$\boxed{2D}$ Recall!

Function Notation

eg. $y = 2x+3$ → linear

$f(x) = 2x+3$ → Quadratic

one variable ← one variable

$\boxed{3D}$ dependent: z

independent: x & y

eg. $ax+by+cz=d$

$z = f(x,y)$

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

Chapter 5

y: dependent variable

Multivariable Calculus

x: independent variable

5.1 Functions of several variables

We are all familiar with simple functions such as $y = x^3$. And we all know the answers to questions such as

- What is the domain and range of the function?
- What does the function look line as a plot in the $xy-$ plane?
- What is the derivative of the function?

Single variable calculus encompasses functions such as $y = x^3$ where y is a function of the single (independent) variable x. The graph of $y = f(x)$ is a *curve* in the $xy-$ plane. In nature, many physical quantities depend on *more than one* independent variable. We are now going to explore how to answer similar questions to the above for functions such as $z = x^3 + y^2$. This is just one example of what we call functions of several variables. We can have as many variables as we want; $z = x^3 + y^2$ is a function of *two* independent variables (x and y), $w = x^3 + y^2 - z^2$ is a function of *three* independent variables (x, y and z), and so forth. Just as we would write $f(x) = x^3$ we can write $f(x, y) = x^3 + y^2$ and $f(x, y, z) = x^3 + y^2 - z^2$ and so on. For the remainder of this course we will focus on functions involving *two* independent variables, but bear in mind that the lessons learnt here will be applicable to functions of any number of variables.

5.1.1 Definition

A function f of two (independent) variables (x, y) is a single valued mapping of a subset of R^2 into a subset of R.

$f: \mathbb{R} \to \mathbb{R}$, $f(x) = x^2 - 3$

Domain $\mathbb{R}^2 \to \mathbb{R}$

$f: \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = x^2y + x - y^2 + 3$

restriction!

What does this mean? Simply that for any allowed value of x and y we can compute a single value for $f(x, y)$. In a sense f is a process for converting pairs of numbers (x and y) into a single number f.

The notation R^2 means all possible choices of x and y such as all points in the xy-plane. The symbol R denotes all real numbers (for example all points on the real line). The use of the word *subset* in the above definition is simply to remind us that functions have an allowed domain (i.e. a subset of R^2) and a corresponding range (i.e. a subset of R).

Notice that we are restricting ourselves to real variables, that is the function's value and its arguments (x, y) are all real numbers. This game gets very exciting and somewhat tricky when we enter the world of complex numbers. Such adventures await you in later year mathematics (not surprisingly this area is known as *Complex Analysis*).

5.1.2 Notation

Here is a function of two variables

$$f(x, y) = \sin(x + y)$$

We can choose the domain to be R^2 and then the range will be the closed set $[-1, +1]$. Another common way of writing all of this is

$$f : (x, y) \in R^2 \mapsto \sin(x + y) \in [-1, 1]$$

Range

This notation identifies the function as f, the domain as R^2, the range as $[-1, 1]$ and most importantly the rule that (x, y) is mapped to $\sin(x + y)$. For this subject we will stick with the former notation.

You should also note that there is nothing sacred about the symbols x, y and f. We are free to choose what ever symbols takes our fancy, for example we could create the function

$(u-v) > 0$

$\log_e \boxed{(u-v)}$

$u > v$

$u > 0$

$$w(u, v) = \log(u - v)$$

$v > 0$

$u > v$

Example 5.1. What would be a sensible choice of domain for the previous function?

5.1.3 Surfaces

A very common application of functions of two variables is to describe a *surface* in 3-dimensional space. How do we do this? The idea is that we take the value of the function to describe the *height* of the surface above the xy-plane. If we use standard Cartesian coordinates then such a surface could be described by the equation

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

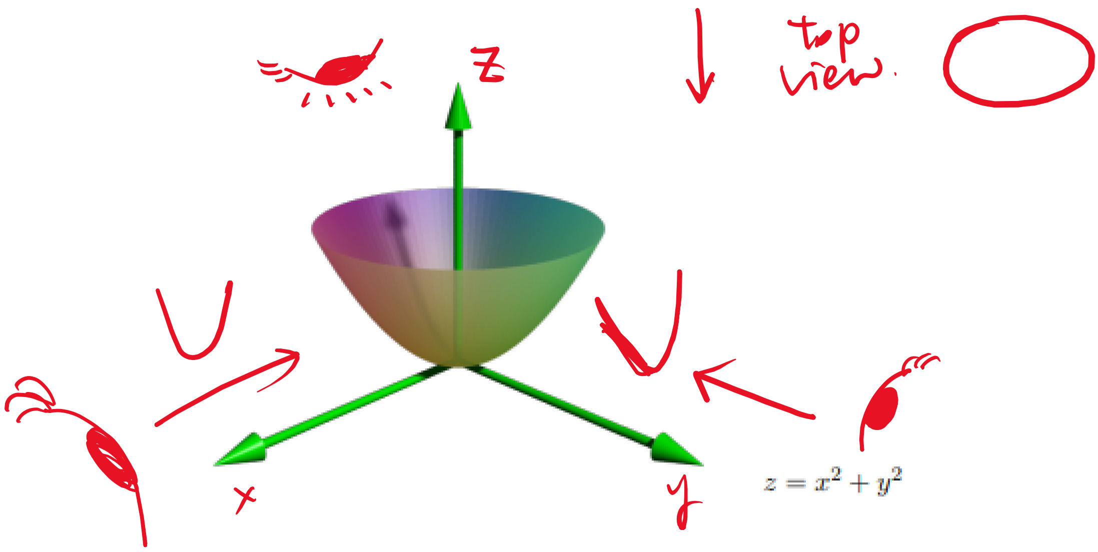
This surface has a height z units above each point (x, y) in the xy-plane.

Just as the equation $y = f(x)$ describes the curve in the $xy-$ plane, the equation $z = f(x, y)$ describes the surface in R^3. Just as the curve $C = f(x)$ is made up of the points (x, y), the surface $S = f(x, y)$ is made up of the points (x, y, z). As $z = f(x, y)$ describes this surface *explicitly* as a height function over a plane, we say that the surface is given in *explicit form*.

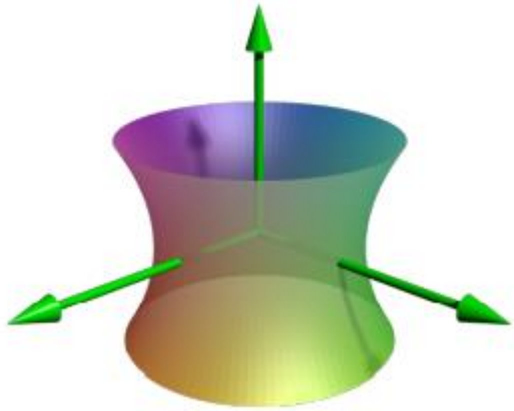
A surface such as $z = f(x, y)$ is also often called the *graph* of the function f.

Here are some simple examples. A very good exercise is to try to convince yourself that the following images are correct (i.e. that they do represent the given equation).

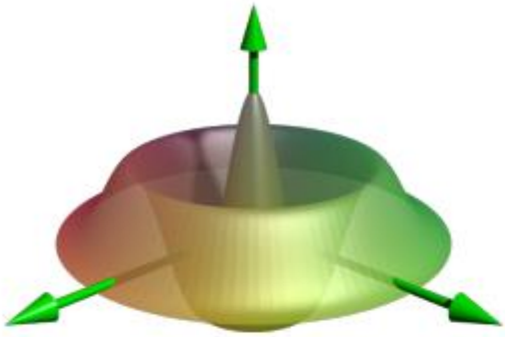
Note that in each of the following r is defined as $r = +\sqrt{(x^2 + y^2)}$.



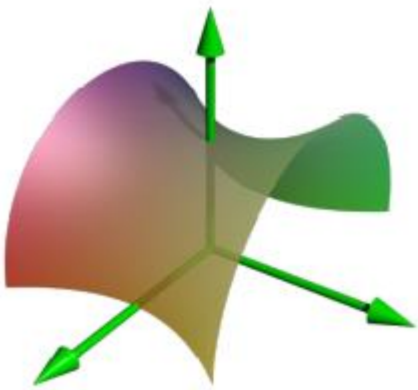
$z = x^2 + y^2$



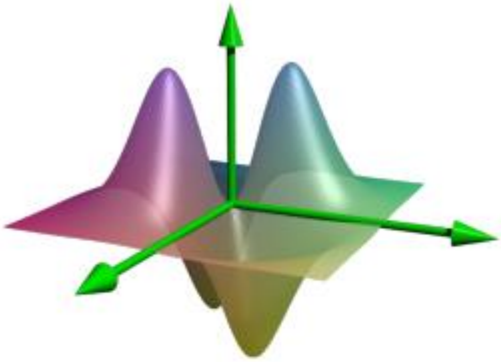
$$1 = x^2 + y^2 - z^2$$



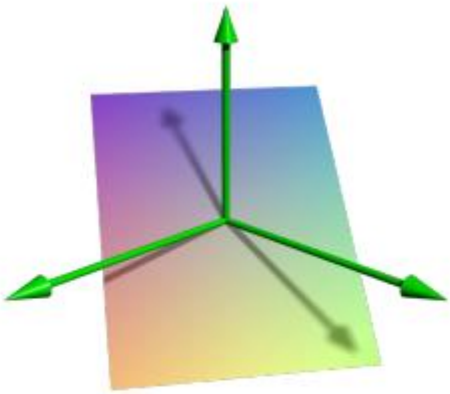
$$z = \cos(3\pi r)\exp\left(-2r^2\right)$$



$$z = \sqrt{1 + y^2 - x^2}$$



$z = -xy\exp\left(-x^2 - y^2\right)$



$1 = x + y + z$

Example 5.2. Sketch and describe the graph of the surface $z = f(x, y) = 6 + 3x + 2y$.

$z = 6 + 3x + 2y$

$3x + 2y - z = -6$ $\qquad ax + by + cz = d$ (plane)

A plane surface in space which passes through the points $(0,0,6)$.

5.1.4 Alternative forms

We might ask are there any other ways in which we can describe a surface? We should be clear that (in this subject) when we say surface we are talking about a 2-dimensional surface in our familiar 3-dimensional space. With that in mind, consider the equation

$$0 = g(x, y, z)$$

eg. $3x+2y-z=-6$

$3x+2y-z-6=0$

$g(x,y,z) = 0$

$g(x,y,z)=3x+2y-z-6$

What do we make of this equation? Well, after some algebra we might be able to re-arrange the above equation into the familiar form

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

for some function f. In this form we see that we have a surface, and thus the previous equation $0 = g(x, y, z)$ also describes a surface. When the surface is described by an equation of the form $0 = g(x, y, z)$ we say that the surface is given in *implicit form*.

Consider all of the points in R^3 (i.e all possible (x, y, z) points). If we now introduce the equation $0 = g(x, y, z)$ we are forced to consider only those (x, y, z) values that satisfy this constraint. We could do so by, for example, arbitrarily choosing (x, y) and using the equation (in the form $z = f(x, y)$ to compute z. Or we could choose say (y, z) and use the equation $0 = g(x, y, z)$ to compute x. Which ever road we travel it is clear that we are free to choose just two of the (x, y, z) with the third constrained by the equation.

Now consider some simple surface and let's suppose we are able to drape a sheet of graph paper over the surface. We can use this graph paper to select individual points on the surface (well as far as the graph paper covers the surface). Suppose we label the axes of the graph paper by the symbols u and v. Then each point on surface is described by a unique pair of values (u, v). This makes sense – we are dealing with a 2-dimensional surface and so we expect we would need 2 numbers ((u, v)) to describe each point on the surface. The parameters (u, v) are often referred to as (local) *coordinates on the surface.*

How does this picture fit in with our previous description of a surface, as an equation of the form $0 = g(x, y, z)$? Pick any point on the surface. This point will have both (x, y, z) and (u, v) coordinates. That means that we can describe the point in terms of either (u, v) or (x, y, z). As we move around the surface all of these coordinates will vary. So given (u, v) we should be able to compute the corresponding (x, y, z) values. That is we should be able to find functions $P(u, v)$, $Q(u, v)$ and $R(u, v)$ such that

$$x = P(u, v) \quad y = Q(u, v) \quad z = R(u, v)$$

parametric form

The above equations describe the surface in **parametric form**.

Example 5.3. Identify (i.e. describe) the surface given by the equations

$$①\ x = 2u + 3v + 1 \quad ②\ y = u - 4v + 2 \quad ③\ z = u + 2v - 1$$

Hint : Try to combine the three equations into one equation involving x, y and z but not u and v.

$① - ② \times 2 \Rightarrow$

$$\begin{aligned} x &= 2u + 3v + 1 \\ (-)\ 2y &= 2u - 8v + 4 \\ \hline x - 2y &= 11v - 3 \quad ④ \end{aligned}$$

$② - ③ \Rightarrow y - z = -6v + 3 \quad ⑤$

$④ \times 6 + ⑤ \times 11 \Rightarrow$

$$\begin{aligned} 6x - 12y &= 66v - 18 \\ (+)\ 11y - 11z &= -66v + 33 \\ \hline 6x - y - 11z &= 15 \end{aligned}$$

plane surface passes through $(0, -15, 0)$.

Example 5.4. Describe the surface defined by the equations

$$①\ x = 3\cos(\phi)\sin(\theta) \quad ②\ y = 4\sin(\phi)\sin(\theta) \quad ③\ z = 5\cos(\theta)$$

for $0 < \phi < 2\pi$ and $0 < \theta < \pi$

From ① $\frac{x}{3} = \cos(\phi)\sin(\theta)$

② $\frac{y}{4} = \sin(\phi)\sin(\theta)$

From ③ $\frac{z}{5} = \cos(\theta)$

$\frac{z^2}{25} = \cos^2(\theta)$ ⑤

$①^2 + ②^2 \Rightarrow \frac{x^2}{9} + \frac{y^2}{16} = \sin^2(\theta)\left[\cos^2(\phi) + \sin^2(\phi)\right]$

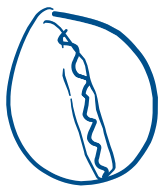
$\frac{x^2}{9} + \frac{y^2}{16} = \sin^2(\theta)$ ④

④ + ⑤

$\frac{x^2}{9} + \frac{y^2}{16} + \frac{z^2}{25} = \sin^2(\theta) + \cos^2(\theta)$

$\frac{x^2}{9} + \frac{y^2}{16} + \frac{z^2}{25} = 1$

Example 5.5. How would your answer to the previous example change if the domain for θ was $0 < \theta < \pi/2$?



Equations for surfaces

A 2-dimensional surface in 3-dimensional space may be described by any of the following forms.

Explicit	$z = f(x, y)$
Implicit	$0 = g(x, y, z)$
Parametric	$x = P(u, v),\ y = Q(u, v),\ z = R(u, v)$

5.2 Partial derivatives

5.2.1 First partial derivatives

We are all familiar with the definition of the derivative of a function of one variable

$$\frac{df}{dx} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

small change in x value, Δx

$x \to x + \Delta x$

The natural question to ask is: Is there similar rule for functions of more than one variable? The answer is yes, and we will develop the necessary formulas by a simple generalisation of the above definition.

Let us suppose we have a function, say $f(x, y)$. Suppose for the moment that we pick a particular value of y, say $y = 3$. Then only x is allowed to vary and in effect we now have a function of just one variable. Thus we can apply the above definition for a derivative which we write as

partial derivative →

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

y is treat as a constant

Notice the use of the symbol ∂ rather than d. This is to remind us that in computing this derivative all other variables are held constant (which in this instance is just y).

Of course we could do the same again but with x held constant. This gives us the derivative in y

x is treated as a constant

partial derivative →

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

$f_x = \frac{\partial f}{\partial x}$

$f_y = \frac{\partial f}{\partial y}$

Each of these derivatives, $\partial f / \partial x$ and $\partial f / \partial y$ are known as **first order partial derivatives** of f (the derivative of a function of one variable is often called an *ordinary derivative*).

We can also look at this in terms of the *rate of change*, as we did for single variable functions. If z is a function of two independent variables x and y (i.e. $z = f(x, y)$) then there are two independent *rates of change*. One of these is the **rate of change of f with respect to the variable x**, and the other is the **rate of change of f with respect to the variable y**.

You might think that we would now need to invent new rules for the (partial) derivatives of products, quotients and so on. But our definition of partial derivatives is built upon the definition of an ordinary derivative of a function of one variable. Thus all the familiar rules carry over without modification. For example, the product rule for partial derivatives is

$$\frac{\partial(fg)}{\partial x} = g\frac{\partial f}{\partial x} + f\frac{\partial g}{\partial x}$$
$$\frac{\partial(fg)}{\partial y} = g\frac{\partial f}{\partial y} + f\frac{\partial g}{\partial y}$$

Computing partial derivatives is no more complicated than computing ordinary derivatives. Hooray for us!

Rules for finding partial derivatives

- To find $\frac{\partial f}{\partial x}$, treat y as a constant and differentiate $f(x, y)$ with respect to x only.
- To find $\frac{\partial f}{\partial y}$, treat x as a constant and differentiate $f(x, y)$ with respect to y only.

Example 5.6. If $f(x, y) = x^3 + x^2y^3 - 2y^2$, find $f_x(2, 1)$ and $f_y(2, 1)$.

$\frac{\partial f}{\partial x} = f_x = 3x^2 + 2xy^3 + 0 = 3x^2+2xy^3$ $\quad f_x(2,1) = 3(2)^2+2(2)(1)^3 = 16$

$\frac{\partial f}{\partial y} = f_y = 0 + 3x^2y^2 - 4y = 3x^2y^2-4y$ $\quad f_y(2,1) = 3(2)^2(1)^2-4(1) = 8$

Example 5.7. If $f(x, y) = \sin(x)\cos(y)$ then

$$\begin{aligned}\frac{\partial f}{\partial x} &= \frac{\partial \sin(x)\cos(y)}{\partial x} \\ &= \cos(y)\frac{\partial \sin(x)}{\partial x} \\ &= \cos(y)\cos(x)\end{aligned}$$

Also find $\frac{\partial f}{\partial y}$. $= \sin(x)\,[-\sin(y)] = -\sin(x)\sin(y)$

Example 5.8. If $g(x,y,z) = e^{-x^2-y^2-z^2}$ then

$$\begin{aligned}\frac{\partial g}{\partial z} &= \frac{\partial e^{-x^2-y^2-z^2}}{\partial z}\\ &= e^{-x^2-y^2-z^2}\frac{\partial(-x^2-y^2-z^2)}{\partial z}\\ &= -2ze^{-x^2-y^2-z^2}\end{aligned}$$

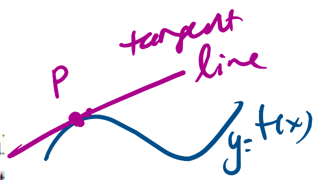
Also find $\dfrac{\partial g}{\partial x}$ and $\dfrac{\partial g}{\partial y}$.

$$\frac{\partial g}{\partial x} = (-2x+0+0)\,e^{-x^2-y^2-z^2} = -2x\,e^{-x^2-y^2-z^2}$$

$$\frac{\partial g}{\partial y} = -2y\,e^{-x^2-y^2-z^2}$$

A word on notation: An alternative notation for $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ is f_x and f_y respectively. You will find both versions are commonly used.

5.3 The tangent plane → 2D surface curve



For functions of one variable we found that a *tangent line* provides a useful means of approximating the function. It is natural to ask how we might generalise this idea to functions of several variables.

Constructing a tangent line for a function of a single variable, $f = f(x)$, is quite simple. (This should be revision!) First we compute the function's value f and its gradient $\dfrac{df}{dx}$ at some chosen point. We then construct a *straight line* equation ($y = mx + c$) with these values at the chosen point. This line is the *tangent line* of the function f at the given point.

gradient of tangent line

$y - y_1 = m_t(x - x_1)$

Example 5.9. Find the tangent line to the function $f(x) = \sin x$ at $x = \dfrac{\pi}{4}$

$f'(x) = \cos(x)$

$f'(\frac{\pi}{4}) = \cos(\frac{\pi}{4}) = \frac{\sqrt{2}}{2}$ (m_t)

when $x = \frac{\pi}{4}$

$y = \sin(\frac{\pi}{4}) = \frac{\sqrt{2}}{2}$

point $(\frac{\pi}{4}, \frac{\sqrt{2}}{2})$

Eqn tangent line is

$y - \frac{\sqrt{2}}{2} = \frac{\sqrt{2}}{2}(x - \frac{\pi}{4})$

$y = \frac{\sqrt{2}}{2}x + \frac{\sqrt{2}}{2}(1 - \frac{\pi}{4})$

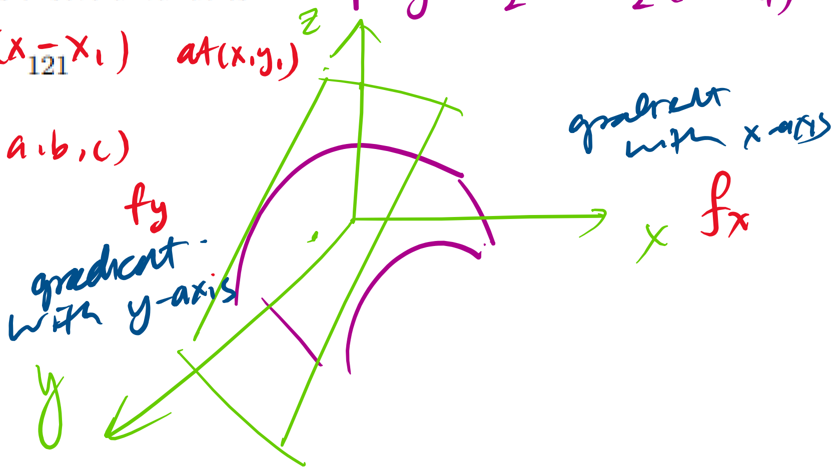
How do we relate this to functions of several variables?

tangent line is $y - y_1 = m_t(x - x_1)$ at (x_1, y_1)

$y = f(x)$

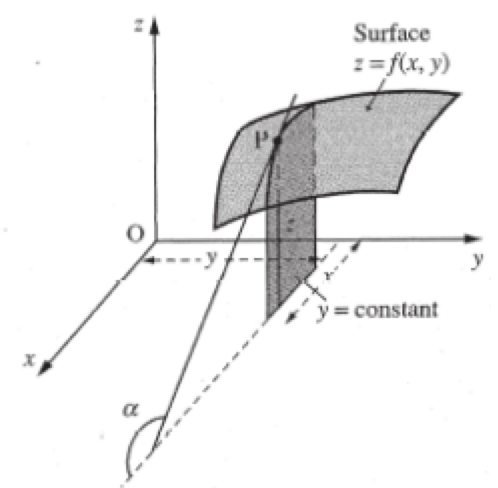
tangent plane $z = f(x,y)$, at (a,b,c)

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



5.3.1 Geometric interpretation

Earlier we noted that the partial derivative $\dfrac{\partial f}{\partial x}$ of the function of two variables $z = f(x, y)$ is the *rate of change* of f in the x-direction, keeping y fixed. To visualise $\dfrac{\partial f}{\partial x}$ as the *slope* (or gradient) of a straight line, consider the diagram below.



This diagram shows the intersection of the vertical plane y = 'constant' with the smooth differentiable surface $z = f(x, y)$ in R^3.

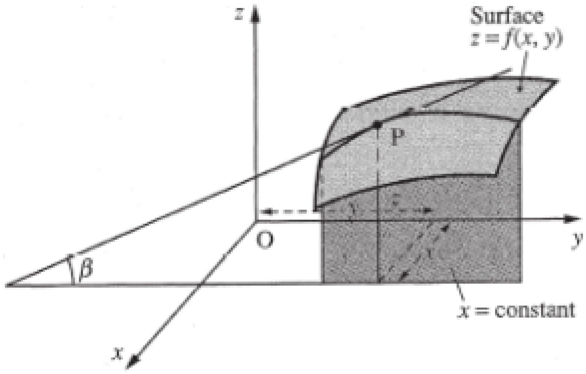
The intersection of the plane with this surface is a curve, C_1 say. On C_1, x can vary but y stays constant. We now draw the *tangent line* to the surface at the point P that also lies in the vertical plane y = constant. This tangent line, T_1 say, strikes the xy-plane at angle α as shown.

This tangent line has slope $\tan\alpha$ which equals the rate of change of the height z of the surface $z = f(x, y)$ in the x-direction at the point P. We thus have

f_x $$\frac{\partial f}{\partial x} = \tan\alpha = \text{slope of the tangent line to the surface } z = f(x, y)$$

in x-direction.

Similarly, the diagram below illustrates the intersection of the vertical plane $x =$ 'constant' with the surface $z = f(x, y)$.

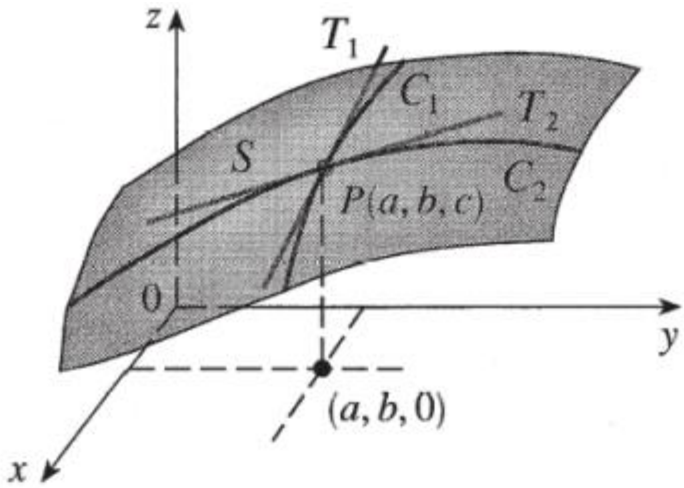


This intersection is a smooth curve, C_2 say, on the surface. On C_2, y can vary but x stays fixed. As we move along this curve, the height z to the surface changes only with respect to the change in the independent variable y. Now we can draw the tangent line to the curve C_2 at the point P. This strikes the xy-plane at angle β. The slope of this tangent line, namely $\tan\beta$, gives the rate of change of f with respect to y at the point P. That is

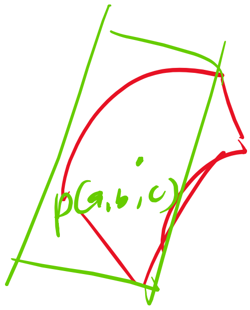
f_y $$\frac{\partial f}{\partial y} = \tan\beta = \text{slope of the tangent line to the surface } z = f(x, y)$$

in y-direction.

What happens if we consider the rate of change of f with respect to x and with respect to y together? It is helpful to look at the diagram below, which shows a section of the curve S of a differentiable function $z = f(x, y)$ at a point $P(a, b, c)$, where $c = f(a, b)$.



If we zoom in onto the surface at P it becomes locally flat. We can then draw the two tangent lines T_1 and T_2 to the surface at P which are tangential to the two curves C_1 and C_2 that lie in the vertical planes $y = b$ and $x = a$. These tangent lines, which have the slopes $\dfrac{\partial f}{\partial x} = \tan\alpha$ and $\dfrac{\partial f}{\partial y} = \tan\beta$ shown previously, give the *rate of change* of $f(x, y)$ in both the x and y directions.



The **tangent plane** to the surface at the point P is the plane that contains both of the tangent lines T_1 and T_2. Let us now find the *equation* of this plane. We know that the general equation of a plane that passes through the point $P(a, b, c)$ is

$$z - c = \overset{f_x}{m}(x - a) + \overset{f_y}{n}(y - b) \tag{5.1}$$

Here m and n are the slopes of the lines of intersection of the general plane with the two vertical planes $y = b$ and $x = a$ that are parallel to the principal coordinate planes (the $xz-$ plane and the $yz-$ plane respectively). If we now put $y = b$ in equation (5.1), we have $z - c = m(x - a)$. This is the equation of the line of intersection of our general plane with the plane $y = b$. It clearly has slope m. Next we put $x = a$ in equation (5.1) and this yields $z - c = n(y - b)$. This is the equation of the line of intersection of our general plane with the plane $x = a$. It clearly has slope n. Lastly, if we choose

$$m = \tan\alpha = \frac{\partial f}{\partial x} = f_x(a, b) \text{ and } = n = \tan\beta = \frac{\partial f}{\partial y} = f_y(a, b)$$

then the equation of the *tangent plane* to the surface $z = f(x, y)$ at the vector form

$\underset{\sim}{r} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} + \begin{pmatrix} p \\ q \\ r \end{pmatrix} t + \begin{pmatrix} l \\ m \\ n \end{pmatrix} s$

gradient vector with x-axis

gradient vector with y-axis

can be in different position

point $(x, y) = (a, b)$ is

$$z = f(a,b) + f_x(a,b) \cdot (x - a) + f_y(a,b) \cdot (y - b) \tag{5.2}$$

Our work is done!

Example 5.10. Find the equation of the tangent plane to the surface $z = 2x^2 + y^2$ at the point $(a, b) = (1, 1)$.

Solution: Here $f(x, y) = 2x^2 + y^2$. Thus $f(a, b) = f(1, 1) = 2 \cdot 1^2 + 1^2 = 3$. $(1,1,3)$
Next,

$$\frac{\partial f}{\partial x} = 4x \qquad \text{therefore } \frac{\partial f}{\partial x}(1, 1) = 4(1) = 4 \quad (m)\ (f_x)$$

$$\frac{\partial f}{\partial y} = 2y \qquad \text{therefore } \frac{\partial f}{\partial y}(1, 1) = 2(1) = 2 \quad (n)\ (f_y)$$

Using equation (5.2) the equation of the tangent plane is

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

$z - 3 = 4(x-1) + 2(y-1)$

$z = 4x + 2y - 4 - 2 + 3 \iff 4x + 2y - z = 3$

5.3.2 Linear approximations

We have done the hard work, and now it is time to enjoy the fruits of our labour. Just as we used the tangent line in approximations for functions of one variable, we can use the *tangent plane* as a way to estimate the original function $f(x, y)$ in a region close to the chosen point.

The equation of the tangent plane to the surface $z = f(x, y)$ at the point (a, b) is also the equation for the **linear approximation** to $z = f(x, y)$ for points (x, y) near (a, b). We can regard the tangent plane equation (5.2) as the natural extension to functions of two variables (x, y) of the **Taylor polynomial of degree one** equation

$$y = T_1(x; a) = f(a) + f'(a) \cdot (x - a).$$

This is the linear approximation equation for functions of one variable, namely $y = f(x)$, for x near a.

Hence we call

$$z = T_1(x, y) = f(a, b) + f_x(a, b) \cdot (x - a) + f_y(a, b) \cdot (y - b)$$

$\boxed{x=a}$

$T_n(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f^{(2)}(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n$

$T_1(x) = f(a) + \frac{f'(a)}{1!}(x-a) \leftarrow$ one variable

$x=a,\ y=b$

$T_n(x,y) = f(a,b) + \frac{f_x(a,b)}{1!}(x-a) + \frac{f_y(a,b)}{1!}(y-b) + \frac{f_{xx}(a,b)}{2!}(x-a)^2 + \frac{f_{xy}(a,b)}{2!}(x-a)(y-b)$

$\boxed{T_1(x,y) = f(a,b) + \frac{f_x(a,b)}{1!}(x-a) + \frac{f_y(a,b)}{1!}(y-b)} + \frac{f_{yx}(a,b)}{2!}(x-a)(y-b) + \frac{f_{yy}(a,b)}{2!}(y-b)^2$

$+ \dots$

$T_1(x,y) = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b)$

Tangent plane $\xrightarrow{\uparrow}$ $z = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b)$

$L(x,y)$ $\rightarrow$

$= T_1(x,y)$

$=$ tangent plane eqn

the **linear approximation** to $f(x, y)$ for points (x, y) near (a, b). Please note, we will omit the centring point (a, b) from the argument of T_1 as the notation becomes too cumbersome! You will need to understand from context which centring point is being utilised.

Example 5.11. Derive the linear approximation function $T_1(x, y)$ for the function $f(x, y) = \sqrt{3x - y}$ at the point $(4, 3)$.

$f(x,y) = (3x-y)^{\frac{1}{2}}$ $\quad f(4,3) = \sqrt{12-3} = 3$ $\quad$ $\boxed{\text{Point}}$ $(4,3,3)$

$f_x = \frac{1}{2}(3x-y)^{-\frac{1}{2}}(3) = \frac{3}{2\sqrt{3x-y}}$ $\quad f_x(4,3) = \frac{3}{2(3)} = \frac{1}{2}$

$f_y = \frac{1}{2}(3x-y)^{-\frac{1}{2}}(-1) = \frac{-1}{2\sqrt{3x-y}}$ $\quad f_y(4,3) = \frac{-1}{2(3)} = -\frac{1}{6}$

$L(x,y) = T_1(x,y) = 3 + \frac{1}{2}(x-4) - \frac{1}{6}(y-3)$

Tangent plane $\Rightarrow$ $z - 3 = \frac{1}{2}(x-4) - \frac{1}{6}(y-3)$

Example 5.12. Use the result of example 5.9 to estimate $\sin(x)\sin(y)$ at $(\frac{5\pi}{16}, \frac{5\pi}{16})$. $\rightarrow$ $(\frac{\pi}{4}, \frac{\pi}{4}) = (\frac{4\pi}{16}, \frac{4\pi}{16})$

$f(x,y) = \sin(x)\sin(y)$ $\quad f(\frac{\pi}{4}, \frac{\pi}{4}) = \frac{1}{2}$ $\quad (\frac{\pi}{4}, \frac{\pi}{4}, \frac{1}{2})$ $\boxed{\text{point}}$

$f_x = \cos(x)\sin(y)$ $\quad f_x(\frac{\pi}{4}, \frac{\pi}{4}) = \frac{1}{2}$

$f_y = \sin(x)\cos(y)$ $\quad f_y(\frac{\pi}{4}, \frac{\pi}{4}) = \frac{1}{2}$

$L(x,y) = T_1(x,y) = \frac{1}{2} + \frac{1}{2}(x - \frac{\pi}{4}) + \frac{1}{2}(y - \frac{\pi}{4})$ $\Rightarrow$ at point $(\frac{\pi}{4}, \frac{\pi}{4}, \frac{1}{2})$ for $f(x,y)$

Now to Evaluate at $(\frac{5\pi}{16}, \frac{5\pi}{16})$

$L(\frac{5\pi}{16}, \frac{5\pi}{16}) \approx \sin(\frac{5\pi}{16}, \frac{5\pi}{16})$

$= \frac{1}{2} + \frac{1}{2}(\frac{5\pi}{16} - \frac{\pi}{4}) + \frac{1}{2}(\frac{5\pi}{16} - \frac{\pi}{4})$

$= \frac{1}{2} + \frac{\pi}{16} \approx 0.6963$

original value $\sin(\frac{5\pi}{16}, \frac{5\pi}{16})$

≈ 0.69134

Chain Rule

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

5.4 Chain rule

In a previous lecture we saw how we could compute (partial) derivatives of functions of several variables. The trick we employed was to reduce the number of independent variables to just one (which we did by keeping all but one variable constant). There is another way in which we can achieve this reduction, which involves parametrising the function.

$z = f(x,y)$

$x(t) \qquad y(t)$

$z(t)$

Consider a function of two variables $f(x, y)$ and let's suppose we are given a smooth (continuous, with derivatives which are also continuous) curve in the xy-plane. Each point on this curve can be characterised by its distance from some arbitrary starting point on the curve. In this way we can imagine that the (x, y) pairs on this curve are given as functions of one variable, let's call it s. That is, our curve is described by the parametric equations

$$x = x(s) \qquad y = y(s)$$

for some functions $x(s)$ and $y(s)$. The values of the function $f(x, y)$ on this curve are therefore given by

$f(x(t), y(t))$

$f(x,y)$

$x(s) \quad y(s)$

$$f = f(x(s), y(s)) = f(s) \longrightarrow \text{differentiation } \frac{df}{ds}$$

and this is just a function of one variable s. Thus we can compute its derivative df/ds. We will soon see that df/ds can be computed in terms of the partial derivatives.

Chain Rule $\frac{df}{ds} = \left(\frac{\partial f}{\partial x} \times \frac{dx}{ds}\right) + \left(\frac{\partial f}{\partial y} \times \frac{dy}{ds}\right)$

$f(x,y) \to \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$

$x(s) \to \frac{dx}{ds}$

$y(s) \to \frac{dy}{ds}$

Example 5.13. Given the curve

$$x(s) = 2s, \qquad y(s) = 4s^2 \qquad -1 < s < 1$$

and the function

$$f(x, y) = 5x - 7y + 2$$

compute $\frac{df}{ds}$ at $s = 0$.

Chain Rule

$\frac{dx}{ds} = 2 \qquad \frac{dy}{ds} = 8s$

$\frac{\partial f}{\partial x} = 5 \qquad \frac{\partial f}{\partial y} = -7$

$\frac{df}{ds} = (5)(2) + (-7)(8s)$

$\frac{df}{ds} = 10 - 56s$

Direct method $f(x,y) \longrightarrow f(s) \longrightarrow \frac{df}{ds}$

$f(x,y) \Rightarrow f(s) = 5(2s) - 7(4s^2) + 2$

$f(s) = 10s - 28s^2 + 2$

$\frac{df}{ds} = 10 - 56s$

Example 5.14. Show that for the curve $x(s) = s$, $y(s) = 2$ we get $\frac{df}{ds} = \frac{\partial f}{\partial x}$.

$f(x,y) \to \frac{\partial f}{\partial x}$, $\to \frac{\partial f}{\partial y}$

$x(s) = s \qquad y(s) = 2$

$\frac{dx}{ds} = 1 \qquad \frac{dy}{ds} = 0$

Chain Rule

$$\frac{df}{ds} = \left(\frac{\partial f}{\partial x} \times \frac{dx}{ds}\right) + \left(\frac{\partial f}{\partial y} \times \frac{dy}{ds}\right)$$

$$= \frac{\partial f}{\partial x}(1) + 0$$

$$\frac{df}{ds} = \frac{\partial f}{\partial x}$$

$f(x,y) \to \frac{\partial f}{\partial x}$, $f(x,y) \to \frac{\partial f}{\partial y}$

Example 5.15. Show that for the curve $x(s) = -1$, $y(s) = s$ we get $\dfrac{df}{ds} = \dfrac{\partial f}{\partial y}$.

$\frac{dx}{ds} = 0 \quad \frac{dy}{ds} = 1$

$\frac{df}{ds} = \frac{\partial f}{\partial x}(0) + \frac{\partial f}{\partial y}(1)$

$\frac{df}{ds} = \frac{\partial f}{\partial y}$

The last two examples show that df/ds is somehow tied to the partial derivatives of f. The exact link will be made clear in a short while.

What meaning can we assign to this number df/ds? It helps to imagine that we have drawn a graph of $f(x, y)$ (i.e. as a surface over the xy-plane).

Now draw the curve $(x(s), y(s))$ in the xy-plane and imagine walking along that curve, let's call it C. At each point on C, $f(s)$ is the height of the surface above the xy-plane. If you walk a short distance Δs then the height might change by an amount Δf. The rate at which the height changes with respect to the distance travelled is then $\Delta f/\Delta s$. In the limit of infinitesimal distances we recover df/ds. Thus we can interpret df/ds as measuring the *rate of change* of f along the curve. This is exactly what we would have expected – after all, derivatives measure rates-of-change.

The first example above showed how you could compute df/ds by first reducing f to an explicit function of s. It was also hinted that it is also possible to evaluate df/ds using partial derivatives.

Let's go back to basics. The derivative df/ds could be calculated as

$$\frac{df}{ds} = \lim_{\Delta s \to 0} \frac{f(x(s + \Delta s), y(s + \Delta s)) - f(x(s), y(s))}{\Delta s} + \frac{f(x(s), y(s+\Delta s))}{\Delta s}.$$

$$- \frac{f(x(s), y(s+\Delta s))}{\Delta s}.$$

We will re-write this by adding and subtracting $f(x(s), y(s+\Delta s))$ just before the minus sign. After a little rearranging we get

$$\begin{aligned} \frac{df}{ds} &= \lim_{\Delta s \to 0} \frac{f(x(s + \Delta s), y(s + \Delta s)) - f(x(s), y(s + \Delta s))}{\Delta s} \\ &\quad + \lim_{\Delta s \to 0} \frac{f(x(s), y(s + \Delta s)) - f(x(s), y(s))}{\Delta s} \end{aligned}$$

Now let's look at the first limit. If we introduce $\Delta x = x(s+\Delta s) - x(s)$ then we can write

$$\lim_{\Delta s \to 0} \frac{f(x(s+\Delta s), y(s+\Delta s)) - f(x(s), y(s+\Delta s))}{\Delta s}$$

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

$$= \frac{\partial f}{\partial x} \frac{dx}{ds}.$$

We can write a similar equation for the second limit. Combining the two leads us to

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

This is an extremely useful and important result. It is an example of what is known as the **chain rule** for functions of several variables.

The Chain Rule

Let $f = f(x, y)$ be a differentiable function. If the function is parametrized by $x = x(s)$ and $y = y(s)$ then the *chain rule* for derivatives of f along a path $x = x(s), y = y(s)$ is

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

Now that we have covered this much, it's rather easy to see an important extension of the above result. Suppose the path was obtained by holding some other parameter constant. That is, imagine that the path $x = x(s), y = y(s)$ arose from some more complicated expressions such as $x = x(s, t), y =$

$f(x,y) \longrightarrow f(s)$?? No!

$x(s,t)$ $y(s,t)$

What will we get?

$f(s,t)$

will we still get $\frac{df}{ds}$?
No!.

$\longrightarrow \frac{\partial f}{\partial s}, \frac{\partial f}{\partial t}$

$$\frac{\partial f}{\partial s} = \left(\frac{\partial f}{\partial x} \times \frac{\partial x}{\partial s}\right) + \left(\frac{\partial f}{\partial y} \times \frac{\partial y}{\partial s}\right) \qquad \frac{\partial f}{\partial t} = \left(\frac{\partial f}{\partial x} \times \frac{\partial x}{\partial t}\right) + \left(\frac{\partial f}{\partial y} \times \frac{\partial y}{\partial t}\right)$$

$y(s,t)$ with t held constant. How would our formula for the chain rule change? Not much other than we would have to keep in mind throughout that t is constant. We encountered this issue once before and that led to partial rather than ordinary derivatives. Clearly the same change of notation applies here, and thus we would write

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s}$$

as the first partial derivative of f with respect to s.

Let's see where we are at so far. We are given a function of two variables $f = f(x, y)$ and we are also given two other functions, also of two variables, $x = x(s,t), y = y(s,t)$. Then $\partial f/\partial s$ can be calculated using the above chain rule.

Of course you could also compute $\partial f/\partial s$ directly by substituting $x = x(s,t)$ and $y = y(s,t)$ into $f(x,y)$ before taking the partial derivatives. Both approaches will give you exactly the same answer.

Note that there is nothing special in the choice of symbols, x, y, s or t. You will often find (u, v) used rather than (s, t).

Example 5.16. Given $f = f(x, y)$ and $x = 2s + 3t, y = s - 2t$ compute $\partial f/\partial t$ directly and by way of the chain rule.

$$\frac{\partial x}{\partial s} = 2 \qquad \frac{\partial y}{\partial s} = 1$$

$$\frac{\partial x}{\partial t} = 3 \qquad \frac{\partial y}{\partial t} = -2$$

$f(x,y)$

$\hookrightarrow \frac{\partial f}{\partial x}$

$\hookrightarrow \frac{\partial f}{\partial y}$

$$f(x,y) \longrightarrow f(s,t)$$

$$\frac{\partial f}{\partial t} = \left(\frac{\partial f}{\partial x} \times \frac{\partial x}{\partial t}\right) + \left(\frac{\partial f}{\partial y} \times \frac{\partial y}{\partial t}\right)$$
$$= 3\frac{\partial f}{\partial x} - 2\frac{\partial f}{\partial y}$$

$$\frac{\partial f}{\partial s} = \left(\frac{\partial f}{\partial x} \times \frac{\partial x}{\partial s}\right) + \left(\frac{\partial f}{\partial y} \times \frac{\partial y}{\partial s}\right)$$
$$= 2\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}$$

The Chain Rule : Episode 2

Let $f = f(x,y)$ be a differentiable function. If $x = x(u,v), y = y(u,v)$ then

$$\frac{\partial f}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u}$$

$$\frac{\partial f}{\partial v} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v}$$

5.5 Gradient and Directional Derivative

Given any differentiable function of several variables we can compute each of its first partial derivatives. Let's do something 'out of the square'. We will assemble these partial derivatives as a *vector* which we will denote by ∇f. So for a function $f(x, y)$ of two variables we define

gradient vector for $f(x,y)$

$f(x,y) \to f(s)$

$\frac{df}{ds}$

$$\nabla f = \frac{\partial f}{\partial x}\underset{\sim}{i} + \frac{\partial f}{\partial y}\underset{\sim}{j}$$

The is known as the **gradient** of f and is often pronounced *grad f*.

This may be pretty but what use is it? If we look back at the formula for the chain rule we see that we can write it out as a vector dot-product

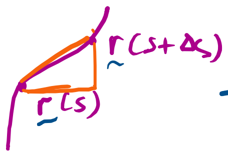
$x(s)$ & $y(s)$ describe a curve

$\underset{\sim}{r} = (x(s), y(s))$

eg $\underset{\sim}{r} = \begin{pmatrix}1\\2\\3\end{pmatrix} + s\begin{pmatrix}2\\1\\3\end{pmatrix} + t\begin{pmatrix}1\\2\\4\end{pmatrix}$

$$\begin{aligned}\frac{df}{ds} &= \frac{\partial f}{\partial x}\frac{dx}{ds} + \frac{\partial f}{\partial y}\frac{dy}{ds} \\ &= \left(\frac{\partial f}{\partial x}\underset{\sim}{i} + \frac{\partial f}{\partial y}\underset{\sim}{j}\right) \cdot \left(\frac{dx}{ds}\underset{\sim}{i} + \frac{dy}{ds}\underset{\sim}{j}\right) \\ &= (\nabla f) \cdot \left(\frac{dx}{ds}\underset{\sim}{i} + \frac{dy}{ds}\underset{\sim}{j}\right)\end{aligned}$$

Claim: $\left(\frac{dx}{ds}, \frac{dy}{ds}\right)$ is tangent to the curve $\underset{\sim}{t}$



$\underset{\sim}{t} = \frac{dr}{ds}$

$\underset{\sim}{t} = \frac{d}{ds}(x(s), y(s))$

$\underset{\sim}{t} = \left(\frac{dx}{ds}, \frac{dy}{ds}\right)$

$= \lim_{\Delta s \to 0}\left[\frac{1}{\Delta s}(\underset{\sim}{r}(s+\Delta s) - \underset{\sim}{r}(s))\right]$

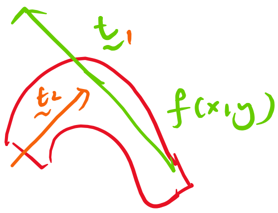
The number that we calculate in this process, i.e. df/ds, is known as the **directional derivative** of f in the direction $\underset{\sim}{t}$. What do we make of the vector on the far right of this equation, i.e. $\frac{dx}{ds}\underset{\sim}{i} + \frac{dy}{ds}\underset{\sim}{j}$? It is not hard to see that it is a *tangent vector* to the curve $(x(s), y(s))$. And if we chose the parameter s to be distance along the curve then we also see that it is a *unit* vector.

Example 5.17. Prove the last pair of statements, i.e. that the vector is a tangent vector and that it is a unit vector.

If s is distance the ratio $\simeq 1$

$|\underset{\sim}{t}| = 1$

Directional derivative



Defn If $|\underset{\sim}{t}| = 1$ and $f(x,y)$

then $\underset{\sim}{t} \cdot \nabla f = \nabla_{\underset{\sim}{t}} f$

is the derivative of f in direction $\underset{\sim}{t}$.

Note: if $\underset{\sim}{t}$ is the tangent vector to $(x(s), y(s))$

then $\nabla_{\underset{\sim}{t}} f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \cdot \left(\frac{dx}{ds}, \frac{dy}{ds}\right)$

$\nabla f \cdot \underset{\sim}{t}$

$\begin{pmatrix}\frac{\partial f}{\partial x}\\ \frac{\partial f}{\partial y}\end{pmatrix} \cdot \begin{pmatrix}\frac{dx}{ds}\\ \frac{dy}{ds}\end{pmatrix} = \left(\frac{\partial f}{\partial x} \cdot \frac{dx}{ds}\right) + \left(\frac{\partial f}{\partial y} \cdot \frac{dy}{ds}\right)$ Chain Rule

It is customary to denote the tangent vector by $\underset{\sim}{t}$ (some people prefer $\underset{\sim}{u}$). With the above definitions we can now write the equation for a **directional derivative** as follows

$$\frac{df}{ds} = \underset{\sim}{t} \cdot \nabla f$$

Yet another variation on the notation is to include the tangent vector as subscript on ∇. Thus we also have

$$\frac{df}{ds} = \nabla_{\underset{\sim}{t}} f$$

Directional derivative

The *directional derivative* df/ds of a function f in the direction $\underset{\sim}{t}$ is given by

$$\frac{df}{ds} = \underset{\sim}{t} \cdot \nabla f = \nabla_{\underset{\sim}{t}} f$$

where the *gradient* ∇f is defined by

$$\nabla f = \frac{\partial f}{\partial x}\underset{\sim}{i} + \frac{\partial f}{\partial y}\underset{\sim}{j}$$

and $\underset{\sim}{t}$ is a unit vector, $\underset{\sim}{t} \cdot \underset{\sim}{t} = 1$.

If given t is not having magnitude 1 then we need find the unit vector $\underset{\sim}{t}$.

$$\hat{\underset{\sim}{t}} = \frac{\underset{\sim}{t}}{|\underset{\sim}{t}|}$$

Example 5.18. Given $f(x, y) = \sin(x)\cos(y)$ compute the directional derivative of f in the direction $\underset{\sim}{t} = (\underset{\sim}{i} + \underset{\sim}{j})/\sqrt{2}$.

$$\frac{df}{ds} = \nabla f \cdot t$$

$$\frac{\partial f}{\partial x} = \cos(x)\cos(y) \qquad \frac{\partial f}{\partial y} = -\sin(x)\sin(y) \qquad \nabla f = \frac{\partial f}{\partial x}\underset{\sim}{i} + \frac{\partial f}{\partial y}\underset{\sim}{j}$$

$$\frac{df}{ds} = \begin{pmatrix} \cos(x)\cos(y) \\ -\sin(x)\sin(y) \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

$$\underset{\sim}{t} = \frac{1}{\sqrt{2}}\underset{\sim}{i} + \frac{1}{\sqrt{2}}\underset{\sim}{j}$$

$$|\underset{\sim}{t}| = \sqrt{\tfrac{1}{2} + \tfrac{1}{2}} = 1 \text{ unit}$$

$$= \frac{1}{\sqrt{2}}\left[\cos(x)\cos(y) - \sin(x)\sin(y)\right]$$

$$\nabla f \begin{pmatrix} 2x \\ 2y \end{pmatrix}$$

$$\underset{\sim}{t} = \frac{dx}{ds}\underset{\sim}{i} + \frac{dy}{ds}\underset{\sim}{j}$$

$$\underset{\sim}{t} = \begin{pmatrix} \cos(0.1) \\ \sin(0.1) \end{pmatrix}$$

$$|\underset{\sim}{t}| = \sqrt{\cos^2(0.1) + \sin^2(0.1)} = 1$$

Example 5.19. Given $\nabla f = 2x\underset{\sim}{i} + 2y\underset{\sim}{j}$ and $x(s) = s\cos(0.1)$, $y(s) = s\sin(0.1)$ compute df/ds at $s = 1$.

$$\frac{df}{ds} = \nabla f \cdot \underset{\sim}{t} = \begin{pmatrix} 2x \\ 2y \end{pmatrix} \begin{pmatrix} \cos(0.1) \\ \sin(0.1) \end{pmatrix}$$

$s = 1$
$x(1) = \cos(0.1)$
$y(1) = \sin(0.1)$

$$= \begin{pmatrix} 2\cos(0.1) \\ 2\sin(0.1) \end{pmatrix} \cdot \begin{pmatrix} \cos(0.1) \\ \sin(0.1) \end{pmatrix}$$

$$= 2\cos^2(0.1) + 2\sin^2(0.1)$$

$$= 2$$

$$\frac{\partial f}{\partial x} = 2xy^2 \qquad \frac{\partial f}{\partial y} = 2x^2y$$

Example 5.20. Given $f(x, y) = (xy)^2$ and the vector $\underset{\sim}{v} = 2\underset{\sim}{i} + 7\underset{\sim}{j}$ compute the directional derivative at $(1, 1)$. Hint: Is $\underset{\sim}{v}$ a unit vector?

unit vector

$$\frac{df}{ds} = \nabla f \cdot \underset{\sim}{t}$$

$$\nabla f = \begin{pmatrix} 2xy^2 \\ 2x^2y \end{pmatrix}$$

$$\underset{\sim}{v} = \sqrt{2^2 + 7^2} = \sqrt{53}$$

$$\underset{\sim}{t} = \hat{\underset{\sim}{v}} = \frac{1}{\sqrt{53}}\begin{pmatrix} 2 \\ 7 \end{pmatrix} = \begin{pmatrix} 2/\sqrt{53} \\ 7/\sqrt{53} \end{pmatrix}$$

$$\frac{df}{ds} = \begin{pmatrix} 2xy^2 \\ 2x^2y \end{pmatrix} \begin{pmatrix} 2/\sqrt{53} \\ 7/\sqrt{53} \end{pmatrix}$$

$$\left.\frac{df}{ds}\right|_{\substack{x=1 \\ y=1}} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 2/\sqrt{53} \\ 7/\sqrt{53} \end{pmatrix} = \frac{18}{\sqrt{53}}$$

We began this discussion by restricting a function of many variables to a function of one variable. We achieved this by choosing a path such as $x = x(s), y = y(s)$. We might ask if the value of df/ds depends on the choice of the path? That is we could imagine many different paths all sharing the one point, call it P, in common. Amongst these different paths might we get different answers for df/ds?

This is a very good question. To answer it let's look at the directional derivative in the form

$$\frac{df}{ds} = \underset{\sim}{t} \cdot \nabla f$$

First we note that ∇f depends only on the values of (x, y) at P. It knows nothing about the curves passing through P. That information is contained solely in the vector $\underset{\sim}{t}$. Thus if a family of curves passing through P share the same $\underset{\sim}{t}$ then we most certainly will get the same value for df/ds for each member of that family. But what class of curves share the same $\underset{\sim}{t}$ at P? Clearly they are all tangent to each other at P. None of the curves cross any other curve at P.

At this point we can dispense with the curves and retain just the tangent vector $\underset{\sim}{t}$ at P. All that we require to compute df/ds is the direction we wish to head in, $\underset{\sim}{t}$, and the gradient vector, ∇f, at P. Choose a different $\underset{\sim}{t}$ and you will get a different answer for df/ds. In each case df/ds measures how rapidly f is changing the direction of $\underset{\sim}{t}$.

$f(x,y)$

1st partial differentiation

$\frac{\partial f}{\partial x}$ / f_x

$\frac{\partial f}{\partial y}$ / f_y

2nd partial derivative

$\frac{\partial f}{\partial x} \rightarrow \frac{\partial^2 f}{\partial x^2}$ (f_{xx})

$\frac{\partial f}{\partial x} \rightarrow \frac{\partial^2 f}{\partial x \partial y}$ (f_{xy})

$\frac{\partial f}{\partial y} \rightarrow \frac{\partial^2 f}{\partial y^2}$ (f_{yy})

$\frac{\partial^2 f}{\partial y \partial x}$ (f_{yx})

same

5.6 Second order partial derivatives

The result of a partial derivative of a function yields another function of one or more variables. We are thus at liberty to take another derivative, generating yet another function. Clearly we can repeat this any number of times (though possibly subject to some technical limitations as noted below, see Exceptions).

Example 5.21. Let $f(x,y) = \sin(x)\sin(y)$. Then we can define $g(x,y) = \dfrac{\partial f}{\partial x}$ and $h(x,y) = \dfrac{\partial g}{\partial x}$.

That is

$\frac{\partial f}{\partial x}$ 1st order

$$g(x,y) = \frac{\partial f}{\partial x} = \frac{\partial\Big(\sin(x)\sin(y)\Big)}{\partial x} = \cos(x)\sin(y)$$

and

$\frac{\partial^2 f}{\partial x^2}$ 2nd order

$$h(x,y) = \frac{\partial g}{\partial x} = \frac{\partial\Big(\cos(x)\sin(y)\Big)}{\partial x} = -\sin(x)\sin(y)$$

Example 5.22. Compute $\dfrac{\partial g}{\partial y}$ for the above example.

$\frac{\partial f}{\partial y} = \sin(x)\cos(y)$

$\frac{\partial g}{\partial y} = \frac{\partial^2 f}{\partial y^2} = -\sin(x)\sin(y)$

From this we see that $h(x,y)$ was computed as follows

$$h(x,y) = \frac{\partial g}{\partial x} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)$$

This is often written as

$$h(x,y) = \frac{\partial^2 f}{\partial x^2}$$

and is known as a **second order partial derivative** of the function $f(x,y)$.

Now consider the case where we compute $h(x,y)$ by first taking a partial derivative in x then followed by a partial derivative in y, that is

$$h(x,y) = \frac{\partial g}{\partial y} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$$

$f(x,y)$ $\qquad \dfrac{\partial^3 f}{\partial x \partial x \partial y} = \dfrac{\partial^3 f}{\partial x \partial y \partial x}$

and this is normally written as

$$h(x,y) = \frac{\partial^2 f}{\partial y \partial x}$$

Note the order on the bottom line – you should read this from right to left. It tells you that to take a partial derivative in x then a partial derivative in y.

The function $z = f(x,y)$ has *two* partial derivatives f_x and f_y. Taking partial derivatives of f_x and f_y yields *four* second order partial derivatives of the function $f(x,y)$.

It's now a short leap to cases where we might try to find, say, the *fifth* partial derivatives, such as

$$P(x,y) = \frac{\partial^5 Q}{\partial x \partial y \partial y \partial x \partial x}$$

Partial derivatives that involve one or more of the independent variables are known as *mixed partial derivatives*.

Example 5.23. Given $f(x,y) = 3x^2 + 2xy$ compute $\dfrac{\partial^2 f}{\partial x \partial y}$ and $\dfrac{\partial^2 f}{\partial y \partial x}$. What do you notice?

$\dfrac{\partial f}{\partial x} = 6x + 2y$ $\qquad \dfrac{\partial f}{\partial y} = 2x$

$\dfrac{\partial^2 f}{\partial x \partial y} = 2$ $\qquad \dfrac{\partial f}{\partial y \partial x} = 2$

> **Order of partial derivatives does not matter: Clairaut's Theorem**
>
> If $f(x,y)$ is a twice-differentiable function whose second order mixed partial derivatives are continuous, then the order in which its mixed partial derivatives are calculated does not matter. Each ordering will yield the same function. For a function of two variables this means
>
> $$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$$

This is not immediately obvious but it can be proved and it is a very useful result.

A quick word on notation: The second order partial derivatives can be written as follows:

$$\frac{\partial^2 f}{\partial x^2} = f_{xx}$$
$$\frac{\partial^2 f}{\partial y^2} = f_{yy}$$
$$\frac{\partial^2 f}{\partial x \partial y} = f_{xy}$$

Example 5.24. Use the above theorem to show that

$$P(x,y) = \frac{\partial^5 Q}{\partial x \partial y \partial y \partial x \partial x} = \frac{\partial^5 Q}{\partial y \partial y \partial x \partial x \partial x} = \frac{\partial^5 Q}{\partial x \partial x \partial x \partial y \partial y}$$

The theorem allows us to simplify our notation, all we need do is record how many of each type of partial derivative are required, thus the above can be written as

$$P(x,y) = \frac{\partial^5 Q}{\partial x^3 \partial y^2} = \frac{\partial^5 Q}{\partial y^2 \partial x^3}$$

Example 5.25. Show that the function $u(x,y) = e^{-x}\cos y$ is a solution of Laplace's equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

Solution:

$$u_x = \frac{\partial}{\partial x}(e^{-x}\cos y) = -e^{-x}\cos y \text{ and } u_y = \frac{\partial}{\partial y}(e^{-x}\cos y) = -e^{-x}\sin y$$

Then

$$u_{xx} = \frac{\partial}{\partial x}\big(-e^{-x}\cos y\big) = e^{-x}\cos y$$

and

$$u_{yy} = \frac{\partial}{\partial y}\big(-e^{-x}\sin y\big) = -e^{-x}\cos y$$

Hence

$$u_{xx} + u_{yy} = e^{-x}\cos y + (-e^{-x}\cos y) = 0$$

5.6.1 Taylor polynomials of higher degree

In earlier lectures we discovered that the **linear approximation** function $T_1(x, y)$ to a function of two variables $f(x, y)$ near the point (a, b) is the *same* as the equation of the **tangent plane** at the point (a, b). In other words, for (x, y) near (a, b) we have

$$f(x,y) \approx T_1(x,y) = f(a,b) + f_x(a,b)\cdot(x-a) + f_y(a,b)\cdot(y-b) \tag{5.3}$$

The function $T_1(x, y)$ is also known as the **Taylor polynomial of degree one** for $f(x, y)$ near (a, b), and clearly uses *first* partial derivatives of $f(x, y)$. Now the tangent plane provides a good fit to $f(x, y)$ only if (x, y) are sufficiently close to (a, b). But this is obviously not always going to be the case. If we want to obtain a more accurate polynomial approximation to the graph of the surface $z = f(x, y)$, we need to take into account the local **curvature** of the surface at (a, b). This is done by including the *second* partial derivatives of $f(x, y)$, namely f_{xx}, f_{xy} and f_{yy}. Using these gives us $T_2(x, y)$ which is the **Taylor polynomial of degree two**. $T_2(x, y)$ is also known as the **quadratic approximation** function:

$$T_2(x,y) = f(a,b) + f_x(a,b)\cdot(x-a) + f_y(a,b)\cdot(y-b) \leftarrow \text{linear approximation}$$
$$+\frac{1}{2!}\left[f_{xx}(a,b)(x-a)^2 + 2f_{xy}(a,b)(x-a)(y-b) + f_{yy}(a,b)(y-b)^2\right] \tag{5.4}$$

f_{xy}
f_{yx} $\updownarrow$ same

$T_1(x) \rightarrow$ linear approx. for $f(x)$

$T_2(x) =$ Quadratic approx for $f(x)$

Example 5.26. Derive the Taylor polynomial of degree two for the function $f(x, y) = e^{-x}\cos y$ near the point $(a, b) = (0, 0)$.

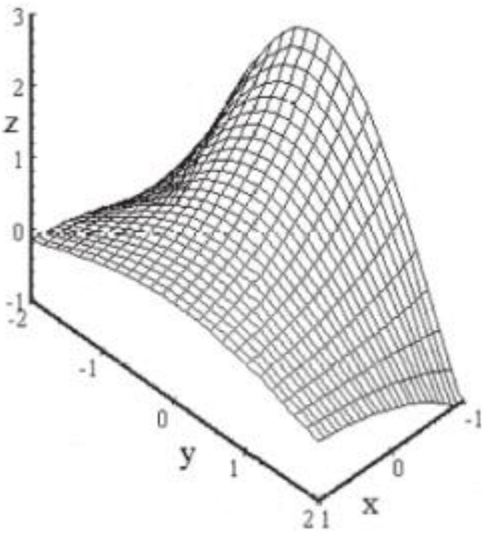
Solution:

Function	Value at $(0,0)$ point $(0,0,1)$
$f(x,y) = e^{-x}\cos y$	$f(0,0) = e^{-0}\cos 0 = 1\cdot 1 = 1$
$f_x(x,y) = \frac{\partial f}{\partial x} = \frac{\partial}{\partial x}\left(e^{-x}\cos y\right) = -e^{-x}\cos y$	$f_x(0,0) = -e^{-0}\cos 0 = -1$
$f_y(x,y) = \frac{\partial f}{\partial y} = \frac{\partial}{\partial y}\left(e^{-x}\cos y\right) = -e^{-x}\sin y$	$f_y(0,0) = -e^{-0}\sin 0 = 0$
$f_{xx}(x,y) = \frac{\partial}{\partial x}(f_x) = \frac{\partial}{\partial x}\left(-e^{-x}\cos y\right) = e^{-x}\cos y$	$f_{xx}(0,0) = e^{-0}\cos 0 = 1$
$f_{yx}(x,y) = f_{xy}(x,y) = \frac{\partial}{\partial y}(f_x) = \frac{\partial}{\partial y}\left(-e^{-x}\cos y\right) = e^{-x}\sin y$	$f_{xy}(0,0) = e^{-0}\sin 0 = 0$
$f_{yy}(x,y) = \frac{\partial}{\partial y}(f_y) = \frac{\partial}{\partial y}\left(-e^{-x}\sin y\right) = -e^{-x}\cos y$	$f_{yy}(0,0) = -e^{-0}\cos 0 = -1$

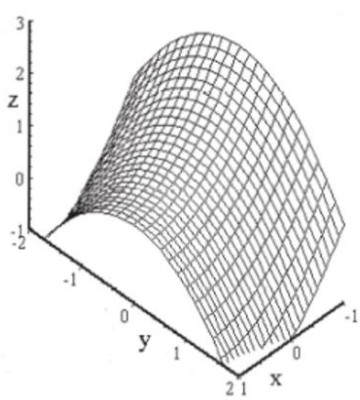
Collecting terms and substituting them into equation (5.4) we obtain: $(0,0,1)$

$$T_2(x,y) = f(0,0) + f_x(0,0)\cdot(x-0) + f_y(0,0)\cdot(y-0)$$
$$+\frac{1}{2}f_{xx}(0,0)\cdot(x-0)^2 + f_{xy}(0,0)\cdot(x-0)(y-0) + \frac{1}{2}f_{yy}(0,0)\cdot(y-0)^2$$
$$= 1 - 1\cdot x + 0\cdot y + \tfrac{1}{2}\cdot 1\cdot x^2 + 0\cdot xy - \tfrac{1}{2}y^2$$
$$e^{-x}\cos y \approx = 1 - x + \tfrac{1}{2}(x^2 - y^2)$$

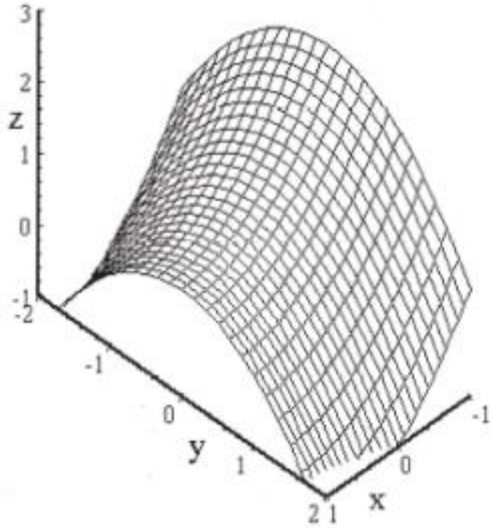
Lastly we can graph the surface $z = f(x,y) = e^{-x}\cos y$ and the quadratic approximation function $T_2(x,y)$.



Graph of $z = e^{-x}\cos y$



Graph of $z = 1 - x + \frac{1}{2}(x^2 - y^2)$



Graph of $z = 1 - x + \frac{1}{2}(x^2 - y^2)$

Looking at these two graphs we see that the Taylor polynomial of degree two, namely $T_2(x, y)$, does a good job in mimicking the shape of the surface $z = f(x, y)$ for points (x, y) **close** to $(0, 0)$. In the plane $x = 1$, the quadratic approximation $z = T_2(x, y)$ falls off too steeply along the y axis as we move away from $y = 0$, while in the plane $x = -1$ it falls away too slowly along the y axis as we move away from $y = 0$.

5.6.2 Exceptions: when derivatives do not exist

In earlier lectures we noted that at the very least a function must be continuous if it is to have a meaningful derivative. When we take successive derivatives we may need to revisit the question of continuity for each new function that we create.

If a function fails to be continuous at some point then we most certainly can not take its derivative at that point.

Example 5.27. Consider the function

$$f(x) = \begin{cases} 0, & -\infty < x < 0, \\ \\ 3x^2, & 0 < x < \infty. \end{cases}$$

It is easy to see that something interesting might happen at $x = 0$. It's also not hard to see that the function is continuous over its whole domain, and thus we can compute its derivative everywhere, leading to

$$\frac{df(x)}{dx} = \begin{cases} 0, & -\infty < x < 0, \\ \\ 6x, & 0 < x < \infty. \end{cases}$$

This too is continuous and we thus attempt to compute its derivative,

$$\frac{d^2 f(x)}{dx^2} = \begin{cases} 0, & -\infty < x < 0, \\ \\ 6, & 0 < x < \infty. \end{cases}$$

Now we notice that this second derivative is not continuous at $x = 0$. We thus can not take any more derivatives at $x = 0$. Our chain of differentiation has come to an end.

We began with a continuous function $f(x)$ and we were able to compute only its first two derivatives over the domain $x \in R$. However, as we noted, its second derivative was not continuous at $x = 0$. We call such a function a C^1 function, meaning that the function has a first derivative which is *also* continuous. The symbol C reminds us that we are talking about continuity and the superscript 1 tells us how many derivatives we can apply before we encounter a non-continuous function. The clause 'over R' just reminds us that the domain of the function is the set of real numbers $(-\infty, \infty)$. Despite the function being twice differentiable, we would not call it a C^2 function, since the second derivative is not continuous.

We should always keep in mind that a function may only posses a finite number of derivatives before we encounter a discontinuity. The tell-tale signs to watch out for are sharp edges, holes or singularities in the graph of the function.

Given.
Recall $y = f(x)$
stationary point $\rightarrow \frac{dy}{dx} = 0$
solve (critical value) x-value
$x = a$
$y = f(a)$

5.7 Stationary points

5.7.1 Finding stationary points

Suppose you run a commercial business and that by some means you have constructed the following formula for the profit of one of your lines of business

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

$z = f(x,y)$
stationary point still occurs at $\frac{\partial f}{\partial x} = 0$ and $\frac{\partial f}{\partial y} = 0$

Clearly the profit f depends on two variables x and y. Sound business practice suggest that you would like to maximise your profits. In mathematical terms this means find the values of (x, y) such that f is a maximum. A simple plot of the graph of f shows us that the maximum occurs at $(0, 0)$ (corresponding to a maximum profit of 4 units). We may not be able to do this so easily for other functions, and thus we need some systematic way of computing the points (x, y) at which f is maximised.

You have seen similar problems for the case of a function of one variable. And from that you may expect that for the present problem we will be making a statement about the *derivatives* of f in order that we have a maximum (i.e. that the derivatives should be zero). Let's make this precise.

Let's denote the (as yet unknown) point at which the function is a maximum by P. Now if we have a maximum at this point, then moving in any direction from this point should see the function decrease. That is the *directional derivative* must be non-positive in every direction from P. In other words we must have

$$\frac{df}{ds} = \underset{\sim}{t} \cdot (\nabla f)_p \leq 0$$

for every choice of $\underset{\sim}{t}$. Let us assume (for the moment) that $(\nabla f)_p \neq 0$ then we should be able to compute $\lambda > 0$ so that $\underset{\sim}{t} = \lambda (\nabla f)_p$ is a unit vector. If you now substitute this into the above you will find

Determine Nature stationary pt-
$y = f(x) \rightarrow$ 1st Derivative Test
2nd Derivative Test

$z = f(x,y)$
Determine nature of stationary point
$\rightarrow$ 2nd Derivative Test

$$\lambda (\nabla f)_p \cdot (\nabla f)_p \le 0$$

Look carefully at the left hand side. Each term is positive (remember $\underset{\sim}{a} \cdot \underset{\sim}{a}$ is the *squared* length of a vector $\underset{\sim}{a}$) yet the right hand side is either zero or negative. Thus this equation does not make sense and we have to reject our only assumption, that $(\nabla f)_p \neq 0$.

We have thus found that if f is to have a maximum at P then we must have

$$0 = (\nabla f)_p$$

This is a vector equation and thus each component of ∇f is zero at P, that is

$$0 = \frac{\partial f}{\partial x}, \quad \text{and} \quad 0 = \frac{\partial f}{\partial y} \quad \text{at } P$$

It is from these equations that we would then compute the (x, y) coordinates of P.

Of course we could have posed the related question of finding the points at which a function is minimised. The mathematics would be much the same except for a change in words (maximum to minimum) and a corresponding change in $\pm$ signs. The end result is the same though, the gradient ∇f must vanish at P.

Example 5.28. Find the points at which $f = 4 - x^2 - y^2$ attains its maximum.

$f_x = -2x$ $\qquad$ $f_y = -2y$

For stationary point, $f_x = 0$ $\qquad$ and $\qquad$ $f_y = 0$

$-2x = 0$ $\qquad$ $-2y = 0$

$x = 0$ $\qquad$ $y = 0$

point $\Rightarrow$ $f(x, y) = 4$

$(0, 0, 4) \rightarrow$ Maximum point

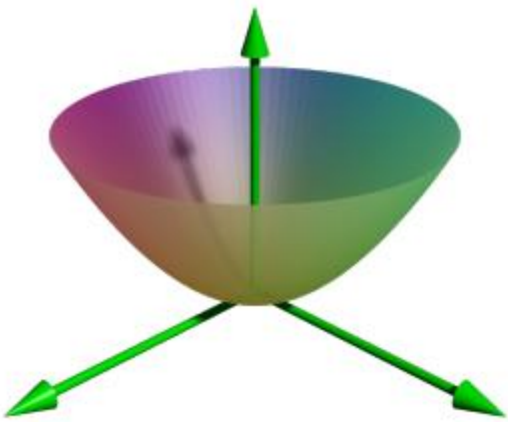
Recall that *stationary points* were found in functions of one variable by setting the derivative to zero and solving for x. We either found a local maximum, a local minimum, or an inflection point (for example $f(x) = x^3$ at $x = 0$). As we have just seen, for functions of two variables, the stationary points are found similarly, and we obtain the following types:

- A local minimum
- A local maximum
- A saddle point

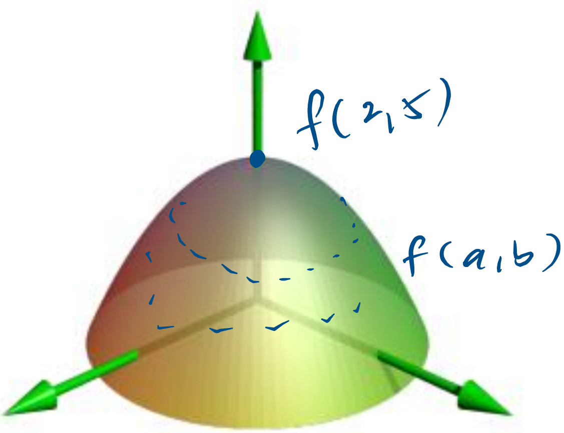
When we solve the equations

$$0 = (\nabla f)_p$$

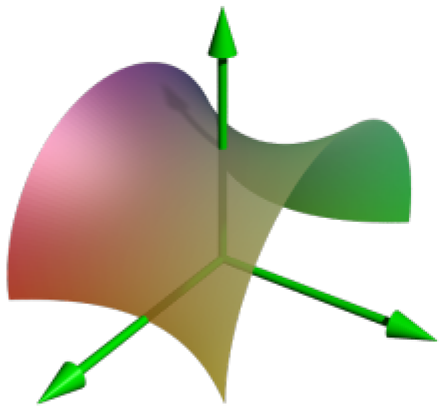
we might get more than one point P. What do we make of these points? Some of them might correspond to minimums while others might correspond to maximums of f, and others still may correspond to saddle points. The three options are shown in the following graphs.



A typical local minimum



A typical local maximum



A typical saddle point

A typical case might consist of any number of points like the above.

5.7.2 Notation

Rather than continually having to qualify the point as corresponding to a local minimum or local maximum of f we commonly lump these into the one term *local extrema*. Please note that although saddle points are a type of stationary point, we do not call them extrema (there is nothing 'extreme' going on with a saddle point!).

Note when we talk of minima, maxima and extrema we are talking about the (x, y) points at which the function has a local minimum, maximum or extremum respectively.[1]

5.7.3 Minima, Maxima or Saddle point

We have just seen that a function of two variables has stationary points when $\dfrac{\partial f}{\partial x} = \dfrac{\partial f}{\partial y} = 0$. If (a, b) are the (x, y) coordinates of a stationary point for $f(x, y)$, then we can say

- A local **maximum** occurs when $f(x, y) \leq f(a, b)$ for all (x, y) close to (a, b)

[1] In fact, like the one variable case, two variable functions can possess points of singularity, i.e., where the partial derivatives do not exist, and these can correspond to local extrema as well. As you'd expect, the collection of all stationary points and singularity points are called the critical points. However, we will not encounter two variable functions with points of singularity in this course!

- A local **minimum** occurs when $f(x, y) \geq f(a, b)$ for all (x, y) close to (a, b)
- A **saddle point** if it is neither a maximum or minimum

There is another way we can classify the stationary points of a function of several variables. You should recall that for a function of one variable, $f = f(x)$, that its extrema could be characterised simply by evaluating the *sign of the second derivative*. In other words, for $y = f(x)$, the extrema are where

$$\frac{df}{dx} = 0$$

Then for these values of x (where $\frac{df}{dx} = 0$) we examine the second derivative. If

- $\frac{d^2f}{dx^2} > 0$ then this corresponds to a **local minima**
- $\frac{d^2f}{dx^2} < 0$ then this corresponds to a **local maxima**
- $\frac{d^2f}{dx^2} = 0$ then no decision can be made (e.g. x^3 or x^4).

Now we want to take this idea to functions of several variables. Can we do this? Yes, but with some modifications. Without going into the details (these are covered in a different course) we state the following test.

Characterising stationary points - second derivative test

If $0 = \nabla f$ at a point P then, at P compute

$$D = \frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2$$

Using D we can now classify the stationary points P.

A **local minima** when	$D > 0$	and	$\frac{\partial^2 f}{\partial x^2} > 0$
A **local maxima** when	$D > 0$	and	$\frac{\partial^2 f}{\partial x^2} < 0$
A **Saddle point** when	$D < 0$		
Inconclusive when	$D = 0$		

Example 5.29. Classify the stationary points of $f(x, y) = x^2 + y^2 - 2x - 6y + 14$.

$f_x = 2x - 2$ $\quad$ $f_y = 2y - 6$

$f_{xx} = 2$ $\quad$ $f_{yy} = 2$ $\quad$ $f_{xy} = 0$

critical point

$f_x = 0$ $\quad$ $f_y = 0$

$2x - 2 = 0$ $\quad$ $2y - 6 = 0$

$x = 1$ $\quad$ $y = 3$

$f(1,3) = 1 + 9 - 2 - 18 + 14 = 4$

$(1, 3, 4)$

$D = f_{xx} f_{yy} - (f_{xy})^2$ $\quad$ $f_{xx} = 2$

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

$D > 0$, $f_{xx} > 0$

$\Rightarrow$ local minumum at $(1, 3, 4)$

Example 5.30. Classify the stationary points of $f(x, y) = y^2 - x^2$.

$f_x = -2x$ $\quad$ $f_y = 2y$

$f_{xx} = -2$ $\quad$ $f_{yy} = 2$ $\quad$ $f_{xy} = 0$

Statory point

$f_x = 0$ $\quad$ $f_y = 0$

$-2x = 0$ $\quad$ $2y = 0$

$x = 0$ $\quad$ $y = 0$

$f(0,0) = 0$ $\quad$ $(0, 0, 0)$

$D = f_{xx} f_{yy} - (f_{xy})^2$

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

$D < 0 \Rightarrow$ saddle point at $(0, 0, 0)$

5.7.4 Application of extrema

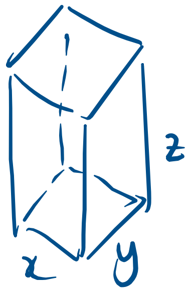
As a final note, we will now turn to some applications of the use of extrema.

Example 5.31. We are required to build a rectangular box with a volume of 12 cubic centimetres. Since we are trying to economise on building costs, we also require the box to be made out of the smallest amount of material. What are the dimensions of the box that will satisfy these requirements?

surface area.

Solution: First we need to set up our equations. Let the dimensions of the box be x, y and z for the length, width and height respectively. The volume of the box is given by

$$V = xyz = 12$$



The total surface area of the box is

$$A = 2xy + 2xz + 2yz$$

Rearranging the equation for the volume will give

$$z = \frac{12}{xy}$$

Substituting this into our equation for the surface area will give

$$A = 2xy + 2x\frac{12}{xy} + 2y\frac{12}{xy} = \boxed{2xy + \frac{24}{y} + \frac{24}{x}}$$

Now we have a function $A = f(x, y)$ which we can minimise. Taking partial derivatives of $A = f(x, y)$ with respect to x and y gives

$$\frac{\partial A}{\partial x} = 2y - \frac{24}{x^2}$$

and

$$\frac{\partial A}{\partial y} = 2x - \frac{24}{y^2}$$

Now let $\frac{\partial A}{\partial x} = 0$ and $\frac{\partial A}{\partial y} = 0$ and solve for both x and y.

$$2y - \frac{24}{x^2} = 0$$

$$y = \frac{12}{x^2}$$

substitution

$$2x - \frac{24}{y^2} = 0$$

$$2xy^2 = 24$$

$$2x\left(\frac{144}{x^4}\right) = 24$$

$$x^3 = 12$$

$$x = \sqrt[3]{12}$$

$$y = \frac{12}{(\sqrt[3]{12})^2}$$

$$y = \sqrt[3]{12}$$

$$z = \frac{12}{\sqrt[3]{12}\,\sqrt[3]{12}}$$

$$z = \sqrt[3]{12}$$

objective lowest cost $\Rightarrow$ least material (minimum)

Thus the dimensions of the box are $x = \sqrt[3]{12}$, $y = \sqrt[3]{12}$ and $z = \sqrt[3]{12}$.

The second partial derivatives (for the above values of x and y) are

$x = \sqrt[3]{12}$

$y = \sqrt[3]{12}$

$\frac{\partial A}{\partial x} = 2y - \frac{24}{x^2}$

$$\frac{\partial^2 A}{\partial x^2} = \frac{48}{y^3}$$

$$D = \left(\frac{48}{(\sqrt[3]{12})^3}\right)\left(\frac{48}{(\sqrt[3]{12})^3}\right) - (2)^2$$

$\frac{\partial A}{\partial y} = 2x - \frac{24}{y^2}$

$$\frac{\partial^2 A}{\partial y^2} = \frac{48}{x^3}$$

$$\frac{\partial^2 A}{\partial x \partial y} = 2$$

$$D = 8 > 0$$

$$\left.\frac{\partial^2 A}{\partial x^2}\right|_{x=\sqrt[3]{12}} = 4 > 0$$

Using the Second Derivative Test,

$$D = \frac{\partial^2 A}{\partial x^2}\frac{\partial^2 A}{\partial y^2} - \left(\frac{\partial^2 A}{\partial x \partial y}\right)^2$$

Since $D =$ 8 the dimensions of the box of volume $12cm^3$ that uses the *minimum* amount of material are $x = \sqrt[3]{12}$ $y = \sqrt[3]{12}$ $z = \sqrt[3]{12}$.

Earlier we looked at methods of finding the shortest distance from a point to a plane. We can now use extrema to answer questions such as these.

Example 5.32. A plane has the equation $2x + 3y + z = 12$. Find the point on the plane closest to the origin.

Solution: To answer this question we clearly want to *minimise* the distance from the origin to the point on the plane. Let (x, y, z) be the point on the plane. The distance between two points is given by

$$d = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}$$

Note that if we let $G = d^2$ and minimise G this is the same as minimising d. Since $(x_0, y_0, z_0) = (0, 0, 0)$ we can now write

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

Using the equation of the plane, where $z = 12 - 2x - 3y$ we now have

$$G = x^2 + y^2 + (12 - 2x - 3y)^2$$

We now take partial derivatives of G with respect to x and y.

$$\frac{\partial G}{\partial x} = 2x + 2(12-2x-3y)(-2)$$
$$= 10x + 12y - 48$$

and

$$\frac{\partial G}{\partial y} = 2y + 2(12-2x-3y)(-3)$$
$$= 12x + 20y - 72$$

Let both $\frac{\partial G}{\partial x} = 0$ and $\frac{\partial G}{\partial y} = 0$ and solve for x and y (simultaneous equations is handy here).

$10x + 12y - 48 = 0$ | $12x + 20y - 72 = 0$

$(\div 2) \Rightarrow 5x + 6y = 24$ — 1 | $(\div 4)$ $3x + 5y = 18$ — 2

Solve 1 & 2 $\Rightarrow$ 1 $\times 3$ − 2 $\times 5$ $\Rightarrow -7y = -18$

$y = \frac{18}{7}$, $x = \frac{12}{7}$

$z = -\frac{1}{7}$

Thus $x =$ and $y =$. We can substitute these into the equation of the plane to find $z =$.

point $\left(\frac{18}{7}, \frac{12}{7}, -\frac{1}{7}\right)$

The second partial derivatives, using the above values for x and y, are

$$\frac{\partial^2 G}{\partial x^2} = 10$$

$$\frac{\partial^2 G}{\partial y^2} = 12$$

and

$$\frac{\partial^2 G}{\partial x \partial y} = 12$$

Using the Second Derivative Test,

$$D = \frac{\partial^2 G}{\partial x^2}\frac{\partial^2 G}{\partial y^2} - \left(\frac{\partial^2 G}{\partial x \partial y}\right)^2 = (10)(12) - (12)^2 = 56 > 0$$

$$\frac{\partial^2 G}{\partial x^2} = 10 > 0 \quad \Rightarrow \text{minimum}$$

Thus the point on the plane that gives the *minimum* distance to the origin is $(x, y, z) = (\quad)$.

point $\left(\frac{18}{7}, \frac{12}{7}, -\frac{1}{7}\right)$ gave mini distance from $(0,0,0)$

Example 5.33. You are given three positive numbers. The *product* of the three numbers is P. The *sum* of the three numbers is 10.

(a) Find the three numbers that will give a *maximum* product?
(b) Show that this gives a maximum product.
(c) If it was required that the three numbers be *whole* numbers, can you find the three (non-zero)numbers that sum to 10 and give a maximum product? Does your answer to (a) help you find this?

Let 3 positive number be x, y, z

$P = xyz \qquad x + y + z = 10$

$z = 10 - x - y$

subs $z = 10 - x - y$ into P

$$P = xy(10 - x - y) = 10xy - x^2y - xy^2$$

objective: max product

$P_x = 10y - 2xy - y^2$ $\qquad$ $P_y = 10x - x^2 - 2xy$

$P_x = 0$ $\qquad$ and $P_y = 0$

$y(10 - 2x - y) = 0$ $\qquad$ $x(10 - x - 2y) = 0$

$y = 0$ (rejected) $\qquad$ $10 - 2x - y = 0$ —①

$x = 0$ (rejected) $\qquad$ $10 - x - 2y = 0$ —②

From ① $y = 10 - 2x$

subs into ② $\quad 10 - x - 2(10 - 2x) = 0$

$$10 - x - 20 + 4x = 0$$

$$3x = 10$$

$$x = \frac{10}{3}$$

$$y = 10 - 2\left(\frac{10}{3}\right) = \frac{10}{3}$$

$$z = 10 - x - y$$

$$z = 10 - \frac{10}{3} - \frac{10}{3} = \frac{10}{3}$$

$x = y = z = \frac{10}{3}$

(b) $P_{xx} = -2y$ $\qquad$ $P_{yy} = -2x$ $\qquad$ $P_{xy} = 10 - 2x - 2y$

$x = y = z = \frac{10}{3}$

$P_{xx} = -\frac{20}{3}$ $\qquad$ $P_{yy} = -\frac{20}{3}$ $\qquad$ $P_{xy} = -\frac{10}{3}$

$$D = \left(-\frac{20}{3}\right)\left(-\frac{20}{3}\right) - \left(-\frac{10}{3}\right)^2 = \frac{100}{3} > 0$$

$P_{xx} = -\frac{20}{3} < 0$

$\Rightarrow$ Maximum product!

(c) Whole number

Whole number

	(a)	Case I	Case II	Case III
x	$\frac{10}{3}$	3	2	8
y	$\frac{10}{3}$	3	3	1
z	$\frac{10}{3}$	4	5	1
$P=xyz$	$\frac{1000}{27}$ ≈ 37.03	36	30	8

✓