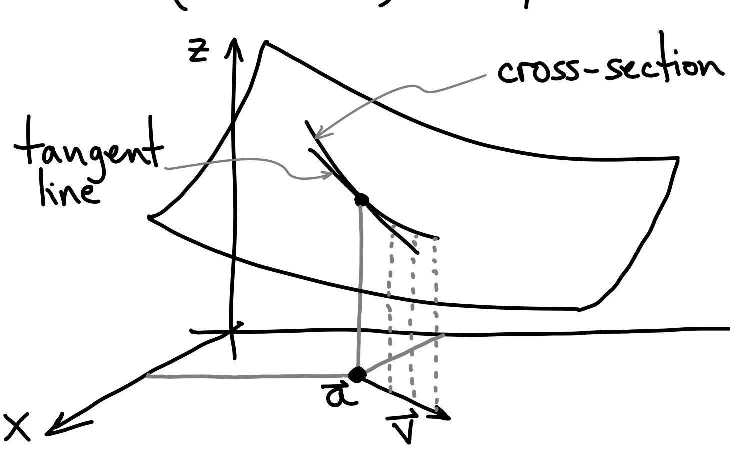
If you replace $\vec{e}_i$ above with any other <u>unit vector</u> $\vec{v}$, you get all of the same properties! This is called the "directional derivative of f in the $\vec{v}$-direction".

$$D_{\vec{v}} f(\vec{a}) = \frac{d}{dt}\Big|_{t=0} f(\vec{a} + t\vec{v}) = \nabla f(\vec{a}) \cdot \vec{v}$$

It is:

① the <u>slope</u> of the graph of f in the $\vec{v}$-direction

② the factor you multiply by ds (or ds/dt) to get df (or df/dt) as you move in the $\vec{v}$-direction.



Think:

$$df = \frac{df}{ds}\, ds$$

$$\frac{df}{dt} = \frac{df}{ds}\frac{ds}{dt}$$

Ex:) Say $f(x,y) = x^2+y^2$. What is the slope of the graph of f at $\vec{a} = (3,1)$, in the direction of the vector $\vec{v} = (3,4)$?

The unit vector in the $\vec{v}$-direction is $\vec{u} = \vec{v}/5$.

$$\begin{aligned} \text{slope} &= D_{\vec{u}} f(\vec{a}) = \nabla f(\vec{a}) \cdot \vec{u} \\ &= (2x, 2y)\Big|_{\vec{a}} \cdot (3,4)/5 \\ &= (6,2) \cdot (3,4)/5 = 26/5 \end{aligned}$$

(NB – don't conflate phrasing with "...toward the point..."!)

Ex:) Temperature is $T(x,y,z) = 60 + 25e^{-(x^2+y^2+z^2)}$. At $\vec{a} = (1,0,2)$, what is dT/ds in the direction of $\vec{v} = (2,4,3)$?

$$\nabla T = 25(-2x,-2y,-2z)\, e^{-(x^2+y^2+z^2)}, \quad \nabla T(\vec{a}) = -50e^{-5}(1,0,2)$$

$$\begin{aligned} \frac{dT}{ds} &= D_{\vec{u}} T(\vec{a}) = \nabla T(\vec{a}) \cdot \vec{u} \\ &= -50e^{-5}(1,0,2) \cdot (2,4,3)/\sqrt{29} \\ &= \frac{-400\, e^{-5}}{\sqrt{29}} \end{aligned}$$

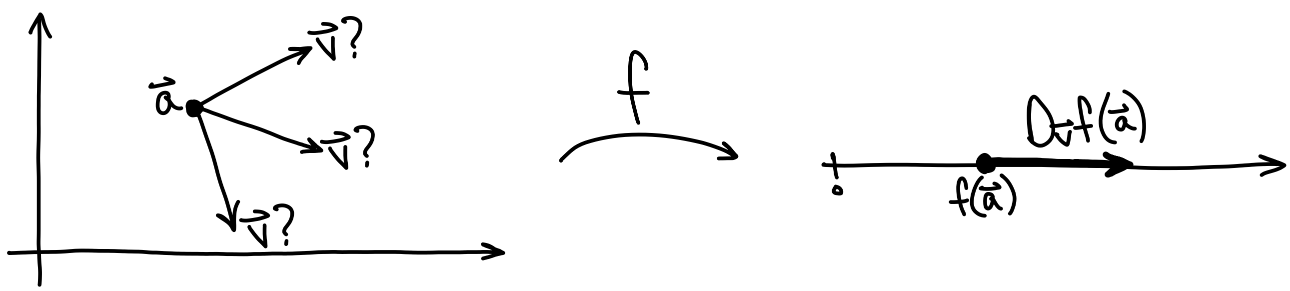
<u>Warning</u> – in some contexts (not this course!), the "unit vector" requirement is removed. Then you don't get ① and ②!

But of course, $\nabla f(\vec{a}) \cdot \vec{v}$ is still interesting...

Ex:) For $\underset{(t)}{\mathbb{R}} \xrightarrow{\vec{x}} \mathbb{R}^n \xrightarrow{f} \mathbb{R}$, note that $\frac{df}{dt} = \nabla f \cdot \frac{d\vec{x}}{dt}$.

Fastest increase / steepest ascent

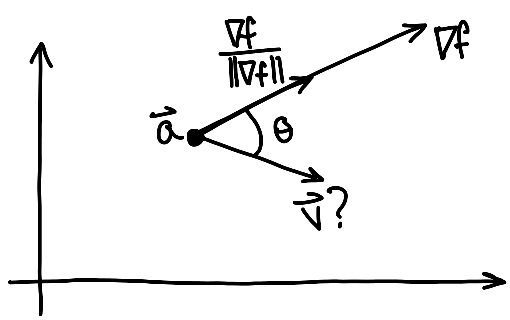
Say $f: \mathbb{R}^n \to \mathbb{R}^1$ is differentiable at $\vec{a}$. Of all unit vectors $\vec{v}$, (directions), which one maximizes $D_{\vec{v}} f(\vec{a})$?



- $D_{\vec{v}} f(\vec{a}) = df/ds$, so $\vec{v}$ is the "direction of fastest increase."
- $D_{\vec{v}} f(\vec{a}) = \text{slope}$, so $\vec{v}$ is the "direction of steepest ascent".

How do we find this $\vec{v}$?

$$D_{\vec{v}} f = \nabla f \cdot \vec{v} = \|\nabla f\| \, \|\vec{v}\| \cos\theta = \|\nabla f\| \cos\theta$$



This is maximized when $\theta = 0$, and $\vec{v}$ points the same direction as ∇f.

So $\vec{v} = \dfrac{\nabla f}{\|\nabla f\|}$.

And what is $D_{\vec{v}} f$ in that direction?

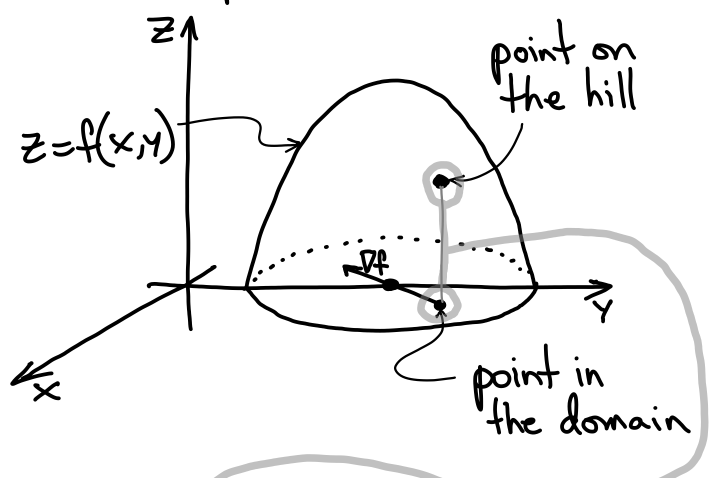
$$D_{\vec{v}} f = \nabla f \cdot \vec{v} = \nabla f \cdot \left(\frac{\nabla f}{\|\nabla f\|} \right) = \frac{\nabla f \cdot \nabla f}{\|\nabla f\|} = \frac{\|\nabla f\|^2}{\|\nabla f\|} = \|\nabla f\|$$

The above give us a geometric characterization of ∇f — it is the unique vector in the domain with:

① the <u>direction of ∇f</u> is the direction of fastest increase (a.k.a. "steepest ascent")

② the <u>magnitude of ∇f</u> is that maximum $D_{\vec{v}} f = \frac{df}{ds}$ = slope (a.k.a. the "steepness" of the graph).

<u>Ex:</u>) A hill is the shape of the graph of $f(x,y) = 12 - x^2 - y^2 + 4y$. At the point sitting above $(1,3)$, which way is "uphill" and how steep is it there?



$$\nabla f = (-2x, -2y+4)$$
$$\nabla f(\vec{a}) = (-2,-2)$$
$$\|\nabla f(\vec{a})\| = 2\sqrt{2}$$

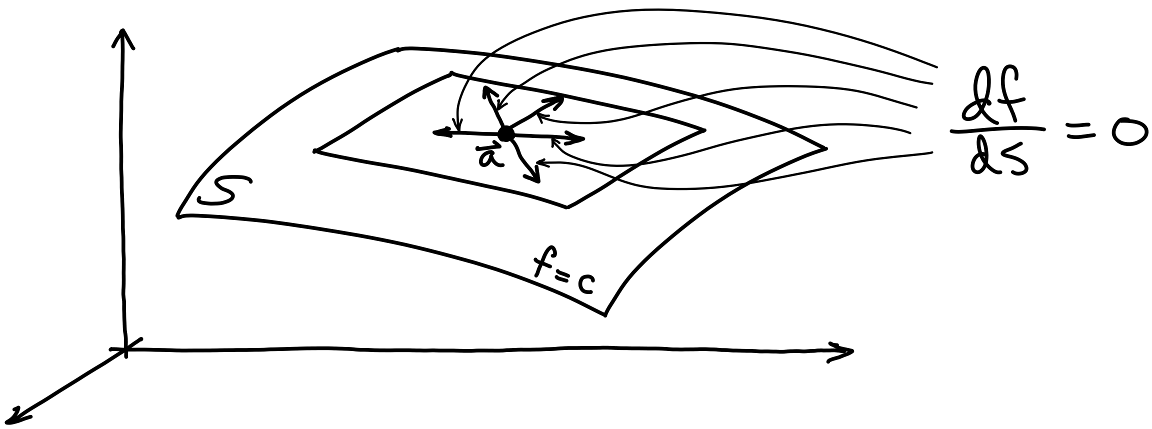
So uphill is "southwest", with slope $= 2\sqrt{2}$.

(NB — For convenience of language we sometimes conflate these two points. Be sure to interpret as intended!)

Level sets

Say $f: \mathbb{R}^n \to \mathbb{R}$ is differentiable and $\vec{a} \in \mathbb{R}^n$ is on the level set $S = \{f = c\}$.

Note that as you move along S, with $f = c$, we have $\frac{df}{ds} = 0$.

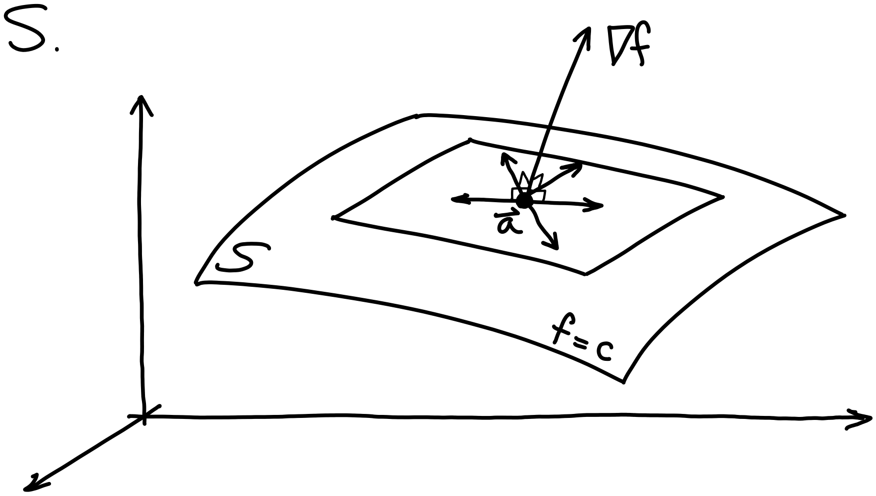


So for all tangent vectors $\vec{v}$, we have $D_{\vec{v}} f = 0$.

But of course also $D_{\vec{v}} f = \nabla f \cdot \vec{v}$.

$$\Longrightarrow \nabla f \cdot \vec{v} = 0$$
$$\Longrightarrow \nabla f \perp \vec{v}$$

This is true for all $\vec{v}$ tangent to S, so ∇f must be $\perp$ to S.

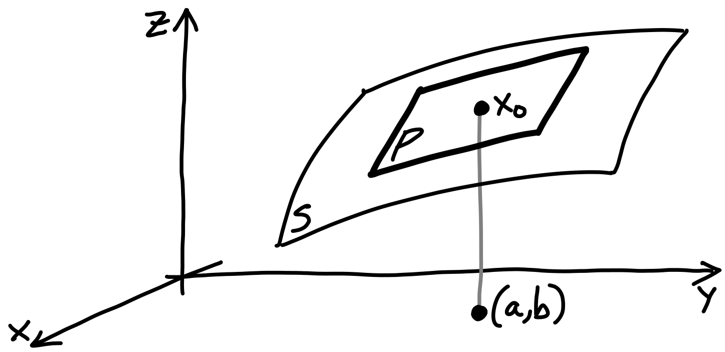


So, <u>gradients</u> are always $\perp$ to <u>level sets</u>.

Tangent planes

Previously we viewed the tangent plane as the graph of the linear approximation. Instead...

Say S is the graph $z=f(x,y)$, and P is the tangent plane at $\vec{x}_0=(a,b,f(a,b))$ sitting above (a,b) in the domain.



Notice:

$$\frac{\text{graph of } f}{z=f(x,y)} \quad = S = \quad \frac{\text{level set of } g(x,y,z)}{0=z-f(x,y)}$$

$$\Longleftrightarrow$$

So we can use the "calculus of level sets" on g!

$$\vec{n}=\nabla g=\begin{pmatrix}-f_x(a,b)\\-f_y(a,b)\\1\end{pmatrix} \text{ is } \perp \text{ to } S \text{ at } (a,b)\text{, and so also } P.$$

So the equation of P is

$$\vec{n}\cdot(\vec{x}-\vec{x}_0)=0 \Rightarrow \begin{pmatrix}-f_x(a,b)\\-f_y(a,b)\\1\end{pmatrix}\cdot\begin{pmatrix}x-a\\y-b\\z-f(a,b)\end{pmatrix}=0$$

Solving for z gives the familiar tangent plane,

$$z=f(a,b)+f_x(a,b)(x-a)+f_y(a,b)(y-b)$$

the graph of the familiar linear approximation.

Ex:) Find the tangent plane to the graph of $f(x,y) = x^3 - 4y^2$ at the point $\vec{p}$ above $(2,1)$.

graph : $z = x^3 - 4y^2$

level set : $\underbrace{z - x^3 + 4y^2}_{g(x,y,z)} = 0$

$\nabla g = (-3x^2, 8y, 1)$

$\vec{x}_0 = \vec{p} = (2,1,4)$

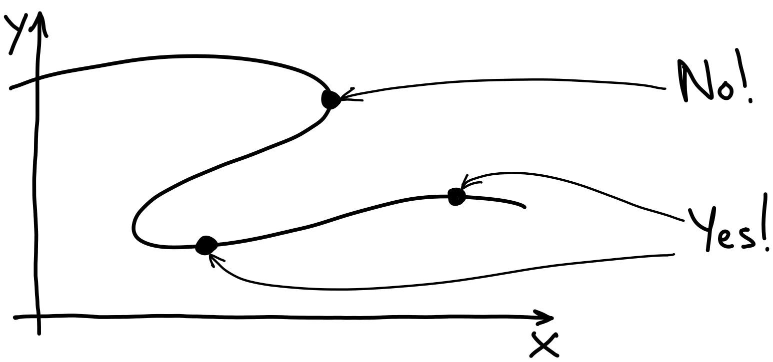
$\vec{n} = \nabla g(\vec{p}) = (-12, 8, 1)$

Then $\vec{n}\cdot\vec{x} = \vec{n}\cdot\vec{x}_0$ becomes $-12x + 8y + z = -12$.

NB – the surface doesn't have to be a graph for something like this to work!

<u>Implicit function theorem, implicit differentiation</u>

Suppose $F(x,y) = 0$. Can I view y as a (local) function of x? Sometimes...

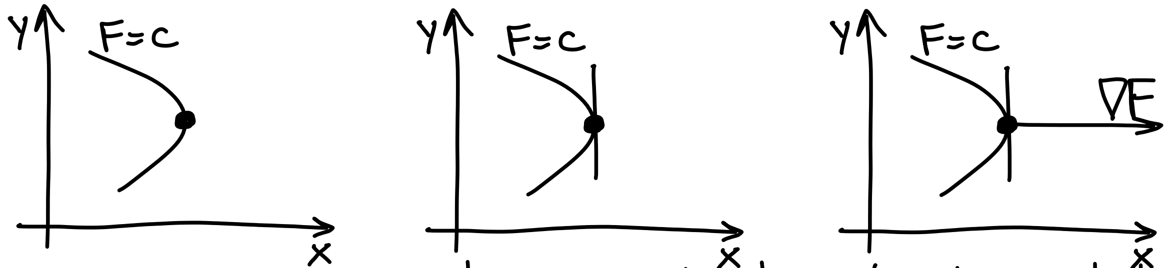


Who cares?

Well, if y is not a function, then "$\frac{dy}{dx}$" is <u>meaningless</u>.

And until you <u>know</u>, you can't reasonably use $\frac{dy}{dx}$ <u>at all</u>!

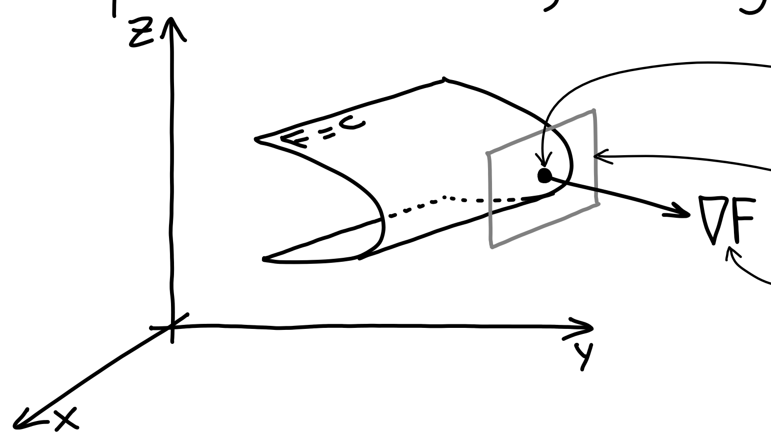
If y is not locally a function:



curve "wraps under" itself — has a vertical tangent line — has horizontal gradient $\Rightarrow \frac{\partial F}{\partial y} = 0$

So, "problem" $\Rightarrow \frac{\partial F}{\partial y} = 0$; or, $\frac{\partial F}{\partial y} \neq 0 \Rightarrow$ "no problem!"

Similarly for 3 variables, viewing z as a function of x, y:



wraps under

vertical tangent plane

horizontal gradient

$$\Rightarrow \frac{\partial F}{\partial z} = 0$$

<u>Thm:</u> Say $f: \mathbb{R}^n \to \mathbb{R}$ is C^1, and at a point on $F(x_1, \ldots, x_n) = 0$ we know $\frac{\partial F}{\partial x_i} \neq 0$. Then we can view x_i locally as a function of the other variables.

> Take the partial w.r.t. the variable that you want to view as a function.

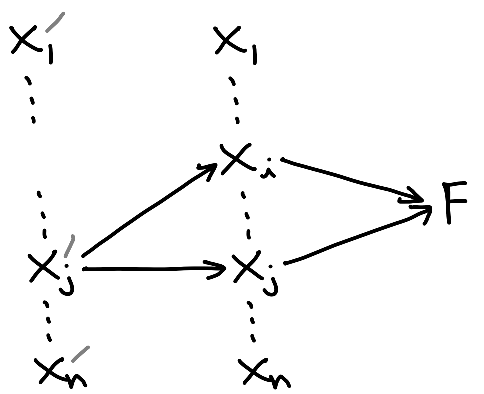
Ex:) Consider the surface $x^3 + x^2y - xz + z^3 - 27 = 0$, near $(1,2,3)$. Is x locally a function of y, z?

$$\frac{\partial F}{\partial x} = 3x^2 + 2xy - z \qquad \left.\frac{\partial F}{\partial x}\right|_{(1,2,3)} = 4 \underbrace{\neq 0}_{}$$

$\hookrightarrow$ so, yes!

How do we then take derivatives of this new function?

$F(x_1, \ldots, x_n) = 0$, $\frac{\partial F}{\partial x_i} \neq 0 \Rightarrow x_i$ is a function, what is $\frac{\partial x_i}{\partial x_j}$?



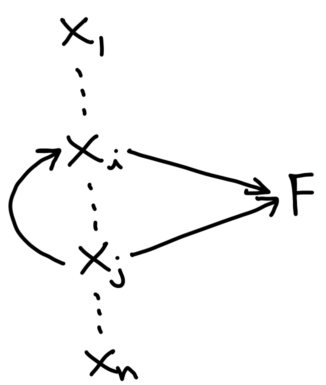
All but x_i, x_j are constants!

And $F = c \Rightarrow \frac{\partial F}{\partial x_j'} = 0$. So

$$0 = \frac{\partial F}{\partial x_j'} = \frac{\partial F}{\partial x_i}\frac{\partial x_i}{\partial x_j'} + \frac{\partial F}{\partial x_j}\underbrace{\frac{\partial x_j}{\partial x_j'}}_{=1}$$

$$\Longrightarrow \boxed{\frac{\partial x_i}{\partial x_j'} = -\frac{\partial F/\partial x_j}{\partial F/\partial x_i}}$$

Alternate diagram:



$$0 = \frac{\partial F}{\partial x_i}\frac{\partial x_i}{\partial x_j} + \frac{\partial F}{\partial x_j}$$

$$\Longrightarrow \boxed{\frac{\partial x_i}{\partial x_j} = -\frac{\partial F/\partial x_j}{\partial F/\partial x_i}}$$

NB — The "$-$" here is very important!

Ex: Consider the surface $x^3 + x^2y - xz + z^3 - 27 = 0$, near $(1,2,3)$, where x is locally a function of y,z. What is $\frac{\partial x}{\partial z}$?

$$\frac{\partial x}{\partial z} = -\frac{\partial F/\partial z}{\partial F/\partial x} = -\frac{-x+3z^2}{3x^2+2xy-z}$$

$$\frac{\partial x}{\partial z}(1,2,3) = -\frac{26}{4} = \frac{-13}{2}$$

NB – In most questions involving implicit differentiation, it is not given that the relevant variable is (locally) a function!

$$\boxed{\text{Checking that/if it is is an important part of the reasoning in a valid solution!}}$$