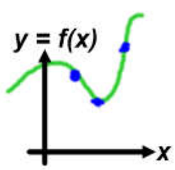
14.1: Functions of Several Variables

Multivariable Functions

2D functions

One numerical value in the domain maps to one numerical value in the range

x	$y = f(x)$
2	8
3	5
6	12



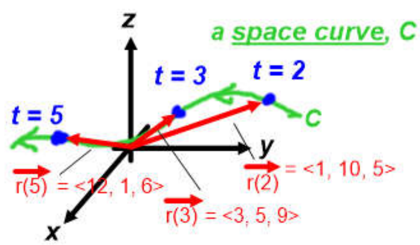
3D vector functions

One numerical value in the domain (parameter t) maps to a vector, $\vec{r}(t)$ in the range

t = 'the parameter'

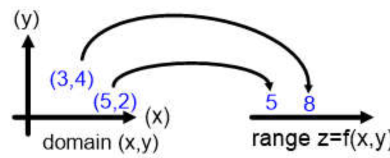
t	$\vec{r}(t) = \langle x(t), y(t), z(t) \rangle$
2	<1, 10, 5>
3	<3, 5, 9>
5	<12, 1, 6>

x(t), y(t), z(t) are the parametric equations of C

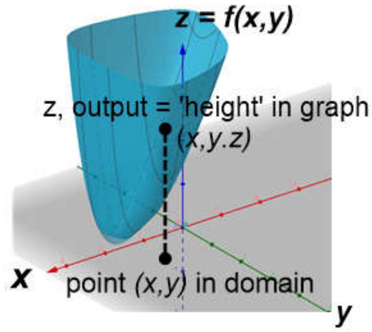


3D multivariable functions

Two numerical values (x,y) in the domain map to one numerical output value in the range

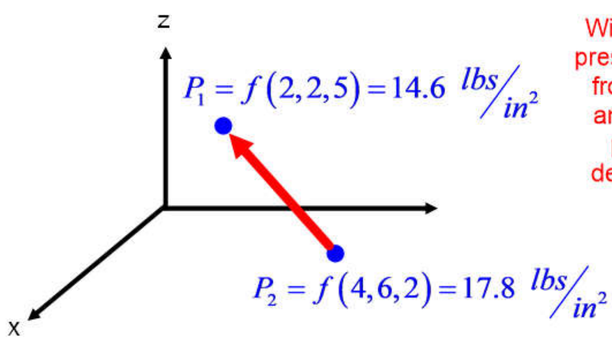


(x,y)	$z = f(x,y)$
(3,4)	8
(5,2)	5



All of this could extend to any number of input variables, but always one output variable (if you need more than one output variable, then use vector functions).

For example, atmospheric pressure P could be a function of x, y, z, which are spatial dimensions, P = f(x,y,z):



Wind might then be caused by a difference in pressure between any two points...wind flowing from higher pressure to lower pressure, with amount of wind proportional to difference in pressure in a specific direction. (Need a derivative in a direction, covered later in this chapter.)

Now, the 3D space is the domain and pressure, P is a numerical variable with values (but no way to show it on this graph, except as numbers).

Sketching multivariable functions

Impractical to sketch surfaces by hand, can use software: **https://www.geogebra.org/3d?lang=en**

For planes, can use intercepts to sketch...

Ex) Sketch $z = f(x, y) = 6 - 3x - 2y$

Domain/range of multivariable functions

Just as with single variable functions, the domain of a multivariable functions represents the set of all input values for which there is a single output value defined and the range is the set of all output values produced.

Things that can limit the domain:

- The function itself might state a domain (whoever created it specifies where it can be used.)
- If the function is defined algebraically, anything that would can an undefined condition must be excluded:
  - Dividing by zero
  - Even roots of negative numbers
  - Logarithms of zero or negative numbers
- If the function is defined graphically in 3D by a surface, then the domain is the part of the x-y plane over which the surface exists.
- If the function is modeling a physical phenomenon, values which make no sense must be excluded (negative liters of volume, negative time, shapes with negative dimensions, etc.)

Ex) Find the domain and range of $h(x, y) = 4x^2 + y^2$

Level curves, level surfaces, contour maps

For a function z = f(x,y) we can find a **level curve** by setting z equal to a constant and tracing out the path in the (x,y) domain which produces this z value.

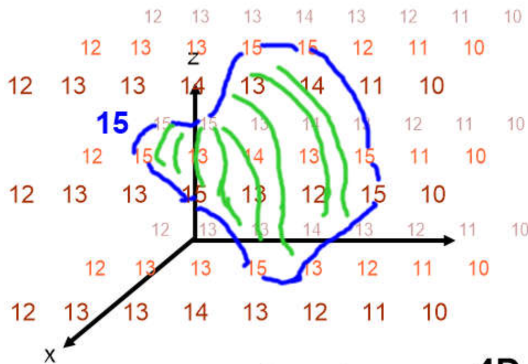
For a function w = f(x,y,z) we can find a **level surface** by setting w equal to a constant and highlight the surface of all points in the (x,y,z) domain which produces this w value.

A set of level curves or level surfaces is called a **contour map**.

Level curves

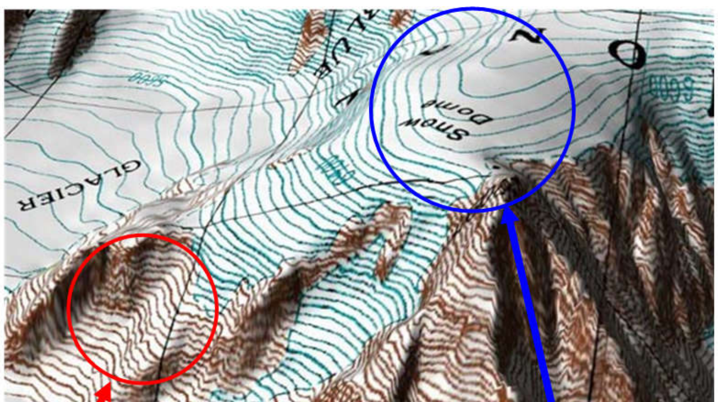
$z = f(x,y)$ (height) **3D**

Level surface



$P = f(x,y,z)$ (pressure) **4D**

Note that how closely spaced the level curves are is determined by how rapidly the output value is changing...fast change = close curve spacing:



In this area, level curves spacing is very close, so height is changing very rapidly (more 'steep')

In this area, level curves spacing is farther, so height is changing less rapidly (less 'steep')

Ex) Draw a contour map showing several curves for $f(x,y) = x^3 - y$

Limits of Multivariable Functions

For single-variable functions

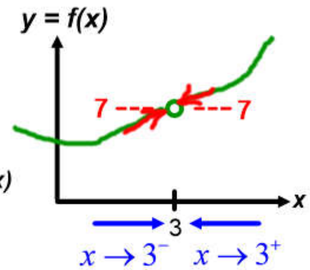
We've studied the idea of a limit in brief calculus and AP Calc BC:

> If, as *x* approaches a value *c* in the domain, *f(x)* approaches value *L* in the range, regardless of the direction of approach, then we say the limit of *f(x)* at *x=c* is *L*.
>
> $$\lim_{x\to c} f(x) = L$$

As we approach x=3 from the left side, y approaches 7, so we say the left-handed limit exists and is 7 (the number being approached).

$$\lim_{x\to 3^-} f(x) = 7$$

x	y = f(x)
2.9	6.87
2.95	6.92
2.99	6.998



x	y = f(x)
3.3	8.12
3.15	7.45
3.01	7.002

As we approach x=3 from the right side, y approaches 7, so we say the right-handed limit exists and is 7 (the number being approached).

$$\lim_{x\to 3^+} f(x) = 7$$

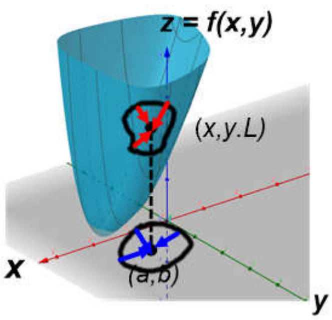
Because the same y value (7) is being approached regardless of direction, we can say the limit as x approaches 3 is 7 (without specifying a direction): $\lim_{x\to 3} f(x) = 7$

Note: it doesn't matter what happens at exactly *x=3*, the limit is about the value being approached.

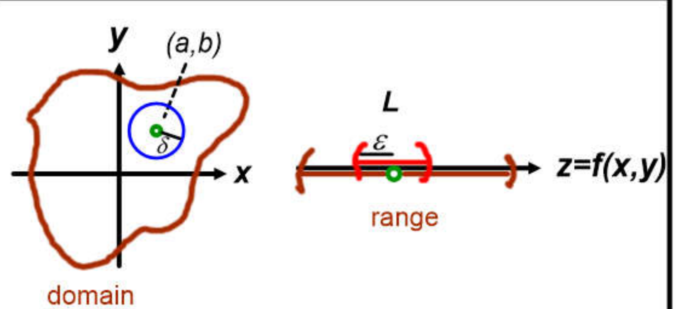
For multivariable functions

> If, as *(x,y)* approaches a point *(a,b)* in the domain, *f(x,y)* approaches value *L* in the range, regardless of the direction of approach, then we say the limit of *f(x,y)* at *x=c* is *L*.
>
> $$\lim_{(x,y)\to(a,b)} f(x,y) = L$$

Ways to picture this:



If you approach *(a,b)* in the domain, from any direction, the same *z* limit value is approached.



For every point *(a,b)* in the domain there is a value *L* in the range and for every $\pm\varepsilon$ region a distance around *L* there exists a circle of radius δ such that

$$0 < \sqrt{(x-a)^2 + (y-b)^2} < \delta \;\; and \; |f(x,y) - L| < \varepsilon$$

Limits of Multivariable Functions

Imagine you have two functions and some values for each are given in the following tables:

Values of $f(x,y)$

x \ y	-3	-2	-1	0	1	2	3
-3	7.3	7.4	7.5	7.5	7.4	7.4	7.3
-2	7.4	7.5	7.5	7.6	7.5	7.5	7.4
-1	7.5	7.6	7.7	7.7	7.7	7.6	7.5
0	7.5	7.7	7.8	7.8	7.7	7.6	7.6
1	7.6	7.8	7.9	7.9	7.7	7.7	7.6
2	7.7	7.9		7.9	7.8	7.8	7.6
3	7.9	8.0	8.1	8.0	7.9	8.0	7.8
4	8.1	8.2	8.3	8.2	8.1	8.1	7.9

Values of $g(x,y)$

x \ y	-3	-2	-1	0	1	2	3
-3	7.3	7.4	4.5	0.5	1.4	3.4	6.3
-2	3.4	5.6	5.5	-1.6	2.5	3.5	5.4
-1	1.5	2.4	6.7	3.7	1.7	2.6	3.5
0	-0.5	-1.4	7.8	-6.8	-5.7	-4.6	-2.6
1	-3.6	-4.5	7.9	-5.9	-7.7	-7.7	-7.6
2	-7.7	-7.8		-7.9	-6.1	-3.2	-5.3
3	-5.9	-2.3	8.1	-3.0	-2.9	1.0	2.8
4	3.1	5.6	8.3	2.2	3.1	4.1	5.9

Both of these functions are undefined at *(2, -1)* but for the limit, that doesn't matter. Instead, we look at what output value is being approached as we approach from different directions.

For *f(x,y)* the output values seem to approach **8.0** as the limit, regardless of direction so we can say:

$$\lim_{(x,y)\to(2,-1)} f(x,y) = 8$$

For *g(x,y)* the output values seem to approach different values depending upon direction so we must say the limit doesn't not exist:

$$\lim_{(x,y)\to(2,-1)} g(x,y) = DNE$$

Computing Limits of Multivariable Functions

Can we find such limits without resorting to listing many values? Yes...we try approaching from directions which are easy to compute (along the x- or y-axes, along the family of lines y=mx, along parabolas, etc.

It is easier to show that the limit ***doesn't*** exist...if we find even one counter-example where different values are being approached, then the whole function does not have a limit at that domain value.

To state a limit exists, in practice, we test a variety of cases and if they all approach the same value we can be reasonably confident the limit exists, but to be certain we must use the *Squeeze Theorem* and find two functions which we can show are above and below the actual value, then show that these functions have the same limit.

This is best to see through examples...

Ex) Find the limits:

$$\lim_{(x,y)\to(2,4)} \frac{x^2 - y^2}{x^2 + y^2}$$

Finding limits algebaically

- Try plugging in (if well behaved, this often works)
- If plugging in results in dividing by zero you can try:
  - factoring
  - synthetic division
  - rationalizing (if there is a radical)
- Use L'Hopital's Rule

We can't use our tactics for single variable functions, because the function can be well-behaved from one direction but discontinuous from other directions.

But we do know that polynomials and rational functions are continuous over their domains.

Computing Limits of Multivariable Functions

Ex) Find the limits:

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

If we are evaluating a limit at a value not in the domain of the function or for any other reason are not sure about the continuity of the function at this domain location, we need to switch tactics: Instead, we evaluate along different functional paths by defining relationships between the function variables to establish a path, then evaluate the limits along many different paths.

We try the following paths, in order of difficulty:

- **The x and y-axes: $y = 0$, $x = 0$.**
- **All lines of form: $y = mx + b$.**
- **All parabolic paths of form: $y = ax^2+bx+c$.**

If a limit does not exist, usually by the time we get through parabolic paths, one of these limit will be DNE...as soon as we encounter one DNE, we declare the whole limit DNE.

If all the limits through parabolas exist, then we still haven't shown conclusively that the limit exists along all paths, so we switch tactics again and use The Squeeze Theorem for Multivariable Functions (will show in later example).

Ex) Find the limits:

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

Ex) Find the limit: $\displaystyle\lim_{(x,y)\to(0,0)} \frac{xy}{x^2 + y^2}$

Ex) Find the limit: $\lim_{(x,y)\to(0,0)} \frac{3x^2y}{x^2+y^2}$

At this point, we believe the limit may exist but haven't shown this conclusively, so we switch tactics to using the Squeeze Theorem for Multivariable Functions.

Consider our function: $\frac{3x^2y}{x^2+y^2}$

This can written as two factors: $\frac{3x^2y}{x^2+y^2} = \frac{x^2}{x^2+y^2}3y$ ...but the left factor will always be less than 1:

$\frac{x^2}{x^2+y^2} < 1$ ...because if y = 0 this = 1, and if y is anything else, the denominator increase. As y approaches infinity, this factor approaches zero.

So depending upon the value of y, this factor will always be somewhere between 0 and 1. That means...

$$(0)3y \le \frac{x^2}{x^2+y^2}3y \le (1)3y$$

$$0 \le \frac{3x^2y}{x^2+y^2} \le 3y \quad \text{if } y>0$$

$$3y \le \frac{3x^2y}{x^2+y^2} \le 0 \quad \text{if } y<0$$

To accomodate both of these cases, we can use an absolute value like this: $0 \le \frac{3x^2|y|}{x^2+y^2} \le 3|y|$

Now for the Squeeze Theorem: if we evaluate the limits on the ends and they evaluate to the same number, then the limit in the middle is also the same value:

$$\lim_{(x,y)\to(0,0)} 0 \le \lim_{(x,y)\to(0,0)} \frac{3x^2|y|}{x^2+y^2} \le \lim_{(x,y)\to(0,0)} 3|y|$$

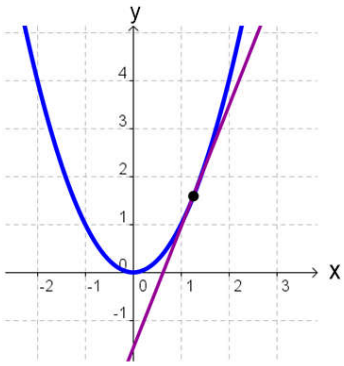
$$0 \le \lim_{(x,y)\to(0,0)} \frac{3x^2|y|}{x^2+y^2} \le 0$$

Therefore, our limit's value is zero (for any path): $\lim_{(x,y)\to(0,0)} \frac{3x^2|y|}{x^2+y^2} = 0$

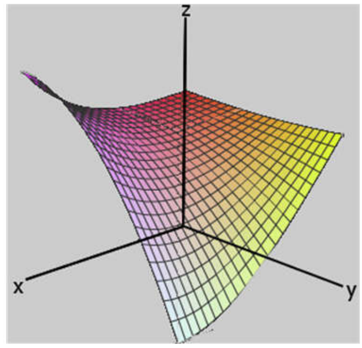
14.3: Partial Derivatives

Partial Derivatives

$y = f(x)$



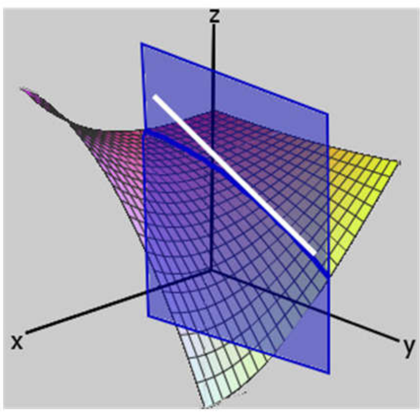
$z = f(x,y)$



Derivative = slope of line tangent to curve

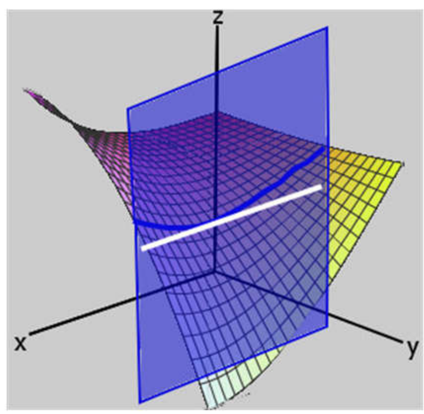
What would derivative mean for a surface?

$z = f(x,y)$



If we hold x constant (on plane x=6) derivative could equal slope of tangent line of z=f(y) curve.

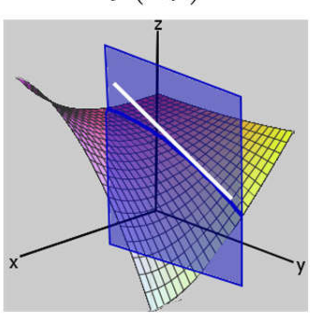
$z = f(x,y)$



If we hold y constant (on plane y=8) derivative could equal slope of tangent line of z=f(x) curve.

These 'directional' derivatives are called partial derivatives.

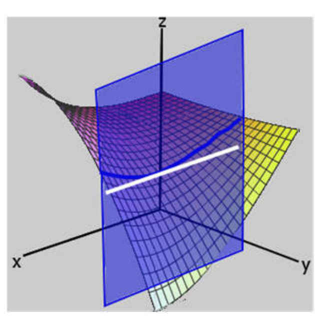
$z = f(x,y)$



$f_y(x,y) = \dfrac{\partial z}{\partial y}$ means:

"Differentiate f **with respect to y** while treating **x as a constant.**"

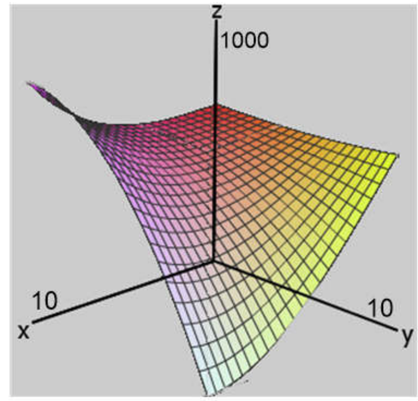
$z = f(x,y)$



$f_x(x,y) = \dfrac{\partial z}{\partial x}$ means:

"Differentiate f **with respect to x** while treating **y as a constant.**"

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



$f_x(x,y) =$

$f_y(x,y) =$

$f_x(5,5) =$

$f_y(5,5) =$

Try it: Find $f_x,\ f_y,\ f_x(2,-1),\ f_y(-2,3)$

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

Higher order partial derivatives

$f_x(x,y)$ *and* $f_y(x,y)$ are known as **'first-order'** partial derivatives.
But these partial derivatives are also function of x and y, so they form their own 'surfaces'. We could take partial derivatives of these, which are called **'second-order'** partial derivatives:

Starting with $f_x(x,y)$	Starting with $f_y(x,y)$
$f_{xx}(x,y) = \dfrac{\partial}{\partial x}\dfrac{\partial z}{\partial x} = \dfrac{\partial^2 z}{\partial x^2}$	$f_{yy}(x,y) = \dfrac{\partial}{\partial y}\dfrac{\partial z}{\partial y} = \dfrac{\partial^2 z}{\partial y^2}$
$f_{xy}(x,y) = \dfrac{\partial}{\partial y}\dfrac{\partial z}{\partial x} = \dfrac{\partial^2 z}{\partial y \partial x}$	$f_{yx}(x,y) = \dfrac{\partial}{\partial x}\dfrac{\partial z}{\partial y} = \dfrac{\partial^2 z}{\partial x \partial y}$

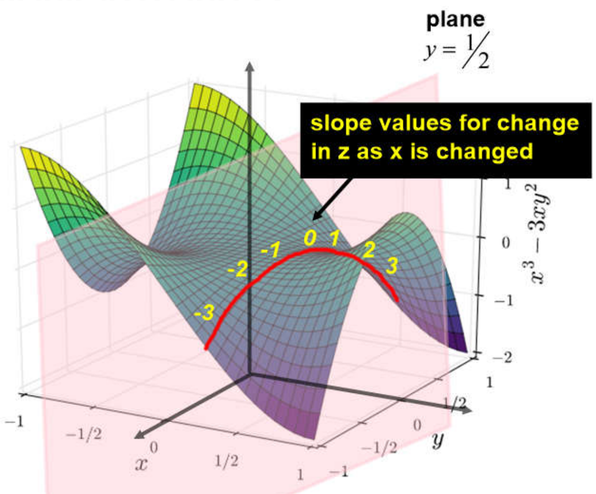
$f_{yx}(x,y)$ $\qquad$ $\dfrac{\partial^2 z}{\partial x \partial y}$

Interpretation of 2nd partial derivatives

f_{xx} is the 2nd derivative with x always as the variable and y constant (on a plane)

This represents how quickly the first derivative is changing as x is increased, meaning how quickly the 'slope' is changing.

Therefore f_{xx} represents the **concavity** of the trace of the curve on the constant y-plane at this point.

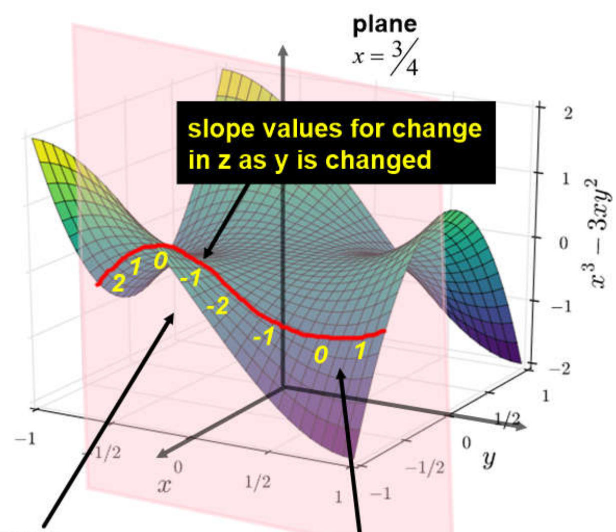


As x increases, slope decreases, so $f_{xx} < 0$ and surface is **concave down** in the x direction.

f_{yy} is the 2nd derivative with x always as the variable and y constant (on a plane)

This represents how quickly the first derivative is changing as y is increased, meaning how quickly the 'slope' is changing.

Therefore f_{yy} represents the **concavity** of the trace of the curve on the constant x-plane at this point.



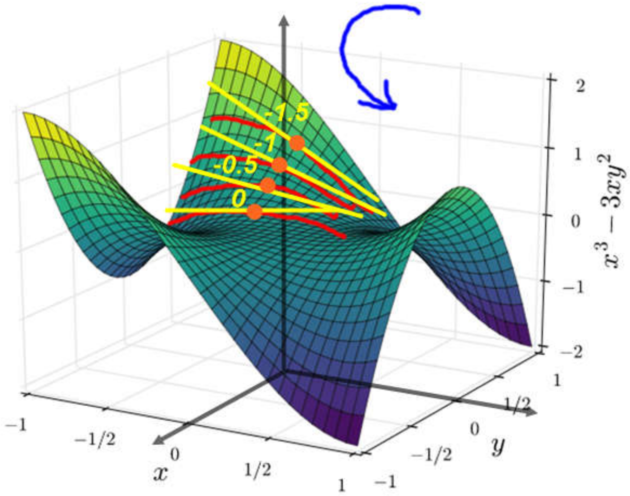
As y increases, slope decreases, so $f_{yy} < 0$ and surface is **concave down** in the y direction here.

As y increases, slope increase, so $f_{yy} > 0$ and surface is **concave up** in the y direction here.

Interpretation of 2nd partial derivatives

f_{yx} represents how quickly the
This represents how quickly the first derivative in the y direction is changing if we move in the x direction.

I've never seen a stated official meaning for the mixed derivative, but to me it seems like it represents how rapidly the surface is **'twisting'** in the region of a given point.

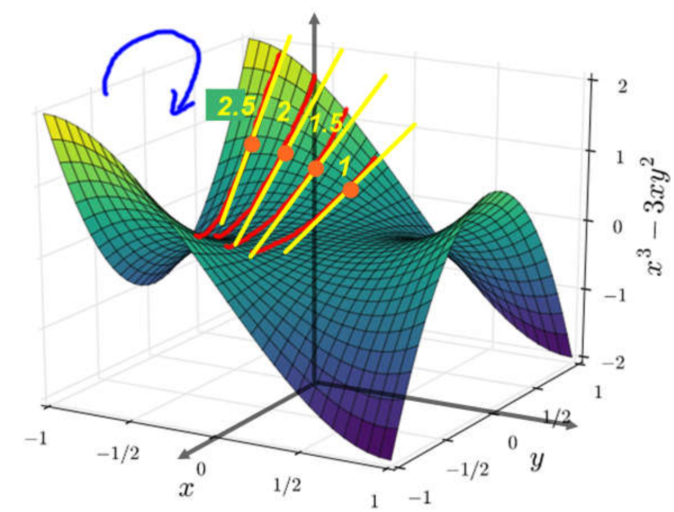


$f_{yx} \approx -0.5$ *per unit in x*

Surface is twisting slowly as we move in x direction.

f_{xy} represents how quickly the
This represents how quickly the first derivative in the x direction is changing if we move in the y direction.

I believe this represents rate of **'twisting',** but now as we move in the y direction.



$f_{xy} \approx -0.5$ *per unit in y*

Surface is twisting slowly as we move in y direction.

Clairaut's Theorem

Suppose f is defined on a disk D that contains the point (a,b).
If the functions f_{xy} and f_{yx} are both continuous on D, then

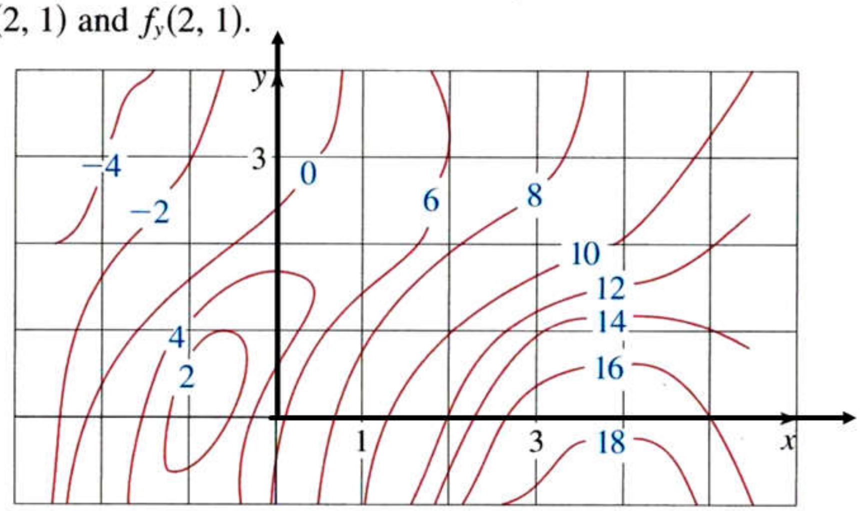
$$f_{xy}(a,b) = f_{yx}(a,b)$$

If our interpretations of mixed derivative are correct, this means that in a small region the rate of twisting in x and y directions is always the same (which makes sense because it is a smooth surface).

Ex) Find $f_{xx}, f_{yy}, f_{xy}, f_{yx}$ for $f(x, y) = 2x^3 - 3y + x^2$

Ex) Find f_x, f_y, f_z for $f(x,y,z) = e^{xy} \ln z$

10. A contour map is given for a function f. Use it to estimate $f_x(2, 1)$ and $f_y(2, 1)$.



4. The wave heights h in the open sea depend on the speed v of the wind and the length of time t that the wind has been blowing at that speed. Values of the function $h = f(v, t)$ are recorded in feet in the following table.

Duration (hours)

Wind speed (knots) v \ t	5	10	15	20	30	40	50
10	2	2	2	2	2	2	2
15	4	4	5	5	5	5	5
20	5	7	8	8	9	9	9
30	9	13	16	17	18	19	19
40	14	21	25	28	31	33	33
50	19	29	36	40	45	48	50
60	24	37	47	54	62	67	69

(a) What are the meanings of the partial derivatives $\partial h/\partial v$ and $\partial h/\partial t$?

(b) Estimate the values of $f_v(40, 15)$ and $f_t(40, 15)$. What are the practical interpretations of these values?

(c) What appears to be the value of the following limit?

$$\lim_{t\to\infty} \frac{\partial h}{\partial t}$$

Ex) Find the first partial derivatives of $f(x,y) = \int_y^x \cos(t^2)\,dt$

14.4: Tangent Planes, Linear Approximations, and Differentials

Tangent plane to a surface

Starting with an equation of a plane:

$$a(x-x_0)+b(y-y_0)+c(z-z_0)=0$$

(the constant on the right is set by the point)

$$\begin{pmatrix} ax+by+cz-(ax_0+by_0+cz_0)=0 \\ ax+by+cz=(ax_0+by_0+cz_0) \\ ax+by+cz=d \end{pmatrix}$$

Divide by c and solve for $z-z_0$:

$$z-z_0=-\frac{a}{c}(x-x_0)-\frac{b}{c}(y-y_0)$$

If we stay on plane $y=y_0$:

$$z-z_0=-\frac{a}{c}(x-x_0)$$

$$z-z_0=m_x(x-x_0)$$ the multiplier is a slope in the x-direction for point-slope form line

If we make this slope the partial derivative, then this line is tangent to curve at the point:

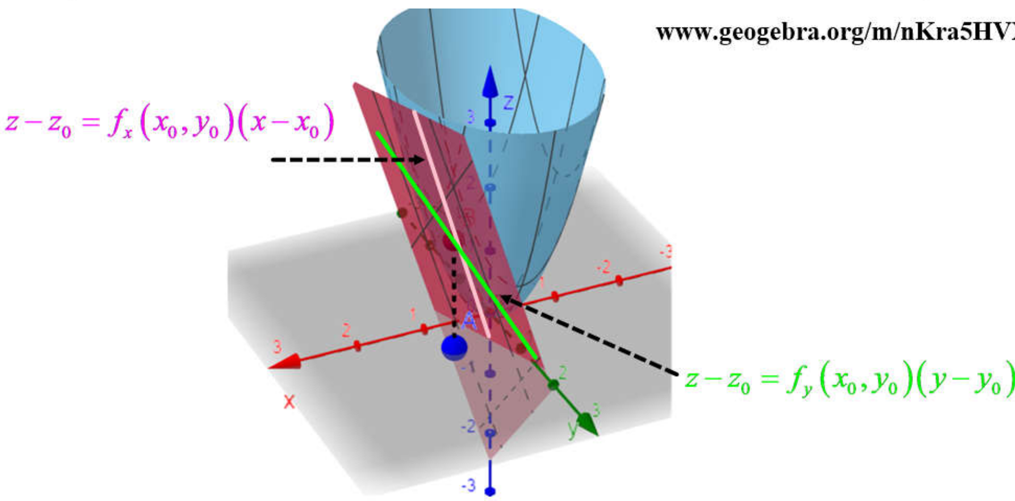
$$z-z_0=f_x(x_0,y_0)(x-x_0)$$

If we so the same for the other term on plane $x=x_0$ we get:

$$\boxed{z-z_0=f_x(x_0,y_0)(x-x_0)+f_y(x_0,y_0)(y-y_0)}$$

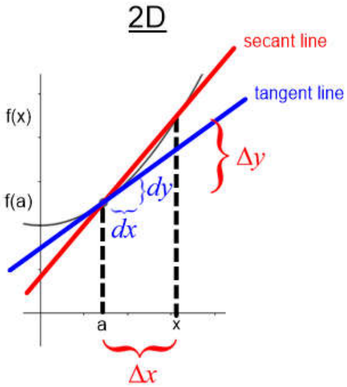
Equation of tangent plane to surface $z=f(x,y)$ at point (x_0,y_0,z_0)

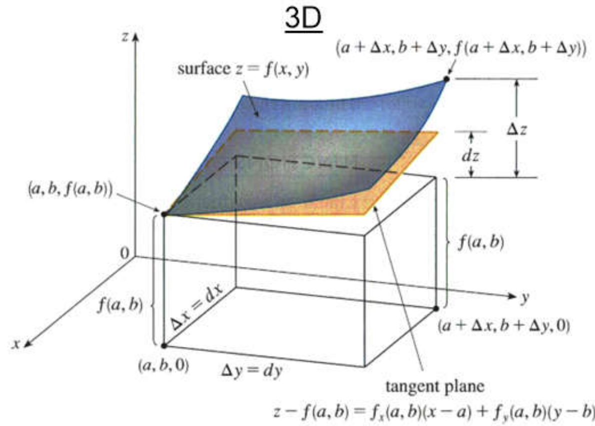
www.geogebra.org/m/nKra5HVX



Ex) Find the tangent plane to $z=2x^2+y^2$ at (1, 1, 3).

Differentials





A **differential** is a very small (infinitesimal) change in a direction:

$$\frac{\Delta y}{\Delta x} = \textit{slope of secant line}$$

$$\frac{dy}{dx} = f'(x) = \textit{slope of tangent line}$$

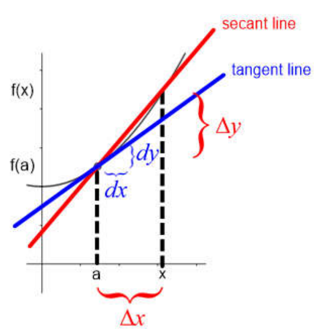
$$\boxed{dy = f'(x)\,dx}$$

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

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

$$\boxed{\begin{aligned} dz &= f_x(x,y)\,dx + f_y(x,y)\,dy \\ dz &= \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy \end{aligned}}$$

Difference between dz and Δz



Remember in brief calculus when we talked about the difference between 'average speed' and 'instantaneous speed'?

$$\textit{avg speed} = \frac{s_2 - s_1}{t_2 - t_1} = \frac{\Delta s}{\Delta t}$$

(Use 'algebra' for 'average')

$$\textit{instantaneous speed} = \lim_{\Delta t \to 0} \frac{s_2 - s_1}{t_2 - t_1} = s'(t)$$

(Use 'derivative' for the instantaneous speed at t_1)

Similarly, dz is calculated using derivatives: $dz = f_x(x,y)\,dx + f_y(x,y)\,dy$

But Δz is calculated using algebra: $\Delta z = f(x_2, y_2) - f(x_1, y_1)$

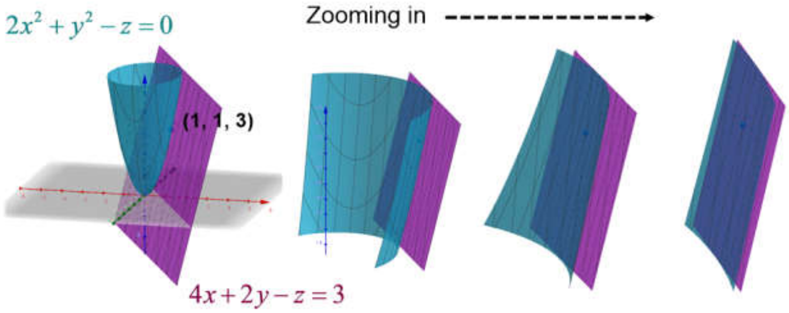
Ex) If $f(x,y) = x^2 + 3xy - y^2$ if x changes from 2 to 2.05 and y changes from 3 to 2.96

a) Find dz

b) Find Δz

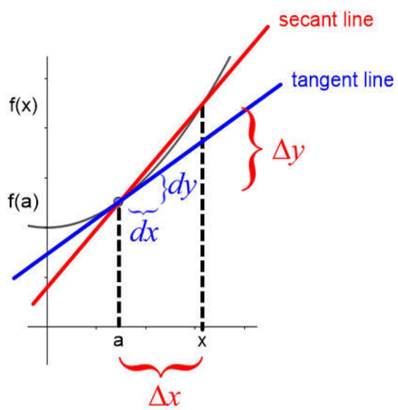
Linear Approximation

www.geogebra.org/m/prqajfgb



In the vicinity of the point of tangency, the tangent plane is a good approximation of the surface.

In 2-D, if we are not too far away from x, then the tangent line is a good approximation to the curve f(x):



For the secant line (on the actual curve):

$$(y - y_a) = m_{secant}(x - x_a)$$
$$(f(x) - f(a)) = \frac{\Delta y}{\Delta x}(x - a)$$
$$f(x) = f(a) + \frac{\Delta y}{\Delta x}(x - a)$$

For the tangent line (with slope which is the limit as change in x approaches zero = the derivative):

$$(y - y_a) = m_{tangent}(x - x_a)$$
$$(f(x) - f(a)) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}(x - a)$$
$$(f(x) - f(a)) = f'(x)(x - a)$$
$$f(x) = f(a) + f'(x)(x - a)$$

Then the f(x) using the tangent line is an approximation of the actual f(x) on the curve:

$$f(x)_{tangent} \approx f(x)_{secant}$$

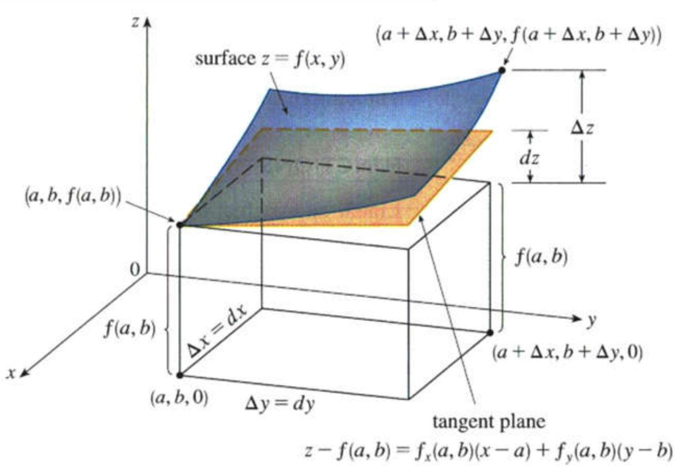
$$\boxed{f(x) \approx f(a) + f'(x)(x - a)}$$

Similarly, in 3D we can approximate a surface using the tangent plane.

Equation of tangent plane to surface $z = f(x, y)$ at point (x_0, y_0, z_0):

$$z - z_0 = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$
$$z - f(a,b) = f_x(a,b)(x - a) + f_y(a,b)(y - b)$$

$$\boxed{z = f(x,y) \approx f(a,b) + f_x(a,b)(x - a) + f_y(a,b)(y - b)}$$



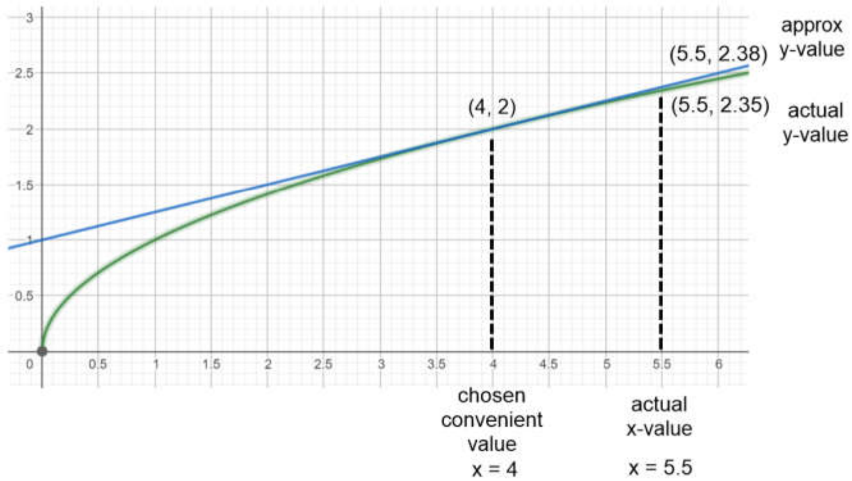
Linear Approximation

Examples: Find the linearization of the surface, then use it to approximate.

2D

$$\textit{approximate } f(5.5) \textit{ for } f(x)=\sqrt{x}$$

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

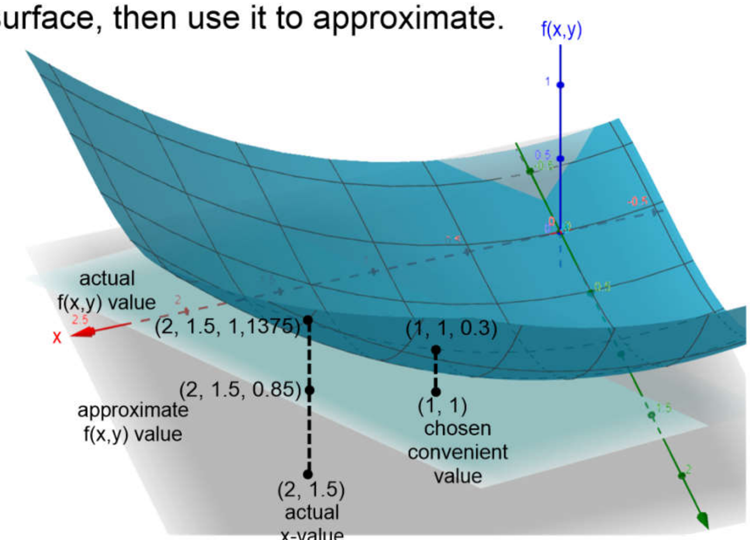


Examples: Find the linearization of the surface, then use it to approximate.

3D

$$\textit{approximate } f(2,1.5) \textit{ for } f(x,y)=0.2x^2+0.1y^3$$

$$f(x,y)\approx f(a,b)+f_x(a,b)(x-a)+f_y(a,b)(y-b)$$



14.5: The Chain Rule

Chain rule for single variable functions

Remember the Chain Rule from brief calculus? $\frac{dy}{dx}=\frac{dy}{du}\frac{du}{dx}$

(Almost like 'unit analysis' / cancelation)

It enabled things like this:

$$y=e^{3x^2} \quad Find\ \frac{dy}{dx}$$

$$define\ u=3x^2 \quad then\ y=e^u$$

$$\frac{du}{dx}=6x \qquad \frac{dy}{du}=e^u=e^{3x^2}$$

$$and\ \frac{dy}{dx}=\frac{dy}{du}\frac{du}{dx}$$

$$=e^{3x^2}(6x)$$

...which extended all our derivative shortcuts

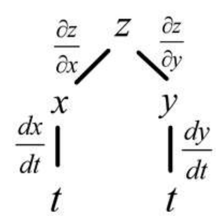
Chain rule for multivariable functions

For multivariable functions, the situation is a little more complex, but still based on this 'unit analysis' idea of multiplying derivatives to get 'cancelation'. There are two main 'cases':

__Case 1__: When z is a function of two variables *x* and *y*, and these two variables each functions of the same single variable, *t* : $z=f(x,y)=f(x(t),y(t))$

$$\boxed{\frac{dz}{dt}=\frac{\partial z}{\partial x}\frac{dx}{dt}+\frac{\partial z}{\partial y}\frac{dy}{dt}}$$

Could draw this tree to help remember:



Ex) Find $\frac{dz}{dt}$ when *t* = *0*, if $z=x^2y+3xy^4$

$$x=\sin(2t)$$

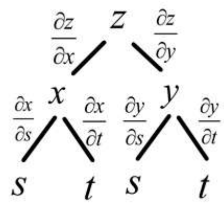
$$y=\cos(t)$$

Chain rule for multivariable functions

<u>Case 2</u>: When z is a function of two variables *x* and *y*, and these two variables are each functions of the same <u>two</u> variables, *s* and *t*: $z = f(x,y) = f\left(x(s,t), y(s,t)\right)$

$$\boxed{\begin{aligned}\frac{\partial z}{\partial s} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s}\\ \frac{\partial z}{\partial t} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t}\end{aligned}}$$

Could draw this tree to help remember:



Ex) Find $\frac{\partial z}{\partial s}$ *and* $\frac{\partial z}{\partial t}$ when *t* = *0*, if $z = x^2 y + 3xy^4$

$x = e^s \sin(2t)$

$y = s^3 \cos(t)$

This idea is extendible for any number of first or second level variables.

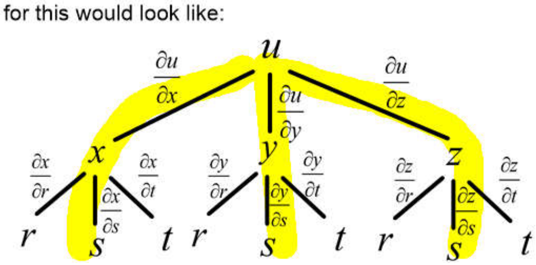
Ex) Find $\frac{\partial u}{\partial s}$ when *r* = *2*, *s* = *1*, *t* = *0*, if $u = x^4 y + y^2 z^3$

$x = rse^t$

$y = rs^2 e^{-t}$

$z = r^2 s \sin(t)$

Tree for this would look like:



$$\frac{\partial u}{\partial s} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial s}$$

Implicit Differentiation

We first encountered implicit differentiation in brief calculus: Find $\frac{dy}{dx}$ if $x^3 + xy + y^3 = 5$

When it wasn't easy or possible to first solve for y, we just took derivatives in place, using the chain rule to multiply by a dy/dx for terms where the variable wasn't x:

$$\frac{d}{dx}\left[x^3\right]+\frac{d}{dx}[xy]+\frac{d}{dx}\left[y^3\right]=\frac{d}{dx}[5]$$

(product rule)

$$\frac{d}{dx}\left[x^3\right]+(x)\frac{d}{dx}[y]+(y)\frac{d}{dx}[x]+\frac{d}{dx}\left[y^3\right]=\frac{d}{dx}[5]$$

$$\frac{d}{dx}\left[x^3\right]+(x)\frac{d}{dx}[y]+(y)\frac{d}{dx}[x]+\frac{d}{dx}\left[y^3\right]=\frac{d}{dx}[5]$$

$$3x^2+x\left(1\frac{dy}{dx}\right)+y(1)+3y^2\frac{dy}{dx}=0$$

$$\frac{dy}{dx}\left(x+3y^2\right)=-3x^2-y$$

$$\frac{dy}{dx}=\frac{-3x^2-y}{x+3y^2}$$

Now we can extend this more generally. If an equation written as $F(x,y)=0$ defines y implicitly as a function of x, then we can apply the Chain Rule to both sides:

$$\frac{\partial F}{\partial x}\frac{dx}{dx}+\frac{\partial F}{\partial y}\frac{dy}{dx}=0$$

and since $\frac{dx}{dx}=1$

$$\frac{\partial F}{\partial x}(1)+\frac{\partial F}{\partial y}\frac{dy}{dx}=0$$

$$\boxed{\frac{dy}{dx}=-\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}}=-\frac{F_x}{F_y}}$$

Find $\frac{dy}{dx}$ if $x^3 + xy + y^3 = 5$

$$F = x^3 + xy + y^3 - 5$$

$$F_x = 3x^2 + y$$

$$F_y = x + 3y^2$$

$$\frac{dy}{dx}=-\frac{F_x}{F_y}=-\frac{3x^2+y}{x+3y^2}$$

Not only is this easier, we can extend it for higher numbers of variables.

For $z=f(x,y,z)$*, express as* $F(x,y,z)=0$ *Then for* $\frac{\partial z}{\partial x}$*, by the Chain Rule:*

$$\frac{\partial F}{\partial x}\frac{\partial x}{\partial x}+\frac{\partial F}{\partial y}\frac{\partial y}{\partial x}+\frac{\partial F}{\partial z}\frac{\partial z}{\partial x}=0$$

Also, $\frac{\partial x}{\partial x}=1$ *and* $\frac{\partial y}{\partial x}=0$

(because *x* and *y* are orthogonal input variables)

$$\frac{\partial F}{\partial x}+\frac{\partial F}{\partial z}\frac{\partial z}{\partial x}=0$$

Similarly, for $\frac{\partial z}{\partial y}$:

$$\boxed{\frac{\partial z}{\partial x}=-\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}}-\frac{F_x}{F_z} \qquad \frac{\partial z}{\partial y}=-\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial z}}-\frac{F_y}{F_z}}$$

Higher order partial derivatives (not going over in class, see ex 7 in book and notes)

We can use these ideas to take derivatives of derivatives (higher-order derivatives):

Ex) Find $\frac{\partial z}{\partial r}$ *and* $\frac{\partial^2 z}{\partial r^2}$ if $z = f(x, y)$ is some function with continuous 2nd partial derivatives and $x(r,s) = r^2 + s^2, \quad y(r,s) = 2rs$

1st derivative by Chain Rule: $\frac{\partial z}{\partial r} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial r}$

$$\boxed{\frac{\partial z}{\partial r} = \frac{\partial z}{\partial x}(2r) + \frac{\partial z}{\partial y}(2s)}$$

Now taking the 2nd derivative:

$$\frac{\partial^2 z}{\partial r^2} = \frac{\partial}{\partial r}\left[\frac{\partial z}{\partial x}(2r) + \frac{\partial z}{\partial y}(2s)\right]$$

$$= \frac{\partial}{\partial r}\left[\frac{\partial z}{\partial x}(2r)\right] + \frac{\partial}{\partial r}\left[\frac{\partial z}{\partial y}(2s)\right]$$

Each term requires product rule:

$$= \left(\frac{\partial z}{\partial x}\frac{\partial}{\partial r}[(2r)] + (2r)\frac{\partial}{\partial r}\left[\frac{\partial z}{\partial x}\right]\right) + \left(\frac{\partial z}{\partial y}\frac{\partial}{\partial r}[(2s)] + (2s)\frac{\partial}{\partial r}\left[\frac{\partial z}{\partial y}\right]\right)$$

$$= 2\frac{\partial z}{\partial x} + (2r)\underline{\frac{\partial}{\partial r}\left[\frac{\partial z}{\partial x}\right]} + \left(\frac{\partial z}{\partial y}(0) + (2s)\underline{\frac{\partial}{\partial r}\left[\frac{\partial z}{\partial y}\right]}\right)$$

But these terms are not simple 2nd derivatives...they require using Chain Rule:

$$\frac{\partial}{\partial r}\left[\frac{\partial z}{\partial x}\right] = \frac{\partial}{\partial x}\left[\frac{\partial z}{\partial x}\right]\frac{\partial x}{\partial r} + \frac{\partial}{\partial y}\left[\frac{\partial z}{\partial x}\right]\frac{\partial y}{\partial r} = \frac{\partial^2 z}{\partial x^2}(2r) + \frac{\partial^2 z}{\partial y \partial x}(2s)$$

$$\frac{\partial}{\partial r}\left[\frac{\partial z}{\partial y}\right] = \frac{\partial}{\partial x}\left[\frac{\partial z}{\partial y}\right]\frac{\partial x}{\partial r} + \frac{\partial}{\partial y}\left[\frac{\partial z}{\partial y}\right]\frac{\partial y}{\partial r} = \frac{\partial^2 z}{\partial x \partial y}(2r) + \frac{\partial^2 z}{\partial y^2}(2s)$$

Substitution these back in:

$$= 2\frac{\partial z}{\partial x} + (2r)\left(\frac{\partial^2 z}{\partial x^2}(2r) + \frac{\partial^2 z}{\partial y \partial x}(2s)\right) + (2s)\left(\frac{\partial^2 z}{\partial x \partial y}(2r) + \frac{\partial^2 z}{\partial y^2}(2s)\right)$$

$$= 2\frac{\partial z}{\partial x} + 4r^2\frac{\partial^2 z}{\partial x^2} + 4rs\underline{\frac{\partial^2 z}{\partial y \partial x}} + 4rs\underline{\frac{\partial^2 z}{\partial x \partial y}} + 4s^2\frac{\partial^2 z}{\partial y^2}$$

same by Clairaut's Theorem

$$\boxed{\frac{\partial^2 z}{\partial r^2} = 2\frac{\partial z}{\partial x} + 4r^2\frac{\partial^2 z}{\partial x^2} + 8rs\frac{\partial^2 z}{\partial x \partial y} + 4s^2\frac{\partial^2 z}{\partial y^2}}$$

14.6: Directional Derivatives, Gradient Vector

Note: These lesson notes are presented in the order that makes it easiest to understand and use, but is not the order in which things are usually defined, derived, and proved.

No derivations will be included here, but there is a separate PDF showing all the derivations on www.mrfelling.com if you are interested in the details.

Definition of the Gradient Vector

For $z = f(x,y)$ the **Gradient Vector** is defined to be:

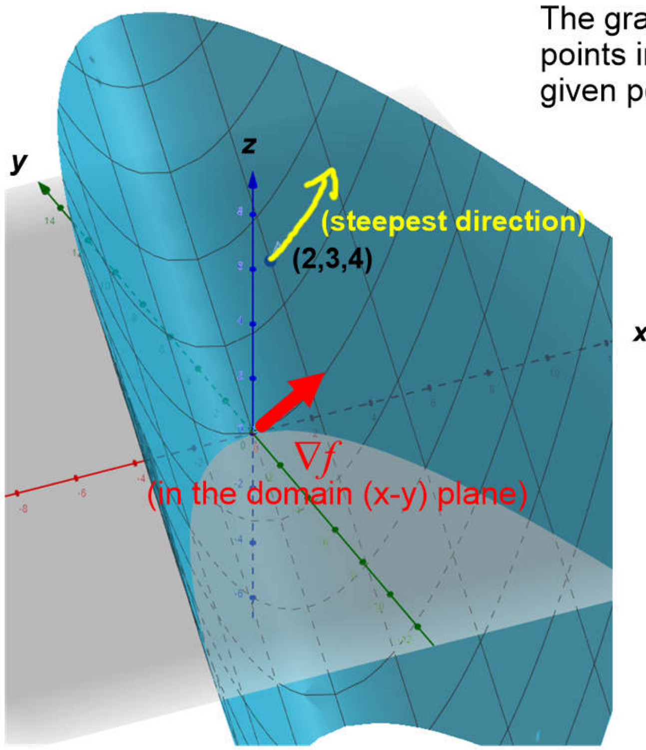
$$\boxed{\nabla f(x,y) = \left\langle f_x(x,y),\ f_y(x,y) \right\rangle}$$

$$\nabla f(x,y) = \left\langle f_x(x,y),\ f_y(x,y) \right\rangle$$

this symbol is called 'del'

This is understood to always mean a vector, so an arrow is implied and typically not included over the ∇ .

Geometric meaning of the Gradient Vector



The gradient vector is <u>in the domain (x-y) plane</u> and points in the direction of fastest change in z at the given point.

$$z = f(x,y) = \frac{1}{4}x^2 + y$$

$$\nabla f = \left\langle \frac{1}{2}x, 1 \right\rangle$$

At (x,y) = (2,3)

$$z = f(2,3) = \frac{1}{4}(2)^2 + (3) = 4$$

$$\nabla f = \langle 1,1 \rangle$$

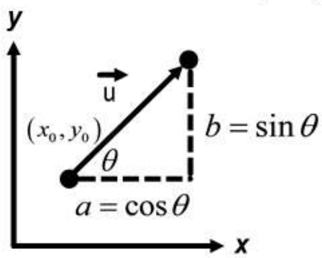
Directional Derivatives

We know that the derivative in the x and y directions are given by the partial derivatives, f_x and f_y.

What about derivatives in an arbitrary direction? These are called directional derivatives.

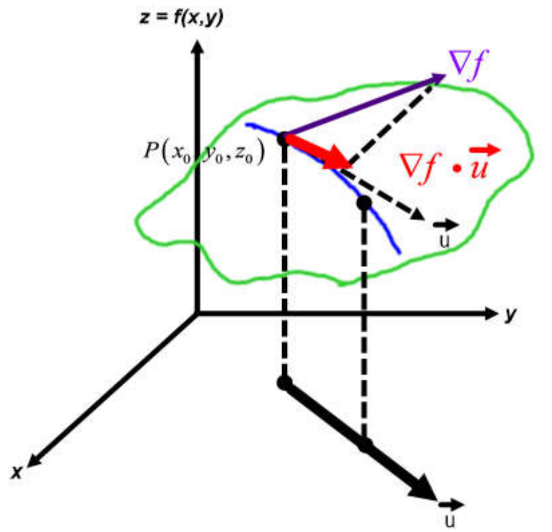
First, we need to define a direction, and this is done in the domain (x,y-plane) because direction doesn't depend upon what the output value is. We define a **unit direction vector:**

unit direction vector, $\vec{u} = \langle a, b \rangle$



Then it can be shown that the derivative in the arbitrary direction is the dot product of the gradient vector with the unit direction vector:

$$\boxed{D_u f(x,y) = \nabla f(x,y) \bullet \vec{u}}$$



You can think of this as the gradient representing the derivative in the direction where it is highest (fastest change in z), and by taking the dot product we are finding the 'projection' of this gradient vector in a given direction (the amount of the maximum derivative which is in this direction).

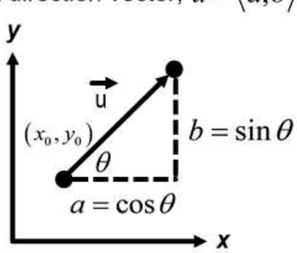
In fact, if you use the magnitude and angle representation of the dot product:

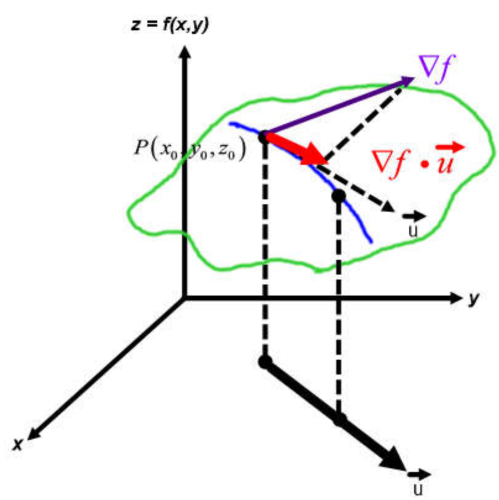
$$D_u f(x,y) = \nabla f(x,y) \bullet \vec{u}$$

$$D_u f(x,y) = \left|\nabla f(x,y)\right| \left|\vec{u}\right| \cos\theta$$

$\cos\theta$ is maximum (1) when $\theta = 0$ which occurs when $\vec{u}$ is in the direction of ∇f and is minimum (0) when the vectors are perpendicular.

unit direction vector, $\vec{u} = \langle a, b \rangle$





Note: Very important!

The direction vector <u>must</u> be a unit vector!

In higher dimensions

These ideas are all extendable for a higher number of input variables.

$$\text{if } w = f(x,y,z)$$

$$\nabla f(x,y,z) = \langle f_x(x,y,z),\ f_y(x,y,z),\ f_z(x,y,z) \rangle$$

$$D_u f(x,y,z) = \nabla f(x,y,z) \cdot \vec{u}$$

...except that now the domain is 3D, so the direction vector would be a vector in $\mathbb{R}^3$

Gradient and relationship to level curves

A great example from the book to illustrate: If $f(x,y) = xe^y$

Find the rate of change of *f* at the point *P(2,0)* in the direction from *P* to *Q(1/2, 2)*.
In what direction does *f* have the maximum rate of change? What is this maximum rate of change?

Let's look at these results and their relationship to the level curves...

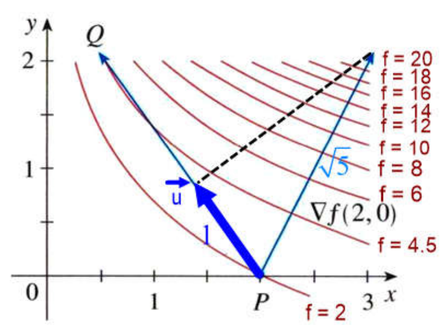
$\overrightarrow{PQ} = \left\langle -\frac{3}{2}, 2 \right\rangle$ ...is the direction vector between the points.

$\vec{u} = \left\langle -\frac{3}{5}, \frac{4}{5} \right\rangle$ ...is direction normalized to unit vector length.

$\nabla f = \langle 1, 2 \rangle$ ...is gradient vector which shows direction of fastest change in output values.

$|\nabla f| = \sqrt{5}$ ...is the magnitude of the gradient vector f changes $\sqrt{5}$ units for every one unit change in this maximum direction.

$D_u f(2,0) = 1$ ...is the directional derivative value, f changes only 1 unit for every one unit change in direction from P to Q.

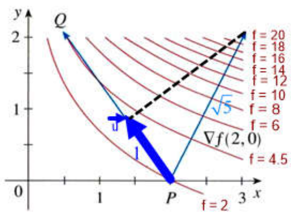


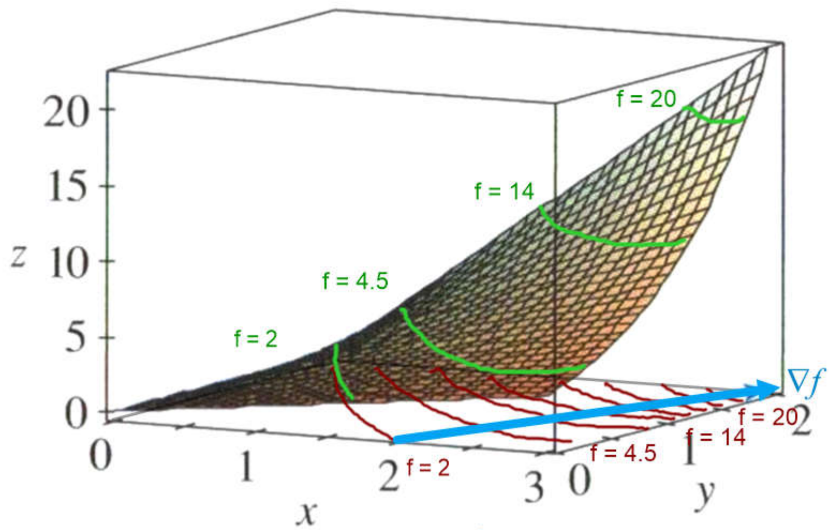
Notice how the gradient points in a way where the level curves get most rapidly more crowded (f values increasing rapidly).

Also, note that the gradient vector appears to be perpendicular to the level curves.

Gradient and relationship to level curves

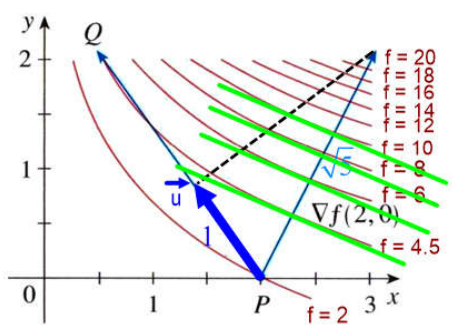
Remember, the gradient vector is in the domain (xy-plane). Here is how this looks in 3D...





∇f is in the xy-plane in direction of maximum rate of height change

Tangent lines



In the domain, if you look at how the gradient vector crosses the level curves, they appear to be perpendicular, and it can be shown that they are perpendicular (derived in the separate document on www.mrfelling.com)

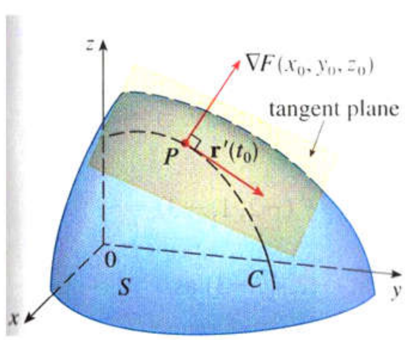
We can define **tangent lines** to the level curves at the points of contact with the gradient vector.

Gradient for *f(x,y,z)*, level surfaces, tangent planes

If you add another input variable, the domain because 3D, but gradient and directional derivative are defined in a similar way:

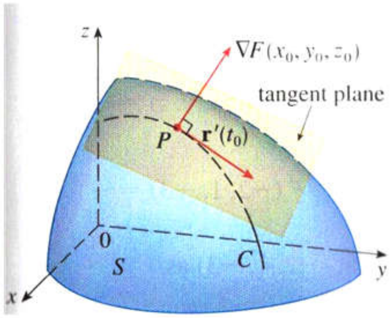
$$w = F(x,y,z) \qquad \nabla F = \langle F_x, F_y, F_z \rangle$$
$$\vec{u} = \langle a,b,c \rangle \qquad D_u F(x,y,z) = \nabla F(x,y,z) \cdot \vec{u}$$

We can't draw a representation of the function anymore (would be 4 dimensional), but we can draw the domain (now 3D):



Now we have a **level surface** (blue surface in this diagram) which corresponds to the maroon level curves in the previous example for some particular value *k* where $F(x,y,z) = k$

Gradient for *f(x,y,z)*, level surfaces, tangent planes



Just as the gradient vector was perpendicular to the level curves before, now the gradient vector is perpendicular to the level surfaces. Before, we defined tangent lines at the points of intersection, here, we define **tangent planes** instead.

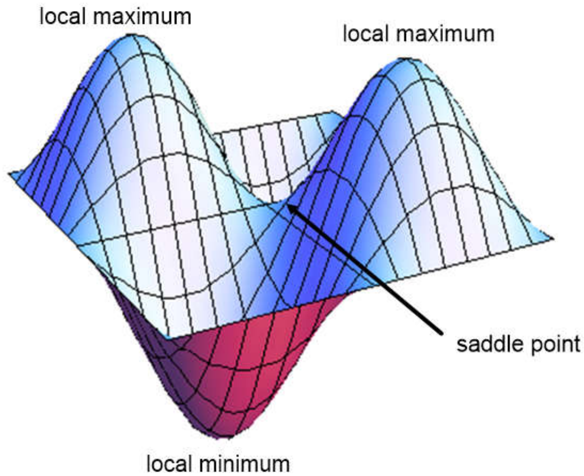
The **equation of a tangent plane** at point $P(x_0,y_0,z_0)$ can be shown to be:

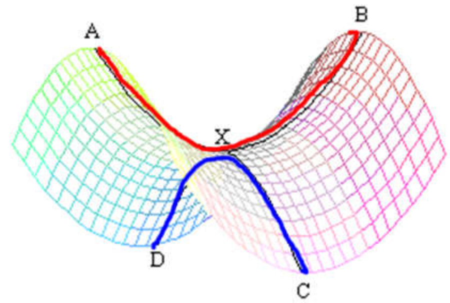
$$\boxed{F_x(x-x_0)+F_y(y-y_0)+F_z(z-z_0)=0}$$

where the partial derivatives are evaluated at the point.

14.7: Maximum and Minimum Values

Geometric meaning of local minimum, maximum, and saddle point

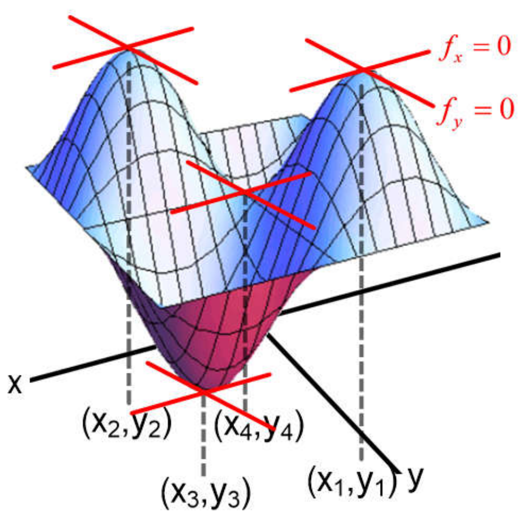




A **saddle point** occurs when concavity is up in one direction and down in the other, and corresponds to an inflection point, but in 3D.

A **local maximum (or minimum)** occurs when this is the highest (or lowest) output value within a small local region.

Local minimum, maximum, and saddle points occur at <u>critical points</u>



Critical points

Critical points occur when

$$f_x(x, y) = 0 \ \ and \ \ f_y(x, y) = 0$$

...which corresponds to:

$$\nabla f = \langle f_x, f_y \rangle = \langle 0, 0 \rangle$$

2nd Derivative Test

Determining whether a critical point is a minimum, maximum, or saddle point:

Find $f_{xx}(x,y),\ f_{yy}(x,y),\ and\ f_{xy}(x,y)$

For each critical point...

Compute the Hessian Determinant:

$$D = f_{xx}(x_0,y_0) \bullet f_{yy}(x_0,y_0) - \left[f_{xy}(x_0,y_0)\right]^2$$

Hessian Determinant :

$$D = \begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix}$$

$$D = f_{xx}f_{yy} - f_{xy}f_{yx}$$

$D < 0$

(x_0, y_0) *is a saddle point*

$D > 0$

- $f_{xx}(x_0,y_0) < 0$: (x_0, y_0) *is a local maximum*
- $f_{xx}(x_0,y_0) > 0$: (x_0, y_0) *is a local minimum*

$D = 0$

inconclusive

So either graph or try various (x,y) pairs around the point to see if min or max

Optimization without constraints

- Find all critical points
- Find an expression for *D*
- Evaluate *D* at each critical point to determine what is happening in the output

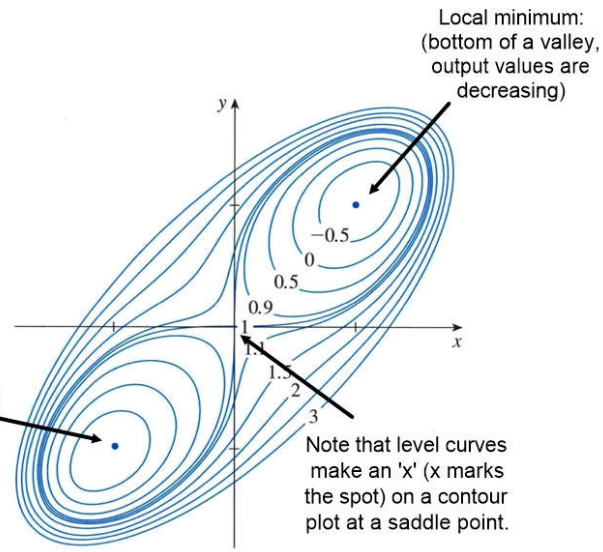
Ex) Find and classify the critical points for $z = x^4 + y^4 - 4xy + 1$

$z = x^4 + y^4 - 4xy + 1$

Results...

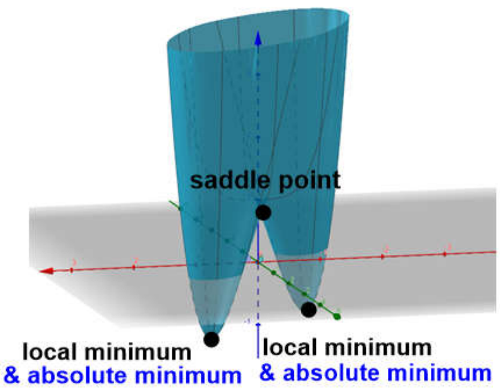
Saddle point at (0,0)

Local minimums at (1,1), (-1,-1)



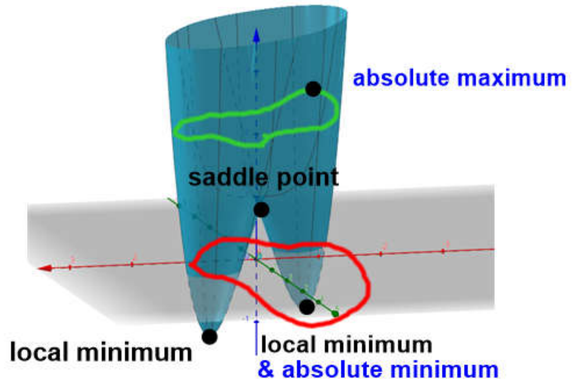
Optimization with a domain boundary constraint (absolute min and max)

Over the entire xy-plane domain
(without constraint)



These local minimum points would also be the absolute minimum points (if they were both equally negative). This surface has no absolute maximum over the entire domain.

Over just a portion of the domain
(with a boundary constraint)



If we add a surrounding boundary in the domain and require staying within in, then we must also check the output value on the surface for all points above this boundary. The absolute min or max may occur at critical points or it may occur above the constraint boundary.

Procedure for finding absolute max and min:

- Find all local max, min and add them to a list of (x,y) pairs
- Graph the domain and express the edge of the domain as a series of edge curves
- For each edge curve, determine its equation and substitute into f(x,y), then determine for what (x,y) pairs on this edge will f(x,y) be either a min or max...add each (x,y) pair to the list.
- Plug all of the (x,y) pairs on the list into f(x,y) to determine where the absolute min and max occur.

Ex) #32. Find the absolute maximum and minimum values of *f* on the set *D*

$$f(x,y) = 4x + 6y - x^2 - y^2 \qquad D = \{(x,y) \mid 0 \le x \le 4, 0 \le y \le 5\}$$

- Find all local max, min and add them to a list of (x,y) pairs:

possible locations of absolute max, min

(x,y)	
(2,3)	

$$\begin{cases} f_x = 4 - 2x = 0 \\ f_y = 6 - 2y = 0 \end{cases}$$

$$2x = 4 \qquad 2y = 6$$

$$x = 2 \qquad y = 3$$

one critical point at (2,3)

use 2nd derivative test...

$$f_{xx} = -2$$

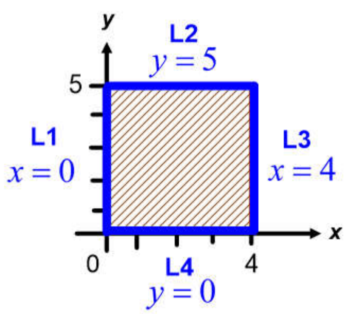
$$f_{xy} = 0$$

$$f_{yy} = -2$$

$$D = (-2)(-2) - (0)^2 = 4$$

(2,3) is a local max or min, so it should be on the list

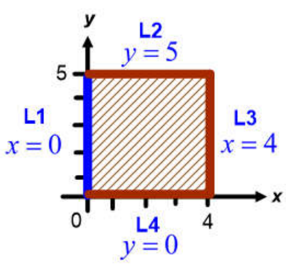
- Graph the domain and express the edge of the domain as a series of edge curves:



- For each edge curve, determine its equation and substitute into f(x,y), then determine for what (x,y) pairs on this edge will f(x,y) be either a min or max...add each (x,y) pair to the list:

possible locations of absolute max, min

(x,y)	
(2,3)	
(0,3)	
(0,0)	
(0,5)	

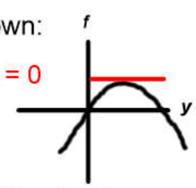


$L1$:

$$f(x,y) = 4x + 6y - x^2 - y^2$$

$$f(0,y) = 6y - y^2$$

This is a parabola opening down:



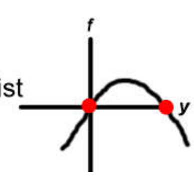
max will occur when 1st derivative = 0

$$f_y = 6 - 2y = 0$$

$$y = 3$$

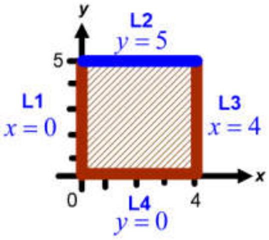
so (0,3) should be on the list to check

min will occur at either y=0 or y=5 (must stay in this part of the domain)



so just put both of these points on the list (0,0) and (0,5)

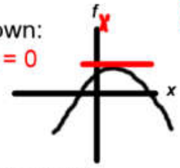
- For each edge curve, determine its equation and substitute into f(x,y), then determine for what (x,y) pairs on this edge will f(x,y) be either a min or max...add each (x,y) pair to the list:



$L2:$

$$f(x,y)=4x+6y-x^2-y^2$$
$$f(x,5)=4x+30-x^2-5^2$$
$$f(x,5)=-x^2+4x+5$$

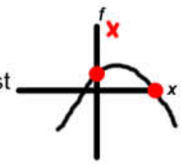
This is also a parabola opening down:
max will occur when 1st derivative = 0



$$f_x=-2x+4=0$$
$$x=2$$

so (2,5) should be on the list to check

min will occur at either x=0 or x=4 (must stay in this part of the domain)

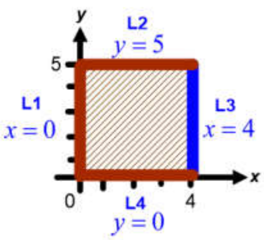


so just put both of these points on the list
(0,5) and (4,5)

possible locations of absolute max, min

(x,y)
(2,3)
(0,3)
(0,0)
(0,5)
(2,5)
(4,5)

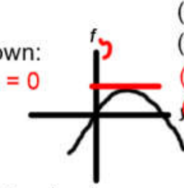
- For each edge curve, determine its equation and substitute into f(x,y), then determine for what (x,y) pairs on this edge will f(x,y) be either a min or max...add each (x,y) pair to the list:



$L3:$

$$f(x,y)=4x+6y-x^2-y^2$$
$$f(4,y)=4(4)+6y-(4)^2-y^2$$
$$f(4,y)=6y-y^2$$

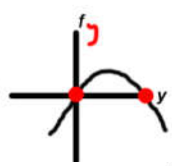
This is also a parabola opening down:
max will occur when 1st derivative = 0



$$f_y=6-2y=0$$
$$y=3$$

so (4,3) should be on the list to check

min will occur at either y=0 or y=5 (must stay in this part of the domain)

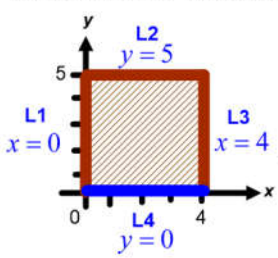


so just put both of these points on the list
(4,0) and (4,5)

possible locations of absolute max, min

(x,y)
(2,3)
(0,3)
(0,0)
(0,5)
(2,5)
(4,5)
(4,3)
(4,0)

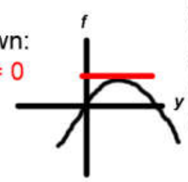
- For each edge curve, determine its equation and substitute into f(x,y), then determine for what (x,y) pairs on this edge will f(x,y) be either a min or max...add each (x,y) pair to the list:



$L4:$

$$f(x,y)=4x+6y-x^2-y^2$$
$$f(x,0)=4x-x^2$$

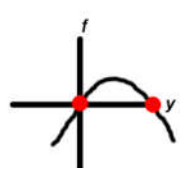
This is also a parabola opening down:
max will occur when 1st derivative = 0



$$f_x=4-2x=0$$
$$x=2$$

so (2,0) should be on the list to check

min will occur at either x=0 or x=4 (must stay in this part of the domain)



so just put both of these points on the list
(0,0) and (4,0)

possible locations of absolute max, min

(x,y)
(2,3)
(0,3)
(0,0)
(0,5)
(2,5)
(4,5)
(4,3)
(4,0)
(2,0)

- Plug all of the (x,y) pairs on the list into f(x,y) to determine where the absolute min and max occur:

possible locations of absolute max, min (x,y)	$f(x,y)=4x+6y-x^2-y^2$
(2,3)	$f(2,3)=4(2)+6(3)-(2)^2-(3)^2=13$
(0,3)	9
(0,0)	0
(0,5)	5
(2,0)	4
(4,5)	5
(4,3)	9
(4,0)	0
(2,0)	4

absolute maximum of f =13 at (2,3)
absolute minimum of f = 0 at (0,0) and (4,0)

If you're thinking, "there must be a better way"...you're right
...stayed tuned: Method of Lagrange Multipliers

Optimization a geometric scenario

Ex) #40. Find the point on the plane *x - y + z = 4* that is closest to the point (1,2,3).

'Closest' means minimum distance. 3D distance formula:

$$d=\sqrt{(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2}$$

Distance from general point *(x,y,z)* to *(1,2,3)*:

$$d=\sqrt{(x-1)^2+(y-2)^2+(z-3)^2}$$

But if point must be on the plane, then $z=4-x+y$:

$$d=\sqrt{(x-1)^2+(y-2)^2+(4-x+y-3)^2}$$

$$d=\sqrt{x^2-2x+1+y^2-4y+4+(1-x+y)(1-x+y)}$$

$$d=\sqrt{x^2-2x+1+y^2-4y+4+(1-x+y)(1-x+y)}$$

$$d=\sqrt{x^2-2x+1+y^2-4y+4+1-x+y-x+x^2-xy+y-xy+y^2}$$

$$d=\sqrt{2x^2+2y^2-2xy-4x-2y+6}$$

We need to minimize d, but could choose to minimize d^2, so define new function f(x,y):

$$f(x,y)=2x^2+2y^2-2xy-4x-2y+6$$

Now, just minimize this (without constraints)

$$f(x,y)=2x^2+2y^2-2xy-4x-2y+6$$

Now, just minimize this (without constraints)

$$\begin{cases} f_x=4x-2y-4=0 \\ f_y=4y-2x-2=0 \end{cases}$$

$$\begin{cases} 4x-2y=4 \\ -2x+4y=2 \end{cases}$$

$$\left[\begin{array}{cc|c} 4 & -2 & 4 \\ -2 & 4 & 2 \end{array}\right] rref \rightarrow \left[\begin{array}{cc|c} 1 & 0 & \frac{5}{3} \\ 0 & 1 & \frac{4}{3} \end{array}\right]$$

one critical point at $\left(\frac{5}{3},\frac{4}{3}\right)$...

...to make sure it is a min:

$f_{xx}=4$

$f_{xy}=-2$

$f_{yy}=4$

$D=(4)(4)-(-2)^2=12$ *so* max *or* min

checking $f_{xx}=4>0$ *means concave up*

Therefore, minimum distance occurs at $\left(\frac{5}{3},\frac{4}{3}\right)$

$x=\frac{5}{3},\ y=\frac{4}{3}$

on plane, $z=4-x+y=4-\left(\frac{5}{3}\right)+\left(\frac{4}{3}\right)=\frac{11}{3}$

So point is at $\left(\frac{5}{3},\frac{4}{3},\frac{11}{3}\right)$

The minimum distance is: $d=\sqrt{\left(\frac{5}{3}-1\right)^2+\left(\frac{4}{3}-2\right)^2+\left(\frac{11}{3}-3\right)^2}\approx 1.1547$

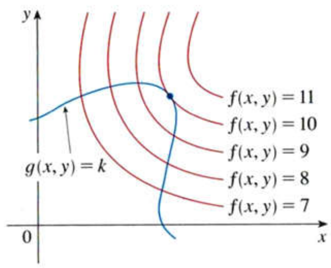
14.8: Lagrange Multipliers

Optimization with constraints using Method of Lagrange Multipliers

In the last section, we said that, if a function has a constraint in the domain, to establish a function's absolute maximum or minimum we must find all local extrema but also find the possible max and min values around all pieces of the constraint boundary. This can be very time consuming, and luckily there is another way that is easier and also much more flexible, for checking finding the max/min on the boundary: using the **Method of Lagrange Multiplers**.

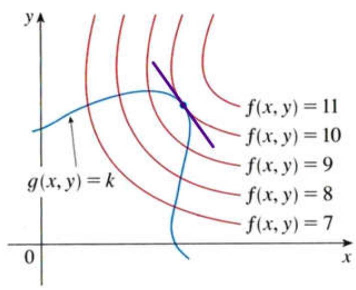
Here is the idea: Suppose we have a function $z = f(x, y)$ and we also have a constraint $g(x, y) = k$ (which represents staying on some curve in the domain).

In the domain, you could graph level curves for particular values of the output variable...

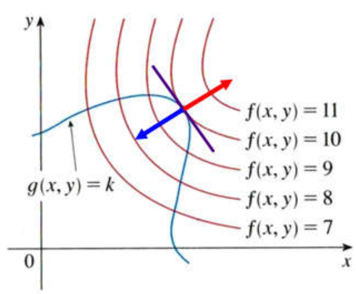


...and as you move along the constraint curve, there is some point at which you've reached the maximum output value (here, that value is 10).

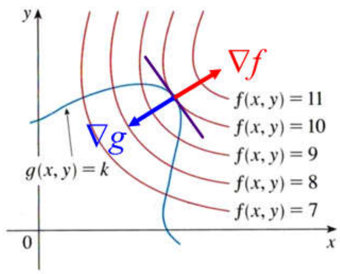
When you are at this point of maximum output, the level curve and the constraint curve are tangent to each other:



...and this means that the normal vectors for the curves at this point are pointing in the same (opposite) direction:



These normal vectors are also the gradients of the curves...



...and these gradients are parallel, meaning the one is a scalar multiple of the other. In other words, there exists a scalar, λ , for which this is true:

$$\nabla f = \lambda \nabla g$$

(at the minimum or maximum point)

This is equally true if we are at a minimum output value. The scalar 'lambda' λ is called a **Lagrange Multiplier.**

Optimization with constraints using Method of Lagrange Multipliers

We can use this fact to locate the max or min values of a function subject to a constraint, and this can be extended to any number of input variables:

Method of Lagrange Multipliers:

To find the max or min of $f(x, y, z)$ subject to the constraint $g(x, y, z) = k$
(if extreme values exist, and $\nabla g \neq 0$ *anywhere on* $g(x, y, z) = k$)

- Find all values of $x, y, z,$ *and* λ such that

$$\nabla f(x, y, z) = \lambda \nabla g(x, y, z)$$

- Evaluate all the points *(x,y,z)* found to identify which is the max and which the min.

Ex) Find the extreme values of the function $f(x, y, z) = x^2 + 2y^2$
subject to the constraint $x^2 + y^2 = 1$

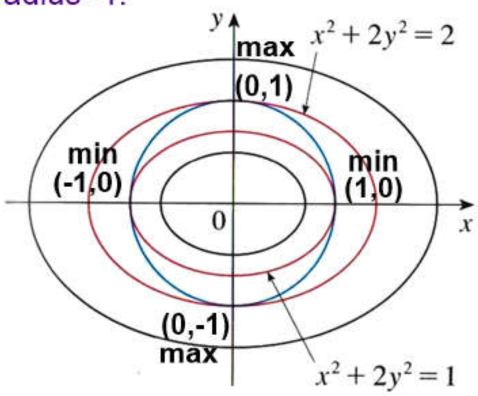
Geometric interpretation

Ex) Find the extreme values of the function $f(x,y,z) = x^2 + 2y^2$

subject to the constraint $x^2 + y^2 = 1$

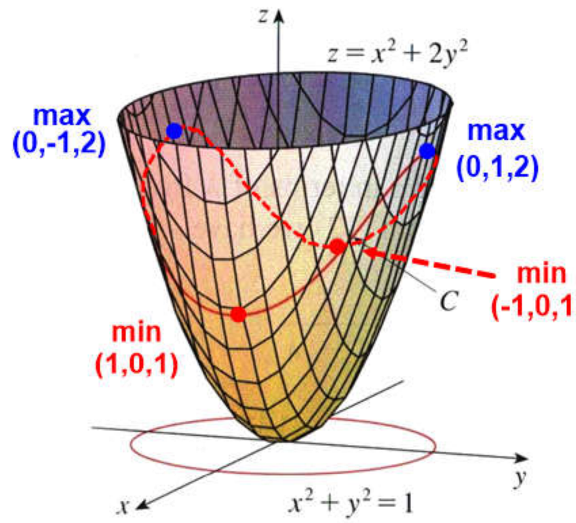
Result: max f = 2 at (0,1) and (0,-1)
min f = 1 at (1,0) and (-1,0)

In the domain, the *f* function has level curves, and the constraint is a circle centered at the origin with radius=1.



The constraint 'grazes' two level curves, the max at *f* = *2*, and the min at *f* = *1*.

This is what is happening in 3D...



Optimization with multiple constraints

If you took brief calculus, you saw a variation on this procedure which used lambda and partial derivatives (a special case which we had to use because we didn't know about gradients), but the procedure as described here is more intuitive and generalized.

In fact, you can use it to optimize a function subject to multiple constraints. Each constrain gets it's own Lagrange Multiplier. If we have two constraints, *g* and *h*, we use:

$$\boxed{\nabla f(x,y,z) = \lambda \nabla g(x,y,z) + \mu \nabla g(x,y,z)}$$

Ex) Find the maximum and minimum volumes of a rectangular box whose surface area is 1500 cm^2 and whose total edge length is 200 cm.