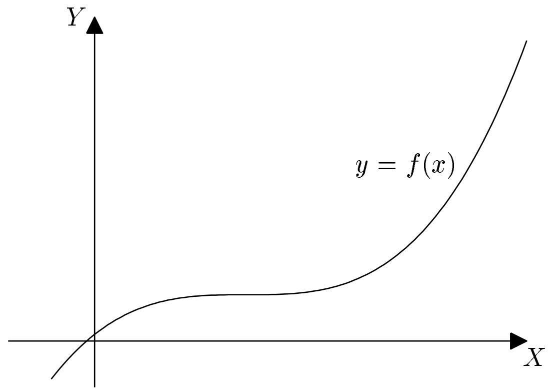
SECTION 2 — PARTIAL DERIVATIVES

Functions of two variables

Functions of one variable ($y = f(x)$) have one independent variable (the x) and one dependent variable (the y). We can represent such functions simply on a graph. For example



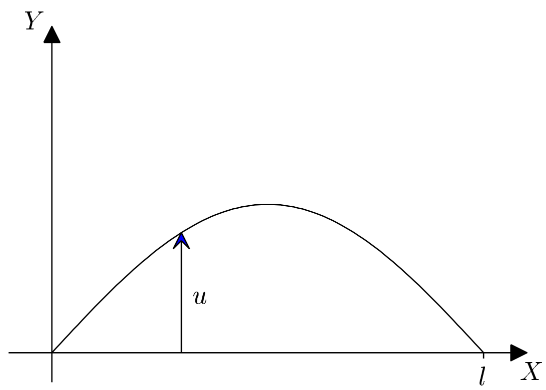
The slope of the tangent at any point is the derivative of the function.

Functions of two variables will have the form $z = f(x, y)$. They have two independent variables (x and y) and one dependent variable (z).

For example, the volume of a cone,

$$V = \frac{1}{3}\pi r^2 h,$$

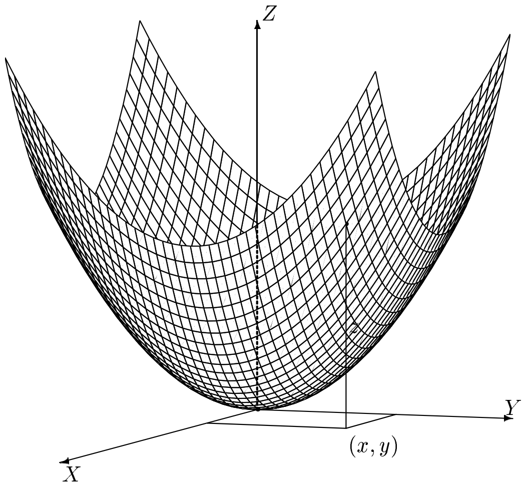
depends on the height h and the radius of the base r. The volume V is the dependent variable. We can vary r, keeping h fixed, or vary h keeping r fixed, or vary both. In each case, V may (will) change.



As a second example, consider a vibrating string that is fixed at $x = 0$ and $x = l$. Its displacement u will depend on x and t. For example

$$u(x, t) = \sin\frac{\pi x}{l}\cos\frac{\pi ct}{l}.$$

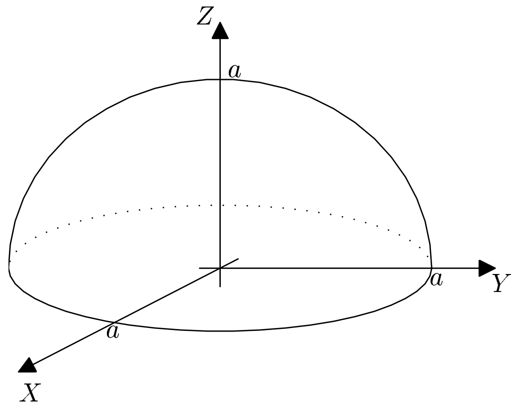
To represent $z = f(x, y)$ graphically, we need three dimensions. The function actually represents a two dimensional surface in three dimensions. Consider the function $z = x^2 + y^2$. At each point in the x–y plane, calculate z and plot the point (x, y, z).

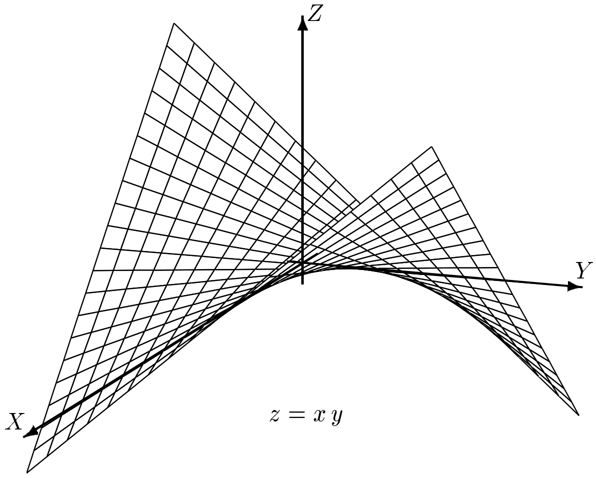


$z = f(x, y) = x^2 + y^2$

The points make up a surface called a paraboloid. It is actually the parabola $z = x^2$ rotated about the z axis.

Another example is the function $z = \sqrt{a^2 - x^2 - y^2}$. This represents a hemisphere of radius a.



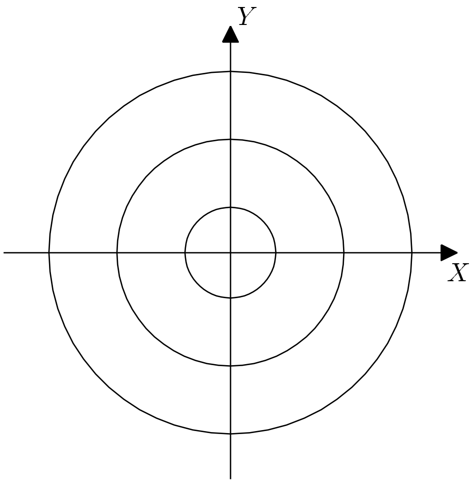


$z = x\,y$

$z = xy$ represents a 'saddle'

Finally, we have already seen that the function $z = ax + by + c$ represents a plane.

Another way to represent these functions is to use 'level curves'. For example, if $z = f(x, y) = x^2 + y^2$, the level curves are $z =$ constant or $x^2 + y^2 =$ constant $= a^2$.

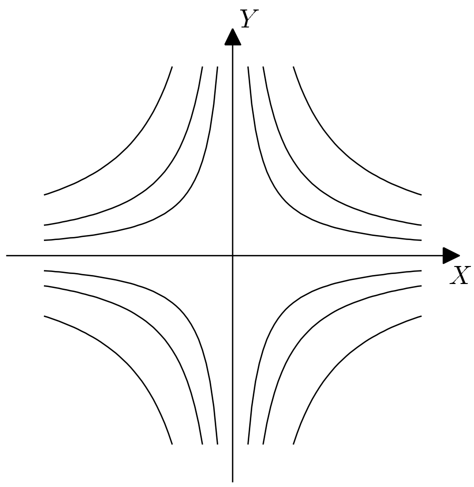


This will be a family of circles.

If $z = xy$, the level curves are $xy = \text{constant}$. i.e.

$$y = \frac{\text{constant}}{x}.$$

This will be a family of hyperbolae.



Families of level curves can be seen as contour maps of the surface.

Derivatives

The derivative is just the rate of change of the function as x and y change. *But* the change in (x, y) can be in any direction.

For example, for the cone, $V = \frac{1}{3}\pi r^2 h$, what happens as $h \to h + \delta h$? We have

$$V \to V + \delta V = \frac{1}{3}\pi r^2 (h + \delta h)$$

so $\delta V = \frac{1}{3}\pi r^2 \delta h$. The rate of change is $\dfrac{\delta V}{\delta h} = \dfrac{1}{3}\pi r^2$ so

$$\lim_{\delta h \to 0} \frac{\delta V}{\delta h} = \frac{1}{3}\pi r^2.$$

This is called $\dfrac{\partial V}{\partial h}$. This is the partial derivative with respect to h (keeping r constant). The ∂ indicates that the other variable is kept constant.
On the other hand, if we change r,

$$V + \delta V = \frac{1}{3}\pi (r + \delta r)^2 h$$

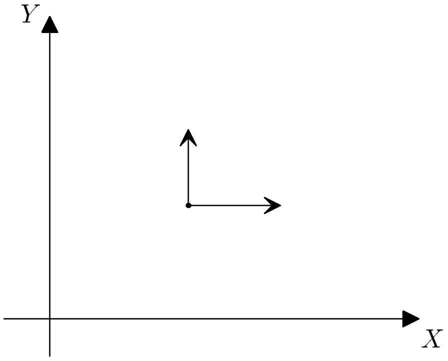
so $\delta V = \frac{1}{3}\pi(2r\delta r + (\delta r)^2)h$. The rate of change is $\dfrac{\delta V}{\delta r} = \dfrac{1}{3}\pi(2r + \delta r)h$ and

$$\frac{\partial V}{\partial r} = \lim_{\delta r \to 0} \frac{\delta V}{\delta r} = \frac{2}{3}\pi r h.$$

Note that we do each differentiation keeping the other variable fixed.
Definition of partial derivative.

If $f(x, y)$ is a function of two variables,

$$\frac{\partial f}{\partial x}(x_0, y_0) = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x, y_0) - f(x_0, y_0)}{\Delta x}$$
$$\frac{\partial f}{\partial y}(x_0, y_0) = \lim_{\Delta y \to 0} \frac{f(x_0, y_0 + \Delta y) - f(x_0, y_0)}{\Delta y}$$



This gives the rate at which f is changing as we move in the x or y direction respectively.

If we think in terms of the surface, it tells if we will be moving uphill or downhill and how fast.

Example 1. $f(x, y) = 2x^2 + 2x^3y^4 - 2xy$.
$\therefore$ $\partial f/\partial x = 4x + 6x^2y^4 - 2y$,
and $\partial f/\partial y = 8x^3y^3 - 2x$.

Example 2 $f(x, y) = e^x \sin(xy)$.
$\therefore$ $\partial f/\partial x = e^x \sin(xy) + e^x y \cos(xy)$,
and $\partial f/\partial y = e^x x \cos(xy)$.

Note that $\left.\frac{\partial f}{\partial x}\right|_y$ and $\left.\frac{\partial f}{\partial y}\right|_x$ is the notation used when it is not clear which variable is to be held constant.

That is, to calculate $\left.\frac{\partial f}{\partial y}\right|_x$, we first express f as a function of x and y, and not some related variables.

We can calculate higher derivatives similarly.

$$\frac{\partial^2 f}{\partial x^2}, \quad \text{(differentiate twice with respect to } x \text{ keeping } y \text{ fixed.)}$$
$$\frac{\partial^2 f}{\partial x \partial y}, \quad \text{(differentiate with respect to } y \text{ then } x.\text{)}$$
$$\frac{\partial^2 f}{\partial y \partial x}, \quad \text{(differentiate with respect to } x \text{ then } y.\text{)}$$

etc.

For $f(x,y) = 2x^2 + 2x^3y^4 - 2xy$,

$$\frac{\partial f}{\partial x} = 4x + 6x^2y^4 - 2y, \qquad \frac{\partial f}{\partial y} = 8x^3y^3 - 2x$$
$$\frac{\partial^2 f}{\partial x^2} = 4 + 12xy^4, \qquad \frac{\partial^2 f}{\partial x \partial y} = 24x^2y^3 - 2$$
$$\frac{\partial^2 f}{\partial y \partial x} = 24x^2y^3 - 2, \qquad \frac{\partial^2 f}{\partial y^2} = 24x^3y^2$$

Note that $\dfrac{\partial^2 f}{\partial y \partial x} = \dfrac{\partial^2 f}{\partial x \partial y}$. This is true for many functions.

Differentials

We already know how a function of two variables changes if we change the value of x *or* y by a small amount. The rate of change is given by the corresponding partial derivative. We can now get an idea of how the function changes if we change the values of *both* x and y by small amounts. Thus, for a function, $z = f(x,y)$, we want to find what happens to z if we move from (x,y) to $(x + \Delta x, y + \Delta y)$. The change in z will be given by

$$\Delta z = f(x + \Delta x, y + \Delta y) - f(x,y).$$

We can write this as

$$\Delta z = \Big[f(x + \Delta x, y + \Delta y) - f(x + \Delta x, y)\Big] + \Big[f(x + \Delta x, y) - f(x,y)\Big].$$

The first term is the change due to the change in y and the second term is the change due to the change in x. We can approximate these separately as

$$f(x + \Delta x, y + \Delta y) - f(x + \Delta x, y) \simeq f_y \Delta y,$$
$$f(x + \Delta x, y) - f(x,y) \simeq f_x \Delta x.$$

Therefore,

$$\Delta z \simeq f_x \Delta x + f_y \Delta y.$$

(Note that we might expect that f_y would need to be evaluated at the point $(x + \Delta x, y)$ rather than the point (x,y). However, the difference that would result is small and, to the accuracy of the approximations we have made, we can evaluate both f_x and f_y at (x,y).)

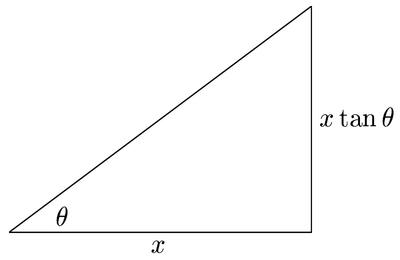
The formula for Δz is called the *increment formula* and it is analogous to the situation for functions of one variable where $\delta y \simeq \dfrac{dy}{dx}\delta x$. We simply have an extra term for the independent variable. It is valid for any *finite* changes, Δx and Δy. If we consider changes in x and y that are infinitesimally small we usually write these changes as dx and dy. The corresponding change in z is

$$dz = f_x dx + f_y dy.$$

dz is called the *differential*.

Example

The area of a right-angled triangle with base x and angle θ is $A = \frac{1}{2}x^2 \tan\theta$.



What is the change in the area if x changes from 1 to 0.95 and θ changes from 45° to 50°?

The increment formula is

$$\Delta A = \frac{\partial A}{\partial x}\Delta x + \frac{\partial A}{\partial \theta}\Delta\theta.$$

Now $\partial A/\partial x = x\tan\theta = 1$ (if $\theta = 45°$) and $\partial A/\partial\theta = \frac{1}{2}x^2\sec^2\theta = 1$. Therefore,

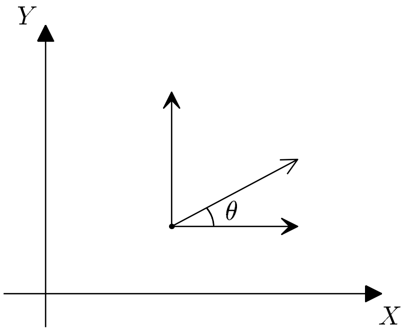
$$\begin{aligned}\Delta A &\simeq \frac{\partial A}{\partial x}\Delta x + \frac{\partial A}{\partial \theta}\Delta\theta \\ &= \Delta x + \Delta\theta \\ &= -0.05 + 5\pi/180 \\ &= 0.0372\end{aligned}$$

If x changes from 1 to 0.95, what change in θ is needed to keep A the same? In this case,

$$\begin{aligned}\Delta A = 0 &= \frac{\partial A}{\partial x}\Delta x + \frac{\partial A}{\partial \theta}\Delta\theta \\ \therefore\ 0 &= -0.05 + \Delta\theta. \\ \therefore\ \Delta\theta &= 0.05 \text{ radians} \\ &= 2.86°.\end{aligned}$$

Directional derivatives

For functions of two variables, $z = f(x, y)$, we know the derivative in the x direction is f_x and the derivative in the y direction is f_y. What about other directions?



Represent $f(x, y)$ by its level curves and consider the point (x_0, y_0). A displacement, $(\Delta x, \Delta y)$, from that point will make an angle θ with the x–axis.

The unit vector in that direction is

$$\hat{\underset{\sim}{u}} = \cos\theta\,\underset{\sim}{i} + \sin\theta\,\underset{\sim}{j}$$

where $\cos\theta = \dfrac{\Delta x}{\Delta s}$, $\sin\theta = \dfrac{\Delta y}{\Delta s}$ and $\Delta s = \sqrt{(\Delta x)^2 + (\Delta y)^2}$. If we move in the direction of $\hat{\underset{\sim}{u}}$, the change in f is

$$\Delta f = \Delta z \simeq f_x \Delta x + f_y \Delta y.$$

The *rate* of change is $\dfrac{\Delta f}{\Delta s}$. The *directional derivative* is

$$\begin{aligned}\frac{df}{ds} &= \lim_{\Delta s \to 0} \frac{\Delta f}{\Delta s} \\ &= \lim_{\Delta s \to 0} \left(f_x \frac{\Delta x}{\Delta s} + f_y \frac{\Delta y}{\Delta s} \right) \\ &= f_x \cos\theta + f_y \sin\theta.\end{aligned}$$

The directional derivative, $\dfrac{df}{ds}$, depends on the direction, θ (or $\hat{\underset{\sim}{u}}$), as well as the position (x_0, y_0). This is the quantity that tells us if we are going uphill or downhill and how fast.

For displacements parallel to the x–axis, $\theta = 0$ and the directional derivative is $\dfrac{df}{ds} = f_x$ (the partial derivative). For displacements parallel to the y–axis, $\theta = \pi/2$ and the directional derivative is $\dfrac{df}{ds} = f_y$

The directional derivative is sometimes written as $D_{\hat{\underset{\sim}{u}}} f$.

Example

Calculate $D_{\hat{\underset{\sim}{u}}} f$ in the direction of $\underset{\sim}{i} + 2\underset{\sim}{j}$ for the function $f(x, y) = 2x^2 + 2x^3y^4 - 2xy$ at the point $(3, 1)$.

Now, $f_x = 4x + 6x^2y^4 - 2y = 64$ at $(3, 1)$ and $f_y = 8x^3y^3 - 2x = 210$ at $(3, 1)$.
$\hat{\underset{\sim}{u}}$ is in the direction $\underset{\sim}{i} + 2\underset{\sim}{j}$ so

$$\begin{aligned}\hat{\underset{\sim}{u}} &= \frac{1}{\sqrt{5}}\underset{\sim}{i} + \frac{2}{\sqrt{5}}\underset{\sim}{j} \\ \therefore\ D_{\hat{\underset{\sim}{u}}} f &= 64 \times \frac{1}{\sqrt{5}} + 210 \times \frac{2}{\sqrt{5}} \\ &= 216.45.\end{aligned}$$

The directional derivative can also be written in terms of vectors. For example,

$$D_{\hat{\underset{\sim}{u}}} f = \frac{df}{ds} = f_x \cos\theta + f_y \sin\theta = (f_x \underset{\sim}{i} + f_y \underset{\sim}{j}) \cdot (u_x \underset{\sim}{i} + u_y \underset{\sim}{j}).$$

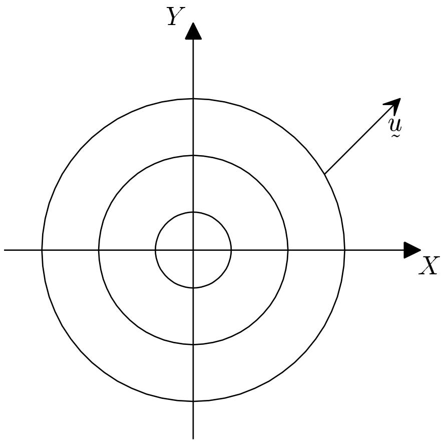
The vector on the right is just $\hat{\underset{\sim}{u}}$. The other vector is a new vector called grad f or $\underset{\sim}{\nabla} f$. Thus

$$D_{\hat{\underset{\sim}{u}}} f = \underset{\sim}{\nabla} f \cdot \hat{\underset{\sim}{u}}.$$

The symbol $\underset{\sim}{\nabla}$ represents a *vector differential operator* which can be written as $\underset{\sim}{i}\dfrac{\partial}{\partial x} + \underset{\sim}{j}\dfrac{\partial}{\partial y}$. Note that $\underset{\sim}{\nabla} f$ contains information about the function and the position and $\hat{\underset{\sim}{u}}$ gives the direction in which we want the derivative.

Example

Consider the function $z = x^2 + y^2$. For this function, the level curves are circles.



Find the directional derivative at $(\sqrt{3}/2, 1/2)$ in the direction of $\underset{\sim}{u} = (1, 1)$.

First we calculate f_x and f_y.

$\partial f/\partial x = 2x = \sqrt{3}$.

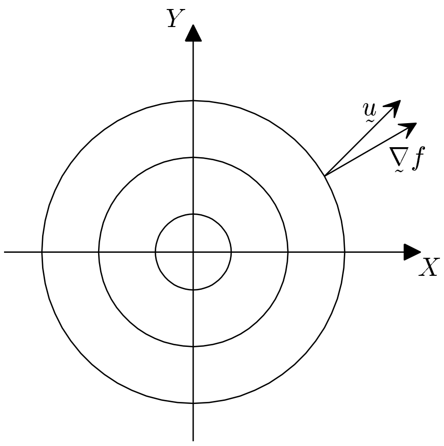
$\partial f/\partial y = 1$.

so $\underset{\sim}{\nabla} f = \sqrt{3}\underset{\sim}{i} + \underset{\sim}{j}$. This tells us what the function looks like at the point. Also, $\hat{\underset{\sim}{u}} = (1/\sqrt{2}, 1/\sqrt{2})$. Therefore

$$\underset{\sim}{\nabla} f \cdot \hat{\underset{\sim}{u}} = \sqrt{3}/\sqrt{2} + 1/\sqrt{2} \simeq 1.93.$$

This tells us how fast f is increasing in the direction of $\hat{\underset{\sim}{u}}$. Note that f is increasing in this direction as the diagram suggests.

It also helps to draw the direction of $\underset{\sim}{\nabla} f$ on the diagram.



If it is drawn carefully, it is perpendicular to the level curves.

We expect this since, if we move along the level curve, df/ds will be zero. Therefore, if we chose $\hat{\underset{\sim}{u}}$ to point along the level curve,

$$\underset{\sim}{\nabla} f \cdot \hat{\underset{\sim}{u}} = 0.$$

i.e. a vector parallel to the level curve is perpendicular to $\underset{\sim}{\nabla} f$.

Note that this gives us a quick way of finding the normal to any curve. If the curve is defined by $f(x, y) = \text{const}$, then the normal to the curve is $\underset{\sim}{\nabla} f = f_x\underset{\sim}{i} + f_y\underset{\sim}{j}$. This helps to give an understanding of the 'meaning' of $\underset{\sim}{\nabla} f$.

The other thing we may want to know is the direction in which f increases the most rapidly. Now $df/ds = \underset{\sim}{\nabla} f \cdot \hat{\underset{\sim}{u}} = |\underset{\sim}{\nabla} f|\,|\hat{\underset{\sim}{u}}| \cos\phi$ where ϕ is the angle between $\underset{\sim}{\nabla} f$ and $\hat{\underset{\sim}{u}}$. So $df/ds = |\underset{\sim}{\nabla} f| \cos\phi$. This is a maximum if $\phi = 0$. i.e. if $\hat{\underset{\sim}{u}} \propto \underset{\sim}{\nabla} f$.

Therefore, $\underset{\sim}{\nabla} f$ points in the direction of maximum increase. Note also that the maximum steepness is given by $|\underset{\sim}{\nabla} f|$.

Also note that $\underset{\sim}{\nabla} f$ is a vector function obtained from a scalar function. In many cases, where a force field is obtained from a scalar potential, the relationship is $\underset{\sim}{F} = -\underset{\sim}{\nabla} V$. Vector functions that can be obtained from a scalar function in this way are very important.

Chain rules

For functions of one variable, if $y = y(x)$, and $x = x(t)$ then

$$\frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt}.$$

Similarly, for functions of two variables, $z = f(x, y)$, x and y might both depend on t. For example, $(x(t), y(t))$ might be the coordinates of a moving particle, so $\underset{\sim}{r} = x(t)\underset{\sim}{i} + y(t)\underset{\sim}{j}$. f might be some potential that affects the motion of the particle. The potential 'seen' by the particle is $z = f(x(t), y(t))$. In this case, z really depends on t only.

What is $\dfrac{dz}{dt}$? How rapidly is the potential changing for the particle?
In time δt, the change in f is $\delta f \simeq f_x \delta x + f_y \delta y$. We want

$$\begin{aligned}\lim_{\delta t \to 0} \frac{\delta f}{\delta t} &= \lim_{\delta t \to 0} \left(f_x \frac{\delta x}{\delta t} + f_y \frac{\delta y}{\delta t} \right) \\ &= f_x \frac{dx}{dt} + f_y \frac{dy}{dt} \\ &= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}\end{aligned}$$

Note this is the same as the chain rule for functions of one variable, except that we have two terms instead of one. It can also be written as

$$\frac{df}{dt} = \underset{\sim}{\nabla} f \cdot \frac{d\underset{\sim}{r}}{dt}.$$

Example

A particle moves through a potential field $V = 3x^2y + e^x$ on a path $\underset{\sim}{r} = t\underset{\sim}{i} + \sin t\,\underset{\sim}{j}$. What is dV/dt when $t = \pi$?

Now $\underset{\sim}{\nabla} V = (6xy + e^x)\underset{\sim}{i} + 3x^2\underset{\sim}{j} = (6t \sin t + e^t)\underset{\sim}{i} + 3t^2\underset{\sim}{j}$.
Also, $d\underset{\sim}{r}/dt = \underset{\sim}{i} + \cos t\,\underset{\sim}{j}$. Therefore

$$\begin{aligned}\frac{dV}{dt} &= (6t \sin t + e^t) + 3t^2 \cos t \\ &= e^{\pi} - 3\pi^2 \text{ at } t = \pi.\end{aligned}$$

Functions of Three Variables

What happens if we have a function of three variables? i.e. $w = f(x, y, z)$. Most of the formulae generalize quite easily. The main difficulty is in visualizing what is happening. For example, consider

$$w = x^2 + y^2 + z^2.$$

We can't 'plot' this as a surface, as we would need four dimensions. So, we can't visualize it in this way.

We can try to look at the the 'level curves'. i.e. $f(x, y, z) = \text{const}$ or

$$x^2 + y^2 + z^2 = \text{constant}.$$

Note that these are now level surfaces. In fact they are spheres with centre at the origin. They need three dimensions rather than two dimensions.

There will be three partial derivatives. The differential is now

$$dw = f_x dx + f_y dy + f_z dz.$$

The increment formula is

$$\delta w = f_x \delta x + f_y \delta y + f_z \delta z.$$

The directional derivative is

$$D_{\hat{\underset{\sim}{u}}} f = \underset{\sim}{\nabla} f \cdot \hat{\underset{\sim}{u}},$$

where $\underset{\sim}{\nabla} f = f_x \underset{\sim}{i} + f_y \underset{\sim}{j} + f_z \underset{\sim}{k}$ and $\hat{\underset{\sim}{u}}$ is a three dimensional unit vector in the direction we are interested in.

There is also a chain rule. If $V(t) = f(x(t), y(t), z(t))$, then

$$\frac{dV}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt} = \underset{\sim}{\nabla} f \cdot \frac{d\underset{\sim}{r}}{dt}.$$

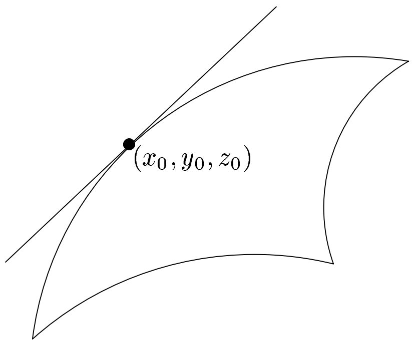
As for functions of two variables, $\underset{\sim}{\nabla} f$ is perpendicular to the level curves, only in this case they are level surfaces. This gives us a way to find the normal of a surface. Suppose we have a surface $z = g(x, y)$. If

$$f(x, y, z) = z - g(x, y)$$

then the equation $f(x, y, z) = 0$ will define the same surface. This is a level surface of f. The normal to this surface is given by $\underset{\sim}{\nabla} f$. i.e.

$$\begin{aligned} \underset{\sim}{n} &\propto \underset{\sim}{\nabla} f \\ &= -\frac{\partial g}{\partial x}\underset{\sim}{i} - \frac{\partial g}{\partial y}\underset{\sim}{j} + \underset{\sim}{k}. \end{aligned}$$

This is the same as the formula we would get if we calculated the normal to the surface $z = g(x, y)$ using other methods. Once we know the normal to the surface, we can calculate the equation of the tangent plane.



The tangent plane at (x_0, y_0, z_0) will be perpendicular to the normal to the surface at that point.

Excercise

Show that the equation the tangent plane to the surface $z = g(x, y)$ at the point (x_0, y_0, z_0) can be expressed as

$$-(x - x_0)g_x - (y - y_0)g_y + (z - z_0) = 0$$

or

$$-g_x x - g_y y + z = (-g_x x_0 - g_y y_0 + z_0) = \text{const.}$$