Chapter 2 Differentiation

2.1 geometry of functions of several variables

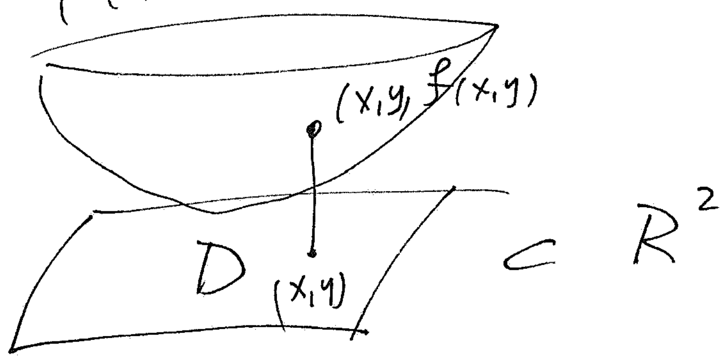
As in one variable, a function of many variables is a "rule" (usually a formula) assigning to a pt in its domain (e.g $f(x+y) = x+y$, $(x,y) \in \mathbb{R}^2$) the value of the function.

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

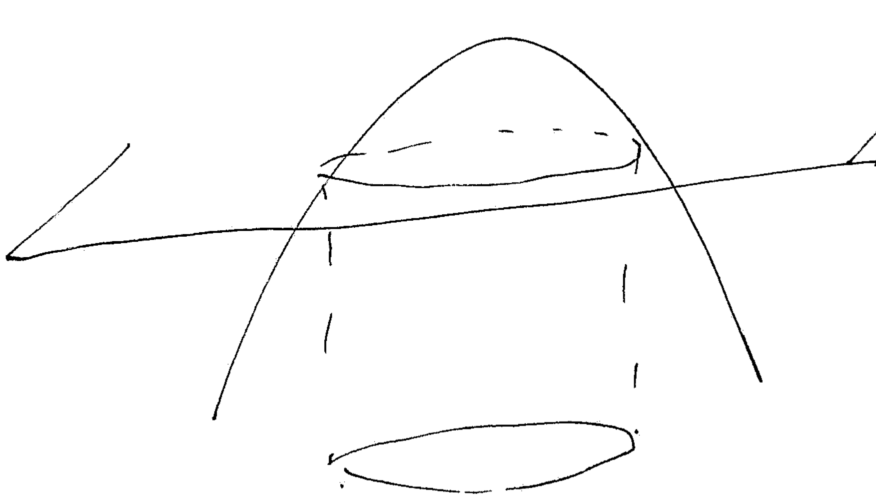
domain of f, a function of the variables $x = (x_1, \dots, x_n)$

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

$n = 2$



Since a graph of a function of 3 or more variables is very hard to visualize, so we often look at its <u>level sets</u>

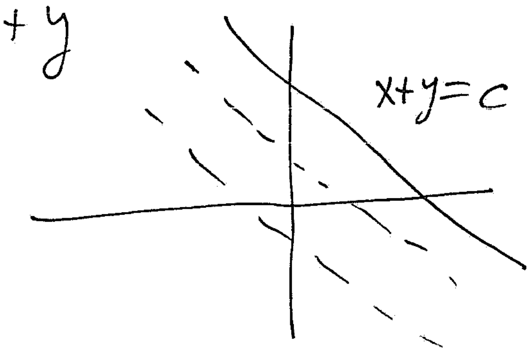


$n=2$ $\quad f^{-1}(c) = \{(x,y) : f(x,y) = c\}$

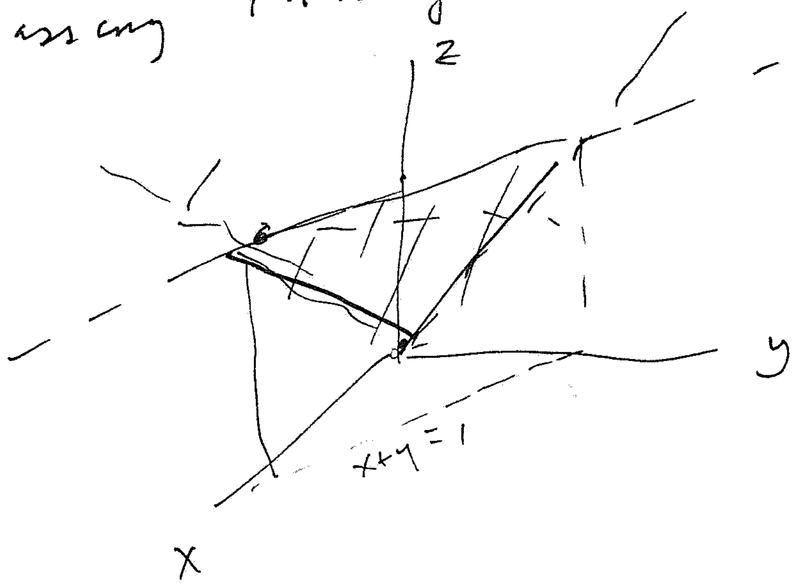
$n=3$ $\quad f^{-1}(c) = \{(x,y,z) : f(x,y,z) = c\}$

<u>Ex</u> $\quad f(x,y) = x+y$

level sets (curves)



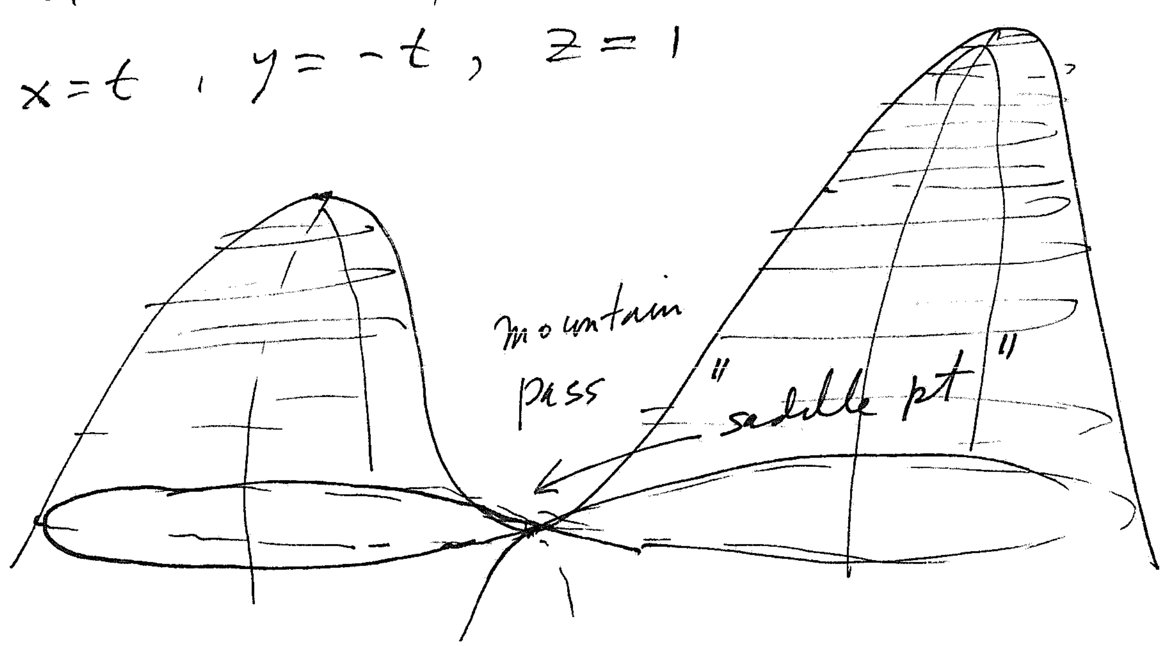
Note the graph of f is the plane $z = x+y$ ($x+y-z=0$) passing through the origin



with normal vector $(1, 1, -1)$

(or containing the line $z=1$ $x+y=0$

$x=t$, $y=-t$, $z=1$

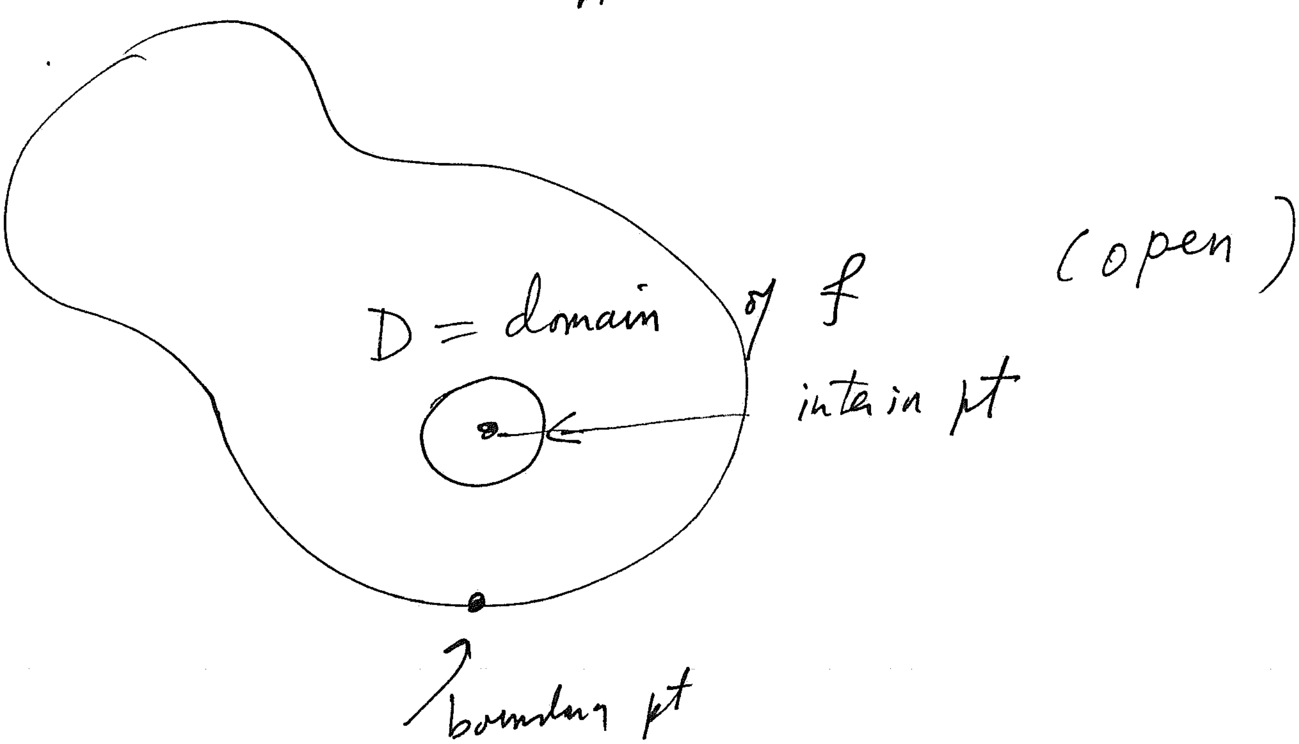


2.2 limits and continuity

$$\lim_{\vec{x}\to\vec{x}_0} f(\vec{x}) = \ ?$$

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

$$n = 2, 3, \ \dots$$



In preparation for our discussion of partial derivatives and directional derivatives, which are

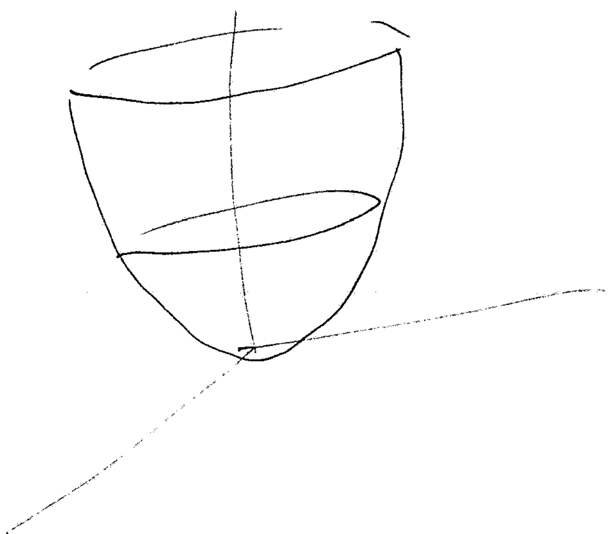
$$f^{-1}(c) = \{ \vec{x} \in D : f(\vec{x}) = c \}$$

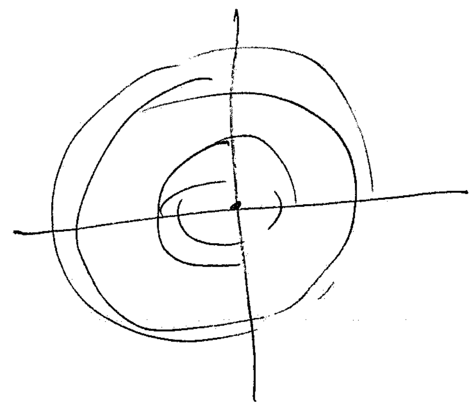
Ex $f(x,y) = x^2 + y^2$

$f^{-1}(c) = \phi \qquad c < 0$

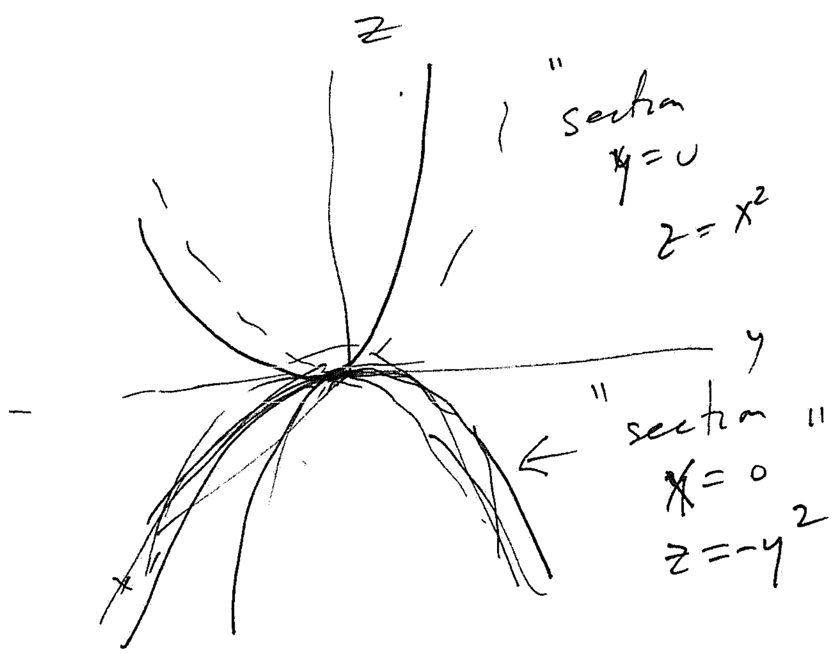
$f^{-1}(c) = \vec{0} \qquad c = 0$

$f^{-1}(c) =$ circle of radius $\sqrt{c}$ centered at $(0,0)$ if $c > 0$





$f(x,y) = x^2 - y^2$



Ex The graph of $f(x,y,z)$

$= x^2+y^2-z^2$

sits in R^4 !

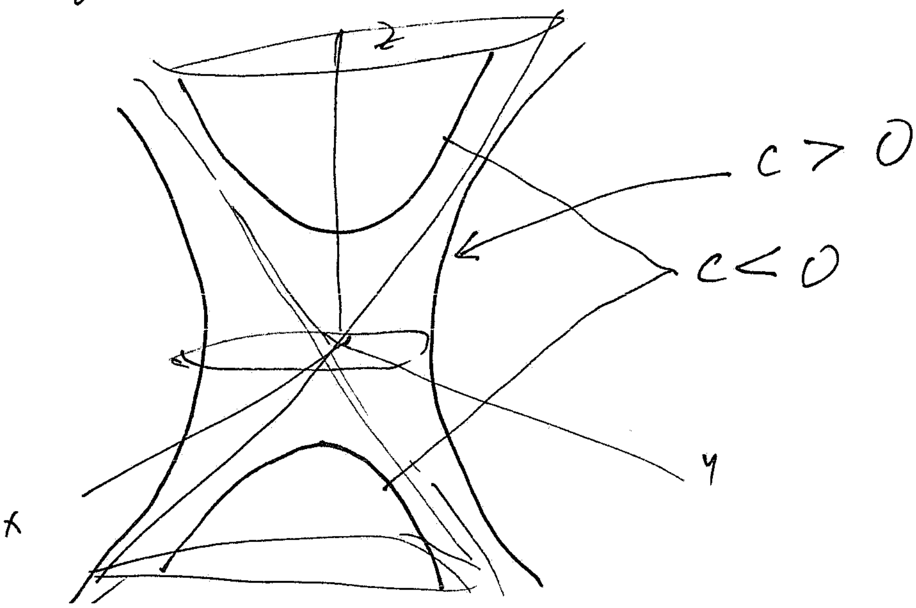
Its level sets (surfaces) $f^{-1}(c) =$

$\{ (x,y,z) \in \mathbb{R}^3 : x^2+y^2-z^2 = c \}$

are still complicated to visualize

see example 6 p 83 text.

$c=0$ double cone



defined in terms of limits
the definition of a limit is
formally almost the same as
for the $n=1$ case.

Note f need not be apriori
defined at $\vec{x}_0$. Then for $\vec{x}_0 \in R^n$
(f defined in a "punctured nhbd of $\vec{x}_0$")

$\lim_{x \to \vec{x}_0} f(x) = L$ if given
$\varepsilon > 0$ $\exists\, \delta > 0$ so that
$|\vec{x} - \vec{x}_0| < \delta \implies |f(\vec{x}) - f(\vec{x}_0)| < \varepsilon$

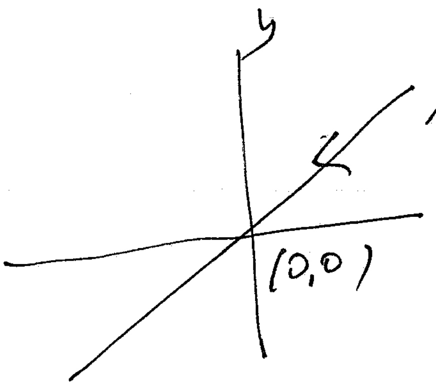
__However__, when $n = 2, 3, \ldots$
it is usually "harder to have a limit"

Example 1 (nice rational function with denominator $\neq 0$)

$$\lim_{(x,y,z)\to(0,0,0)} \frac{2x^2+y^2}{z+1} = \lim_{(x,y)\to(0,0)} \frac{2x^2+y^2}{1}$$

$$= 0$$

Example 2 $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^6}{x^6+y^6}$ does not exist



line $y = \lambda x$

note $0 \leq \dfrac{x^6}{x^6+y^6} \leq 1$

x

approach $(0,0)$ along the line $y = \lambda x$

$$\lim_{x\to 0} \frac{x^6}{x^6+\lambda^6 x^6} = \lim_{x\to 0} \frac{1}{1+\lambda^6} = \frac{1}{1+\lambda^6}$$

so the limit is different depending on the slope λ

Example 3 $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^2}{\sqrt{x^2+y^2}} = 0$

(intuitively the numerator tends to zero "faster" than the denominator)

analytic proof for $(x,y) \neq (0,0)$

$$0 \leq \frac{x^2}{\sqrt{x^2+y^2}} \leq \frac{x^2+y^2}{\sqrt{x^2+y^2}} = \sqrt{x^2+y^2}$$

$$= \underbrace{\|(x,y)-(0,0)\|}_{\|\vec{x}-\vec{0}\|} = \text{distance}$$

from $\vec{x}=(x,y)$ to origin $\vec{0}$. Hence in defn, we may take $\delta = \varepsilon$

(2.3) Partial derivatives and directional derivatives

This is the simplest notion of derivative to understand

given a function of many variables $f(x_1, \ldots, x_n)$ and a pt $\vec{x}_0 = (x_1^0, \ldots, x_n^0)$ in the domain of f

$$\frac{\partial f}{\partial x_i}(\vec{x}_0) = \lim_{h \to 0} \frac{f(x_1^0, \ldots, x_{i-1}^0, x_i^0 + h, x_{i+1}^0, \ldots x_n^0) - f(\vec{x}_0)}{h}$$

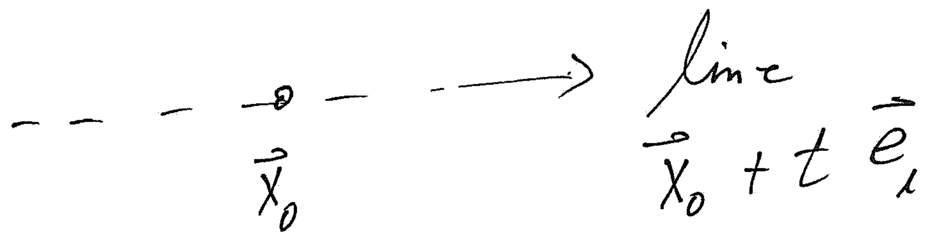
(in vector notation)

$$= \lim_{h \to 0} \frac{f(\vec{x}_0 + h\vec{e}_i) - f(\vec{x}_0)}{h}$$

So we "freeze all the variables except for x_i"

and consider the function of one variable $g(x_i) = f(x_1^0, x_2^0, \ldots, x_{i-1}^0, x_i, x_{i+1}^0, \ldots, x_n^0)$

The partial derivative $\frac{\partial f}{\partial x_i}(\vec{x}_0)$ is then just the ordinary derivative of g at x_i^0



More geometrically, we restrict f to the line through $\vec{x}_0$ in the direction $\vec{e}_i$

and consider $g(t) = f(\vec{x}_0 + t\,\vec{e}_i)$

Then $\frac{\partial f}{\partial x_i}(\vec{x}_0) = g'(0)$ (if it exists)

Example

$$f(x,y,z) = x^3 - 2y e^{z^2}$$

$$\frac{\partial f}{\partial x} = 3x^2, \quad \frac{\partial f}{\partial y} = -2e^{z^2}$$

$$\frac{\partial f}{\partial z} = -2y \cdot 2z e^{z^2}$$

$$= -4yz e^{z^2}$$

Warning The existence of partial derivatives at a pt is a very weak condition and does not imply continuity (see the example in my notes)

Closely related to partial derivative (but more geometrical) is the notion of directional derivative

<u>Def'n</u> (directional derivative in the direction of a unit vector $\vec{v}$, i.e. $\|\vec{v}\| = 1$)

<u>Notation</u> $D_{\vec{v}} f(\vec{x}_0) =$

$$\lim_{t \to 0} \frac{f(\vec{x}_0 + t\,\vec{v}) - f(\vec{x}_0)}{t}$$

when $\vec{v} = \vec{e}_i$ we recover $\frac{\partial f}{\partial x_i}(\vec{x}_0)$ (if it exists)

$$= \frac{d}{dt} f(\vec{x}_0 + t\vec{v}) \Big|_{t=0}$$

16

<u>Example</u> $f(x,y,z) = x^2 e^{-yz}$

Find the "rate of change"

of f at $(1,0,0)$ in the

direction $\vec{v} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$

$\vec{x}_0 = (1,0,0)$ $\quad \vec{x}_0 + t\vec{v}$

$= \left(1+\frac{t}{\sqrt{3}}, \frac{t}{\sqrt{3}}, \frac{t}{\sqrt{3}}\right)$

$$f(\vec{x}_0 + t\vec{v}) = \left(1+\frac{t}{\sqrt{3}}\right)^2 e^{-t^2/3}$$

$$\frac{d}{dt} f(\vec{x}_0 + t\vec{v})\Big|_{t=0} = e^{-t^2/3}\left(2\left(1+\frac{t}{\sqrt{3}}\right)\frac{1}{\sqrt{3}} + \left(1+\frac{t}{\sqrt{3}}\right)^2 \cdot -\frac{2t}{3}\right)\Big|_{t=0}$$

$$= 2/\sqrt{3}$$

We will return to directional derivatives later and find a simpler formula for $D_{\vec{v}} f$

Example (pathology)

Let $f(x,y) = (xy)^{\frac{1}{3}} = x^{\frac{1}{3}} y^{\frac{1}{3}}$

For $x>0, y>0$ ~~$(x,y) \neq (0,0)$~~ both $x^{\frac{1}{3}}$ and $y^{\frac{1}{3}}$ are differentiable functions of one variable. Hence

$$\frac{\partial f}{\partial x}(x,y) = \frac{1}{3} x^{-\frac{2}{3}} y^{1/3}, \quad \frac{\partial f}{\partial y} = \frac{1}{3} x^{\frac{1}{3}} y^{-2/3}$$

It is clear that the limits as $(x,y) \to (0,0)$ of $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ do not exist and are not even bounded.

what about $\frac{\partial f}{\partial x}(0,0)$, $\frac{\partial f}{\partial y}(0,0)$ —

do they exist?

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0} \frac{f(h,0)-f(0,0)}{h}$$

$$= \lim_{h\to 0} \frac{0-0}{h} = 0.$$

Similarly $\frac{\partial f}{\partial y}(0,0) = 0$

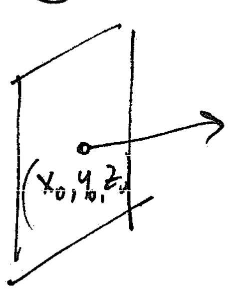
Thus partial derivatives of f exist everywhere. However there is no tangent plane at $(0,0)$

The linear approximation and differentiability of $f(x,y)$

Suppose $f(x,y)$ is a nice (smooth) function and consider ~~that~~ the graph of f near a point (x_0, y_0). In R^3, a non-vertical plane is of the form

$$z = ax + by + c$$

(convince yourself that a vertical plane



i.e horizontal normal vector has the form

$$ax + by + c = 0 \, ! \;)$$

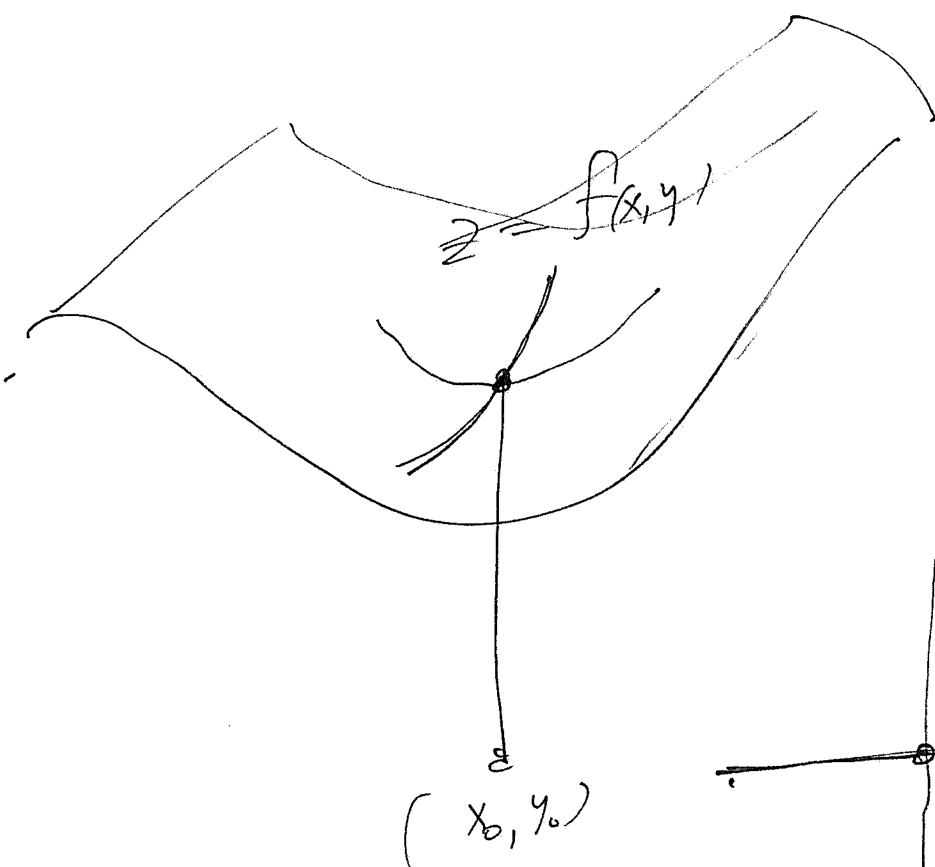
If this non-vertical plane is to be the tangent plane

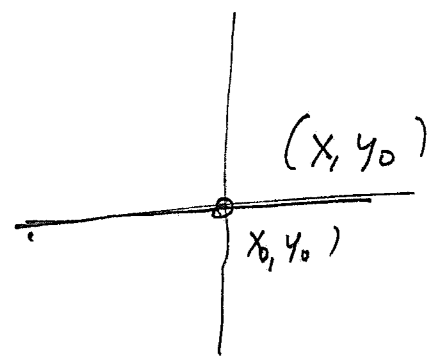
(formally the tangent plane should be the 2-dim'l plane spanned by the "tangent vectors" to curves on the graph $z = f(x,y)$ passing through $(x_0, y_0, f(x_0, y_0))$)

then $a = \frac{\partial f}{\partial x}(x_0, y_0)$, $b = \frac{\partial f}{\partial y}(x_0, y_0)$

Thus the linear approximation (or equation of the tangent plane to the graph(f) at x_0, y_0))

21.1





curves on graph f

$$\gamma_1(x) = (x, y_0, f(x, y_0)) \qquad x \text{ near } x_0$$

$$\gamma_2(y) = (x_0, y, f(x_0, y)) \qquad y \text{ near } y_0$$

with tangent vectors at (x_0, y_0)

$$\gamma_1'(x_0) = (1, 0, f_x(x_0, y_0))$$

$$\gamma_2'(y_0) = (0, 1, f_y(x_0, y_0))$$

21.2

normal direction

$$N = \cancel{0,1,0} \ (1, 0, f_x(x_0, y_0)) \times (0, 1, f_y(x_0, y_0))$$

$$= \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 1 & 0 & f_x \\ 0 & 1 & f_y \end{vmatrix}$$

$$= (-f_x(x_0, y_0), -f_y(x_0, y_0), 1)$$

<u>tangent plane</u> !

$$-(x-x_0) f_x(x_0, y_0) - (y-y_0) f_y(x_0, y_0) + z - f(x_0, y_0)$$

$$= 0$$

or

$$z = f(x_0, y_0) + \cancel{f_x}(x_0, y_0)(x-x_0) + f_y(x_0, y_0)(y-y_0)$$

is $z = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0)$

Defn' We say $f(x, y)$ is differentiable at (x_0, y_0) if $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ exist at (x_0, y_0) and if

$$\frac{\left| f(x,y) - \left\{ f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) \right\} \right|}{\sqrt{(x-x_0)^2 + (y-y_0)^2}}$$

$\to 0$ as $(x, y) \to (x_0, y_0)$ (i.e $\sqrt{(x-x_0)^2 + (y-y_0)^2} \to 0$)

This is analogous to

$$\frac{\left| f(x) - (f(x_0) + f'(x_0)(x - x_0)) \right|}{|x - x_0|} \to 0 \quad \text{as} \quad x \to x_0$$

for a function of one variable.

The derivative as the best linear approximation

Let $f: \mathbb{R} \to \mathbb{R}$ be differentiable. Then near $x=a$

$$f(x) \approx f(a) + f'(a)(x-a) = l(x)$$

$l(x)$ is good approximation if

$$\lim_{h\to 0} \frac{\varepsilon(h)}{h} = \lim_{h\to 0} \frac{f(a+h) - f(a) - f'(a)\,h}{h} = 0$$

Suppose $\vec{\gamma}(t) : \mathbb{R} \to \mathbb{R}^n$ is a parametric curve in $\mathbb{R}^n$. Then $\vec{\gamma}$ is differentiable at $t=a$ if

$$\vec{\gamma}(a+h) \approx \vec{\gamma}(a) + \vec{\gamma}'(a)\,h$$

i.e

$$\lim_{h\to 0} \frac{\vec{\gamma}(a+h) - \vec{\gamma}(a) - \vec{\gamma}'(a)\,h}{h} = \vec{0}$$

Now suppose $f : \vec{\mathbb{R}}^n \to \mathbb{R}$

and $\vec{a} = (a_1, \ldots, a_n) \in \mathbb{R}^n$

Then f is diff. at a

if $f(\vec{a}+\vec{h}) \approx f(\bar{a}) + Df_{\vec{a}} \cdot \vec{h}$

$Df_{\vec{a}} = \nabla f(a)$

$\vec{h} = (0, \ldots \underset{i^{th}}{1} \ldots 0)$

$\Rightarrow Df_{\vec{a}} \cdot \vec{h} = \dfrac{\partial f}{\partial x_i}(\bar{a})$ so necessarily $\dfrac{\partial f}{\partial x_i}(\bar{a})$ must exist (weak condition)

$\vec{a} = (a_1, a_2) \in \mathbb{R}^2$

$n = 2$

$$\lim_{(h_1,h_2)\to(0,0)} \frac{\left| f(a_1+h_1, a_2+h_2) - \big(f(a_1,a_2) + f_x(a_1,a_2)h_1 + f_y(a_1,a_2)h_2 \big) \right|}{\sqrt{h_1^2+h_2^2}} = 0$$

<u>Theorem</u> If $f : U \subset \mathbb{R}^n \to \mathbb{R}$ has continuous partial derivatives in a neighborhood of $x_0 \in U$, then f is differentiable at x_0 and $Df(x_0) = \nabla f(x_0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} e_i$

Ch. 2

2.1-2.6

p 76-144

The differential ($D_{\vec a} f$) of $f: R^n \to R^m$

= a linear map ~~with~~ $L: R^n \to R^m$

so that $f(\vec a + \vec h) \approx f(\vec a) + L(\vec h)$

in sense that

$$\lim_{\vec h \to \vec 0} \frac{f(\vec a+\vec h) - f(\vec a) - L(\vec h)}{\|\vec h\|} = \vec 0$$

↑ vector in R^n

$m=1$ $\quad L(\vec h) = D_{\vec a} f(h)$

$$= \frac{\partial f}{\partial x_1}(\vec a)\, h_1 + \dots + \frac{\partial f}{\partial x_n}(\vec a)\, h_n$$

26.5

Let $f : R^n \to R^m$

If we introduce $e_1, \dots, e_n$ o.n. basis of R^n (coordinates if you like)

and $\bar{e}_1, \dots, \bar{e}_m$ o.n basis of R^m

Then we may think of f as having m components $f(x) = (f^1(x), \dots f^m(x))$

where each f^j is a function of $x \in R^n$

Then $Df(x_0)$ is an $m \times n$ matrix; called the Jacobian matrix

$$Df(x_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \cdots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \vdots & & \\ \frac{\partial f_m}{\partial x_1}(x_0) & \cdots & \frac{\partial f_m}{\partial x_n}(x_0) \end{pmatrix} = \left(\frac{\partial f_i}{\partial x_j}\right)_{m \times n}$$

$1 \le i \le m$

$1 \le j \le n$

Note that the linear approximation is

"small vector in R^n"

$$\vec{f}(x_0+\vec{h}) \approx \vec{f}(x_0) + \left(Df(x_0)\right)\vec{h}$$

$m \times n$ matrix

small column vector $n \times 1$ ie

vectors in R^m

"matrix multiplication" i^{th} row $\cdot$ j^{th} column

$(m \times n)(n \times 1) = m \times 1 =$ (column) vector in $\mathbb{R}^m$

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

$$f(x,y) = (x+y^2, x^3+5y)$$

$\vec{a} = (1,1)$. Then $Df_{\vec{a}}(h_1, h_2) = L(h_1, h_2)$

$$= (h_1 + 2h_2, \; 3h_1 + 5h_2)$$

Check $\displaystyle \lim_{(h_1,h_2)\to(0,0)} \frac{f(1+h_1, 1+h_2) - f(1,1) - L(h_1,h_2)}{\|h\|}$

$$= \lim_{(h_1,h_2)\to(0,0)} \frac{\left(\cancel{1}+\cancel{h_1}+(\cancel{1}+h_2)^2 - \cancel{2} - (\cancel{h_1}+2h_2),\; (\cancel{1}+h_1)^3 + 5(\cancel{1}+\cancel{h_2}) - \cancel{6} - (3h_1+5\cancel{h_2})\right)}{\sqrt{h_1^2+h_2^2}}$$

$$= \lim_{(h_1,h_2)\to(0,0)} \frac{\left(h_2^2,\; 3h_1^2 + h_1^3\right)}{\sqrt{h_1^2+h_2^2}} = (0,0) \quad \text{(exercise)}$$

2.4 Paths and curves

We have already discussed

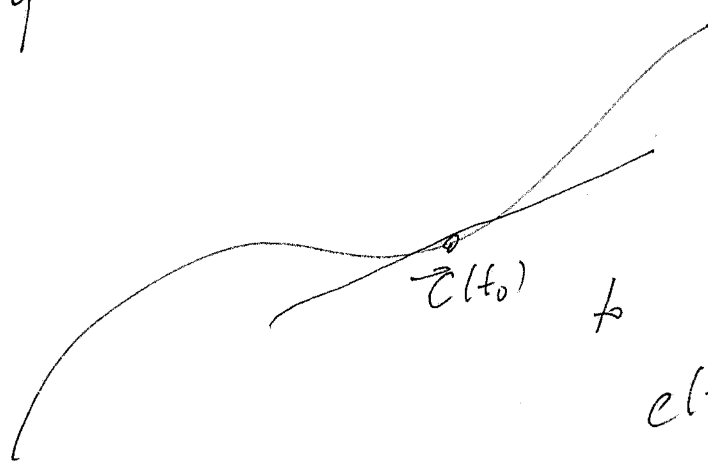
vector functions $\vec{c} : [a, b] \to \mathbb{R}^n$

We may think of $\vec{c}(t)$ as a "path"

a curve in $\mathbb{R}^n$

We say $\vec{c}$ is regular at t_0 if $\vec{c}'(t_0) \neq 0$

Then



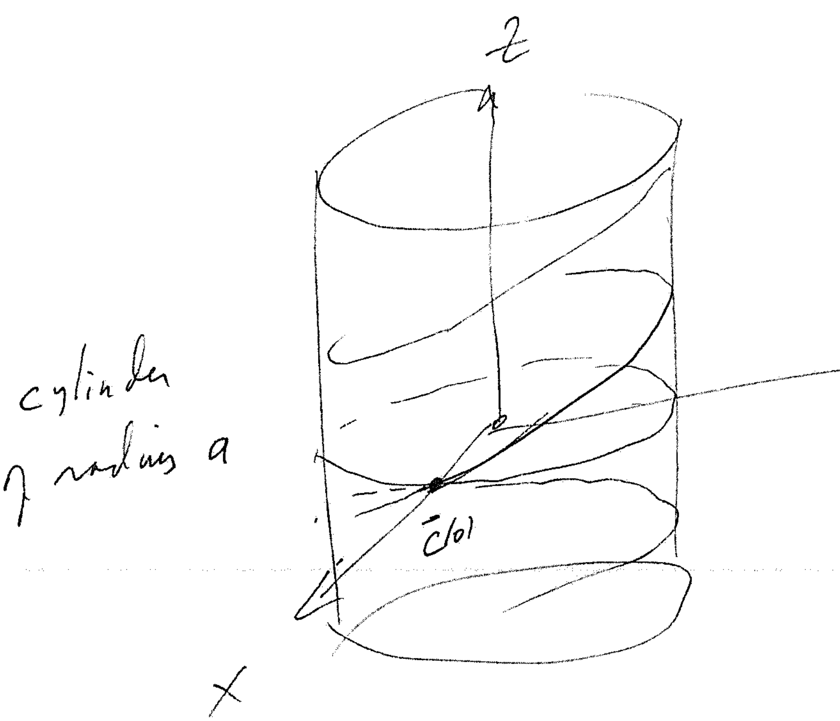
tangent (velocity) vector to path $\vec{c}$ at $c(t_0)$

tangent line $\vec{X}(s) = \vec{c}(t_0) + s\,\vec{c}'(t_0)$ "lives in $\mathbb{R}^n$"

Example 1 (helix)

$$\vec{c}(t) = (a\cos t, a\sin t, bt) \qquad a, b > 0$$

$$-\infty < t < \infty$$



moves on the cylinder of radius a with the z-component increasing at constant speed b

$$\vec{c}\,'(t) = (-a\sin t, a\cos t, b)$$

Example 4. (circular motion)

$$\vec{X}(t) = R\cos\omega t\, \vec{e}_1 + R\sin\omega t\, \vec{e}_2$$

$\vec{i}$ $\vec{j}$

$R > 0$

$$= (R\cos\omega t, R\sin\omega t)$$

lies on the circle of radius R (center $(0,0)$)

and the angle between the positive x axis

and $\vec{X}(t)$ is ωt

If $\omega > 0$ $\vec{X}(t)$ moves about the circle

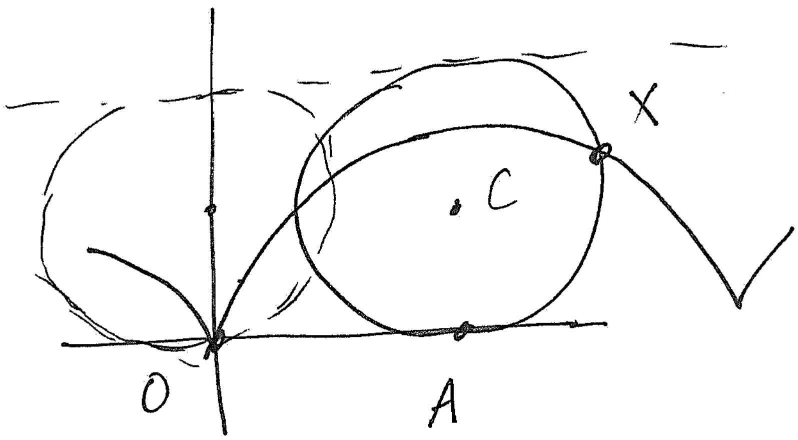
counterclockwise. Since ω is the rate

of change of this angle, it is called

the angular velocity of the motion.

Example 3 (the cycloid)

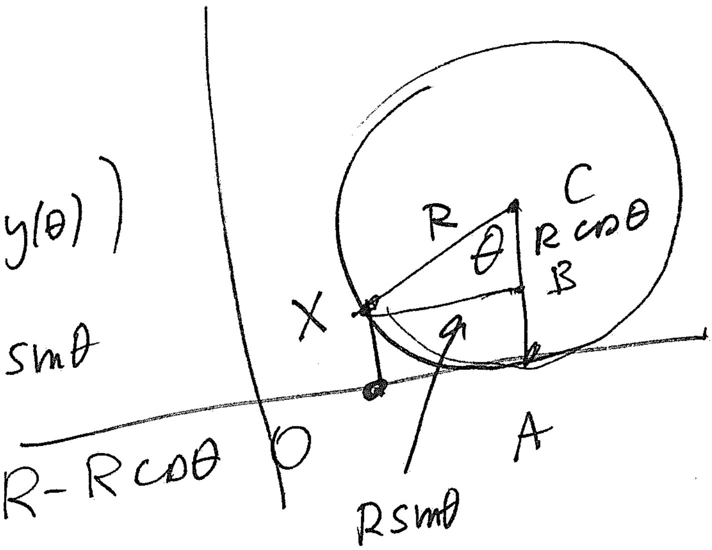
The cycloid is the curve obtained by rolling a wheel of radius R on the ground and mark + follow the point on the wheel that touches the ground. Let X denote this point; then $X = X(\theta)$ where θ is the angle the wheel has turned since it was on the ground.



$X(\theta) = (x(\theta), y(\theta))$

$x(\theta) \overset{R\theta}{=} R\sin\theta$

$y(\theta) = R - R\cos\theta$



$|OA| = |\overset{\frown}{AX}| = R\theta$

so x coord. of A, B, C is $R\theta$

hence

$$\vec{X}(\theta) = R\left((\theta - \sin\theta)\,\vec{\imath} + (1-\cos\theta)\,\vec{\jmath}\right)$$

$$\vec{X}'(\theta) = R\left((1-\cos\theta)\,\vec{\imath} + \cos\theta\,\vec{\jmath}\right)$$

The Chain rule - a simple case

given a path in the plane $x = x(t)$ $y = y(t)$ and a differentiable function $f(x,y)$,

define $g(t) = f(x(t), y(t))$

(the restriction of f to the path)

what is $g'(t_0)$?

$$\frac{g(t_0+h) - g(t_0)}{h} = \frac{f(x(t_0+h), y(t_0+h)) - f(x(t_0), y(t_0))}{h}$$

We suppose the path is differentiable,

So

$$x'(t_0) \approx \frac{x(t_0+h) - x(t_0)}{h} \quad \text{linear approx.}$$

$$y'(t_0) \approx \frac{y(t_0+h) - y(t_0)}{h}$$

so

$$x(t_0+h) \approx x(t_0) + h\, x'(t_0)$$

$$y(t_0+h) \approx y(t_0) + h\, y'(t_0)$$

Now $f(x,y)$ is assumed differentiable

so $f(x,y)$ is well approximated by its linear approximation (tangent plane)

$$f(\underset{x_0}{x(t_0)} + \underset{\Delta x}{h\, x'(t_0)},\ \underset{y_0}{y(t_0)} + \underset{\Delta y}{h\, y'(t_0)})$$

$$\approx f(x(t_0), y(t_0)) + f_x(x(t_0), y(t_0))\, \underbrace{h\, x'(t_0)}_{\Delta x}$$

$$+ f_y(x(t_0), y(t_0))\, \underset{\Delta y}{h\, y'(t_0)}$$

error terms are $o(h)$ (ie go to zero faster than h)

Hence

$$\frac{g(t_0+h) - g(t_0)}{h} \approx f_x(x(t_0), y(t_0))\, x'(t_0) + f_y(x(t_0), y(t_0)) \cdot y'(t_0)$$

and in fact

$$g'(t_0) = \lim_{h \to 0} \frac{g(t_0+h) - g(t_0)}{h} = f_x(x(t_0), y(t_0))\, x'(t_0) + f_y(x(t_0), y(t_0))\, y'(t_0)$$

Sometimes this is written

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

g

"total derivative"

One "should really write"

$$\frac{df}{dt}(x(t), y(t)) = \frac{\partial f}{\partial x}(x(t), y(t))\, \cancel{\frac{dx}{dt}}\, x'(t) + \frac{\partial f}{\partial y}(x(t), y(t))\, y'(t)$$

__Example__ The temperature at a point (x,y) in the plane is $T(x,y)$ and an ant is walking along the parametric curve

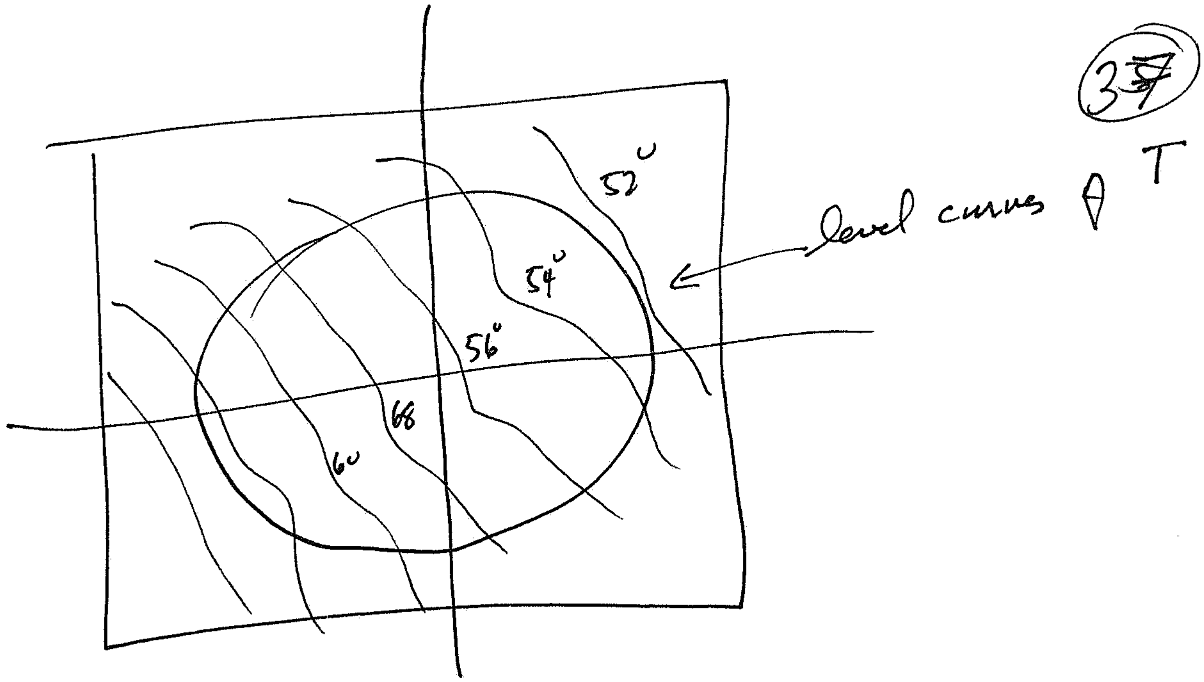
$$x(t) = R\cos\omega t \quad y(t) = R\sin\omega t$$

(on a circle of radius R with angular velocity ω)

How fast is the temperature changing?

Since no explicit fn $T(x,y)$ is given, we must use the chain rule:

$$\frac{dT}{dt}(x(t),y(t)) = -\frac{\partial T}{\partial x}(R\cos\omega t, R\sin\omega t)\cdot \omega R\sin\omega t$$
$$+ \frac{\partial T}{\partial y}(R\cos\omega t, R\sin\omega t)\,\omega R\cos\omega t$$


52°
54°
56°
68
60
level curves of T

The gradient + level curves

If we write $\vec{X}(t) = (x(t), y'(t))$

for the vector form of the path,

then we can write

$$\frac{d}{dt} f(\vec{X}(t)) = \nabla f(\vec{X}(t)) \cdot \vec{X}'(t)$$

where $\nabla f(x,y) = f_x(x,y)\,\vec{i} + f_y(x,y)\,\vec{j}$

is the "differential of f"

Recall $D_{\vec{v}} f$ (the directional derivative)

is given by $D_{\vec{v}} f = \nabla f \cdot \vec{v}$

If $\vec{v}$ has length 1, i.e. $\|\vec{v}\| = 1$

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

$$\theta = \measuredangle(\nabla f, \vec{v})$$

Hence $D_{\vec{v}} f$ is greatest when

$\theta = 0$ or $\vec{v} = \dfrac{\nabla f}{\|\nabla f\|}$ (assuming $\nabla f \neq 0$)

and smallest when $\theta = \pi$ i.e

$$\vec{v} = -\frac{\nabla f}{\|\nabla f\|}$$

Another (even more) important

interpretation of $\nabla f(x,y)$ comes

from looking at the "level

curves" of f

Let (x_0, y_0) be a point

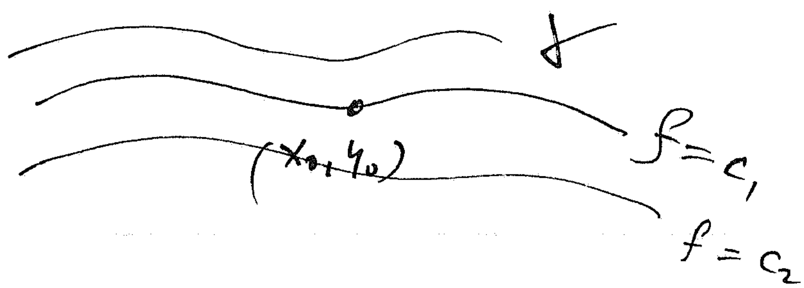
where $f(x_0, y_0) = c_1$

and suppose there is a differentiable

(regular) curve γ, $\vec{X}(t) = (x(t), y(t))$ passing through

(x_0, y_0) on which

$f(x(t), y(t)) = c_1$ and say $\vec{X}(t_0) = (x_0, y_0)$



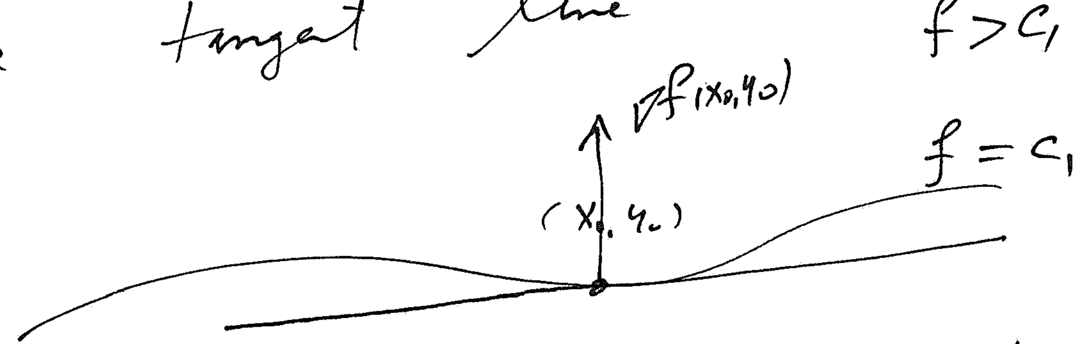
Then

$$0 = \frac{d}{dt} f(x(t), y(t)) \Big|_{t=t_0} = \nabla f(x_0, y_0) \cdot \vec{X}'(t_0)$$

i.e the tangent (velocity) vector

$\bar{X}'(t_0)$ is perpendicular to $\nabla f(x_0, y_0)$

Assuming $\nabla f(x_0, y_0) \neq 0$ we say $\nabla f(x_0, y_0)$ is "normal" to the level curve, which just means that is it the normal to the tangent line



and "points to the side where $f > c_1$"

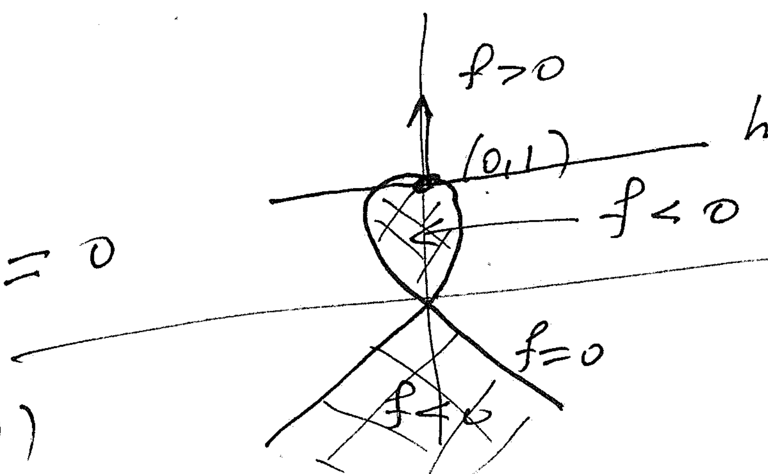
Example Consider

$$f(x,y) = x^2 - y^2 + y^3$$

near the point $(0,1)$

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

The level curve $f=0$ thru $(0,1)$



~~which~~ looks something like this.

Note that $f(0,0) = 0$ but we run into trouble since

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

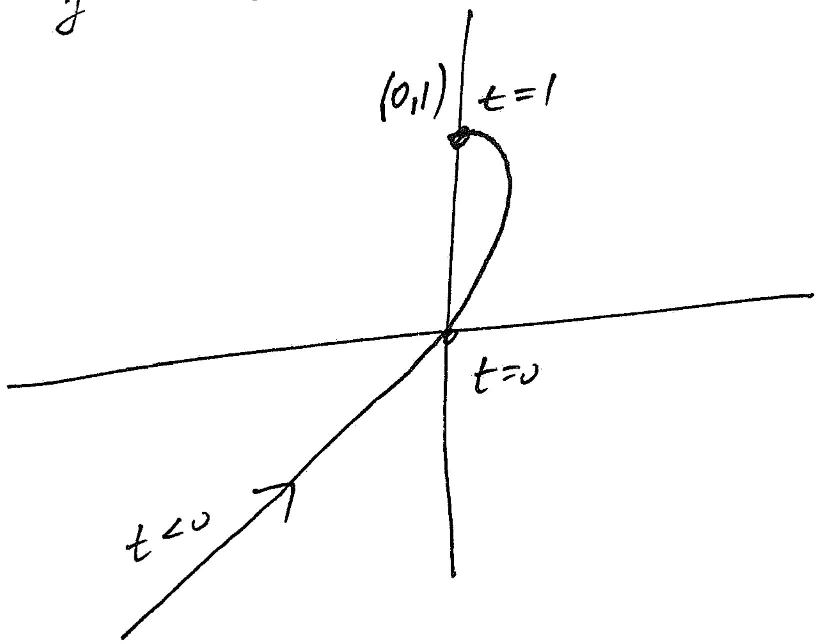
4B

we see from the plot
(which can be done by hand
using polar coordinates $x = r\cos\theta$
$y = r\sin\theta$
but I cheated and used
Wolfram Alpha level curve grapher)

that the level curve has a
"self-intersection" at $(0,0)$ and
there are in fact two tangent
lines. We can "parametrize"
"each piece" of the level curve by choosing

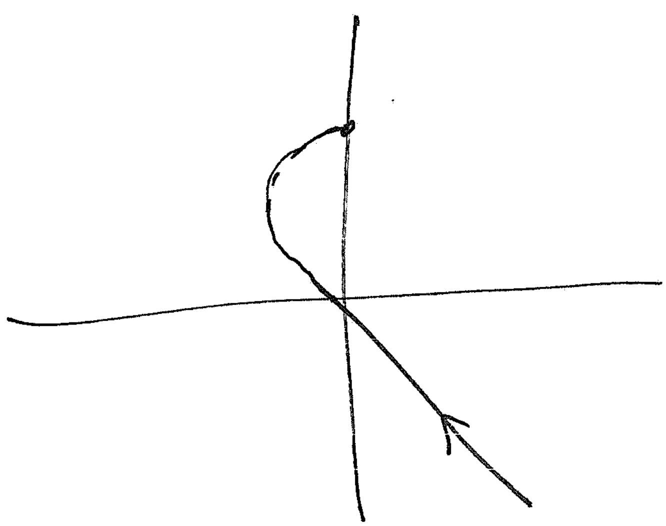
44

$$x = t\sqrt{1-t} \qquad -\infty < t \leq 1$$
$$y = t$$



$$-\infty < t \leq 1$$

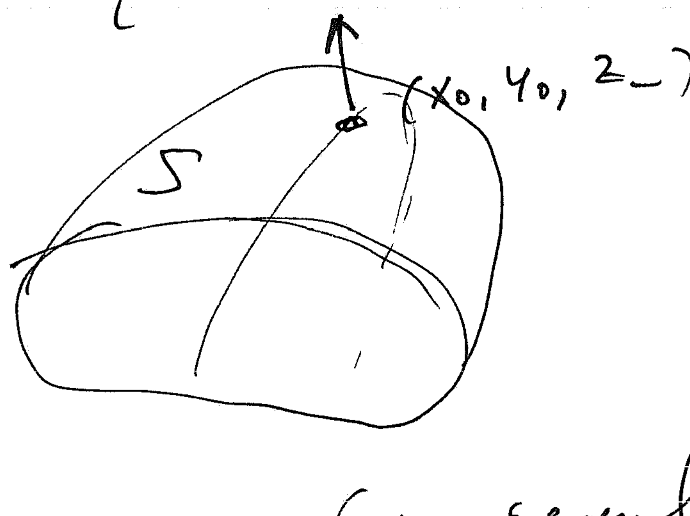
n $$x = -t\sqrt{1-t}$$
$$y = t$$



<u>The chain rule</u> for
$f(x, y, z)$ or $f(\vec{x}) = f(x_1, \ldots, x_n)$

If we look for a
<u>level surface</u> (suppose
$f(x_0, y_0, z_0) = c$)
passing thru (x_0, y_0, z_0)

$$S = \{ (x, y, z) : f(x, y, z) = c \}$$



S is now (in general two dimensional)

then the linear approximation to f at (x_0, y_0, z_0) is

$$f(x,y,z) \simeq f(x_0, y_0, z_0) + \frac{\partial f}{\partial x}()\,(x - x_0) + \frac{\partial f}{\partial y}()\,(y-y_0) + \frac{\partial f}{\partial z}\,(z - z_0)$$

Assuming that $\nabla f(x_0, y_0, z_0) \neq 0$ this suggests that the tangent plane to S should be given

by $\quad \nabla f(\vec{X}_0) \cdot (\vec{X} - \vec{X}_0) = 0$

$$\vec{X}_0 = (x_0, y_0, z_0)$$

45

The mathematical theorem that justifies this is called *the Implicit Function theorem*

Then we can argue (as before) that $\nabla f(\vec{X}_0)$ is normal to the tangent plane as follows.

Take any parametric curve $\vec{X}(t)$ on S passing through (x_0, y_0, z_0) at $t = t_0$

Then $0 = f(\vec{X}(t)) \quad \forall t$

So

$$0 = \frac{d}{dt} f(\vec{X}(t))\Big|_{t=t_0} \overset{\text{chain rule}}{=}$$

$$\nabla f(\vec{X}_0) \bullet \vec{X}'(t_0)$$

velocity vector (tangent vector)

to an arbitrary curve on S passing

thru (x_0, y_0, z_0)

These tangent vectors span the

tangent plane so $\nabla f(\vec{X}_0)$ is

normal to this plane.

<u>Example</u> Find the tangent plane to $x^2+y^2+z^2=16$ at the point $(1, 3, \sqrt{6}) = \vec{X_0}$

Take $f(x,y,z) = x^2+y^2+z^2$

$c = 16$. Then $\nabla f(\vec{X}) = 2\vec{X}$

$= (2, 6, 2\sqrt{6})$ at $(1, 3, \sqrt{6})$

Hence the equation of the tangent plane is

$$(\vec{X} - \vec{X_0}) \cdot (2, 6, 2\sqrt{6}) = 0$$

or $$(x-1) + 3(y-3) + \sqrt{6}(z-\sqrt{6}) = 0$$

i.e $$x + 3y + \sqrt{6}z = 16$$

Properties of the derivative

1. Let $f, g : U \subset \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at $\vec{x}_0$. Then

linearity

a. $D(cf)(x_0) = c\,Df(x_0)$

b. $D(f+g)(x_0) = Df(x_0) + Dg(x_0)$

2. Let $f, g : U \subset \mathbb{R}^n \to \mathbb{R}$ be differentiable at $\vec{x}_0$. Then

product a. $\nabla(fg)(x_0) = g(x_0)\nabla f(x_0) + f(x_0)\nabla g(x_0)$

quotient b. $\nabla(f/g)(x_0) = \frac{1}{g(x_0)^2}\left(g(x_0)\nabla f(x_0) - f(x_0)\nabla g(x_0)\right)$

The general Chain rule

Recall from calculus, that

$$\frac{d}{dx}(g\circ f)(x) = g'(f(x))\, f'(x)$$

$$\qquad\qquad\qquad\quad '' \; g'(u)\Big|_{u=f(x)}$$

General Chain rule

Let $f : U \subset R^n \to R^m$ and

$\quad\;\; g : f(U) \subset R^n \to R^p$

be differentiable functions

Then the composition

$$g\circ f : U \subset R^n \to R^p$$

is differentiable and

$$D(g\circ f)(x_0)\vec{v} = Dg(f(x_0))\big(Df(x_0)\vec{v}\big)$$

where the rhs is the composition of linear maps acting on $\vec{v}$

In terms of the matrix representations if $Jf(x_0)$ is the $m \times n$ Jacobian matrix of f at x_0 and $Jg(f(x_0))$ is the $p \times m$ Jacobian matrix of g at $f(x_0)$, then

$$J(g \circ f)(x_0) = \underset{p \times m}{Jg(f(x_0))} \cdot \underset{m \times n}{Jf(x_0)}$$

so is a $p \times n$ matrix

(as it should be)

$$f : R^3 \longrightarrow R^3$$

con. x, y, z coord. (u, v, w)

(58)

$$f(x,y,z) = (u(x,y,z), v(x,y,z), w(x,y,z))$$

$$g = g(u,v,w) : R^3 \to R$$

Then $g \circ f : R^3 \to R$

and $\nabla(g \circ f) = \left(\frac{\partial g}{\partial u}, \frac{\partial g}{\partial v}, \frac{\partial g}{\partial w}\right) \begin{pmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{pmatrix}$

$$= \left(\frac{\partial g}{\partial u} u_x + \frac{\partial g}{\partial v} v_x + \frac{\partial g}{\partial w} w_x, \ \frac{\partial g}{\partial u} u_y + \frac{\partial g}{\partial v} v_y + \frac{\partial g}{\partial w} w_y, \ \frac{\partial g}{\partial u} u_z + \frac{\partial g}{\partial v} v_z + \frac{\partial g}{\partial w} w_z\right)$$

This is usually computed

as follows:

~~Let $h(x,y,z) = f(u(x,y,z), v(x,y,z), w(x,y,z))$~~

Let $h(x,y,z) = g(u(x,y,z), v(x,y,z), w(x,y,z))$

Then $\frac{\partial h}{\partial x} = \frac{\partial g}{\partial u} \cdot u_x + \frac{\partial g}{\partial v} \cdot v_x + \frac{\partial g}{\partial w}\, \partial w_x$

etc

<u>Example</u> Let $a(u,v,w) = u^2 w + v^2$
$b(u,v,w) = uvw + u - w$

and $u = x - yz$
$v = x^2 y - z$
$w = xy^3 + z$

Find $\frac{\partial b}{\partial x}$

Consider $g \circ f$ where

$$f : \underset{(x,y,z)}{R^3} \longrightarrow \underset{(u,v,w)}{R^3}$$

$$f(x,y,z) = \big(u(x,y,z),\ v(x,y,z),\ w(x,y,z)\big)$$

$$g : \underset{(u,v,w)}{R^3} \longrightarrow \underset{(a,b)}{R^2}$$

$$g(u,v,w) = \big(a(u,v,w),\ b(u,v,w)\big)$$

Then $g \circ f : R^3 \longrightarrow R^2$

$$J_g = \begin{pmatrix} 2uw & 2v & u^2 \\ vw+1 & uw & uv-1 \end{pmatrix}$$

$$J_f = \begin{pmatrix} 1 & -1 & -1 \\ 2xy & x^2 & -1 \\ y^3 & 3xy^2 & 1 \end{pmatrix}$$

$J_{g\circ f}$

56

But remember J_g must be evaluated at $f(x,y,z)$ so

$$J_g(f(x,y,z)) = \begin{pmatrix} 2(x-yz)(xy^3+z) & 2(x^2y-z) & (x-yz)^2 \\ (x^2y-z)(xy^3+z)+1 & (x-yz)(xy^3+z) & (x-yz)(x^2y-z)-1 \end{pmatrix}$$

So this is quite tedious

Instead we compute what we need

$$\frac{\partial b}{\partial x} = \frac{\partial b}{\partial u} u_x + \frac{\partial b}{\partial v} v_x + \frac{\partial b}{\partial w} w_x$$

$$= (vw+1) \cdot 1 + uw \cdot 2xy + (uw-1)\, y^3$$

$$= \left[(x^2y-z)(xy^3+z) + 1\right] \cdot 1 + (x-yz)(xy^3+z) \cdot 2xy$$
$$+ \left[(x-yz)(xy^3+z) - 1\right] y^3$$