PARTIAL DIFFERENTIATION

M. J. Grannell. E .J. Halton.
D. A. Parker. [1]
Department of Mathematics and Statistics,
The University of Central Lancashire.

Second Edition, March 1994.
(Reprinted 2022)

[1]The authors were supported by Enterprise funding.

Contents

1 Preamble

1.1 About this package

This package is for people who need to differentiate functions of more than one variable and apply the results to a variety of problems. It doesn't contain a lot of theory. It isn't really designed for pure mathematicians who require a course discussing conditions for the existence of partial derivatives.

You will find that you need a background knowledge of differentiation of a function of one variable in order to get the most out of this package. In particular, you need to be able to differentiate using the product, quotient and function-of-a-function rules. You will also find it helpful to be able to analyse stationary points of a function of one variable and have some familiarity with Taylor series for functions of one variable. If you are a bit rusty, don't worry - but it would be sensible to do some revision either at the start or as the need arises. Reasonable revision texts are given in the bibliography (Section 13).

If you complete the whole package you should be able to

- understand what is meant by the **partial derivatives** of a function of several variables,
- obtain the **first order partial derivatives** of a function of several variables,
- interpret **geometrically** the first order partial derivatives of a function of two variables,
- obtain **higher derivatives** of a function of several variables,
- understand the **small increments formula** for functions of several variables,
- apply the small increments formula to **approximations and errors**,
- use the **chain rule for partial derivatives** to relate the derivatives of a function with respect to transformed variables to its partial derivatives with respect to its original variables,
- carry out **implicit differentiation** by the methods of partial differentiation,
- obtain the **Taylor series** for functions of two variables,

- determine the **stationary points** of a function of two variables and investigate their nature,
- apply the **method of Lagrange multipliers** to constrained maxima and minima problems,
- determine the **regression line** for a set of data points,
- fit **simple polynomial curves** to a set of data points.

Depending on your own programme of study you may not need to cover everything in this package. Your tutor will advise you what, if anything, can be omitted.

1.2 How to use this package

You MUST do examples! Doing lots of examples for yourself is generally the most effective way of learning the contents of this package and covering the objectives listed above. We recommend that you

- first read the theory - make your own notes where appropriate,
- then work through the worked examples - compare your solutions with the ones in the notes,
- finally do similar examples yourself in a workbook.

The original printing of these notes leaves every other page blank. Use the spare space for your own comments, notes and solutions. You will see certain symbols appearing in the right hand margin from time to time:

◯ denotes the end of a worked example,

□ denotes the end of a proof,

V denotes a reference to videos (see below for details),

EX highlights a point in the notes where you should try examples.

By the time you have reached a package like this one you will probably have realised that learning mathematics rarely goes smoothly! When you get stuck, use your accumulated wisdom and cunning to get around the problem. You might try:

- re-reading the theory/worked examples,

- putting it down and coming back to it later,
- reading ahead to see if subsequent material sheds any light,
- talking to a fellow student,
- looking in a textbook (see the bibliography),
- watching the appropriate video (see the video summaries),
- raising the problem at a tutorial.

1.3 Videos, tutorials and self-help

The videos cover the main points in the notes. The areas covered are indicated in the notes, usually at the ends of sections and subsections. To resolve a particular difficulty you may not need to watch a whole video (they are each about 40 minutes long). They are broken up into sections prefaced with titles which can be read on fast scan. In addition, a summary of the videos associated with this package appears as an appendix to these notes.

Your tutor will tell you about the arrangements for viewing the videos. Try the worked examples **before** watching the solution unfold on the screen. Make notes of any points you cannot follow so that you can explain the difficulty in a subsequent tutorial session. If you are viewing a video individually, remember the rewind button! Unlike a lecture you can get instant and 100 percent accurate replay of what was said.

Your tutor will tell you about tutorial arrangements. These may be related to assessment arrangements. If attendance at tutorials is compulsory then make sure you know the details! The tutorials provide you with individual contact with a tutor. Use this time wisely - staff time is the most expensive of all our resources.

You should come to tutorials in a prepared state. This means that you should have read the notes and the worked examples. You should have tried appropriate examples for yourself. If you have had difficulty with a particular section then you should watch the corresponding video. If your tutor finds that you haven't done these things then s/he may refuse to help you. Your tutor will find it easier to assist you if you can make any queries as specific as possible.

Your fellow students are an excellent form of self-help. Discuss problems with one another and compare solutions. Just be careful that

1. any assessed coursework submitted by you is yours alone,

2. you yourself do really understand solutions worked out jointly with colleagues.

Familiarize yourself with the layout and contents of these notes; scan them before reading them more carefully. The contents page will help you find your way about - use it. The bibliography will point you to textbooks covering the same material as these notes.

When you graduate, your future employer will be just as interested in your capacity for learning as in what you already know. If you can learn mathematics from this package and from textbooks then you will not only have learnt a particular mathematical topic. You will also (and more importantly) have learnt **how to learn** mathematics.

2 Introduction

Many of the most commonly occurring functions in practice, depend on more than one independent variable. For example the volume of a cylinder of radius r and height h is $V = \pi r^2 h$; the distance of a point, (x, y, z) from the origin is $d = \sqrt{x^2 + y^2 + z^2}$; the combined resistance of two resistors, R_1 and R_2, placed in parallel is $R = \dfrac{R_1 R_2}{R_1 + R_2}$. The ideas of calculus, and in particular differentiation and integration need to be extended to deal with functions of several independent variables. In this chapter we concentrate on the differentiation of functions of more than one variable.

We look first at the geometrical representation of a function of two independent variables. Let u be a function of x and y, two independent variables. This is written as

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

The pairs, (x, y),will be members of a domain,D. They will be restricted to be real numbers so that D is a subset of the Oxy plane. To each pair (x, y) there corresponds a value of the function u.

To represent the function geometrically, three mutually perpendicular coordinate axes are needed, Ox, Oy and Ou. For each pair (x_0, y_0) in the domain D we calculate $u_0 = f(x_0, y_0)$ and plot the point (x_0, y_0, u_0). The result is shown in Figure 1.

The function defines a **surface** in a 3-dimensional space.The projection of the surface in the Oxy plane is the domain D. Note that the axes we have drawn in the diagram form a right-handed set. A left-handed set is the mirror-image of a right-handed set (see Figure 2).

To check whether a set is right-handed, imagine a corkscrew with its axis along the u-axis. Rotating the corkscrew from the x-axis to the y-axis moves the corkscrew in the positive u direction (see Figure 3). If your axes are left-handed then the corkscrew will move in the negative u direction. Of course, all of this depends on all corkscrews being right-handed, which they are to the best of our knowledge!

It is a well-established mathematical convention that we always use a right-handed set of axes.

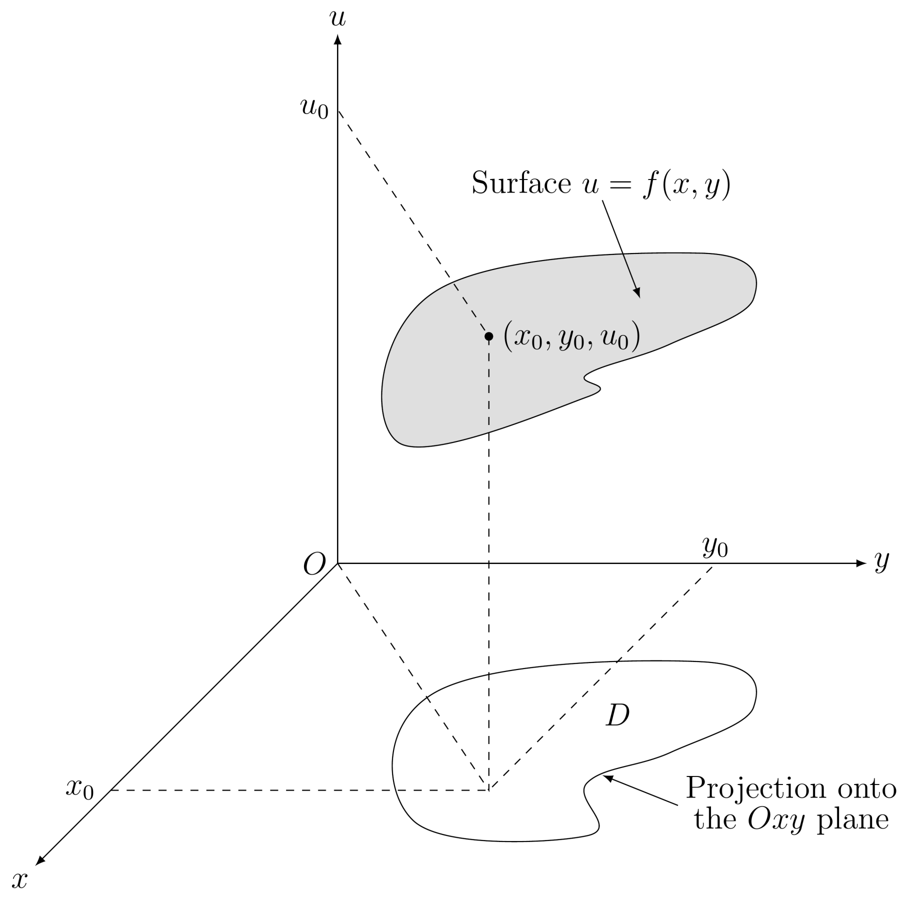

Figure 1: Representation of $u = f(x, y)$.

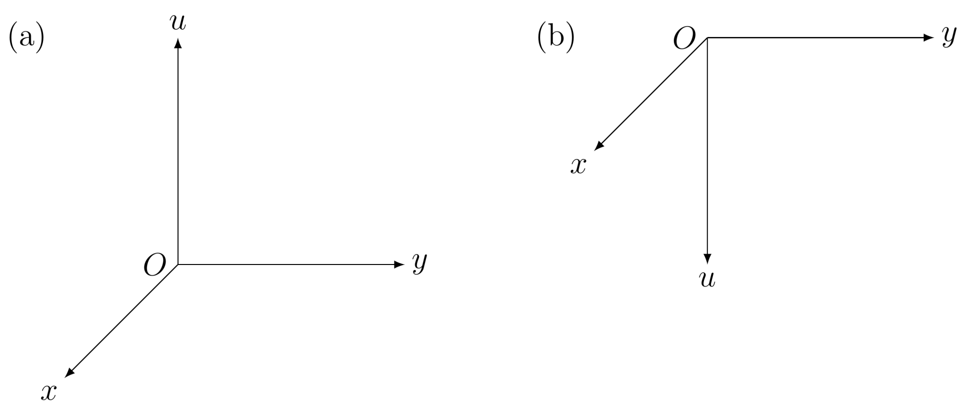

Figure 2: x, y, u axes (a) right-handed, (b) left-handed.

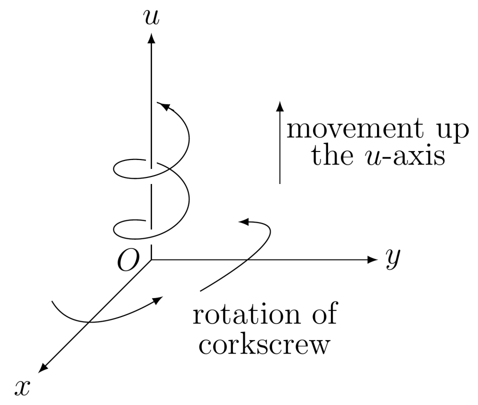

Figure 3: The corkscrew rule.

As an example of a surface consider the function

$$u = 100 - x^2 - y^2$$

with domain

$$\{(x, y), x^2 + y^2 \leq 100\}.$$

This surface, a paraboloid, is plotted in Figure 4.

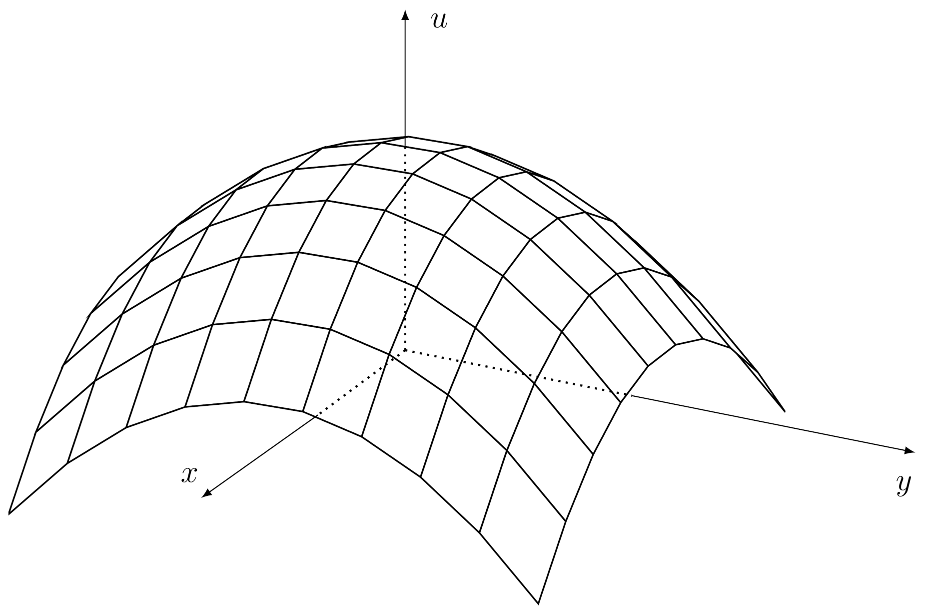

Figure 4: The paraboloid $u = 100 - x^2 - y^2$.

Another way of representing the function is to sketch the curves $f(x, y) = C$ in the Oxy plane for various values of the constant C. These curves are called **level curves** and along any one of them the value of the function is constant (equal to C).In the example above the level curves are defined by

$$100 - x^2 - y^2 = C$$

and putting $C = 0, 51, 75$ generates the circles $x^2 + y^2 = 100, x^2 + y^2 = 49$ and $x^2 + y^2 = 25$. The level curves are shown in Figure 5.

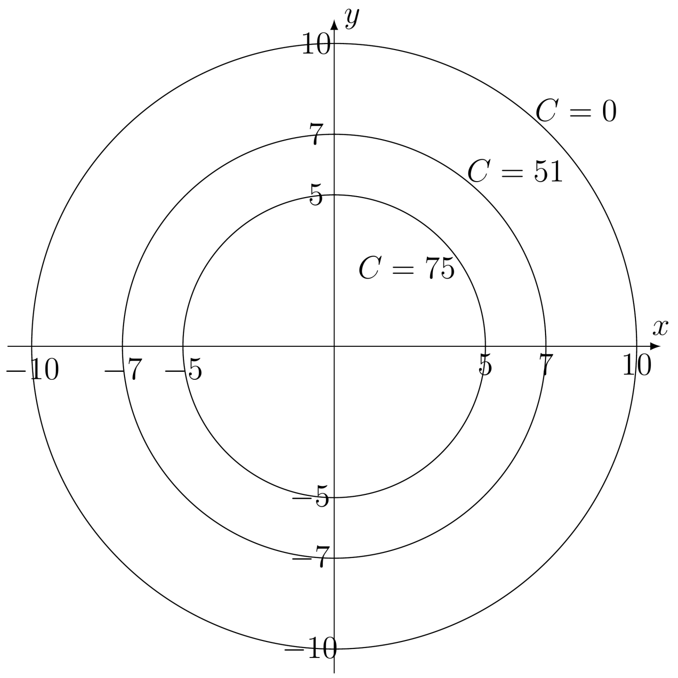

Figure 5: The level curves of $u = 100 - x^2 - y^2$.

The process of sketching level curves is the same as using contours on a map to represent height above sea level. The contours join points of equal height. The height here is the dependent variable u and the coordinates (x, y) are the independent variables.

For functions of more than three variables geometrical intuition usually fails, though the concept of **level surfaces** can be used to represent a function of three variables.

(The video gives examples of functions of several variables and shows how to represent functions of two variables geometrically) V

3 First Order Partial Derivatives

The first order partial derivative of a function of several variables is the derivative obtained when all except one of the independent variables are held constant, and the function is differentiated (once) with respect to that one variable. Thus, $u = f(x, y)$ has two first order partial derivatives denoted by $\dfrac{\partial u}{\partial x}$ and $\dfrac{\partial u}{\partial y}$. The **partial derivative** of u with respect to x is defined as

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

where the limit exists. Here $f(x+\delta x, y) - f(x, y)$ is the change in u produced by changing x to $x + \delta x$ without changing y. This partial derivative may also be written as $\dfrac{\partial f}{\partial x}$,$f_x$ or u_x.

Example 3.1 If $u = x^2 + y^3$ obtain $\dfrac{\partial u}{\partial x}$.

Solution We hold y constant and differentiate with respect to x. So y^3 is also a constant and

$$\frac{\partial u}{\partial x} = 2x.$$

○

Similarly the **partial derivative** of u with respect to y is defined as

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

when the limit exists. Here $f(x, y+\delta y) - f(x, y)$ is the change in u produced by changing y to $y + \delta y$ without changing x. This partial derivative may also be written as $\dfrac{\partial f}{\partial y}$,$f_y$ or u_y.

Example 3.2 If $u = x^2 + y^3$ obtain $\dfrac{\partial u}{\partial y}$.

Solution We hold x constant and differentiate with respect to y. So x^2 is also a constant and

$$\frac{\partial u}{\partial y} = 3y^2.$$

○

All the rules for ordinary differentiation e.g. sum, product, quotient, function of a function apply equally well to partial differentiation.

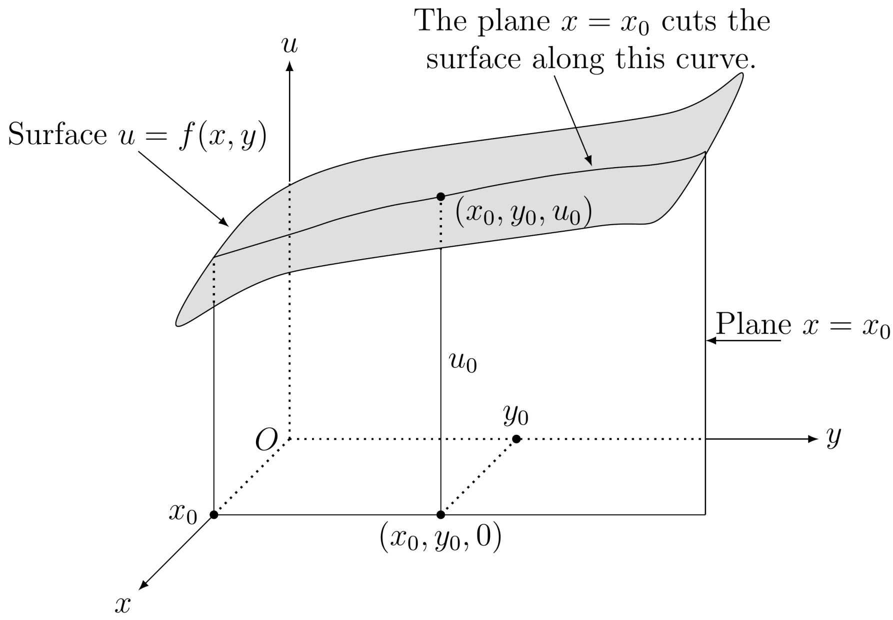

Figure 6: Intersection of $u = f(x, y)$ with $x = x_0$.

For a function of 2 variables, the partial derivatives may be interpreted geometrically. Consider the intersection of the surface $u = f(x, y)$ with the plane $x = x_0$. This is shown in Figure 6.

If we examine the intersection curve in the region of (x_0, y_0, u_0), where $u_0 = f(x_0, y_0)$, it looks like Figure 7.

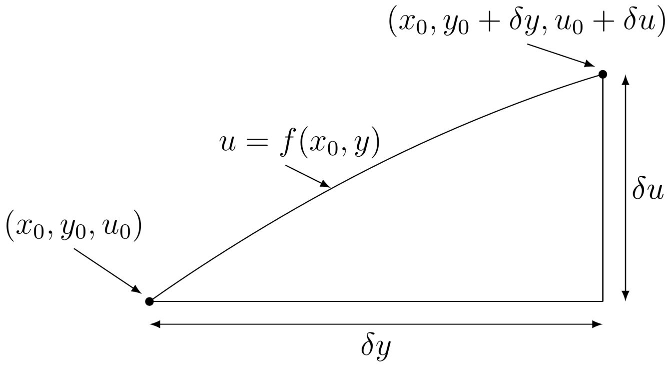

Figure 7: The intersection curve near (x_0, y_0, u_0).

We have drawn not only the point (x_0, y_0, u_0) but also a neighbouring point $(x_0, y_0+\delta y, u_0+\delta u)$. Note that x does not change along the intersection curve because it is fixed at the value x_0. It is y that has altered. The change in y is δy and this change has resulted in a change δu in the value of u. In fact δu is the change in height of the surface above the Oxy plane as we moved from the point vertically above (x_0, y_0) to the point vertically above $(x_0, y_0 + \delta y)$ i.e.

$$\delta u = f(x_0, y_0 + \delta y) - f(x_0, y_0).$$

Hence

$$\frac{\delta u}{\delta y} = \frac{f(x_0, y_0 + \delta y) - f(x_0, y_0)}{\delta y}.$$

We now let δy tend to zero. From the previous diagram you will see that the ratio $\frac{\delta u}{\delta y}$ will approach the gradient of the intersection curve at the point (x_0, y_0, u_0). However, the expression

$$\frac{f(x_0, y_0 + \delta y) - f(x_0, y_0)}{\delta y}$$

is that used to define the partial derivative, and so the limiting value of it as δy tends to zero will be $\frac{\partial u}{\partial y}$ (evaluated at $x = x_0, y = y_0$).

Thus $\frac{\partial u}{\partial y}$ gives the gradient of the intersection curve. But the intersection curve lies in the y-direction on the surface. Hence

$\frac{\partial u}{\partial y}$ represents the gradient of the surface in the y- direction.

Similarly

$\frac{\partial u}{\partial x}$ represents the gradient of the surface in the x- direction.

In two dimensions we have the concept of a tangent line to a curve. The counterpart in three dimensions is the concept of a tangent plane. A function $u = f(x, y)$ is said to be **differentiable** at (x_0, y_0) if it has a tangent plane at (x_0, y_0, u_0) where $u_0 = f(x_0, y_0)$. If the tangent plane to $u = f(x, y)$ actually exists at (x_0, y_0, u_0) then its equation will be

$$(u - u_0) = a(x - x_0) + b(y - y_0)$$

where $a = \frac{\partial f}{\partial x}$ and $b = \frac{\partial f}{\partial y}$ (both derivatives being evaluated at the point (x_0, y_0)). This uses the simple fact that the tangent plane will have the same gradients in the x and y directions as the original surface. Unfortunately it is possible for both of these gradients to exist without there being a tangent plane. To see this think of a surface which is horizontal in the x and y directions at $(0, 0)$ but along the line $y = x$, say, slopes at $45°$ to the horizontal. Therefore, the existence of the partial derivatives does not of itself guarantee the existence of a tangent plane. In more advanced texts this point is considered in some detail. We will simply assume that all the functions which we consider are differentiable in the sense of having a tangent plane whenever the partial derivatives exist.

Example 3.3 If $u = x^3 - 2x^2y + 3xy^2 - y^4$, find $\dfrac{\partial u}{\partial x}$ and $\dfrac{\partial u}{\partial y}$.

Solution Note that y is treated as a constant and therefore y^2 and y^4 are also treated as constants. Hence

$$\frac{\partial u}{\partial x} = 3x^2 - 4xy + 3y^2$$

$$\frac{\partial u}{\partial y} = -2x^2 + 6xy - 4y^3.$$ ○

Example 3.4 If $u = (x^2 + 4y^2)^{1/2}$ find $\dfrac{\partial u}{\partial x}$ and $\dfrac{\partial u}{\partial y}$.

Solution Applying the function of a function rule,

$$\frac{\partial u}{\partial x} = \frac{1}{2}(x^2 + 4y^2)^{-1/2}.2x = \frac{x}{(x^2 + 4y^2)^{1/2}}$$

$$\frac{\partial u}{\partial y} = \frac{1}{2}(x^2 + 4y^2)^{-1/2}.8y = \frac{4y}{(x^2 + 4y^2)^{1/2}}$$ ○

Example 3.5 If $u = \tan^{-1}\left(\dfrac{x - y}{x + y}\right)$, prove that $x\dfrac{\partial u}{\partial x} + y\dfrac{\partial u}{\partial y} = 0$.

Solution Since

$$u = \tan^{-1}\left(\frac{x - y}{x + y}\right), \qquad \tan u = \left(\frac{x - y}{x + y}\right).$$

Differentiating implicitly with respect to x,

$$\sec^2 u \frac{\partial u}{\partial x} = \frac{(x + y) - (x - y)}{(x + y)^2} = \frac{2y}{(x + y)^2} \qquad (1)$$

and with respect to y,

$$\sec^2 u \frac{\partial u}{\partial y} = \frac{-(x+y)-(x-y)}{(x+y)^2} = \frac{-2x}{(x+y)^2} \tag{2}$$

From (1) and (2)

$$\sec^2 u (x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y}) = \frac{2xy}{(x+y)^2} - \frac{2xy}{(x+y)^2} = 0.$$

Hence

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

since $\sec^2 u = \dfrac{1}{\cos^2 u} \neq 0.$ ○

(The video defines partial derivatives and shows how to obtain them. It also discusses their geometrical interpretation for a function of two variables. V

At this point you should try examples which involve determining the first order partial derivatives of functions of several variables.) EX

4 Higher Order Partial Derivatives

Given $u = f(x, y)$, the partial derivatives $\dfrac{\partial u}{\partial x}$ and $\dfrac{\partial u}{\partial y}$ are themselves functions of the two independent variables x and y and so each may be differentiated with respect to either x or y to produce the four **second-order partial derivatives** of u:-

$$\frac{\partial}{\partial x}(\frac{\partial u}{\partial x}) = \frac{\partial^2 u}{\partial x^2} \qquad \text{(or } (u_x)_x = u_{xx})$$

$$\frac{\partial}{\partial y}(\frac{\partial u}{\partial x}) = \frac{\partial^2 u}{\partial y \partial x} \qquad \text{(or } (u_x)_y = u_{xy}) \quad [\textbf{not } u_{yx}]$$

$$\frac{\partial}{\partial x}(\frac{\partial u}{\partial y}) = \frac{\partial^2 u}{\partial x \partial y} \qquad \text{(or } (u_y)_x = u_{yx}) \quad [\textbf{not } u_{xy}]$$

$$\frac{\partial}{\partial y}(\frac{\partial u}{\partial y}) = \frac{\partial^2 u}{\partial y^2} \qquad \text{(or } (u_y)_y = u_{yy})$$

Example 4.1 Find the four second order partial derivatives of u in Example 3.4.

Solution

$$\frac{\partial u}{\partial x} = x(x^2 + 4y^2)^{-1/2}.$$

Therefore

$$\frac{\partial^2 u}{\partial x^2} = x\left\{-\frac{1}{2}(x^2 + 4y^2)^{-3/2}.2x\right\} + (x^2 + 4y^2)^{-1/2}$$

and so

$$\frac{\partial^2 u}{\partial x^2} = \frac{-x^2}{(x^2 + 4y^2)^{3/2}} + \frac{1}{(x^2 + 4y^2)^{1/2}} = \frac{4y^2}{(x^2 + 4y^2)^{3/2}}.$$

Also

$$\frac{\partial^2 u}{\partial y \partial x} = x\left\{-\frac{1}{2}(x^2 + 4y^2)^{-3/2}.8y\right\} = \frac{-4xy}{(x^2 + 4y^2)^{3/2}}$$

and since

$$\frac{\partial u}{\partial y} = 4y(x^2 + 4y^2)^{-1/2}$$

$$\frac{\partial^2 u}{\partial x \partial y} = 4y\left\{-\frac{1}{2}(x^2 + 4y^2)^{-3/2}.2x\right\} = \frac{-4xy}{(x^2 + 4y^2)^{3/2}}$$

and

$$\frac{\partial^2 u}{\partial y^2} = 4y\left\{-\frac{1}{2}(x^2 + 4y^2)^{-3/2}.8y\right\} + 4(x^2 + 4y^2)^{-1/2}.$$

Thus

$$\frac{\partial^2 u}{\partial y^2} = \frac{-16y^2}{(x^2+4y^2)^{3/2}} + \frac{4}{(x^2+4y^2)^{1/2}} = \frac{4x^2}{(x^2+4y^2)^{3/2}}.$$ ○

Note in this example that $\dfrac{\partial^2 u}{\partial y \partial x} = \dfrac{\partial^2 u}{\partial x \partial y}$. This is not a coincidence. For suitably behaved functions it can be shown that

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

This result is the **mixed derivatives theorem** and will be assumed to be true for all our work on partial differentiation. (The theorem is valid whenever the function u and its first and second order partial derivatives are continuous).

(The video covers the definitions and notation for second order partial derivatives. It also shows how to obtain them and discusses the mixed derivative theorem. V

At this point you should try some examples involving the determination of higher order derivatives.) EX

5 Functions of more than two variables

(This section is not covered in the videos.)

The ideas of partial differentiation extend to functions of any number of independent variables. Let $u = f(x_1, x_2, \ldots, x_n)$ where $x_1, x_2, \ldots, x_n$ are n independent variables. Then u has n first order partial derivatives,

$$\frac{\partial u}{\partial x_i}, \quad i = 1, 2. \ldots, n.$$

Each of these will generate n second order partial derivatives,

$$\frac{\partial^2 u}{\partial x_k \partial x_i}, \quad k = 1, 2, \ldots, n.$$

However the mixed derivative theorem applies and so

$$\frac{\partial^2 u}{\partial x_i \partial x_k} = \frac{\partial^2 u}{\partial x_k \partial x_i}, \quad i, k = 1, 2, \ldots, n.$$

The mixed derivative theorem also applies to higher order derivatives so that, for example,

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

and

$$\frac{\partial^3 u}{\partial x_i \partial x_k \partial x_j} = \frac{\partial^3 u}{\partial x_j \partial x_i \partial x_k}.$$

[Again, continuity of u and all the derivatives involved is sufficient to ensure the validity of these equations]

6 Increments

6.1 Small increments formula

Let $u = f(x, y)$ and suppose changes or **increments** are made in the values of x and y. Let x change from x_0 to $x_0 + \delta x$ and y from y_0 to $y_0 + \delta y$. Then the resulting increment in u is δu, where

$$\delta u = f(x_0 + \delta x, y_0 + \delta y) - f(x_0, y_0)$$

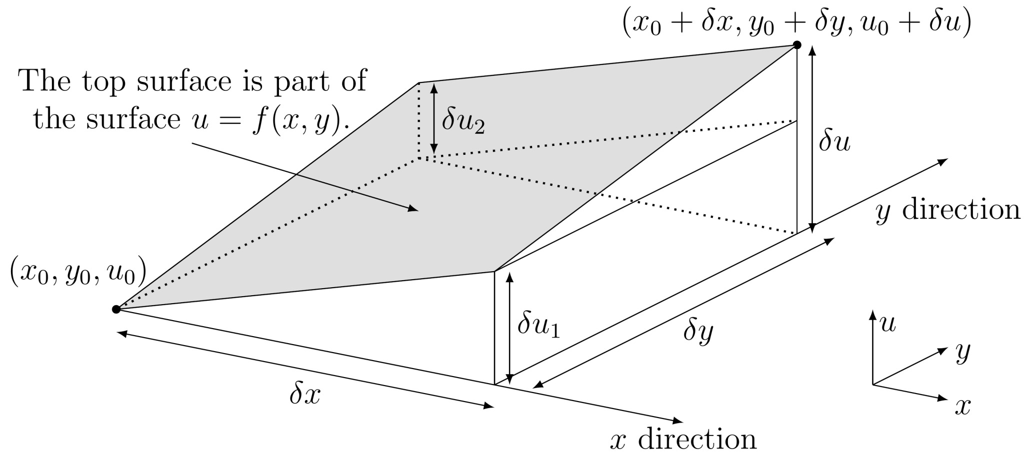

Figure 8: Increment in u generated by increments in x and y.

Figure 8 gives a geometrical interpretation (for ease of viewing we have adopted a slightly different eye position from the normal one - look at the axes to get the perspective right). Assuming the increments are small so that the upper surface is approximately flat (i.e. it coincides with the tangent plane at (x_0, y_0)) then δu is (approximately) given by

$$\delta u \approx \delta u_1 + \delta u_2 \tag{3}$$

But $\dfrac{\delta u_1}{\delta x}$ is (approximately) $\dfrac{\partial u}{\partial x}$ evaluated at (x_0, y_0). Similarly $\dfrac{\delta u_2}{\delta y}$ is given (approximately) by $\dfrac{\partial u}{\partial y}$. Hence (3) gives the approximate equation

$$\boxed{\delta u \approx \frac{\partial u}{\partial x}\delta x + \frac{\partial u}{\partial y}\delta y} \tag{4}$$

We can read this as saying that u changes for two reasons (the two terms on the right hand side). The first is that x changes by δx and this contributes

$\dfrac{\partial u}{\partial x}\delta x$ to the change in u. The second is that y changes by δy and this contributes $\dfrac{\partial u}{\partial y}\delta y$ to the change in u.

Equation (4) is really only an approximation but, for reasonable functions $f(x, y)$ we would expect the accuracy to improve as δx and δy tend to zero. When $\delta x, \delta y$ are very small, the tangent plane at (x_0, y_0, u_0) will lie very close to the surface. The errors in the approximation will then be small as a proportion of δu. Sometimes (4) is known as the **small increments formula.** It extends easily to functions of more than two variables. If

$$u = f(x_1, x_2, \ldots, x_n)$$

then small increments $\delta x_1, \delta x_2, \ldots, \delta x_n$ in the independent variables will produce an increment δu in u given (approximately) by

$$\delta u \approx \frac{\partial u}{\partial x_1}\delta x_1 + \frac{\partial u}{\partial x_2}\delta x_2 + \cdots + \frac{\partial u}{\partial x_n}\delta x_n.$$

(The small increments formula is discussed in the video) V

6.2 Application to Approximations

We can use the small increments formula to estimate the value of a function of two variables at points near to a point with a known value.

Example 6.1 If $u = (x^2 + 4y^2)^{1/2}$, find an approximate value of u when $x = 2.97$ and $y = 2.01$ using partial derivatives. Compare the approximate value with the exact value.

Solution From example 3.4

$$\frac{\partial u}{\partial x} = \frac{x}{(x^2 + 4y^2)^{1/2}}, \qquad \frac{\partial u}{\partial y} = \frac{4y}{(x^2 + 4y^2)^{1/2}}.$$

Let $(x_0, y_0) = (3, 2)$ and $(\delta x, \delta y) = (-0.03, 0.01)$ Then, using the small increments formula:

$$\delta u \approx \frac{\partial u}{\partial x}\delta x + \frac{\partial u}{\partial y}\delta y$$

we have

$$\delta u \approx \frac{3}{(3^2 + 4.2^2)^{1/2}}(-0.03) + \frac{8}{(3^2 + 4.2^2)^{1/2}}(0.01).$$

i.e.

$$\delta u \approx \frac{-0.09}{5} + \frac{0.08}{5} = -0.002.$$

Therefore

$$u(2.97, 2.01) \approx 5 - 0.002 = 4.998$$

The exact value is $((2.97)^2 + 4(2.01)^2)^{1/2} = 4.99813$ to 5 decimal places. ○

(The application of the small increments formula to approximations of functions is covered in the video V

You should now try some examples on approximating functions with the small increments formula) EX

6.3 Application to Errors

If a measured quantity, u, is subject to an error of $p\%$, then this means that the measured quantity lies in the range $u \pm \delta u$ where $\left|\dfrac{\delta u}{u}\right| \leq \dfrac{p}{100}$. The results of partial differentiation may be used to estimate percentage errors in functions of two or more variables.

Example 6.2 The period of a simple pendulum is given by $T = 2\pi\sqrt{\dfrac{l}{g}}$. If the length, l, is subject to a 1% error and the period, T, is subject to a 2% error, estimate the approximate maximum percentage error in the calculated value of g.

Solution Here $\left|\dfrac{\delta l}{l}\right| \leq \dfrac{1}{100}$ and $\left|\dfrac{\delta T}{T}\right| \leq \dfrac{2}{100}$. Since

$$T = 2\pi\sqrt{\frac{l}{g}},$$

then

$$T^2 = \frac{4\pi^2 l}{g}$$

and

$$g = \frac{4\pi^2 l}{T^2}.$$

We have now expressed g as a function of the two independent variables, l and T.

So

$$\delta g \approx \frac{\partial g}{\partial l}\delta l + \frac{\partial g}{\partial T}\delta T$$

Therefore

$$\delta g \approx \frac{4\pi^2}{T^2}\delta l - \frac{8\pi^2 l}{T^3}\delta T.$$

So

$$\frac{\delta g}{g} \approx \frac{4\pi^2}{gT^2}\delta l - \frac{8\pi^2 l}{gT^3}\delta T.$$

i.e.

$$\frac{\delta g}{g} \approx \frac{\delta l}{l} - \frac{2\delta T}{T}$$

and

$$\left|\frac{\delta g}{g}\right| \approx \left|\frac{\delta l}{l} - \frac{2\delta T}{T}\right|.$$

Using the triangle inequality, we have (approximately)

$$\left|\frac{\delta g}{g}\right| \leq \left|\frac{\delta l}{l}\right| + \left|\frac{2\delta T}{T}\right| \leq \frac{1}{100} + \frac{4}{100} = \frac{5}{100}.$$

Therefore the maximum possible error in g is estimated as 5%. ○

(The video discusses the application of the small increments formula to error estimation and covers the above example. V

At this point you should attempt some examples on error estimation.) EX

7 Chain rules for partial derivatives

7.1 One independent variable

Suppose $u = f(x, y)$, and both x and y are functions of a single variable t, i.e. $x = g(t), y = h(t)$. We now suppose t changes by a small amount δt and that this gives rise to small changes $\delta x, \delta y$ in x, y and consequently to a small change δu in u.

Then since

$$\delta u \approx \frac{\partial u}{\partial x}\delta x + \frac{\partial u}{\partial y}\delta y,$$

we have

$$\frac{\delta u}{\delta t} \approx \frac{\partial u}{\partial x}\frac{\delta x}{\delta t} + \frac{\partial u}{\partial y}\frac{\delta y}{\delta t}.$$

In the limit as $\delta t \to 0$, the chain rule for one independent variable is obtained:-

$$\frac{du}{dt} = \frac{\partial u}{\partial x}\frac{dx}{dt} + \frac{\partial u}{\partial y}\frac{dy}{dt}$$

Example 7.1 If the radius of a right circular cone increases at 2 cm/minute and the height at 3 cm/minute, find the rate at which the volume is increasing when the radius is 10cm and the height is 15cm.

Solution Let V denote the volume of the cone. Then

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

where r is the radius and h the height. We have

$$\frac{\partial V}{\partial r} = \frac{2\pi r h}{3}, \quad \frac{\partial V}{\partial h} = \frac{\pi r^2}{3}.$$

Using the chain rule,

$$\begin{aligned}
\frac{dV}{dt} &= \frac{\partial V}{\partial h}\frac{dh}{dt} + \frac{\partial V}{\partial r}\frac{dr}{dt} \\
&= \frac{\pi r^2}{3}\frac{dh}{dt} + \frac{2\pi r h}{3}\frac{dr}{dt} \\
&= \frac{\pi r}{3}(r\frac{dh}{dt} + 2h\frac{dr}{dt}) \\
&= \frac{\pi .10}{3}(10 \times 3 + 30 \times 2) \\
&= 300\pi \\
&\approx 942.48 \quad \text{cu cms/minute.}
\end{aligned}$$

○

(The video discusses the chain rule for one independent variable and solves the above example.) V

7.2 Two independent variables

Suppose $u = f(x, y)$, and both x and y are functions of two independent variables α and β. i.e.

$$x = F(\alpha, \beta) \quad \text{and } y = G(\alpha, \beta).$$

Then

$$u = f(F(\alpha, \beta), G(\alpha, \beta)) \equiv g(\alpha, \beta) \quad \text{(say)}.$$

The chain rule can be used to relate the partial derivatives of u with respect to α and β to those with respect to x and y. We imagine that α, β, change by small increments $\delta\alpha, \delta\beta$, and that these generate small increments $\delta x, \delta y$ in x and y. These in turn generate a small increment δu in the value of u. Since

$$\delta u \approx \frac{\partial u}{\partial x}\delta x + \frac{\partial u}{\partial y}\delta y,$$

then

$$\frac{\delta u}{\delta \alpha} \approx \frac{\partial u}{\partial x}\frac{\delta x}{\delta \alpha} + \frac{\partial u}{\partial y}\frac{\delta y}{\delta \alpha}.$$

In the limit as $\delta\alpha \to 0$,

$$\frac{\partial u}{\partial \alpha} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial \alpha} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial \alpha}.$$

Similarly

$$\frac{\partial u}{\partial \beta} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial \beta} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial \beta}.$$

These equations constitute the chain rule for a function of two independent variables.

Example 7.2 If $u = f(x, y)$ and $x = e^{\alpha}\cosh\beta, y = e^{\alpha}\sinh\beta$ prove that

1.
$$\frac{\partial u}{\partial \alpha} = x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y}, \quad \frac{\partial u}{\partial \beta} = y\frac{\partial u}{\partial x} + x\frac{\partial u}{\partial y}$$

2.
$$\frac{\partial^2 u}{\partial\alpha\partial\beta} - \frac{\partial u}{\partial \beta} = xy\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) + (x^2 + y^2)\frac{\partial^2 u}{\partial x \partial y}$$

Solution We have

$$x = e^\alpha \cosh\beta, \quad y = e^\alpha \sinh\beta.$$

So

$$\frac{\partial x}{\partial \alpha} = e^\alpha \cosh\beta = x$$

and

$$\frac{\partial y}{\partial \alpha} = e^\alpha \sinh\beta = y.$$

By the chain rule,

$$\frac{\partial u}{\partial \alpha} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial \alpha} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial \alpha} = x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y}.$$

Similarly,

$$\frac{\partial x}{\partial \beta} = e^\alpha \sinh\beta = y$$

and

$$\frac{\partial y}{\partial \beta} = e^\alpha \cosh\beta = x.$$

By the chain rule,

$$\frac{\partial u}{\partial \beta} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial \beta} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial \beta} = y\frac{\partial u}{\partial x} + x\frac{\partial u}{\partial y}.$$

To get the higher derivatives such as $\dfrac{\partial^2 u}{\partial\alpha\partial\beta}$ we remind ourselves that they are obtained by repeated first order differentiations. The formulae obtained above, such as

$$\frac{\partial u}{\partial \alpha} = x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y}.$$

apply to any "reasonable" function u. It may help if we write this formula as

$$\frac{\partial}{\partial \alpha}(u) = \left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)(u)$$

or even as

$$\frac{\partial}{\partial \alpha} = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}.$$

The two sides of this equation are differential operators. They are equal in the sense that they produce equal answers when applied to any function. If we apply the formula to $\frac{\partial u}{\partial \beta}$ instead of applying it to u, we get

$$\begin{aligned}
\frac{\partial}{\partial \alpha}(\frac{\partial u}{\partial \beta}) &= \left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)(\frac{\partial u}{\partial \beta}) \\
&= \left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)\left(y\frac{\partial u}{\partial x} + x\frac{\partial u}{\partial y}\right).
\end{aligned}$$

Therefore

$$\begin{aligned}
\frac{\partial^2 u}{\partial \alpha \partial \beta} &= xy\frac{\partial^2 u}{\partial x^2} + x^2\frac{\partial^2 u}{\partial x \partial y} + x\frac{\partial u}{\partial y} + \\
& \quad y^2\frac{\partial^2 u}{\partial y \partial x} + y\frac{\partial u}{\partial x} + xy\frac{\partial^2 u}{\partial y^2} \\
&= xy\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) + (x^2 + y^2)\frac{\partial^2 u}{\partial x \partial y} + x\frac{\partial u}{\partial y} + y\frac{\partial u}{\partial x}.
\end{aligned}$$

Therefore

$$\frac{\partial^2 u}{\partial \alpha \partial \beta} - \frac{\partial u}{\partial \beta} = xy\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) + (x^2 + y^2)\frac{\partial^2 u}{\partial x \partial y}.$$

○

The chain rule may be extended to any number of variables. If u is a function of the n variables, $x_1, x_2, \ldots x_n$ and each x_i is a function of the m variables, $\alpha_1, \alpha_2, \ldots \alpha_m$,then

$$\begin{aligned}
\frac{\partial u}{\partial \alpha_1} &= \frac{\partial u}{\partial x_1}\frac{\partial x_1}{\partial \alpha_1} + \frac{\partial u}{\partial x_2}\frac{\partial x_2}{\partial \alpha_1} + \ldots + \frac{\partial u}{\partial x_n}\frac{\partial x_n}{\partial \alpha_1} \\
\frac{\partial u}{\partial \alpha_2} &= \frac{\partial u}{\partial x_1}\frac{\partial x_1}{\partial \alpha_2} + \frac{\partial u}{\partial x_2}\frac{\partial x_2}{\partial \alpha_2} + \ldots + \frac{\partial u}{\partial x_n}\frac{\partial x_n}{\partial \alpha_2} \\
\vdots & \qquad\qquad \vdots \\
\frac{\partial u}{\partial \alpha_m} &= \frac{\partial u}{\partial x_1}\frac{\partial x_1}{\partial \alpha_m} + \frac{\partial u}{\partial x_2}\frac{\partial x_2}{\partial \alpha_m} + \ldots + \frac{\partial u}{\partial x_n}\frac{\partial x_n}{\partial \alpha_m}
\end{aligned}$$

This set of equations can most easily be expressed in matrix form as

$$\begin{pmatrix} \frac{\partial u}{\partial \alpha_1} \\ \frac{\partial u}{\partial \alpha_2} \\ \vdots \\ \frac{\partial u}{\partial \alpha_m} \end{pmatrix} = \begin{pmatrix} \frac{\partial x_1}{\partial \alpha_1} & \frac{\partial x_2}{\partial \alpha_1} & \cdots & \frac{\partial x_n}{\partial \alpha_1} \\ \frac{\partial x_1}{\partial \alpha_2} & \frac{\partial x_2}{\partial \alpha_2} & \cdots & \frac{\partial x_n}{\partial \alpha_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_1}{\partial \alpha_m} & \frac{\partial x_2}{\partial \alpha_m} & \cdots & \frac{\partial x_n}{\partial \alpha_m} \end{pmatrix} \begin{pmatrix} \frac{\partial u}{\partial x_1} \\ \frac{\partial u}{\partial x_2} \\ \vdots \\ \frac{\partial u}{\partial x_n} \end{pmatrix}$$

The large rectangular $(m \times n)$ matrix is called the Jacobian matrix and is sometimes denoted by

$$\frac{\partial(x_1, x_2, \ldots, x_n)}{\partial(\alpha_1, \alpha_2, \ldots, \alpha_m)}$$

or simply by J.

(The video covers the chain rule for a function of two independent variables and explains part 1 of the example in this section. V

Now you should try some examples using the chain rule for partial derivatives.) EX

7.3 Application to implicit differentiation

Suppose x and y are related by the implicit relation

$$f(x, y) = 0$$

then $\frac{dy}{dx}$ may be obtained by implicit differentiation. Alternatively, we can proceed as follows. Forget for a moment that $f(x, y) = 0$. If we write $u = f(x, y)$ and allow x, y to change by small increments $\delta x, \delta y$ then the resulting change in u is given approximately by

$$\delta u \approx \frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial y}\delta y.$$

Now for all x and y that satisfy $f(x, y) = 0$, we have $\delta u = 0$. For such x and y, given δx we have to choose δy to satisfy

$$0 \approx \frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial y}\delta y.$$

Dividing by δx and taking the limit as $\delta x \to 0$, we obtain

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

So

$$\frac{dy}{dx} = -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}} = -\frac{f_x}{f_y}.$$

Example 7.3 Find $\frac{dy}{dx}$ in terms of x and y when

$$y = \sin([2x - y]^2).$$

Solution Let
$$u = y - \sin([2x - y]^2).$$
Then
$$\frac{\partial u}{\partial x} = -2[2x - y]2\cos([2x - y]^2)$$
and
$$\frac{\partial u}{\partial y} = 1 + 2[2x - y]\cos([2x - y]^2).$$
Therefore
$$\frac{dy}{dx} = \frac{4[2x - y]\cos([2x - y]^2)}{1 + 2[2x - y]\cos([2x - y]^2)}.$$ ○

The method extends to functions defined implicitly with any number of variables. For example, if $f(x, y, z) = 0$ defines z implicitly as a function of x and y, then
$$\frac{\partial z}{\partial x} = -\frac{f_x}{f_z}, \quad \frac{\partial z}{\partial y} = -\frac{f_y}{f_z}.$$

(The video covers the application of partial derivatives to implicit differentiation and discusses the example in this section. V

Now try some examples on the use of partial derivatives to perform implicit differentiation.) EX

8 Taylor Series for a function of two variables

Before examining the problem of determining stationary points for a function of two variables, we need to consider expanding a function of two variables in a Taylor series.

We recall that, if $f(x)$ is suitably well behaved, it has a Taylor series expansion:-

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2!}f''(x) + \ldots + \frac{h^n}{n!}f^{(n)}(x) + \ldots.$$

Consider now a function of two variables, $f(x,y)$. We assume that all the partial derivatives of f exist and are suitably well behaved, and try to find an expansion for $f(x+h, y+k)$. First of all, we regard y as a constant and expand as a function of x:-

$$f(x+h,y+k) = f(x,y+k) + h\frac{\partial f}{\partial x}(x,y+k) + \frac{h^2}{2!}\frac{\partial^2 f}{\partial x^2}(x,y+k) + \ldots \quad (5)$$

[An expression such as $\frac{\partial f}{\partial x}(x, y+k)$ means that the derivative $\frac{\partial f}{\partial x}$ is evaluated at the point $(x, y+k)$.]

Now consider x to be a constant and expand $f(x, y+k)$ as a Taylor series in y:-

$$f(x,y+k) = f(x,y) + k\frac{\partial f}{\partial y}(x,y) + \frac{k^2}{2!}\frac{\partial^2 f}{\partial y^2}(x,y) + \ldots \quad (6)$$

Differentiating this with respect to x gives

$$\frac{\partial f}{\partial x}(x,y+k) = \frac{\partial f}{\partial x}(x,y) + k\frac{\partial^2 f}{\partial x \partial y}(x,y) + \ldots \quad (7)$$

and again,

$$\frac{\partial^2 f}{\partial x^2}(x,y+k) = \frac{\partial^2 f}{\partial x^2}(x,y) + \ldots \quad (8)$$

Substituting the expressions (6),(7),(8) into the right hand side of (5) gives

$$\begin{aligned} f(x+h,y+k) &= f(x,y) + k\frac{\partial f}{\partial y}(x,y) + \frac{k^2}{2!}\frac{\partial^2 f}{\partial y^2}(x,y) + \\ &\quad h\frac{\partial f}{\partial x}(x,y) + hk\frac{\partial^2 f}{\partial x \partial y}(x,y) + \frac{h^2}{2!}\frac{\partial^2 f}{\partial x^2}(x,y) + \ldots \end{aligned}$$

Rearranging the order of the terms on the right hand side and, for clarity, omitting the (x,y), we obtain

$$f(x+h,y+k) = f + \left(h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y}\right) + \frac{1}{2!}\left(h^2\frac{\partial^2 f}{\partial x^2} + 2hk\frac{\partial^2 f}{\partial x \partial y} + k^2\frac{\partial^2 f}{\partial y^2}\right) + \ldots \quad (9)$$

This is the Taylor series expansion for $f(x+h, y+k)$. Clearly f and all its partial derivatives must exist at (x, y). Even then the series may not always converge.

The term containing first derivatives may be written

$$\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right) f,$$

whereas, the term containing second derivatives may be written

$$\frac{1}{2!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 f.$$

The general term is

$$\frac{1}{n!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n f,$$

giving the Taylor series as

$$f(x+h, y+k) = \sum_{n=0}^{\infty} \frac{1}{n!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n f \tag{10}$$

where

$$\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^0$$

is defined as 1.

Note that if we replace h and k by small values δx and δy and ignore terms involving $(\delta x)^2$,$(\delta y)^2$, $(\delta x)(\delta y)$ and higher powers then (9) and (10) give

$$f(x+\delta x, y+\delta y) - f(x, y) \approx \frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial y}\delta y.$$

This is just the small increment formula (see equation (4)).

Example 8.1 Expand the function $f(x, y) = \cos(xy)$ as a Taylor series about the point $(1/2, \pi/2)$, up to and including second order terms. Hence estimate $f(0.51, \pi/2 + 0.01)$ and compare with its exact value.

Solution We present two methods. The first is the direct application of Taylor's theorem of a function of two variables. The second method is to use standard series for functions of a single variable; in effect we use Taylor's theorem for functions of one variable to produce the series in two variables just as we did above in the proof of Taylor's theorem.

Method 1 We are given that

$$f(x,y) = \cos(xy).$$

Therefore

$$\begin{aligned}
\frac{\partial f}{\partial x} &= -y\sin(xy), \\
\frac{\partial f}{\partial y} &= -x\sin(xy), \\
\frac{\partial^2 f}{\partial x^2} &= -y^2\cos(xy), \\
\frac{\partial^2 f}{\partial y^2} &= -x^2\cos(xy), \\
\frac{\partial^2 f}{\partial x \partial y} &= -\sin(xy) - yx\cos(xy).
\end{aligned}$$

At $\left(\frac{1}{2}, \frac{\pi}{2}\right)$,

$$\begin{aligned}
f\left(\frac{1}{2}, \frac{\pi}{2}\right) &= \cos\left(\frac{\pi}{4}\right) = \frac{\sqrt{2}}{2}, \\
\frac{\partial f}{\partial x} &= -\frac{\pi}{2}\sin\left(\frac{\pi}{4}\right) = -\frac{\pi\sqrt{2}}{4}, \\
\frac{\partial f}{\partial y} &= -\frac{1}{2}\sin\left(\frac{\pi}{4}\right) = -\frac{\sqrt{2}}{4}, \\
\frac{\partial^2 f}{\partial x^2} &= -\frac{\pi^2}{4}\cos\left(\frac{\pi}{4}\right) = -\frac{\pi^2\sqrt{2}}{8}, \\
\frac{\partial^2 f}{\partial y^2} &= -\frac{1}{4}\cos\left(\frac{\pi}{4}\right) = -\frac{\sqrt{2}}{8}, \\
\frac{\partial^2 f}{\partial x \partial y} &= -\sin\left(\frac{\pi}{4}\right) - \frac{\pi}{4}\cos\left(\frac{\pi}{4}\right) = -\left(1 + \frac{\pi}{4}\right)\frac{\sqrt{2}}{2}.
\end{aligned}$$

So the series gives

$$\frac{\sqrt{2}}{2} - \frac{\pi\sqrt{2}}{4}h - \frac{\sqrt{2}}{4}k + \frac{1}{2!}\left(-\frac{\pi^2\sqrt{2}}{8}h^2 - (1+\frac{\pi}{4})\sqrt{2}hk - \frac{\sqrt{2}}{8}k^2\right) + \ldots$$

With $h = k = 0.01$, this gives

$$\frac{\sqrt{2}}{2} - \frac{(\pi+1)\sqrt{2}}{4}(0.01) - \frac{\sqrt{2}}{2}(0.01)^2\left(\frac{\pi^2}{8} + (1+\frac{\pi}{4}) + \frac{1}{8}\right) + \ldots$$

which is

$$\approx 0.692242 \quad \text{to 6 decimal places.}$$

This compares with

$$f(0.51, \frac{\pi}{2} + 0.01) = \cos(0.51(\frac{\pi}{2} + 0.01)) = 0.6922413 \quad \text{to 7 decimal places.}$$ ○

Method 2 We use the expansion

$$\cos(a + b) = \cos a \cos b - \sin a \sin b$$

together with the standard series

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots.$$

We have

$$\begin{aligned} \cos\left(\left(\frac{1}{2} + h\right)\left(\frac{\pi}{2} + k\right)\right) &= \cos\left(\frac{\pi}{4} + \frac{\pi h}{2} + \frac{k}{2} + kh\right) \\ &= \cos\frac{\pi}{4}\cos\left(\frac{\pi h}{2} + \frac{k}{2} + kh\right) - \sin\frac{\pi}{4}\sin\left(\frac{\pi h}{2} + \frac{k}{2} + kh\right) \end{aligned}$$

But $\cos\frac{\pi}{4} = \sin\frac{\pi}{4} = \frac{\sqrt{2}}{2}$, and so using the standard series

$$\begin{aligned} \cos\left(\left(\frac{1}{2} + h\right)\left(\frac{\pi}{2} + k\right)\right) &= \frac{\sqrt{2}}{2}\left[1 - \frac{1}{2!}\left(\frac{\pi h}{2} + \frac{k}{2} + kh\right)^2 + \ldots\right] \\ &\quad -\frac{\sqrt{2}}{2}\left[\left(\frac{\pi h}{2} + \frac{k}{2} + kh\right) - \frac{1}{3!}\left(\frac{\pi h}{2} + \frac{k}{2} + kh\right)^3 + \ldots\right] \\ &= \frac{\sqrt{2}}{2} - \frac{\pi\sqrt{2}}{4}h - \frac{\sqrt{2}}{4}k \\ &\quad + \frac{1}{2!}\left(-\frac{\pi^2\sqrt{2}}{8}h^2 - (1 + \frac{\pi}{4})\sqrt{2}hk - \frac{\sqrt{2}}{8}k^2\right) + \ldots \end{aligned}$$

(discarding terms of order greater than 2).

The evaluation of $f\left(0.51, \frac{\pi}{2} + 0.01\right)$ now follows as in method 1. ○

(The video explains how to expand a function of two variables as a Taylor series and then demonstrates the method by expanding $f(x, y) = e^{xe^y}$ as a Taylor seies about $(0, 0)$. V

At this point you should try some examples involving the expansion of a function of two variables as a Taylor series.) EX

9 Application of Taylor Series to Stationary Points

Given $u = f(x, y)$, then u has a **local maximum** at (x, y) provided $f(x, y)$ is always greater than $f(x + h, y + k)$ for all arbitrary, small h, k (but not both zero).i.e. $f(x + h, y + k) - f(x, y) < 0$ as h, k vary but remain small. (see Figure 9.)

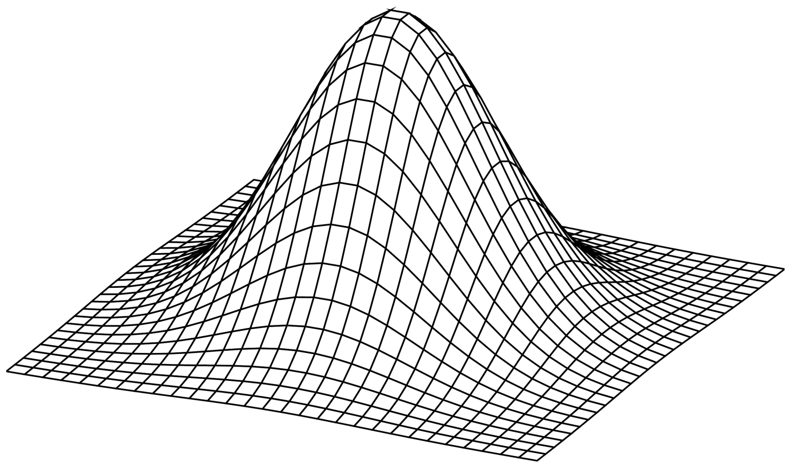

Figure 9: Local maximum of $u = f(x, y)$.

Similarly, u has a **local minimum** at (x, y), provided $f(x, y)$ is always less than $f(x + h, y + k)$ for all arbitrary small h, k (but not both zero).i.e. $f(x + h, y + k) - f(x, y) > 0$ as h, k vary but remain small.

Note that the definitions do not require the function to be differentiable at the point concerned, nor to have a Taylor series expansion about it. Nevertheless we shall assume that $f(x, y)$ has a Taylor series expansion and explore under what circumstances (x, y) can be a local maximum or a local minimum.

Using the Taylor series expansion

$$f(x+h, y+k) - f(x, y) = \left(h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y}\right) + \frac{1}{2!}\left(h^2\frac{\partial^2 f}{\partial x^2} + 2hk\frac{\partial^2 f}{\partial x \partial y} + k^2\frac{\partial^2 f}{\partial y^2}\right) + \dots$$

When h and k are sufficiently small, because h^2, hk and k^2 will be negligible compared with h and k, the sign of the right hand side will normally be determined by the sign of $(h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y})$. This will change sign if the signs of both h and k are changed, **unless**

$$h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y} = 0.$$

Thus for a local maximum or a minimum, we require

$$h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y} = 0 \text{ for arbitrary small } h, k.$$

This implies that, at a local maximum or minimum,

$$\frac{\partial f}{\partial x} = 0 \text{ and } \frac{\partial f}{\partial y} = 0.$$

Geometrically this means that the tangent plane is horizontal (i.e. parallel to the Oxy plane).

Any point at which both $\frac{\partial f}{\partial x} = 0$ and $\frac{\partial f}{\partial y} = 0$ is called a **stationary point** of f. The stationary points of f are those where the tangent plane is horizontal.

Assume that (x, y) is a stationary point of f. Then, provided that the second derivatives are not all zero, the sign of $f(x + h, y + k) - f(x, y)$ will depend, for small values of h and k, on the sign of

$$h^2\frac{\partial^2 f}{\partial x^2} + 2hk\frac{\partial^2 f}{\partial x \partial y} + k^2\frac{\partial^2 f}{\partial y^2}.$$

Let

$$a = \frac{\partial^2 f}{\partial x^2}, b = \frac{\partial^2 f}{\partial x \partial y}, c = \frac{\partial^2 f}{\partial y^2}$$

and assume that $\frac{\partial^2 f}{\partial x^2} \neq 0$ (i.e. $a \neq 0$). Then

$$\begin{aligned} ah^2 + 2bhk + ck^2 &= a\left(h^2 + \frac{2b}{a}hk + \frac{c}{a}k^2\right) \\ &= a\left\{\left(h + \frac{b}{a}k\right)^2 + \frac{c}{a}k^2 - \frac{b^2}{a^2}k^2\right\} \\ &= a\left\{\left(h + \frac{b}{a}k\right)^2 + \frac{(ac - b^2)}{a^2}k^2\right\} \qquad (11) \end{aligned}$$

We reiterate that at a stationary point, equation (11) determines the sign of $f(x + h, y + k) - f(x, y)$ for small h, k.

If $ac > b^2$,the expression in $\{\}$ in equation (11) is always non-negative and it is zero only if $h = k = 0$. Hence if $ac > b^2$ and $a > 0$, the whole expression in equation (11) is always positive (except when $h = k = 0$) and so the point in question is a local minimum.

i.e. if $(\frac{\partial^2 f}{\partial x^2})(\frac{\partial^2 f}{\partial y^2}) > (\frac{\partial^2 f}{\partial x \partial y})^2$ **and** $\frac{\partial^2 f}{\partial x^2} > 0$**,** (x, y) **is a local minimum.**

If $ac > b^2$ and $a < 0$, the expression in equation (11) is always negative (except when $h = k = 0$) and so the point in question is a local maximum.

i.e. if $(\frac{\partial^2 f}{\partial x^2})(\frac{\partial^2 f}{\partial y^2}) > (\frac{\partial^2 f}{\partial x \partial y})^2$ **and** $\frac{\partial^2 f}{\partial x^2} < 0$**,** (x, y) **is a local maximum.**

Finally if $ac < b^2$ then the expression in equation (11) can be made to assume both positive and negative values for arbitrarily small h, k. A stationary point with this property is called a **saddle point** (so called because it resembles a saddle - see Figure 10).

i.e. if $(\frac{\partial^2 f}{\partial x^2})(\frac{\partial^2 f}{\partial y^2}) < (\frac{\partial^2 f}{\partial x \partial y})^2$**, then** (x, y) **is a saddle point.**

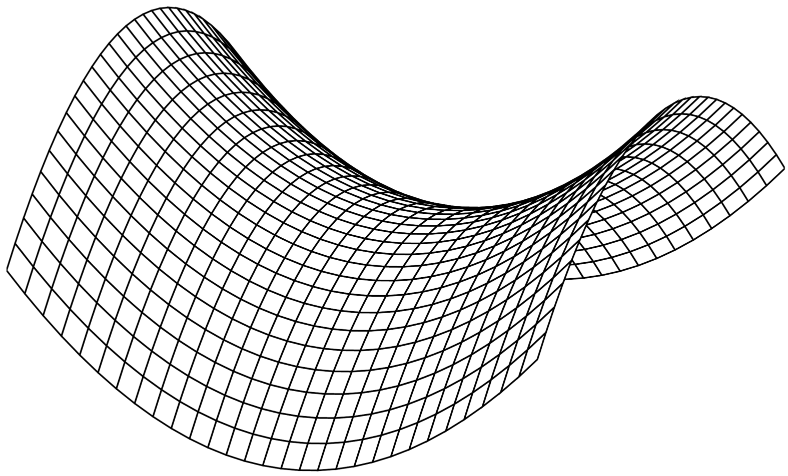

Figure 10: Saddle point of $u = f(x, y)$.

Note that if $ac = b^2$, the test is inconclusive and provides no information. We assumed that $a \neq 0$; if on the other hand $a = 0$ and $ac < b^2$ the point is a saddle point and if $a = 0$ and $ac = b^2$, the test is also inconclusive. (We cannot have $a = 0$ and $ac > b^2$ of course).

The results may be summarised in a table:-

LOCAL MAXIMUM	LOCAL MINIMUM	SADDLE POINT
If (a)$\dfrac{\partial f}{\partial x} = \dfrac{\partial f}{\partial y} = 0$ and (b)$\dfrac{\partial^2 f}{\partial x^2}\dfrac{\partial^2 f}{\partial y^2} > (\dfrac{\partial^2 f}{\partial x \partial y})^2$ and (c) $\dfrac{\partial^2 f}{\partial x^2} < 0$ then the point (x, y) is a local maximum	If (a)$\dfrac{\partial f}{\partial x} = \dfrac{\partial f}{\partial y} = 0$ and (b)$\dfrac{\partial^2 f}{\partial x^2}\dfrac{\partial^2 f}{\partial y^2} > (\dfrac{\partial^2 f}{\partial x \partial y})^2$ and (c) $\dfrac{\partial^2 f}{\partial x^2} > 0$ then the point (x, y) is a local minimum	If (a)$\dfrac{\partial f}{\partial x} = \dfrac{\partial f}{\partial y} = 0$ and (b)$\dfrac{\partial^2 f}{\partial x^2}\dfrac{\partial^2 f}{\partial y^2} < (\dfrac{\partial^2 f}{\partial x \partial y})^2$ then the point (x, y) is a saddle point

The matrix, H, defined by

$$H(x, y) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} \\ \dfrac{\partial^2 f}{\partial x \partial y} & \dfrac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

is called the **Hessian matrix**. The condition that

$$\frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} > \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2$$

is equivalent to the statement that the determinant of the Hessian matrix is positive. So we have $\det(H) > 0$ for a maximum or minimum, $\det(H) < 0$ for a saddle point.

Example 9.1 Find all the stationary points of the function

$$f(x, y) = 8x^3 - 2xy^2 - 9x^2 + 6y^2$$

and determine their nature.

Solution We have

$$\frac{\partial f}{\partial x} = 24x^2 - 2y^2 - 18x, \quad \frac{\partial f}{\partial y} = -4xy + 12y$$

For stationary points,

$$24x^2 - 2y^2 - 18x = 0 \tag{12}$$

and

$$-4xy + 12y = 0 \tag{13}$$

From (13)

$$-4y(x-3) = 0.$$

Therefore

$$y = 0 \quad \text{or } x = 3$$

Putting $y = 0$ in (12) gives

$$24x^2 - 18x = 0$$

$$\text{i.e. } 6x(4x-3) = 0.$$

$$\text{Therefore } x = 0 \text{ or } 3/4.$$

Putting $x = 3$ in (12) gives

$$\begin{aligned} 24 \times 9 - 2y^2 - 18 \times 3 &= 0. \\ \text{i.e. } 2y^2 &= 162 \\ \text{Therefore } y &= \pm 9 \end{aligned}$$

Hence the stationary points are

$$(0,0), (3/4, 0), (3, 9), (3, -9)$$

Now

$$\frac{\partial^2 f}{\partial x^2} = 48x - 18, \quad \frac{\partial^2 f}{\partial x \partial y} = -4y, \quad \frac{\partial^2 f}{\partial y^2} = -4x + 12.$$

Constructing a table:-

POINT	$\frac{\partial^2 f}{\partial x^2}$	$\frac{\partial^2 f}{\partial y^2}$	$\frac{\partial^2 f}{\partial x \partial y}$	$\det(H)$	CONCLUSION
$(0,0)$	-18	12	0	-216	SADDLE POINT
$(3/4,0)$	18	9	0	162	LOCAL MINIMUM
$(3,9)$	126	0	-36	-36^2	SADDLE POINT
$(3,-9)$	126	0	36	-36^2	SADDLE POINT

○

Example 9.2 Show that a rectangular closed box of given surface area, S, which has maximum volume is a cube.

Solution Let x, y and z be the lengths of the sides of the box and let V denote its volume. Then

$$V = xyz \tag{14}$$

and

$$S = 2(xy + yz + zx) \tag{15}$$

where x, y and z are not all independent variables. We take x and y as independent variables and compute z from (15). This gives

$$(y + x)z = \frac{S}{2} - xy.$$

So

$$z = \frac{S - 2xy}{2(y + x)} \tag{16}$$

Substituting for z in (14) gives

$$V = \frac{xy(S - 2xy)}{2(y + x)}.$$

For a stationary point, $\dfrac{\partial V}{\partial x} = 0$ and $\dfrac{\partial V}{\partial y} = 0$. Now

$$\frac{\partial V}{\partial x} = \frac{1}{2}\left\{\frac{(y + x)(Sy - 4xy^2) - xy(S - 2xy)}{(y + x)^2}\right\}.$$

Therefore $\dfrac{\partial V}{\partial x} = 0$ when

$$y(y + x)(S - 4xy) - xy(S - 2xy) = 0.$$

i.e. when

$$y(Sy + Sx - 4xy^2 - 4x^2y - Sx + 2x^2y) = 0.$$

i.e. when

$$y(Sy - 4xy^2 - 2x^2y) = 0.$$

i.e. when

$$y^2(S - 4xy - 2x^2) = 0.$$

Clearly $y = 0$ will not give a maximum volume. Therefore, for a maximum,

$$S - 4xy - 2x^2 = 0. \tag{17}$$

By symmetry, putting $\frac{\partial V}{\partial y} = 0$ will yield

$$S - 4xy - 2y^2 = 0. \tag{18}$$

By subtracting (18) from (17)

$$-2x^2 + 2y^2 = 0.$$

Therefore

$$y^2 = x^2.$$

i.e.

$$y = \pm x.$$

Since x and y are the lengths of the box, they must both be positive and so $y = x$. Putting $y = x$ in (17) gives

$$S - 6x^2 = 0,$$

i.e.

$$x^2 = S/6.$$

Therefore

$$y = x = \sqrt{\frac{S}{6}}$$

Finally, substituting for y and x in (16)

$$z = \frac{S - \frac{2S}{6}}{2\left(2\sqrt{\frac{S}{6}}\right)} = \frac{\frac{S}{6}}{\sqrt{\frac{S}{6}}} = \sqrt{\frac{S}{6}}.$$

Hence $x = y = z = \sqrt{\frac{S}{6}}$ and $V = (\frac{S}{6})^{3/2}$. Is this a maximum volume? Evaluating the second derivatives and applying the standard tests would be very tedious in this case. We can argue on physical grounds that there must be a maximum volume for a box of fixed surface area. Since this is the only candidate it must be the maximum. ○

(The video covers the principles of determining stationary points and solves the first example in this section. V

Now you should try several examples involving the determination of stationary points and the identification of their nature.) EX

10 Use of Lagrange Multipliers

(This section is not covered in the videos.)

In the last example of the previous section we found the maximum volume of a rectangular closed box of given surface area. This reduced to finding the maximum of $V = xyz$ subject to $S - 2(xy + yz + zx) = 0$. This is an example of a constrained maximisation problem and the second equation is called a **constraint**. We were able to eliminate z using the constraint and then solve the problem by putting $\dfrac{\partial V}{\partial x}$ and $\dfrac{\partial V}{\partial y}$ equal to zero.

We can express this type of problem more generally by saying that we wish to find the stationary point of

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

subject to the constraint that $c(x, y, z) = 0$. In principle, we solve the constraint equation for z, so that z is expressed as a function of x and y: i.e. $c(x, y, z) = 0$ gives $z = z(x, y)$ and then

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

The stationary points are given by solving the equations $\dfrac{\partial w}{\partial x} = 0$ and $\dfrac{\partial w}{\partial y} = 0$. These equations give

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial x} = 0 \quad \text{and} \quad \frac{\partial f}{\partial y} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial y} = 0 \tag{19}$$

We have to be very careful with the notation here. $\dfrac{\partial f}{\partial x}$ means differentiate $f(x, y, z)$ with respect to x ignoring the fact that z itself depends on x. On the other hand $\dfrac{\partial w}{\partial x}$ means differentiate $f(x, y, z(x, y))$ with respect to x taking account of the fact that z is a function of x and y.

However, it may not always be easy, or indeed possible, to solve the constraint equation for z (or x or y). The **Method of Lagrange Multipliers** neatly circumvents this problem. Lagrange introduced this method in his famous paper on mechanics, written when he was nineteen.

We let

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

where λ is called a **Lagrange Multiplier**; it is a constant. Then the stationary values of the constrained problem are the solutions of

$$\frac{\partial g}{\partial x} = 0, \quad \frac{\partial g}{\partial y} = 0, \quad \frac{\partial g}{\partial z} = 0 \quad \text{and} \quad c(x, y, z) = 0 \tag{20}$$

In fact these equations also determine the value of λ but this is usually of little interest.

We now prove that these equations have the same solutions for x and y as the equations (19)

Substituting for $g(x, y, z)$ in the equations (20) we have

$$\begin{aligned} \frac{\partial f}{\partial x} + \lambda \frac{\partial c}{\partial x} &= 0 \quad (21) \\ \frac{\partial f}{\partial y} + \lambda \frac{\partial c}{\partial y} &= 0 \quad (22) \\ \frac{\partial f}{\partial z} + \lambda \frac{\partial c}{\partial z} &= 0 \quad (23) \end{aligned}$$

From (23)

$$\lambda = -\frac{\partial f}{\partial z} / \frac{\partial c}{\partial z} \quad (24)$$

Putting equation (24) into equation (21) gives

$$0 = \frac{\partial f}{\partial x} + \left(-\frac{\partial f}{\partial z} / \frac{\partial c}{\partial z} \right) \frac{\partial c}{\partial x} \quad (25)$$

Since $c(x, y, z) = 0$, then differentiating with respect to x we obtain

$$\frac{\partial c}{\partial x} + \frac{\partial c}{\partial z}\frac{\partial z}{\partial x} = 0.$$

From this

$$\frac{\partial c}{\partial z} = -\frac{\partial c}{\partial x} / \frac{\partial z}{\partial x} \quad (26)$$

Now equations (25) and (26) give

$$0 = \frac{\partial f}{\partial x} + \left(\frac{\partial f}{\partial z}\frac{\partial z}{\partial x} / \frac{\partial c}{\partial x} \right) \frac{\partial c}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial x}$$

This is the first of equations (19). In a similar way the second of equations (19) may be obtained. It follows that equations (20) have the same solutions as equations (19). □

Example 10.1 Find the stationary points of $w = x^2 + y^2 + z^2$ subject to $x + 2y = z - 3$.

Solution Firstly, rewrite the constraint as $x + 2y - z + 3 = 0$. Let

$$g = x^2 + y^2 + z^2 + \lambda(x + 2y - z + 3).$$

Then $\dfrac{\partial g}{\partial x} = \dfrac{\partial g}{\partial y} = \dfrac{\partial g}{\partial z} = 0$ give

$$\begin{aligned} 2x + \lambda &= 0 & (27)\\ 2y + 2\lambda &= 0 & (28)\\ 2z - \lambda &= 0 & (29) \end{aligned}$$

We solve these together with the constraint

$$x + 2y - z + 3 = 0 \tag{30}$$

to find the values of x, y and z.

From (29), $\lambda = 2z$ and so (27,28) give

$$2x + 2z = 0$$

and

$$2y + 4z = 0.$$

Thus $x = -z$ and $y = -2z$. Substituting in (30) yields

$$-z - 4z - z + 3 = 0.$$

Therefore

$$6z = 3.$$

Hence

$$z = 1/2, \quad x = -1/2, \quad y = -1$$

○

Note that the numerical value of λ is not of any interest although it is sometimes convenient to obtain it in the elimination process.

In the last example, z can easily be eliminated to give

$$w = x^2 + y^2 + (x + 2y + 3)^2.$$

The equations $\dfrac{\partial w}{\partial x} = \dfrac{\partial w}{\partial y} = 0$ then give

$$2x + 2(x + 2y + 3) = 0$$

and

$$2y + 4(x + 2y + 3) = 0$$

and these equations have the same solution, $x = -1/2$, $y = -1$, that we have already found. From the constraint

$$z = x + 2y + 3,$$

we obtain $z = 1/2$ as before.

The method of Lagrange multipliers extends to a function of n variables

$$w = f(x_1, x_2, \ldots, x_n)$$

subject to m constraints:-

$$\begin{aligned} c_1(x_1, x_2, \ldots, x_n) &= 0 \\ c_2(x_1, x_2, \ldots, x_n) &= 0 \\ &\vdots \\ c_m(x_1, x_2, \ldots, x_n) &= 0 \end{aligned}$$

where $m < n$. The constrained stationary values of w are determined by the solutions of

$$\begin{aligned} \frac{\partial w}{\partial x_1} + \lambda_1 \frac{\partial c_1}{\partial x_1} + \lambda_2 \frac{\partial c_2}{\partial x_1} + \ldots + \lambda_m \frac{\partial c_m}{\partial x_1} &= 0 \\ \frac{\partial w}{\partial x_2} + \lambda_1 \frac{\partial c_1}{\partial x_2} + \lambda_2 \frac{\partial c_2}{\partial x_2} + \ldots + \lambda_m \frac{\partial c_m}{\partial x_2} &= 0 \\ &\vdots \\ \frac{\partial w}{\partial x_n} + \lambda_1 \frac{\partial c_1}{\partial x_n} + \lambda_2 \frac{\partial c_2}{\partial x_n} + \ldots + \lambda_m \frac{\partial c_m}{\partial x_n} &= 0 \end{aligned}$$

together with the constraint equations. Each constraint requires its own Lagrange multiplier.

Note that the method of Lagrange multipliers locates stationary points without determining their nature. (i.e. as local maxima, minima and saddle points). In practical problems it may be necessary to appeal to physical arguments to determine the nature of the points found by the Lagrange method; the second derivative test is frequently too unwieldly to be applied. It remains the case that, if you are asked to locate a maximum, it is not sufficient to simply produce stationary points with no further comment.

(At this point you should try to solve some constrained maximum and minimum problems using Lagrange multipliers.) EX

11 Curve Fitting

(This section is not covered in the videos.)

Suppose we have a set of n points

$$\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$$

which are thought to satisfy a linear relationship of the form $y = a + bx$. In practice, due to experimental or rounding errors for example, the points will not all lie exactly on a straight line. The situation is illustrated by Figure 11

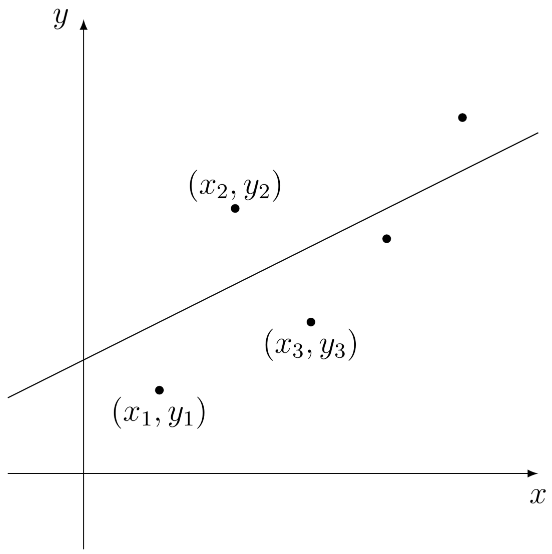

Figure 11: Data points.

The question arises as to how we determine the "best" straight line that fits this data. We can use our techniques from partial differentiation to find values for a and b so that $y = a + bx$ is the "best" straight line. But first we need to define precisely what we mean by "best".

We shall assume that the x-values, $(x_1, x_2, \ldots, x_n)$ are exactly correct and we let the line be $y = f(x) = a + bx$. Because of experimental or rounding errors, the values of y obtained from this formula will differ from the values $y_1, y_2, \ldots, y_n$ obtained experimentally. In other words, $f(x_i) = a + bx_i$ will not be exactly equal to y_i. Let

$$d_i = y_i - f(x_i) \quad i = 1, \ldots, n$$

so that d_i represents the vertical distance from the line to the point (x_i, y_i). See Figure 12.

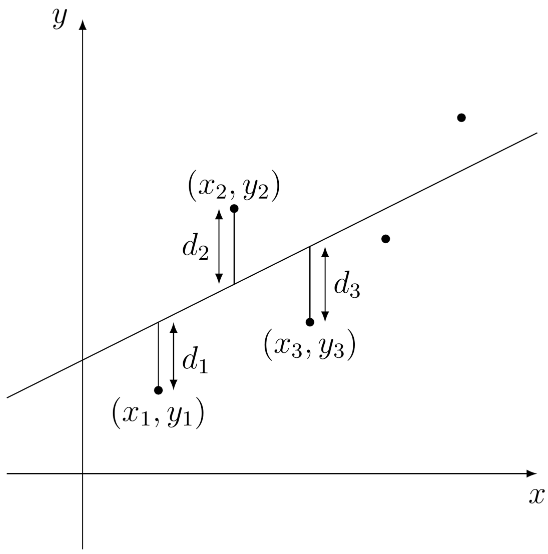

Figure 12: Distances from the line to points (x_i, y_i).

We want to choose a and b to minimise the d_i's. But if we try to minimise the sum of all the d_i's the negative and positive ones will tend to cancel out. Instead we minimise the sum of all the d_i^2's; the **sum of the squared errors**. Let

$$\begin{aligned} S &= \sum_{i=1}^{n} [y_i - f(x_i)]^2 \\ &= \sum_{i=1}^{n} [y_i - a - bx_i]^2 \end{aligned}$$

This expression for S is minimised by setting $\frac{\partial S}{\partial a}$ and $\frac{\partial S}{\partial b}$ equal to 0. Note that x_i and y_i here are fixed numbers, not variables. Carrying out the partial differentiation:-

$$\frac{\partial S}{\partial a} = \sum_{i=1}^{n} 2[y_i - a - bx_i](-1) = 0.$$

i.e.

$$\sum_{i=1}^{n} [y_i - a - bx_i] = 0$$

or

$$na + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \tag{31}$$

[note that

$$\sum_{i=1}^{n} a = \underbrace{a + a + \cdots + a}_{n \text{ terms}} = na]$$

Likewise,

$$\frac{\partial S}{\partial b} = \sum_{i=1}^{n} 2[y_i - a - bx_i](-x_i) = 0.$$

or

$$a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i \tag{32}$$

Equations (31) and (32) are two simultaneous equations, called the **normal equations**, which determine a and b. The method is referred to as the "method of least squares" and the straight line obtained is called the **regression line of y on x.**

Note that we have not shown that the stationary point obtained is a minimum. This can be verified using the second derivative test. (Alternatively, note that there is only one stationary point and S clearly has a local minimum; so this must be it!)

Example 11.1 A rod is heated to various temperatures (measured in deg C) and its length in millimetres is measured:-

Temperature, T	10	20	30	40	50	60	70
Length, l	962.3	962.5	962.6	962.9	963.0	963.2	963.4

Assuming that there is no error in the temperature measurement find the best straight line, $l = a + bT$, which fits this data.

Solution Here T is the independent variable and l the dependent variable. We need to calculate

$$\sum_{i=1}^{7} T_i, \quad \sum_{i=1}^{7} l_i, \quad \sum_{i=1}^{7} T_i^2 \text{ and } \sum_{i=1}^{7} l_i T_i.$$

We draw up a table

	T_i	l_i	l_iT_i	T_i^2
	10	962.3	9623	100
	20	962.5	19250	400
	30	962.6	28878	900
	40	962.9	38516	1600
	50	963.0	48150	2500
	60	963.2	57792	3600
	70	963.4	67438	4900
$\sum$	280	6739.9	269647	14000

The normal equations become

$$7a + 280b = 6739.9 \tag{33}$$

and

$$280a + 14000b = 269647 \tag{34}$$

Eliminating a,

$$(40 \times 280 - 14000)b = 40 \times 6739.9 - 269647$$

and so

$$b = 0.01821 \text{ to 4 s.f.}$$

Eliminating b,

$$(50 \times 7 - 280)a = 50 \times 6739.9 - 269647$$

and so

$$a = 962.1 \text{ to 4 s.f.}$$

Hence, to 4 s.f., the best straight line is

$$l = 962.1 + 0.0182T.$$

○

The method extends to fitting curves other than straight lines. For example to fit the parabola $y = a + bx + cx^2$ to the data points (x_i, y_i) we need to minimise

$$S = \sum_{i=1}^{n} [y_i - a - bx_i - cx_i^2]^2$$

by equating $\dfrac{\partial S}{\partial a}, \dfrac{\partial S}{\partial b}$ and $\frac{\partial S}{\partial c}$ to zero. This gives

$$\begin{aligned}
\frac{\partial S}{\partial a} &= \sum_{i=1}^{n} 2[y_i - a - bx_i - cx_i^2](-1) = 0 \\
\frac{\partial S}{\partial b} &= \sum_{i=1}^{n} 2[y_i - a - bx_i - cx_i^2](-x_i) = 0 \\
\frac{\partial S}{\partial c} &= \sum_{i=1}^{n} 2[y_i - a - bx_i - cx_i^2](-x_i^2) = 0
\end{aligned}$$

The normal equations are then

$$\begin{aligned}
na + b\sum_{i=1}^{n} x_i + c\sum_{i=1}^{n} x_i^2 &= \sum_{i=1}^{n} y_i \\
a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 + c\sum_{i=1}^{n} x_i^3 &= \sum_{i=1}^{n} x_i y_i \\
a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i^3 + c\sum_{i=1}^{n} x_i^4 &= \sum_{i=1}^{n} x_i^2 y_i
\end{aligned}$$

Solution of these simultaneous equations determines a, b and c.

(Now you should try some examples on curvefitting.) EX

12 Summary

When you have completed this package you should be able to do the things listed below.

- understand what is meant by the **partial derivatives** of a function of several variables,
- obtain the **first order partial derivatives** of a function of several variables,
- interpret **geometrically** the first order partial derivatives of a function of two variables,
- obtain **higher derivatives** of a function of several variables,
- understand the **small increments formula** for functions of several variables,
- apply the small increments formula to **approximations and errors**,
- use the **chain rule for partial derivatives** to relate the derivatives of a function with respect to transformed variables to its partial derivatives with respect to its original variables,
- carry out **implicit differentiation** using the methods of partial differentiation,
- obtain the **Taylor series** for functions of two variables,
- determine the **stationary points** of a function of two variables and investigate their nature,
- apply the **method of Lagrange multipliers** to constrained maxima and minima problems,
- determine the **regression line** for a set of data points,
- fit **simple polynomial curves** to a set of data points.

13 Bibliography

The following textbooks cover the basic prerequisites for this package (differentiation, Taylor series and stationary points of a function of one variable) as well as an elementary treatment of partial differentiation

Jeffrey, A. Mathematics for Engineers and Scientists, Van Nostrand, 1989.

Thomas, G. B. and Finney, R. L. Calculus and Analytic Geometry, Addison-Wesley, 1988.

Larson, R. E., Hostetler, R. P. and Edwards, B. H. Calculus, D. C. Heath and Company, 1990.

Gilbert, J. Guide to Mathematical Methods, Macmillan, 1991.

There are many other textbooks which cover partial differentiation and its applications and many of these are in the University library. Your tutor should be able to advise you which textbooks are suitable for your own needs, but you need never be short of an alternative approach or more questions to try!

14 Appendix - Video Summaries

There are four videos associated with the topic of partial differentiation. The presenter is Mike Grannell from the Department of Mathematics at the University of Central Lancashire. We recommend that you read the preamble to these notes which makes some suggestions about how you should approach viewing the videos.

Video title: Partial Differentiation (part 1). (32 minutes)

Summary

1. Examples of functions of several variables.

2. Geometric representation of functions of two variables as surfaces. Right-handed axes. The example

$$u = 100 - x^2 - y^2.$$

3. **Partial Derivatives:** Definition of partial derivatives and how to obtain them. The examples

$$\begin{aligned} u &= x^3 - 2x^2y + 3xy^2 - y^4 \\ u &= (x^2 + 4y^2)^{1/2} \end{aligned}$$

4. Geometric interpretation of the partial derivatives $\dfrac{\partial u}{\partial x}, \dfrac{\partial u}{\partial y}$ as the gradients of the surface $u = f(x, y)$ in the x and y directions.

Video title: Partial Differentiation (part 2). (34 minutes)

Summary

1. Notation for partial derivatives

2. Definitions of and notation for **second order partial derivatives.** The example
$$u = x^3 - 2x^2y + 3xy^2 - y^4$$

3. The **mixed derivative theorem.** Conditions for
$$\frac{\partial^2 u}{\partial x \partial y} = \frac{\partial^2 u}{\partial y \partial x}$$

4. The **small increments formula** for $u = f(x, y)$
$$\delta u \approx \frac{\partial u}{\partial x}\delta x + \frac{\partial u}{\partial y}\delta y.$$

5. Approximations of functions using the small increments formula. The example $u = (x^2 + 4y^2)^{1/2}$ at $x = 2.97$ and $y = 2.01$

6. Use of the small increments formula to estimate errors. The example $g = 4\pi^2 l/T^2$ when l, T are subject to 1% and 2% errors respectively.

Video title: Partial Differentiation (part 3). (29 minutes)

Summary

1. The **chain rule for one independent variable.** The example: Find $\dfrac{dV}{dt}$ when $V = \frac{1}{3}\pi r^2 h$, $\dfrac{dr}{dt} = 2$ and $\dfrac{dh}{dt} = 3$.

2. The **chain rule for two independent variables.** The example: If $u = f(x, y)$, $x = e^\alpha \cosh\beta$, $y = e^\alpha \sinh\beta$ prove that

$$\frac{\partial u}{\partial \alpha} = x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} \text{ and } \frac{\partial u}{\partial \beta} = y\frac{\partial u}{\partial x} + x\frac{\partial u}{\partial y}.$$

3. Use of partial derivatives to perform **implicit differentiation.** The example: Find $\dfrac{dy}{dx}$ in terms of x and y when $y = \sin([2x - y]^2)$.

Video title: Partial Differentiation (part 4). (37 minutes)

Summary

1. **Taylor's theorem** for a function of two variables; notation and relation to small increments formula. The example: Expand $f(x, y) = e^{xe^y}$ as a Taylor series about $(0, 0)$ up to and including second order terms. (Two methods are demonstrated: a) direct use of Taylor's theorem for two variables and b) repeated use of Taylor's theorem for one variable).

2. **Stationary points**; local maxima, minima and saddle points.

3. **Testing stationary points**; statement of the test. The example

$$f(x, y) = 8x^3 - 2xy^2 - 9x^2 + 6y^2.$$