Chapter 8

Differentiation of Functions of Several Variables

We conclude with two chapters which are really left over from last year’s calculus course, and which should help to remind you of the techniques you met then. We shall mainly be concerned with differentiation and integration of functions of more than one variable. We describe

- how each process can be done;
- why it is interesting, in terms of applications; and
- how to interpret the process geometrically.

In this chapter we concentrate on differentiation, and in the last one, move on to integration.

8.1 Functions of Several Variables

Last year, you did a significant amount of work studying functions, typically written as $y = f(x)$, which represented the variation that occurred in some (dependent) variable $y$, as another (independent) variable, $x$, changed. For example you might have been interested in the height $y$ after a given time $x$, or the area $y$ enclosed by a rectangle with sides $x$ and $10 - x$. Once the function was known, the usual rules of calculus could be applied, and results such as the time when the particle hits the ground, or the maximum possible area of the rectangle, could be calculated. In the earlier part of the course, we have extended this work by taking a more rigorous approach to a lot of the same ideas.

We are going to do the same thing now for functions of several variables. For example the height $y$ of a particle may depend on the position $x$ and the time $t$, so we have $y = f(x, t)$; the volume $V$ of a cylinder depends on the radius $r$ of the base and its height $h$, and indeed, as you know, $V = \pi r^2 h$; or the pressure of a gas may depend on its volume $V$ and temperature $T$, so $P = P(V, T)$. Note the trick I have just used: it is often convenient to use $P$ both for the (defendant) variable, and for the function itself: we don’t always need separate symbols as in the $y = f(x)$ example.

When studying the real world, it is unusual to have functions which depend solely on a single variable. Of course the single variable situation is a little simpler to study, which is
why we started with it last year. And just as last year, we shall usually have a “standard”
function name; instead of $y = f(x)$, we often work with $z = f(x, y)$, since most of the extra
complications occur when we have two, rather than one (independent) variable, and we
don’t need to consider more general cases like $w = f(x, y, z)$, or even $y = f(x_1, x_2, \ldots, x_n)$.

**Graphing functions of Several Variables**

One way we tried to understand the function $y = f(x)$ was by drawing its graph, as shown
in Fig 8.1. We then used such a graph to pick out points such as the local minimum at
$x = 3/2$, and to see how we could get the same result using calculus.

Working with two or more independent variables is more complicated, but the ideas are
familiar. To plot $z = f(x, y)$ we think of $z$ as the **height** of the function $f$ at the point
$(x, y)$, and then try to sketch the resulting surface in three dimensions. So we represent a
function as a **surface** rather than a curve.

**8.1. Example.** Sketch the surface given by $z = 2 - x/2 - 2y/3$

**Solution.** We know the surface will be a plane, because $z$ is a linear function of $x$ and $y$.
Thus it is enough to plot three points that the plane passes through. This gives Fig. 8.2.

![Figure 8.2: Sketching a function of two variables](image-url)
8.1. FUNCTIONS OF SEVERAL VARIABLES

difficulties when dealing with more complicated functions, which make sketching and visualisation rather harder than for functions of one variable. And if there are three or more independent variables, there is really no good way of visualising the behaviour of the function directly. But for just two independent variables, there are some tricks.

8.2. Example. Sketch the surface given by \( z = x^2 - y^2 \).

Solution. We can represent the surface directly by drawing it as shown in Fig. 8.3.

![Surface plot of \( z = x^2 - y^2 \).](image)

Such a representation is easy to create using suitable software and Fig. 8.3 shows the resulting surface. We now describe how to looking at similar examples without such a program. One approach is to draw a contour map of the surface, and then use the usual tricks to visualise the surface.

For the surface \( z = x^2 - y^2 \), the points where \( z = 0 \) lie on \( x^2 = y^2 \), so form the lines \( y = x \) and \( y = -x \). We can continue in this way, and look at the points where \( z = 1 \); so \( x^2 - y^2 = 1 \). This is one of the hyperbolae shown Fig. 8.4; indeed, fixing \( z \) at different values shows the contours (lines of constant height or \( z \) value) are all the same shape, but with different constants. We this get the alternative representation as a contour map shown in Fig. 8.4.

A final way to confirm that you have the right view of the surface is to section it in different planes. So far we have looked at the intersection of the planes \( z = k \) with the surface \( z = x^2 - y^2 \) for different values of the constant \( k \). If instead we fix \( x \), at the value \( a \), then \( z = a^2 - y^2 \). Each of these curves is a parabola with its vertex upwards, at the point \( y = 0, z = a^2 \).

8.3. Exercise. By looking at the curves where \( z \) is constant, or otherwise, sketch the surface given by \( z = \sqrt{x^2 + y^2} \).
CHAPTER 8. DIFFERENTIATION OF FUNCTIONS OF SEVERAL VARIABLES

Figure 8.4: Contour plot of the surface $z = x^2 - y^2$. The missing points near the $x$-axis are an artifact of the plotting program.

Continuity

As you might expect, we say that a function $f$ of two variables is \textbf{continuous} at $(x_0, y_0)$ if

$$\lim_{x \to x_0, y \to y_0} f(x, y) = f(x_0, y_0).$$

The only complication comes when we realise that there are many different ways if which $x \to x_0$ and $y \to y_0$. We illustrate with a simple example.

8.4. Example. Investigate the continuity of $f(x, y) = \frac{2xy}{x^2 + y^2}$ at the point $(0, 0)$.

Solution. Consider first the case when $x \to 0$ along the $x$-axis, so that throughout the process, $y = 0$. We have

$$f(x, 0) = \frac{2x \cdot 0}{x^2 + 0} = 0 \to 0 \text{ as } x \to 0.$$  

Next consider the case when $x \to 0$ and $y \to 0$ on the line $y = x$, so we are looking at the special case when $x = y$. We have

$$f(x, x) = \frac{2x^2}{x^2 + x^2} = 1 \to 1 \text{ as } x \to 0.$$  

Of course $f$ is only continuous if it has the same limit however $x \to 0$ and $y \to 0$, and we have now seen that it doesn’t; so $f$ is not continuous at $(0, 0)$.

Although we won’t go into it, the usual “putting together” theorems show that $f$ is continuous everywhere else.
8.2 Partial Differentiation

The usual rules for differentiation apply when dealing with several variables, but we now require to treat the variables one at a time, keeping the others constant. It is for this reason that a new symbol for differentiation is introduced. Consider the function

\[ f(x, y) = \frac{2y}{y + \cos x} \]

We can consider \( y \) fixed, and so treat it as a constant, to get a partial derivative

\[ \frac{\partial f}{\partial x} = \frac{2y \sin x}{(2y + \cos x)^2} \]

where we have differentiated with respect to \( x \) as usual. Or we can treat \( x \) as a constant, and differentiate with respect to \( y \), to get

\[ \frac{\partial f}{\partial y} = \frac{(2y + \cos x)2 - 2y \cdot 2}{(2y + \cos x)^2} = \frac{2 \cos x}{(2y + \cos x)^2}. \]

Although a partial derivative is itself a function of several variables, we often want to evaluate it at some fixed point, such as \((x_0, y_0)\). We thus often write the partial derivative as

\[ \frac{\partial f}{\partial x}(x_0, y_0). \]

There are a number of different notations in use to try to help understanding in different situations. All of the following mean the same thing:-

\[ \frac{\partial f}{\partial x}(x_0, y_0), \quad f_1(x_0, y_0), \quad f_x(x_0, y_0) \quad \text{and} \quad D_1 F(x_0, y_0). \]

Note also that there is a simple definition of the derivative in terms of a Newton quotient:-

\[ \frac{\partial f}{\partial x} = \lim_{\delta x \to 0} \frac{f(x_0 + \delta x, y_0) - f(x_0, y_0)}{\delta x} \]

provided of course that the limit exists.

8.5 Example. Let \( z = \sin(x/y) \). Compute \( x \frac{\partial z}{\partial x} + y \frac{\partial z}{\partial y} \).

Solution. Treating first \( y \) and then \( x \) as constants, we have

\[ \frac{\partial z}{\partial x} = \frac{1}{y} \cos \left( \frac{x}{y} \right) \quad \text{and} \quad \frac{\partial z}{\partial y} = \frac{-x}{y^2} \cos \left( \frac{x}{y} \right). \]

thus

\[ x \frac{\partial z}{\partial x} + y \frac{\partial z}{\partial y} = x \cos \left( \frac{x}{y} \right) - \frac{x}{y} \cos \left( \frac{x}{y} \right) = 0. \]

Note: This equation is an equation satisfied by the function we started with, which involves both the function, and its partial derivatives. We shall meet a number of examples of such a partial differential equation later.
8.6. Exercise. Let \( z = \log(x/y) \). Show that \( x \frac{\partial z}{\partial x} + y \frac{\partial z}{\partial y} = 0 \). The fact that the last two function satisfy the same differential equation is not a coincidence. With our next result, we can see that for any suitably differentiable function \( f \), the function \( z(x, y) = f(x/y) \) satisfies this partial differential equation.

8.7. Exercise. Let \( z = f(x/y) \), where \( f \) is suitably differentiable. Show that \( x \frac{\partial z}{\partial x} + y \frac{\partial z}{\partial y} = 0 \).

Because the definitions are really just version of the 1-variable result, these examples are quite typical; most of the usual rules for differentiation apply in the obvious way to partial derivatives exactly as you would expect. But there are variants. Here is how we differentiate compositions.

8.8. Theorem. Assume that \( f \) and all its partial derivatives \( f_x \) and \( f_y \) are continuous, and that \( x = x(t) \) and \( y = y(t) \) are themselves differentiable functions of \( t \). Let \( F(t) = f(x(t), y(t)) \).

Then \( F \) is differentiable and
\[
\frac{dF}{dt} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt}.
\]

Proof. Write \( x = x(t) \), \( x_0 = x(t_0) \) etc. Then we calculate the Newton quotient for \( F \).
\[
F(t) - F(t_0) = f(x, y) - f(x_0, y_0)
= f(x, y) - f(x_0, y) + f(x_0, y) - f(x_0, y_0)
= \frac{\partial f}{\partial x}(\xi, y)(x - x_0) + \frac{\partial f}{\partial y}(x_0, \eta)(y - y_0)
\]
Here we have used the Mean Value Theorem (5.18) to write
\[
f(x, y) - f(x_0, y) = \frac{\partial f}{\partial x}(\xi, y)(x - x_0)
\]
for some point \( \xi \) between \( x \) and \( x_0 \), and have argued similarly for the other part. Note that \( \xi \), pronounced “Xi” is the Greek letter “x”; in the same way \( \eta \), pronounced “Eta” is the Greek letter “y”. Thus
\[
\frac{F(t) - F(t_0)}{t - t_0} = \frac{\partial f}{\partial x}(\xi, y) \frac{(x - x_0)}{t - t_0} + \frac{\partial f}{\partial y}(x_0, \eta) \frac{(y - y_0)}{t - t_0}
\]
Now let \( t \to t_0 \), and note that in this case, \( x \to x_0 \) and \( y \to y_0 \); and since \( \xi \) and \( \eta \) are trapped between \( x \) and \( x_0 \), and \( y \) and \( y_0 \) respectively, then also \( \xi \to x_0 \) and \( \eta \to y_0 \). The result then follows from the continuity of the partial derivatives.

8.9. Example. Let \( f(x, y) = xy \), and let \( x = \cos t \), \( y = \sin t \). Compute \( \frac{df}{dt} \) when \( t = \pi/2 \).

Solution. From the chain rule,
\[
\frac{df}{dt} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} = -y(t) \sin t + x(t) \cos t = -1 \cdot \sin(\pi/2) = -1.
\]

The chain rule easily extends to the several variable case; only the notation is complicated. We state a typical example
8.10. Proposition. Let \( x = x(u, v), \ y = y(u, v) \) and \( z = z(u, v) \), and let \( f \) be a function defined on a subset \( U \subseteq \mathbb{R}^3 \), and suppose that all the partial derivatives of \( f \) are continuous. Write

\[
F(u, v) = f(x(u, v), y(u, v), z(u, v)).
\]

Then

\[
\frac{\partial F}{\partial u} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial u} + \frac{\partial f}{\partial z} \frac{\partial z}{\partial u}
\quad \text{and} \quad
\frac{\partial F}{\partial v} = = \frac{\partial f}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial v} + \frac{\partial f}{\partial z} \frac{\partial z}{\partial v}.
\]

The introduction of the domain of \( f \) above, simply to show it was a function of three variables is clumsy. We often do it more quickly by saying

Let \( f(x, y, z) \) have continuous partial derivatives.

This has the advantage that you are reminded of the names of the variables on which \( f \) acts, although strictly speaking, these names are not bound to the corresponding places. This is an example where we adopt the notation which is very common in engineering maths. But note the confusion if you ever want to talk about the value \( f(y, z, x) \), perhaps to define a new function \( g(x, y, z) \).

8.11. Example. Assume that \( f(u, v, w) \) has continuous partial derivatives, and that

\[
u = x - y, \quad v = y - z \quad w = z - x.
\]

Let

\[
F(x, y, z) = f(u(x, y, z), v(x, y, z), w(x, y, z)).
\]

Show that

\[
\frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} + \frac{\partial F}{\partial z} = 0.
\]

Solution. We apply the chain rule, noting first that from the change of variable formulae, we have

\[
\frac{\partial u}{\partial x} = 1, \quad \frac{\partial v}{\partial x} = 0, \quad \frac{\partial w}{\partial x} = -1,
\]
\[
\frac{\partial u}{\partial y} = -1, \quad \frac{\partial v}{\partial y} = 1, \quad \frac{\partial w}{\partial y} = 0,
\]
\[
\frac{\partial u}{\partial z} = 0, \quad \frac{\partial v}{\partial z} = -1, \quad \frac{\partial w}{\partial z} = 1.
\]

Then

\[
\frac{\partial F}{\partial x} = \frac{\partial f}{\partial u} \frac{\partial u}{\partial x} 1 + \frac{\partial f}{\partial v} \frac{\partial v}{\partial x} 0 + \frac{\partial f}{\partial w} \frac{\partial w}{\partial x} - 1 = \frac{\partial f}{\partial u} - \frac{\partial f}{\partial w},
\]
\[
\frac{\partial F}{\partial y} = \frac{\partial f}{\partial u} \frac{\partial u}{\partial y} - 1 + \frac{\partial f}{\partial v} \frac{\partial v}{\partial y} 1 + \frac{\partial f}{\partial w} \frac{\partial w}{\partial y} 0 = -\frac{\partial f}{\partial u} + \frac{\partial f}{\partial v},
\]
\[
\frac{\partial F}{\partial z} = \frac{\partial f}{\partial u} \frac{\partial u}{\partial z} 0 + \frac{\partial f}{\partial v} \frac{\partial v}{\partial z} - 1 + \frac{\partial f}{\partial w} \frac{\partial w}{\partial z} 1 = -\frac{\partial f}{\partial v} + \frac{\partial f}{\partial w}.
\]

Adding then gives the result claimed.
8.3 Higher Derivatives

Note that a partial derivative is itself a function of two variables, and so further partial derivatives can be calculated. We write

$$\frac{\partial^2 f}{\partial x \partial x} = \frac{\partial^2 f}{\partial x^2}, \quad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial x \partial y}, \quad \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial y \partial x}, \quad \frac{\partial^2 f}{\partial y \partial y} = \frac{\partial^2 f}{\partial y^2}.$$  

This notation generalises to more than two variables, and to more than two derivatives in the way you would expect. There is a complication that does not occur when dealing with functions of a single variable; there are four derivatives of second order, as follows:

$$\frac{\partial^2 f}{\partial x \partial x}, \quad \frac{\partial^2 f}{\partial x \partial y}, \quad \frac{\partial^2 f}{\partial y \partial x} \quad \text{and} \quad \frac{\partial^2 f}{\partial y \partial y}.$$  

Fortunately, when \( f \) has mild restrictions, the order in which the differentiation is done doesn’t matter.

8.12. Proposition. Assume that all second order derivatives of \( f \) exist and are continuous. Then the mixed second order partial derivatives of \( f \) are equal. i.e.

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

8.13. Example. Suppose that \( f(x, y) \) is written in terms of \( u \) and \( v \) where \( x = u + \log v \) and \( y = u - \log v \). Show that, with the usual convention,

$$\frac{\partial^2 f}{\partial u \partial u} = \frac{\partial^2 f}{\partial x^2} + 2 \frac{\partial^2 f}{\partial x \partial y} + \frac{\partial^2 f}{\partial y^2},$$  

and

$$v^2 \frac{\partial^2 f}{\partial v \partial v} = \frac{\partial f}{\partial y} - \frac{\partial f}{\partial x} + \frac{\partial^2 f}{\partial x \partial x} + 2 \frac{\partial^2 f}{\partial x \partial y} + \frac{\partial^2 f}{\partial y^2}.$$  

You may assume that all second order derivatives of \( f \) exist and are continuous.

Solution. Using the chain rule, we have

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

and

$$\frac{\partial f}{\partial v} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial v} = \frac{1}{v} \frac{\partial f}{\partial x} + \frac{1}{v} \frac{\partial f}{\partial y}.$$  

Thus using both these and their operator form, we have

$$\frac{\partial^2 f}{\partial u^2} = \frac{\partial}{\partial u} \left( \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \right) = \frac{\partial}{\partial x} \frac{\partial f}{\partial x} + \frac{\partial}{\partial x} \frac{\partial f}{\partial y} + \frac{\partial}{\partial y} \frac{\partial f}{\partial x} + \frac{\partial}{\partial y} \frac{\partial f}{\partial y} = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial x \partial y} + \frac{\partial^2 f}{\partial y \partial x} + \frac{\partial^2 f}{\partial y^2},$$  

while differentiating with respect to \( v \), we have

$$\frac{\partial^2 f}{\partial v^2} = \frac{\partial}{\partial v} \left( \frac{1}{v} \frac{\partial f}{\partial x} + \frac{1}{v} \frac{\partial f}{\partial y} \right) = - \frac{1}{v^2} \frac{\partial f}{\partial x} + \frac{1}{v} \frac{\partial}{\partial v} \left( \frac{\partial f}{\partial x} \right) + \frac{1}{v^2} \frac{\partial f}{\partial y} - \frac{1}{v} \frac{\partial}{\partial v} \left( \frac{\partial f}{\partial y} \right) = - \frac{1}{v^2} \frac{\partial f}{\partial x} + \frac{1}{v} \left( \frac{1}{v^2} \frac{\partial^2 f}{\partial x^2} - \frac{1}{v} \frac{\partial^2 f}{\partial x \partial y} \right) + \frac{1}{v^2} \frac{\partial f}{\partial y} - \frac{1}{v} \left( \frac{1}{v} \frac{\partial^2 f}{\partial x \partial y} - \frac{1}{v} \frac{\partial^2 f}{\partial y^2} \right) = \frac{1}{v^2} \left( \frac{\partial f}{\partial y} - \frac{\partial f}{\partial x} \right) + \frac{\partial^2 f}{\partial x \partial y} - \frac{2}{v^2} \frac{\partial^2 f}{\partial x \partial y} + \frac{\partial^2 f}{\partial y^2}.$$
8.4 Solving equations by Substitution

One of the main interests in partial differentiation is because it enables us to write down how we expect the natural world to behave. We move away from 1-variable results as soon as we have properties which depend on e.g. at least one space variable, together with time. We illustrate with just one example, designed to whet the appetite for the whole subject of mathematical physics.

Assume the displacement of a length of string at time $t$ from its rest position is described by the function $f(x, t)$. This is illustrated in Fig 8.4. The laws of physics describe how the string behaves when released in this position, and allowed to move under the elastic forces of the string; the function $f$ satisfies the wave equation

$$\frac{\partial^2 f}{\partial t^2} = c^2 \frac{\partial^2 f}{\partial x^2}.$$

![Figure 8.5: A string displaced from the equilibrium position](image)

8.14. Example. Solve the equation

$$\frac{\partial^2 F}{\partial u \partial v} = 0.$$

Solution. Such a function is easy to integrate, because the two variables appear independently. So $\frac{\partial F}{\partial v} = g_1(v)$, where $g_1$ is an arbitrary (differentiable) function. since when differentiated with respect to $u$ we are given that $\frac{\partial^2 F}{\partial u \partial v} = 0$. Thus we can integrate with respect to $v$ to get

$$F(u, v) = \int g_1(v) dv + h(u) = g(v) + h(u),$$

where $h$ is also an arbitrary (differentiable) function.

8.15. Example. Rewrite the wave equation using co-ordinates $u = x - ct$ and $v = x + ct$.

Solution. Write $f(x, t) = F(u, v)$ and now in principle confuse $F$ with $f$, so we can tell them apart only by the names of their arguments. In practice we use different symbols to help the learning process; but note that in a practical case, all the $F$'s that appear below, would normally be written as $f$'s. By the chain rule

$$\frac{\partial}{\partial x} = \left(\frac{\partial}{\partial u}\right).1 + \left(\frac{\partial}{\partial v}\right).1 \quad \text{and} \quad \frac{\partial}{\partial t} = \left(\frac{\partial}{\partial u}\right).(-c) + \left(\frac{\partial}{\partial v}\right).c.$$
differentiating again, and using the operator form of the chain rule as well,

\[
\frac{\partial^2 f}{\partial t^2} = \left( c \frac{\partial}{\partial v} - c \frac{\partial}{\partial u} \right) \left( c \frac{\partial F}{\partial v} - c \frac{\partial F}{\partial u} \right) = c^2 \frac{\partial^2 F}{\partial v^2} - c^2 \frac{\partial^2 F}{\partial u \partial v} - c^2 \frac{\partial^2 F}{\partial v \partial u} + c^2 \frac{\partial^2 F}{\partial u^2}
\]

and similarly

\[
\frac{\partial^2 f}{\partial x^2} = \left( \frac{\partial^2 F}{\partial v^2} + \frac{\partial^2 F}{\partial u^2} \right) + 2c^2 \frac{\partial^2 F}{\partial u \partial v}.
\]

Substituting in the wave equation, we thus get

\[
4c^2 \frac{\partial^2 F}{\partial u \partial v} = 0,
\]

an equation which we have already solved. Thus solutions to the wave equation are of the form \(f(u) + g(v)\) for any (suitably differentiable) functions \(f\) and \(g\). For example we may have \(\sin(x - ct)\). Note that this is not just any function; it must be constant when \(x = ct\).

8.16. Exercise. Let \(F(x, t) = \log(2x + 2ct)\) for \(x > -ct\), where \(c\) is a fixed constant. Show that

\[
\frac{\partial^2 F}{\partial t^2} - c^2 \frac{\partial^2 F}{\partial x^2} = 0.
\]

Note that this is simply checking a particular case of the result we have just proved.

8.5 Maxima and Minima

As in one variable calculations, one use for derivatives in several variables is in calculating maxima and minima. Again as for one variable, we shall rely on the theorem that if \(f\) is continuous on a closed bounded subset of \(\mathbb{R}^2\), then it has a global maximum and a global minimum. And again as before, we note that these must occur either at a local maximum or minimum, or else on the boundary of the region. Of course in \(\mathbb{R}\), the boundary of the region usually consisted of a pair of end points, while in \(\mathbb{R}^2\), the situation is more complicated. However, the principle remains the same. And we can test for local maxima and minima in the same way as for one variable.

8.17. Definition. Say that \(f(x, y)\) has a critical point at \((a, b)\) if and only if

\[
\frac{\partial f}{\partial x}(a, b) = \frac{\partial f}{\partial y}(a, b) = 0.
\]

It is clear by comparison with the single variable result, that a necessary condition that \(f\) have a local extremum at \((a, b)\) is that it have a critical point there, although that is not a sufficient condition. We refer to this as the first derivative test.
8.5. MAXIMA AND MINIMA

We can get more information by looking at the second derivative. Recall that we gave a number of different notations for partial derivatives, and in what follows we use \( f_x \) rather than the more cumbersome \( \frac{\partial f}{\partial x} \) etc. This idea extends to higher derivatives; we shall use \( f_{xx} \) instead of \( \frac{\partial^2 f}{\partial x^2} \), and \( f_{xy} \) instead of \( \frac{\partial^2 f}{\partial x \partial y} \) etc.

8.18. Theorem (Second Derivative Test). Assume that \((a, b)\) is a critical point for \( f \).

Then

- If, at \((a, b)\), we have \( f_{xx} < 0 \) and \( f_{xx} f_{yy} - f_{xy}^2 > 0 \), then \( f \) has a local maximum at \((a, b)\).
- If, at \((a, b)\), we have \( f_{xx} > 0 \) and \( f_{xx} f_{yy} - f_{xy}^2 > 0 \), then \( f \) has a local minimum at \((a, b)\).
- If, at \((a, b)\), we have \( f_{xx} f_{yy} - f_{xy}^2 < 0 \), then \( f \) has a saddle point at \((a, b)\).

The test is inconclusive at \((a, b)\) if \( f_{xx} f_{yy} - f_{xy}^2 = 0 \), and the investigation has to be continued some other way.

Note that the discriminant is easily remembered as

\[
\Delta = \begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix} = f_{xx} f_{yy} - f_{xy}^2
\]

A number of very simple examples can help to remember this. After all, the result of the test should work on things where we can do the calculation anyway!

8.19. Example. Show that \( f(x, y) = x^2 + y^2 \) has a minimum at \((0, 0)\).

Solution. We have \( f_x = 2x; f_y = 2y \), so \( f_x = f_y \) precisely when \( x = y = 0 \), and this is the only critical point. We have \( f_{xx} = f_{yy} = 2; f_{xy} = 0 \), so \( \Delta = f_{xx} f_{yy} - f_{xy}^2 = 4 > 0 \) and there is a local minimum at \((0, 0)\).

8.20. Exercise. Let \( f(x, y) = xy \). Show there is a unique critical point, which is a saddle point

Proof. We give an indication of how the theorem can be derived — or if necessary how it can be remembered. We start with the two dimensional version of Taylor’s theorem, see section 5.6. We have

\[
f(a + h, b + k) \sim f(a, b) + h \frac{\partial f}{\partial x}(a, b) + k \frac{\partial f}{\partial y}(a, b) + \frac{1}{2} \left( h^2 \frac{\partial^2 f}{\partial x^2} + 2kh \frac{\partial^2 f}{\partial x \partial y} + k^2 \frac{\partial^2 f}{\partial y^2} \right)
\]

where we have actually taken an expansion to second order and assumed the corresponding remainder is small.

We are looking at a critical point, so for any pair \((h, k)\), we have \( \frac{\partial f}{\partial x}(a, b) + k \frac{\partial f}{\partial y}(a, b) = 0 \) and everything hinges on the behaviour of the second order terms. It is thus enough to study the behaviour of the quadratic \( Ah^2 + 2Bhk + Ck^2 \), where we have written

\[
A = \frac{\partial^2 f}{\partial x^2}, \quad B = \frac{\partial^2 f}{\partial x \partial y}, \quad \text{and} \quad C = \frac{\partial^2 f}{\partial y^2}.
\]
Assuming that $A \neq 0$ we can write
\[ Ah^2 + 2Bhk + Ck^2 = A \left( h + \frac{Bk}{A} \right)^2 + \left( C - \frac{B^2}{A} \right)k^2 \]

where we write $\Delta = CA - B^2$ for the discriminant. We have thus expressed the quadratic as the sum of two squares. It is thus clear that
- if $A < 0$ and $\Delta > 0$ we have a local maximum;
- if $A > 0$ and $\Delta > 0$ we have a local minimum; and
- if $\Delta < 0$ then the coefficients of the two squared terms have opposite signs, so by going out in two different directions, the quadratic may be made either to increase or to decrease.

Note also that we could have completed the square in the same way, but starting from the $k$ term, rather than the $h$ term; so the result could just as easily be stated in terms of $C$ instead of $A$.

8.21. Example. Let $f(x, y) = 2x^3 - 6x^2 - 3y^2 - 6xy$. Find and classify the critical points of $f$. By considering $f(x, 0)$, or otherwise, show that $f$ does not achieve a global maximum.

Solution. We have $f_x = 6x^2 - 12x - 6y$ and $f_y = -6y - 6x$. Thus critical points occur when $y = -x$ and $x^2 - x = 0$, and so at $(0, 0)$ and $(1, -1)$. Differentiating again, $f_{xx} = 12x - 12$, $f_{yy} = -6$ and $f_{xy} = -6$. Thus the discriminant is $\Delta = -6(12x - 12) - 36$. When $x = 0$, $\Delta = 36 > 0$ and since $f_{xx} = -12$, we have a local maximum at $(0, 0)$. When $x = 1$, $\Delta = -36 < 0$, so there is a saddle at $(1, -1)$.

To see there is no global maximum, note that $f(x, 0) = 2x^3(1 - 3/x) \to \infty$ as $x \to \infty$, since $x^3 \to \infty$ as $x \to \infty$.

8.22. Exercise. Find the extrema of $f(x, y) = xy - x^2 - y^2 - 2x - 2y + 4$.

8.23. Example. An open-topped rectangular tank is to be constructed so that the sum of the height and the perimeter of the base is 30 metres. Find the dimensions which maximise the surface area of the tank. What is the maximum value of the surface area? [You may assume that the maximum exists, and that the corresponding dimensions of the tank are strictly positive.]

Solution. Let the dimensions of the box be as shown.

Let the area of the surface of the material be $S$. Then
\[ S = 2xh + 2yh + xy, \]

and since, from our restriction on the base and height,
\[ 30 = 2(x + y) + h, \quad \text{we have} \quad h = 30 - 2(x + y). \]

Substituting, we have
\[ S = 2(x + y)\left(30 - 2(x + y)\right) + xy = 60(x + y) - 4(x + y)^2 + xy, \]
and for physical reasons, $S$ is defined for $x \geq 0$, $y \geq 0$ and $x + y \leq 15$.

A global maximum (which we are given exists) can only occur on the boundary of the domain of definition of $S$, or at a critical point, when $\frac{\partial S}{\partial x} = \frac{\partial S}{\partial y} = 0$. On the boundary of the domain of definition of $S$, we have $x = 0$ or $y = 0$ or $x + y = 15$, in which case $h = 0$. We are given that we may ignore these cases. Now

$$S = -4x^2 - 4y^2 - 7xy + 60x + 60y,$$

so

$$\frac{\partial S}{\partial x} = -8x - 7y + 60 = 0,$$

$$\frac{\partial S}{\partial y} = -8y - 7x + 60 = 0.$$

Subtracting gives $x = y$ and so $15x = 60$, or $x = y = 4$. Thus $h = 14$ and the surface area is $S = 16(-4 - 4 - 7 + 15 + 15) = 240$ square metres. Since we are given that a maximum exists, this must be it. [If both sides of the surface are counted, the area is doubled, but the critical proportions are still the same.]

Sometimes a function necessarily has an absolute maximum and absolute minimum — in the following case because we have a continuous function defined on a closed bounded subset of $\mathbb{R}^2$, and so the analogue of 4.35 holds. In this case exactly as in the one variable case, we need only search the boundary (using ad - hoc methods, which in fact reduce to 1-variable methods) and the critical points in the interior, using our ability to find local maxima.

8.24. Example. Find the absolute maximum and minimum values of

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

on the triangular plate in the first quadrant bounded by the lines $x = 0$, $y = 0$ and $y = 9 - x$.

Solution. We know there is a global maximum, because the function is continuous on a closed bounded subset of $\mathbb{R}^2$. Thus the absolute max will occur either in the interior, at a critical point, or on the boundary. If $y = 0$, investigate $f(x, 0) = 2 + 2x - x^2$, while if $x = 0$, investigate $f(0, y) = 2 + 2y - y^2$. If $y = 9 - x$, investigate

$$f(x, 9 - x) = 2 + 2x + 2(9 - x) - x^2 - (9 - x)^2$$
for an absolute maximum. In fact extreme may occur when \((x, y) = (0, 1)\) or \((1, 0)\) or \((0, 0)\) or \((9, 0)\) or \((0, 9)\), or \((9/2, 9/2)\). At these points, \(f\) takes the values \(-41/2, 2, 3, -61\).

Next we seek critical points in the interior of the plate,
\[ f_x = 2 - 2x = 0 \quad \text{and} \quad f_y = 2 - 2y = 0. \]
so \((x, y) = (1, 1)\) and \(f(1, 1) = 4\), so this must be the global maximum. Can check also using the second derivative test, that it is a local maximum.

### 8.6 Tangent Planes

Consider the surface \(F(x, y, z) = c\), perhaps as \(z = f(x, y)\), and suppose that \(f\) and \(F\) have continuous partial derivatives. Suppose now we have a smooth curve on the surface, say \(\phi(t) = (x(t), y(t), z(t))\). Then since the curve lies in the surface, we have
\[ F(x(t), y(t), z(t)) = c, \]
and so, applying the chain rule, we have
\[
\frac{dF}{dt} = \frac{\partial F}{\partial x} \frac{dx}{dt} + \frac{\partial F}{\partial y} \frac{dy}{dt} + \frac{\partial F}{\partial z} \frac{dz}{dt} = 0.
\]
or, writing this in terms of vectors, we have
\[
\nabla F \cdot v(t) = \left( \frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}, \frac{\partial F}{\partial z} \right) \cdot \left( \frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} \right) = 0.
\]
Since the RH vector is the velocity of a point on the curve, which lies on the surface, we see that the left hand vector must be the normal to the curve (i.e. to \(\nabla F\)). Thus \((x - x_0, y - y_0, z - z_0)\) and \(\nabla F\) are perpendicular, and that requirement is the equation which gives the tangent plane.

Note that we have defined the gradient vector \(\nabla F\) associated with the function \(F\) by
\[
\nabla F = \left( \frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}, \frac{\partial F}{\partial z} \right).
\]

#### 8.25. Theorem. The tangent to the surface \(F(x, y, z) = c\) at the point \((x_0, y_0, z_0)\) is given by
\[
\frac{\partial F}{\partial x} (x - x_0) + \frac{\partial F}{\partial y} (y - y_0) + \frac{\partial F}{\partial z} (z - z_0) = 0.
\]

**Proof.** This is a simple example of the use of vector geometry. Given that \((x_0, y_0, z_0)\) lies on the surface, and so in the tangent, then for any other point \((x, y, z)\) in the tangent plane, the vector \((x - x_0, y - y_0, z - z_0)\) must lie in the tangent plane, and so must be normal to the normal to the curve (i.e. to \(\nabla F\)). Thus \((x - x_0, y - y_0, z - z_0)\) and \(\nabla F\) are perpendicular, and that requirement is the equation which gives the tangent plane. \(\square\)

#### 8.26. Example. Find the equation of the tangent plane to the surface
\[ F(x, y, z) = x^2 + y^2 + z - 9 = 0 \]
at the point \(P = (1, 2, 4)\).
8.7. Linearisation and Differentials

We obtained a geometrical view of the function \( f(x, y) \) by considering the surface \( z = f(x, y) \), or \( F(x, y, z) = z - f(x, y) = 0 \). Note that the tangent plane to this surface at the point \( (x_0, y_0, f(x_0, y_0)) \) lies close to the surface itself. Just as in one variable, we used the tangent line to approximate the graph of a function, so we shall use the tangent plane to approximate the surface defined by a function of two variables.

The equation of our tangent plane is

\[
\frac{\partial F}{\partial x}(x - x_0) + \frac{\partial F}{\partial y}(y - y_0) + \frac{\partial F}{\partial z}(z - f(x_0, y_0)) = 0,
\]

and writing the derivatives in terms of \( f \), we have

\[
\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) + (-1)(z - f(x_0, y_0)) = 0,
\]

or, writing in terms of \( z \), the height of the tangent plane above the ground plane,

\[
z = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0).
\]

Our assumption that the tangent plane lies close to the surface is that \( z \approx f(x, y) \), or that

\[
f(x, y) \approx f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0).
\]

We call the right hand side the linear approximation to \( f \) at \( (x_0, y_0) \).

We can rewrite this with \( h = x - x_0 \) and \( k = y - y_0 \), to get

\[
f(x, y) - f(x_0, y_0) \approx h \frac{\partial f}{\partial x}(x_0, y_0) + k \frac{\partial f}{\partial y}(x_0, y_0), \quad \text{or} \quad df \approx h \frac{\partial f}{\partial x}(x_0, y_0) + k \frac{\partial f}{\partial y}(x_0, y_0).
\]

This has applications; we can use it to see how a function changes when its independent variables are subjected to small changes.

8.28. Example. A cylindrical oil tank is 25 m high and has a radius of 5 m. How sensitive is the volume of the tank to small variations in the radius and height.

Solution. Let \( V \) be the volume of the cylindrical tank of height \( h \) and radius \( r \). Then \( V = \pi r^2 H \), and so

\[
dV = \frac{\partial V}{\partial r} \bigg|_{(r_0, h_0)} \, dr + \frac{\partial V}{\partial h} \bigg|_{(r_0, h_0)} \, dh,
\]

\[
= 250\pi dr + 25\pi dh.
\]

Thus the volume is 10 times as sensitive to errors in measuring \( r \) as it is to measuring \( h \).

Try this with a short fat tank!
8.29. Example. A cone is measured. The radius has a measurement error of 3%, and the height an error of 2%. What is the error in measuring the volume?

Solution. The volume $V$ of a cone is given by $V = \frac{\pi r^2 h}{3}$, where $r$ is the radius of the cone, and $h$ is the height. Thus

$$dV = \frac{2}{3} \pi r h \, dr + \frac{1}{3} \pi r^2 \, dh$$

$$= 2 \cdot (0.03) \cdot \frac{\pi r^2 h}{3} + \frac{\pi r^2 h}{3} \cdot (0.02) = V (0.06 + 0.02)$$

Thus there is an 8% error in measuring the volume.

8.30. Exercise. The volume of a cylindrical oil tank is to be calculated from measured values of $r$ and $h$. What is the percentage error in the volume, if $r$ is measured with an accuracy of 2%, and $h$ measured with an accuracy of 0.5%.

8.8 Implicit Functions of Three Variables

Finally in this section we discuss another application of the chain rule. Assume we have variables $x$, $y$ and $z$, related by the equation $F(x, y, z) = 0$. Then assuming that $F_z \neq 0$, there is a version of the implicit function theorem which means we can in principle, write $z = z(x, y)$, so we can “solve for $z$”. Doing this gives

$$F(x, y, z(x, y)) = 0.$$  

Now differentiate both sides partially with respect to $x$. We get

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

and so since $\frac{\partial y}{\partial x} \neq 0$,

$$0 = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial z} \frac{\partial z}{\partial x} \quad \text{and} \quad \frac{\partial z}{\partial x} = -\frac{\partial F}{\partial F} \frac{\partial z}{\partial z}$$

Now assume in the same way that $F_x \neq 0$ and $F_y \neq 0$, so we can get two more relations like this, with $x = x(y, z)$ and $y = y(z, x)$. Then form three such equations,

$$\frac{\partial z}{\partial x} \frac{\partial x}{\partial y} \frac{\partial y}{\partial z} = -\left( \frac{\partial F}{\partial x} \right) \left( \frac{\partial F}{\partial y} \right) \left( \frac{\partial F}{\partial z} \right) = -1!$$

We met this result as Equation 1.1 in Chapter 1, when it seemed totally counter-intuitive!
Chapter 9

Multiple Integrals

9.1 Integrating functions of several variables

Recall that we think of the integral in two different ways. In one way we interpret it as the area under the graph \( y = f(x) \), while the fundamental theorem of the calculus enables us to compute this using the process of “anti-differentiation” — undoing the differentiation process.

We think of the area as

\[
\sum f(x_i) \, dx_i = \int f(x) \, dx,
\]

where the first sum is thought of as a limiting case, adding up the areas of a number of rectangles each of height \( f(x_i) \), and width \( dx_i \). This leads to the natural generalisation to several variables: we think of the function \( z = f(x, y) \) as representing the height of \( f \) at the point \((x, y)\) in the plane, and interpret the integral as the sum of the volumes of a number of small boxes of height \( z = f(x, y) \) and area \( dx_i \, dy_j \). Thus the volume of the solid of height \( z = f(x, y) \) lying above a certain region \( R \) in the plane leads to integrals of the form

\[
\int \int_R = \sum_{i=1}^{n} \sum_{j=1}^{m} f(x_i, y_j) \, dx_i \, dy_j = \lim S_{mn}.
\]

We write such a double integral as \( \int \int_R f(x, y) \, dA \).

9.2 Repeated Integrals and Fubini’s Theorem

As might be expected from the form, in which we can sum over the elementary rectangles \( dx \, dy \) in any order, the order does not matter when calculating the answer. There are two important orders — where we first keep \( x \) constant and vary \( y \), and then vary \( x \); and the opposite way round. This gives rise to the concept of the repeated integral, which we write as

\[
\int \left( \int_R f(x, y) \, dx \right) \, dy \quad \text{or} \quad \int \left( \int_R f(x, y) \, dy \right) \, dx.
\]
Our result that the order in which we add up the volume of the small boxes doesn’t matter is the following, which also formally shows that we evaluate a double integral as any of the possible repeated integrals.

9.1. Theorem (Fubini’s theorem for Rectangles). Let \( f(x, y) \) be continuous on the rectangular region \( R : a \leq x \leq b; c \leq y \leq d \). Then

\[
\int \int_R f(x, y) \, dA = \int_c^d \left( \int_a^b f(x, y) \, dx \right) \, dy = \int_a^b \left( \int_c^d f(x, y) \, dy \right) \, dx
\]

Note that this is something like an inverse of partial differentiation. In doing the first inner (or repeated) integral, we keep \( y \) constant, and integrate with respect to \( x \). Then we integrate with respect to \( y \). Of course if \( f \) is a particularly simply function, say \( f(x, y) = g(x)h(y) \), then it doesn’t matter which order we do the integration, since

\[
\int \int_R f(x, y) \, dA = \int_a^b g(x) \, dx \int_c^d h(y) \, dy.
\]

We use the Fubini theorem to actually evaluate integrals, since we have no direct way of calculating a double (as opposed to a repeated) integral.

9.2. Example. Integrate \( z = 4 - x - y \) over the region \( 0 \leq x \leq 2 \) and \( 0 \leq y \leq 1 \). Hence calculate the volume under the plane \( z = 4 - x - y \) above the given region.

Solution. We calculate the integral as a repeated integral, using Fubini’s theorem.

\[
V = \int_{x=0}^{2} \int_{y=0}^{1} (4 - x - y) \, dy \, dx = \int_{x=0}^{2} \left[ (4y - xy - y^2/2) \right]_{y=0}^{1} \, dx = \int_{x=0}^{2} (4 - x - 1/2) \, dx \text{ etc.}
\]

From our interpretation of the integral as a volume, we recognise \( V \) as volume under the plane \( z = 4 - x - y \) which lies above \( \{(x, y) \mid 0 \leq x \leq 2; 0 \leq y \leq 1\} \).

9.3. Exercise. Evaluate \( \int_0^3 \int_0^2 (4 - y^2) \, dy \, dx \), and sketch the region of integration.

In fact Fubini’s theorem is valid for more general regions than rectangles. Here is a pair of statements which extend its validity.

9.4. Theorem (Fubini’s theorem — Stronger Form). Let \( f(x, y) \) be continuous on a region \( R \)

- if \( R \) is defined as \( a \leq x \leq b; g_1(x) \leq y \leq g_2(x) \). Then

\[
\int \int_R f(x, y) \, dA = \int_a^b \left( \int_{g_1(x)}^{g_2(x)} f(x, y) \, dy \right) \, dx
\]

- if \( R \) is defined as \( c \leq y \leq d; h_1(y) \leq x \leq h_2(y) \). Then

\[
\int \int_R f(x, y) \, dA = \int_c^d \left( \int_{h_1(y)}^{h_2(y)} f(x, y) \, dx \right) \, dy
\]
9.2. REPEATED INTEGRALS AND FUBINI’S THEOREM

Proof. We give no proof, but the reduction to the earlier case is in principle simple; we just extend the function to be defined on a rectangle by making it zero on the extra bits. The problem with this as it stands is that the extended function is not continuous. However, the difficulty can be fixed.

This last form enables us to evaluate double integrals over more complicated regions by passing to one of the repeated integrals.

9.5. Example. Evaluate the integral

\[\int_1^2 \left(\int_x^2 \frac{y^2}{x^2} \, dy\right) \, dx\]

as it stands, and sketch the region of integration.

Reverse the order of integration, and verify that the same answer is obtained.

Solution. The diagram in Fig.9.1 shows the area of integration.

We first integrate in the given order.

\[
\int_1^2 \left(\int_x^2 \frac{y^2}{x^2} \, dy\right) \, dx = \int_1^2 \left[ \frac{y^3}{3x^2} \right]_x^2 \, dx = \int_1^2 \left( \frac{8}{3x^2} - \frac{x}{3} \right) \, dx \\
= \left[ \frac{8}{3x} - \frac{x^2}{6} \right]_1^2 = \left( -\frac{4}{3} + \frac{2}{3} \right) - \left( -\frac{8}{3} - \frac{1}{6} \right) = \frac{5}{6}.
\]

Reversing the order, using the diagram, gives

\[
\int_1^2 \left(\int_1^y \frac{y^2}{x^2} \, dx\right) \, dy = \int_1^2 \left[ -\frac{y^2}{x} \right]_1^y \, dy = \int_1^2 (-y + y^2) \, dy \\
= \left[ -\frac{y^2}{2} - \frac{y^3}{3} \right]_1^2 = \left( -2 + \frac{8}{3} \right) - \left( -\frac{1}{2} + \frac{1}{6} \right) = \frac{5}{6}.
\]

Thus the two orders of integration give the same answer.

Another use for the ideas of double integration just automates a procedure you would have used anyway, simply from your knowledge of 1-variable results.
9.6. Example. Find the area of the region bounded by the curve \( x^2 + y^2 = 1 \), and above the line \( x + y = 1 \).

Solution. We recognise an area as numerically equal to the volume of a solid of height 1, so if \( R \) is the region described, the area is

\[
\int \int_R 1 \, dx \, dy = \int_0^1 \left( \int_{y=1-x}^{y=\sqrt{1-x^2}} dy \right) dx = \int_0^1 \sqrt{1-x^2} - (1-x) \, dx = \ldots .
\]

And we also find that Fubini provides a method for actually calculating integrals; sometimes one way of doing a repeated integral is much easier than the other.

9.7. Example. Sketch the region of integration for

\[
\int_0^1 \left( \int_y^1 x^2 e^{xy} \, dx \right) \, dy.
\]

Evaluate the integral by reversing the order of integration.

Solution. The diagram in Fig.9.2 shows the area of integration.

Interchanging the given order of integration, we have

\[
\int_0^1 \left( \int_y^1 x^2 e^{xy} \, dx \right) \, dy = \int_0^1 \left( \int_0^x x^2 e^{xy} \, dy \right) \, dx
\]

\[
= \int_0^1 \left[ \frac{x^2 e^{xy}}{y} \right]_0^x \, dx
\]

\[
= \int_0^1 x e^{x^2} \, dx - \int_0^1 x \, dx
\]

\[
= \left[ \frac{1}{2} e^{x^2} - \frac{x^2}{2} \right]_0^1 = \frac{1}{2} (e-2).
\]

9.8. Exercise. Evaluate the integral

\[
\int \int_R \left( \sqrt{x} - y^2 \right) \, dy \, dx,
\]

where \( R \) is the region bounded by the curves \( y = x^2 \) and \( x = y^4 \).
### 9.3 Change of Variable — the Jacobian

Another technique that can sometimes be useful when trying to evaluate a double (or triple etc) integral generalise the familiar method of integration by substitution.

Assume we have a change of variable \( x = x(u, v) \) and \( y = y(u, v) \). Suppose that the region \( S' \) in the \( uv \)-plane is transformed to a region \( S \) in the \( xy \)-plane under this transformation. Define the **Jacobian** of the transformation as

\[
J(u, v) = \left| \begin{array}{cc}
\frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\
\frac{\partial y}{\partial u} & \frac{\partial y}{\partial v}
\end{array} \right| = \frac{\partial(x, y)}{\partial(u, v)}.
\]

It turns out that this correctly describes the relationship between the element of area \( dx \, dy \) and the corresponding area element \( du \, dv \).

With this definition, the change of variable formula becomes:

\[
\int \int_S f(x, y) \, dx \, dy = \int \int_{S'} f(x(u, v), y(u, v)) \left| J(u, v) \right| du \, dv.
\]

Note that the formula involves the *modulus* of the Jacobian.

#### 9.9 Example.
Find the area of a circle of radius \( R \).

**Solution.** Let \( A \) be the disc centred at 0 and radius \( R \). The area of \( A \) is thus \( \int \int_A dx \, dy \). We evaluate the integral by changing to polar coordinates, so consider the usual transformation \( x = r \cos \theta, \, y = r \sin \theta \) between Cartesian and polar co-ordinates. We first compute the Jacobian;

\[
\frac{\partial x}{\partial r} = \cos \theta, \quad \frac{\partial y}{\partial r} = \sin \theta, \quad \frac{\partial x}{\partial \theta} = -r \sin \theta, \quad \frac{\partial y}{\partial \theta} = r \cos \theta.
\]

Thus

\[
J(r, \theta) = \left| \begin{array}{cc}
\frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\
\frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta}
\end{array} \right| = \left| \begin{array}{cc}
\cos \theta & -r \sin \theta \\
\sin \theta & r \cos \theta
\end{array} \right| = r(\cos^2 \theta + \sin^2 \theta) = r.
\]

We often write this result as

\[
dA = dx \, dy = r \, dr \, d\theta
\]

Using the change of variable formula, we have

\[
\int \int_A dx \, dy = \int \int_A |J(r, \theta)| \, dr \, d\theta = \int_0^R \int_0^{2\pi} r \, dr \, d\theta = 2\pi R^2.
\]

We thus recover the usual area of a circle.

Note that the Jacobian \( J(r, \theta) = r > 0 \), so we did indeed take the modulus of the Jacobian above.

#### 9.10 Example.
Find the volume of a ball of radius 1.
CHAPTER 9. MULTIPLE INTEGRALS

Solution. Let $V$ be the required volume. The ball is the set $\{(x, y, z) \mid x^2 + y^2 + z^2 \leq 1\}$. It can be thought of as twice the volume enclosed by a hemisphere of radius 1 in the upper half plane, and so

$$V = 2 \int_D \sqrt{1 - x^2 - y^2} \, dx \, dy$$

where the region of integration $D$ consists of the unit disc $\{(x, y) \mid x^2 + y^2 \leq 1\}$. Although we can try to do this integration directly, the natural co-ordinates to use are plane polars, and so we instead do a change of variable first. As in 9.9, if we write $x = r \cos \theta$, $y = r \sin \theta$, we have $dx \, dy = r \, dr \, d\theta$. Thus

$$V = 2 \int_D \sqrt{1 - x^2 - y^2} \, dx \, dy = 2 \int_D (\sqrt{1 - r^2}) \, r \, d\theta$$

$$= 2 \int_0^{2\pi} \int_0^1 \sqrt{1 - r^2} \, dr \, d\theta$$

$$= 4\pi \left[ \frac{(1 - r^2)^{3/2}}{3} \right]_0^1$$

$$= \frac{4\pi}{3}.$$

Note that after the change of variables, the integrand is a product, so we are able to do the $dr$ and $d\theta$ parts of the integral at the same time.

And finally, we show that the same ideas work in 3 dimensions. There are (at least) two co-ordinate systems in $\mathbb{R}^3$ which are useful when cylindrical or spherical symmetry arises. One of these, **cylindrical polars** is given by the transformation

$$x = r \cos \theta, \quad y = r \sin \theta, \quad z = z,$$

and the Jacobian is easily calculated as

$$\frac{\partial (x, y, z)}{\partial (r, \theta, z)} = r \quad \text{so} \quad dV = dx \, dy \, dz = r \, dr \, d\theta \, dz.$$

The second useful co-ordinate system is **spherical polars** with transformation

$$x = r \sin \phi \cos \theta, \quad y = r \sin \phi \sin \theta, \quad z = r \cos \phi.$$

The transformation is illustrated in Fig 9.3.

It is easy to check that Jacobian of this transformation is given by

$$dV = r^2 \sin \phi \, dr \, d\phi \, d\theta = dx \, dy \, dz.$$ 

9.11. **Example.** The moment of inertia of a solid occupying the region $R$, when rotated about the $z$-axis is given by the formula

$$I = \int \int \int_R (x^2 + y^2) \rho \, dV.$$ 

Calculate the moment of inertia about the $z$-axis of the solid of unit density which lies outside the cylinder of radius $a$, inside the sphere of radius $2a$, and above the $x - y$ plane.
9.3. CHANGE OF VARIABLE — THE JACOBIAN

Figure 9.3: The transformation from Cartesian to spherical polar co-ordinates.

Figure 9.4: Cross section of the right hand half of the solid outside a cylinder of radius \( a \) and inside the sphere of radius \( 2a \)

Solution. Let \( I \) be the moment of inertia of the given solid about the \( z \)-axis. A diagram of a cross section of the solid is shown in Fig 9.4.

We use cylindrical polar co-ordinates \((r, \theta, z)\); the Jacobian gives \( dx \, dy \, dz = r \, dr \, d\theta \, dz \), so

\[
I = \int_0^{2\pi} d\theta \int_a^{2a} r \, dr \int_0^{\sqrt{4a^2-r^2}} r^2 \, dz
\]

\[
= 2\pi \int_a^{2a} r^3 \sqrt{4a^2-r^2} \, dr.
\]

We thus have a single integral. Using the substitution \( u = 4a^2 - r^2 \), you can check that the integral evaluates to \( 22\sqrt{3}\pi a^5/5 \).

9.12. Exercise. Show that

\[
\int \int \int z^2 \, dx \, dy \, dz = \frac{4}{15} \pi
\]

(where the integral is over the unit ball \( x^2 + y^2 + z^2 \leq 1 \)) first by using spherical polars, and then by doing the \( z \) integration first and using plane polars.)