Chapter 1

Partial differentiation

1.1 Functions of one variable

We begin by recalling some basic ideas about real functions of one variable. For example, the volume V of a sphere only depends on its radius r and is given by the formula

$$V = \tfrac{4}{3}\pi r^3.$$

We write $V = f(r)$, where $f(r) = \frac{4}{3}\pi r^3$ to emphasise the fact that volume is a function f of the radius (only). Two related ideas should also be recalled.

Domain In general, the domain D is the set of points at which the formula is to be calculated. In the present example, since the radius should be real and cannot be negative, the domain consists of all non-negative real numbers, $[0, \infty)$ (it is debatable whether or not 0 should be included or excluded but this is not an important issue).

To be precise when we define a real function f, we should specify not only the formula but also its domain D by writing $f\colon D \to \mathbb{R}$. If we do not specify the domain, we assume that the domain is the *maximal domain*, that is the set of *all* points at which the formula makes sense. For the present example, f is defined by

$$f\colon [0, \infty) \to \mathbb{R}, \quad \text{where} \quad f(r) = \tfrac{4}{3}\pi r^3.$$

If we were to say simply that f was defined by

$$f(r) = \tfrac{4}{3}\pi r^3,$$

then it would be assumed that the domain of f is the maximal domain which is $\mathbb{R} = (-\infty, \infty)$ since the formula makes sense for all real numbers r.

Example 1.1 What is the maximal domain of the real function g defined by $g(x) = \sqrt{x^2 + 3x + 2}$?

Solution :

Answer: The maximal domain is $(-\infty, -2] \cup [-1, \infty)$. □

Graph In general, this is the set of all ordered pairs $(a, f(a))$ where a is a point in the domain. This is usually shown as a curve in the cartesian plane. In the present example, the graph is the set of point $(r, \frac{4}{3}\pi r^3)$ for all $r \geq 0$ and is illustrated in Figure 1.1.

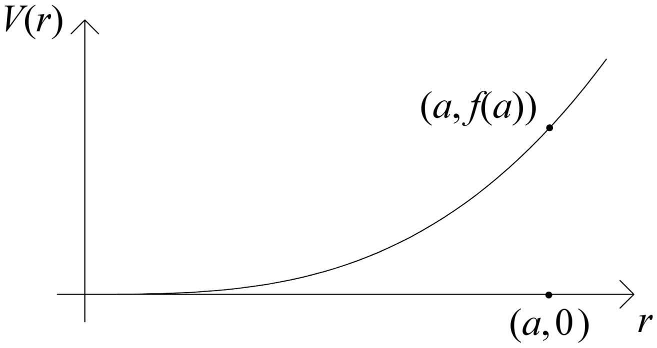

Figure 1.1: Graph of $f\colon D \to \mathbb{R}$

The volume V of a cylinder, on the other hand, depends on two dimensions, the radius r and the height h. In this case we might write $V = g(r, h)$, where $g(r, h) = \pi r^2 h$ defines a *function of two variables*. In the next section we will extend the notions of domain and graph to functions of several variables.

1.2 Functions of several variables

We will only discuss the case of two variables but the main ideas are valid for any number of variables.

Let D be a subset of $\mathbb{R}^2$, that is, a region in a plane. A typical element of D is a point (x, y). A function $f\colon D \to \mathbb{R}$ is a rule which determines a unique real number $z = f(x, y)$ for each $(x, y) \in D$. The graph of g is the set of points $(a, b, c) \in \mathbb{R}^3$ such that $(a, b) \in D$ and $c = f(a, b)$. This is typically represented as a *surface* in (three dimensional) space. Figure 1.2 illustrates this.

Similar definitions exists for functions of any number of variables but the graph of a function of more that two variables cannot be simply represented.

Remark As with real functions of one variable, we often don't give the domain of a function f of several variables explicitly; instead we assume that the domain of f is maximal.

Example 1.2 Determine the maximal domain of the function f defined by $f(x, y) = \sqrt{1 - x^2 - y^2}$.

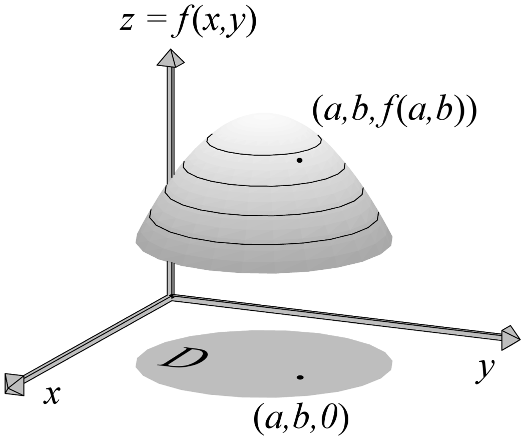

Figure 1.2: Graph of $f\colon D \to \mathbb{R}$

Solution :

Answer: $D = \{(x,y) : x^2 + y^2 \leq 1\}$, which is the unit disk centre $(0,0)$. ☐

Aids to visualisation of surfaces

In several parts of this module, and in module 2Y, it will be important to be able to visualise a surface which is either the graph of a function of two variables $z = f(x,y)$ or, more generally, is a relation $F(x,y,z) = 0$. We will here give several examples illustrating some useful techniques.

Spheres

A sphere of radius r, centre (a,b,c) consists of those points (x,y,z) which are a distance r from (a,b,c). Thus, by Pythagoras's theorem, this sphere is defined by

$$(x-a)^2 + (y-b)^2 + (z-c)^2 = r^2.$$

Furthermore, if we solve for z we get

$$z = c \pm \sqrt{r^2 - (x-a)^2 - (y-b)^2}.$$

Because of this, for any given a, b, c, the graph of a function $f(x,y) = c + \sqrt{r^2 - (x-a)^2 - (y-b)^2}$ is the "northern" hemisphere and $f(x,y) = c - \sqrt{r^2 - (x-a)^2 - (y-b)^2}$ the corresponding "southern" hemisphere.

Given an equation

$$x^2 + y^2 + z^2 + \alpha x + \beta y + \gamma z + \delta = 0,$$

one may always complete the square to write this in the form

$$(x + \tfrac{1}{2}\alpha)^2 + (y + \tfrac{1}{2}\beta)^2 + (z + \tfrac{1}{2}\gamma)^2 = \frac{1}{4}(\alpha^2 + \beta^2 + \gamma^2) - \delta$$

which defines a sphere if and only if $\frac{1}{4}(\alpha^2+\beta^2+\gamma^2)-\delta>0$.

Example 1.3 Sketch the graph of $f(x,y)=-\sqrt{1-2x-x^2-y^2}$.

Solution :

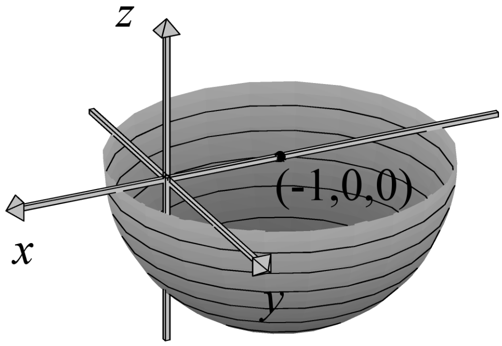

Figure 1.3: Graph of $f(x,y)=-\sqrt{1-2x-x^2-y^2}$

□

Taking cross-sections

The plane x = constant is parallel to the yz-plane and may, or may not, have a non-empty intersection with the surface $F(x,y,z)=0$. This intersection is called a *cross-section* of the surface. Typically, this cross-section will be a curve on the plane can give useful clues to the overall nature of the surface. Similarly, we may take cross-section with the planes y = constant and z = constant.

In particular, for a surface $z=f(x,y)$, the cross-section with the plane $z=c$, where c is a constant, is the curve $f(x,y)=c$ and is called a *level curve* or *contour*. The second name is used because of the close connection with contour lines on a map (lines linking points with the same height above sea-level). In this analogy, $z=f(x,y)$ represents part of the surface of the earth and each level curve represents a particular contour line on a map.

For each choice of c the level curve is denoted L_c and is the set of points (x,y) in D for which $f(x,y)$ has the value c. For different choices of c, L_c may be a curve, a point or points, or the empty set. Note that each point in the domain of f lies on a particular level curve.

Example 1.4 By considering the level curves and the cross-sections $x=0$ and $y=0$, obtain a sketch of $z=\sqrt{x^2+y^2}$.

Solution :

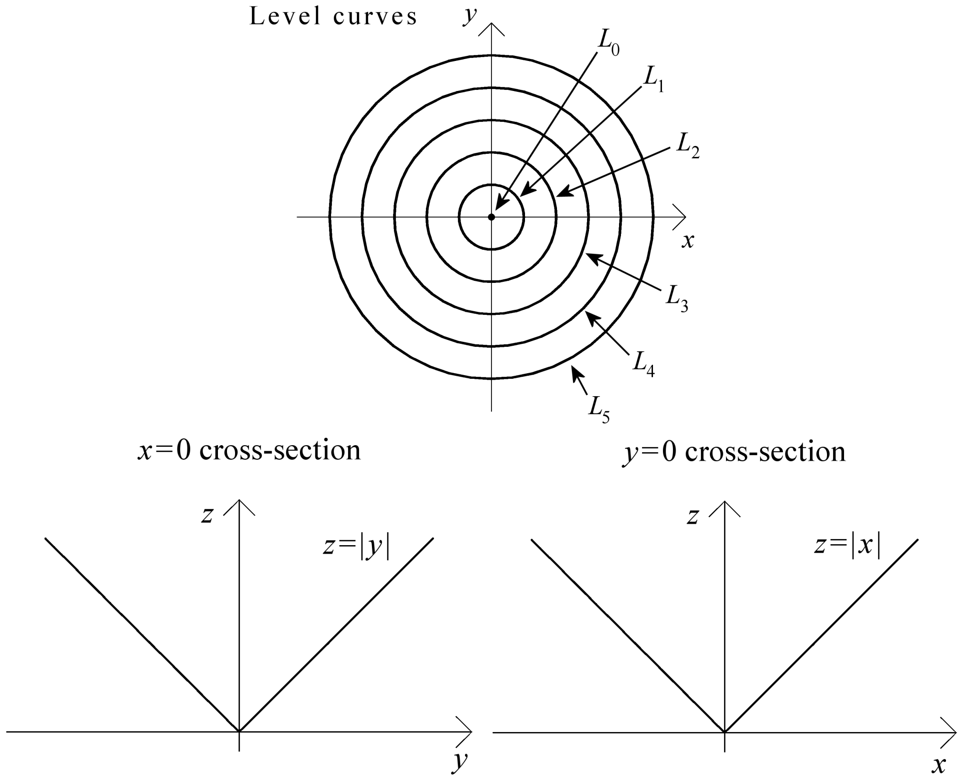

Figure 1.4: Cross sections

Putting this information together, we see that the surface defined by $z = \sqrt{x^2 + y^2}$ is a (circular) *cone* with vertex at $(0, 0)$ (Figure 1.5).

$\square$

Planes

Recall that a plane with normal vector $\mathbf{n} = (\alpha, \beta, \gamma)$ has equation $\alpha x + \beta y + \gamma z = \delta$.

In particular, the graph of $f(x, y) = ax + by + c$ is the plane $z = ax + by + c$ with normal $(a, b, -1)$ passing through the point $(0, 0, c)$. Observe that the cross-sections of a plane are either straight lines (or $\emptyset$.)

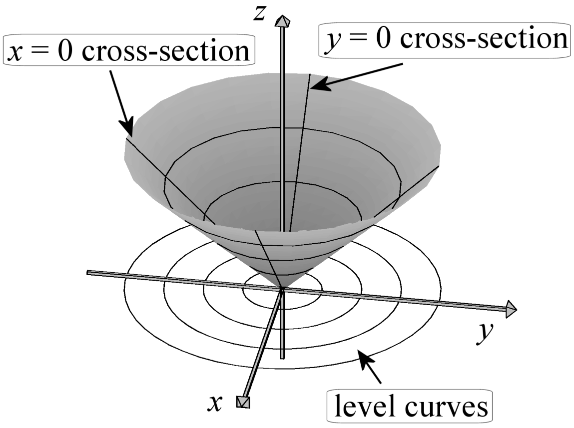

Figure 1.5: The cone $z = \sqrt{x^2 + y^2}$

Example 1.5 Sketch the part of the surface

$$2x + y + 4z = 1,$$

where $x, y, z \geq 0$.

Solution :

Answer: A sketch of the plane is shown in Figure 1.6. ☐

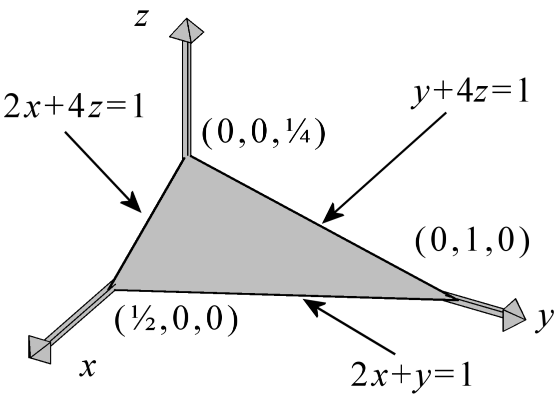

Figure 1.6: The plane $2x + y + 4z = 1$

Other surfaces

Other standard surfaces are shown in *Advanced Caclulus* - Section 138.

1.3 Partial derivatives

In this section we want to generalise, to functions of several variables, the notion of *gradient* as it is understood for functions of one variable. Recall that if the limit

$$\lim_{h\to 0}\frac{g(a+h)-g(a)}{h}$$

exists then this limit is called *derivative* of g at a. This is written as

$$\frac{dg}{dx}(a) \quad \text{or} \quad g'(a),$$

and is the gradient of the tangent to the graph of g at a point $(a, g(a))$.

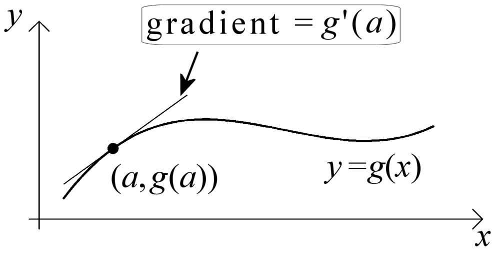

Figure 1.7: Derivative of a function of one variable

Now consider f, a function of two variables. On the surface $z = f(x, y)$, there is no single meaning of gradient. Imagine this surface to be a mountainside. When walking or skiing straight down the mountain the gradient may be very large but traversing the mountain the gradient is much less. Indeed by choosing a direction one may make the gradient have any value in between. For this reason it is necessary to define *two* gradients in terms of vertical cross-section of the surface in the x and y directions.

As in Figure 1.8, consider a point $(a, b, f(a, b))$ on the surface. Taking the cross-sections $x = a$ and $y = b$ through this point we obtain the graphs of two functions of *one* variable; $z = f(x, b) = g(x)$ (say) and $z = f(a, y) = h(y)$ (say). For each of these functions we can (provided the derivatives exist) determine gradients called the *partial x and y derivatives of f at (a, b)* and written as

$$\frac{\partial f}{\partial x}(a,b) = \text{derivative of } f(x,y) \text{ w.r.t. } x \text{ with } y \text{ held constant, evaluated at } (x,y)=(a,b). \text{ This equals } g'(a).$$

and

$$\frac{\partial f}{\partial y}(a,b) = \text{derivative of } f(x,y) \text{ w.r.t. } y \text{ with } x \text{ held constant, evaluated at } (x,y)=(a,b). \text{ This equals } h'(b).$$

For a function f of n variables $x_1, x_2, \ldots, x_n$ we define n *partial derivatives*

$$\frac{\partial f}{\partial x_i} = \text{derivative of } f(x_1,\ldots,x_n) \text{ w.r.t. } x_i \text{ with all other variables held constant.}$$

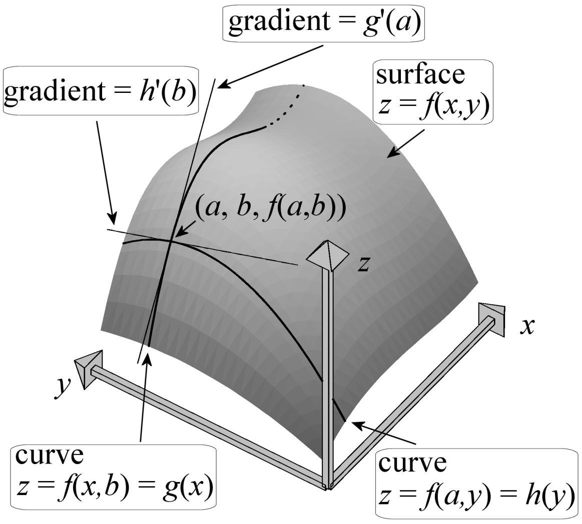

Figure 1.8: Gradients on cross-sections

Remarks

1. It is important to distinguish the notation used for partial derivatives $\dfrac{\partial f}{\partial x}$ from ordinary derivatives $\dfrac{df}{dx}$.
2. We also use subscript notation for partial derivatives. If $f = f(x, y)$ then we may write

$$\frac{\partial f}{\partial x} \equiv f_x \equiv f_1, \text{ and } \frac{\partial f}{\partial y} \equiv f_y \equiv f_2.$$

   In general, the notation f_n, where n is a positive integer, means the derivative of f with respect to its n-th argument, (with all other variables held constant). This notation is the direct analogue of the $'$ notation for ordinary derivatives. Recall we can use the chain rule to calculate

$$\frac{d}{dx} f(x^2) = f'(x^2) \frac{d}{dx}(x^2) = 2x f'(x^2).$$

   Below we carry out similar calculations involving partial derivatives.
3. Like ordinary derivatives, partial derivatives do not always exist at every point. In this module we will always assume that derivatives exist unless it is otherwise stated.
4. If $z = f(x, y)$ then the partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ can be interpreted as the gradients of the tangent lines to the surface $z = f(x, y)$ in the directions parallel to the $x-$ and $y-$axes, respectively.

Formal definition of Partial Derivative

Suppose f is a suitably well behaved function of three variables x, y, z. Then at (a, b, c),

$$\frac{\partial f}{\partial x} = \lim_{h \to 0} \frac{f(a + h, b, c) - f(a, b, c)}{h}.$$

This is by analogy with the definition of ordinary derivatives. Note how the y and z coordinates are unaffected.

Example 1.6 Find $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$ where

$$\text{(a) } f(x,y) = x^3y^2 + x, \quad \text{(b) } f(x,y) = \sin(x + xy).$$

Solution :

(a)

Answer: $f_x = 3x^2y^2 + 1$ and $f_y = 2x^3y$.

(b)

Answer: $f_x = (1 + y)\cos(x + xy)$ and $f_y = x\cos(x + xy)$. □

Example 1.7 Find $\dfrac{\partial z}{\partial x}$ where $z = \sin^{-1}\left(\dfrac{x}{x+y}\right)$ and $x, y > 0$.

[Note that $\sin^{-1} u$ is the inverse sine function (sometimes written as $\arcsin u$), and *not* the reciprocal $1/\sin u$. The domain of $\sin^{-1}$ is $[-1, 1]$ and, since $x, y > 0$, $x/(x + y)$ lies in this domain.]

Solution :

Answer: $z_x = \frac{y}{x+y}\frac{1}{\sqrt{2xy+y^2}}$. □

Example 1.8 Let $u = f(r)$ where $r^2 = x^2 + y^2 + z^2$. Show that

$$xu_x + yu_y + zu_z = rf'(r).$$

Solution :

$\square$

See *Advanced Caclulus* - Section 87 for other examples of implicit partial differentiation.

1.4 Higher order derivatives

Let u be a function of several variables $x, y, \ldots$. Then u_x (if it exists) is also a function of the same variables and so may also have partial derivatives. We define

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial}{\partial x}(u_x) = u_{xx} = u_{11}, \qquad \frac{\partial^2 u}{\partial y \partial x} = \frac{\partial}{\partial y}(u_x) = u_{xy} = u_{12},$$
$$\frac{\partial^2 u}{\partial x \partial y} = \frac{\partial}{\partial x}(u_y) = u_{yx} = u_{21}, \qquad \frac{\partial^2 u}{\partial y^2} = \frac{\partial}{\partial y}(u_y) = u_{yy} = u_{22}, \quad \text{etc.}$$

In general, $u_{xyz\ldots}$ denotes the result of taking the x-derivative, then the y-derivative, then the z-derivative, $\ldots$ of u. The total number of partial derivatives taken is called the *order* of the derivative. For example, $u_{xxy} = u_{112}$ is a third order derivative.

There is no automatic guarantee that, for example, $u_{xy} = u_{yx}$ but the following theorem (the proof of which is omitted) states the conditions under which the order in which the derivatives are taken is unimportant.

Theorem *Let u be a function of x, y such u_{xy} and u_{yx} exist and are continuous at a point (a, b). [Roughly speaking, this means that there are no hole or jump in the graphs of u_{xy} and u_{yx} at (a, b).] Then,*

$$u_{xy}(a, b) = u_{yx}(a, b).$$

Remarks

1. This result extends to functions of any number of variables and to third and higher order derivatives. For example, let u depend on three variables then, provided these derivatives exist and are continuous,

$$u_{1213} = u_{3211} = u_{2113} = \cdots = u_{1123}.$$

2. Unless otherwise stated, functions considered in this module will be assumed to have continuous partial derivatives of all orders. Hence the order in which we take partial derivatives will be unimportant.

Example 1.9 Determine all second order derivatives of $u = \sin xy$ and verify that $u_{xy} = u_{yx}$.

Solution :

Answer: The first derivatives are $u_x = y\cos xy, \quad u_y = x\cos xy$. □

See *Advanced Caclulus* - Section 86 for other examples of the product rule in partial differentiation.

Example 1.10 Let $u = f(x/y)$, where f is an arbitrary (twice differentiable, with continuous second derivative) function of one variable. Show that

$$xu_x + yu_y = 0,$$

and *deduce* that

$$x^2u_{xx} + 2xyu_{xy} + y^2u_{yy} = 0.$$

Solution :

□

1.5 The chain rule for functions of several variables

We have already made extensive use of the chain rule for functions of one variable. This is used to find the derivative of a *composition* of functions; if $F(x) = f(u(x))$ then

$$\frac{dF}{dx} = \frac{du}{dx}\frac{df}{du} = u'(x)f'(u(x)).$$

We now want to extend this technique to functions of several variables.

Theorem *Let* $F(x,y) = f\big(u(x,y), v(x,y)\big)$. *Then*

$$\frac{\partial F}{\partial x} = \frac{\partial u}{\partial x}\frac{\partial f}{\partial u} + \frac{\partial v}{\partial x}\frac{\partial f}{\partial v} \qquad \textit{and} \qquad \frac{\partial F}{\partial y} = \frac{\partial u}{\partial y}\frac{\partial f}{\partial u} + \frac{\partial v}{\partial y}\frac{\partial f}{\partial v}.$$

This is called the chain rule for functions of two variables*.*

Remarks

1. Observe the pattern

$$\frac{\partial F}{\partial x} = \frac{\boxed{\partial u}}{\partial x}\frac{\partial f}{\boxed{\partial u}} + \frac{\boxed{\partial v}}{\partial x}\frac{\partial f}{\boxed{\partial v}},$$

   [all terms on the right have ∂f on top and ∂x on bottom and ∂u or ∂v which "cancels".]

2. The chain rule is extended in an obvious way to functions of any number of variables. For example, if $F(x,y,z) = f\big(u(x,y,z), v(x,y,z), w(x,y,z)\big)$ then

$$\frac{\partial F}{\partial x} = \frac{\partial u}{\partial x}\frac{\partial f}{\partial u} + \frac{\partial v}{\partial x}\frac{\partial f}{\partial v} + \frac{\partial w}{\partial x}\frac{\partial f}{\partial w}.$$

3. There are two special cases of this formula. First, the one variable chain rule that we used above; if $F(x,y) = f(u(x,y))$ then

$$\frac{\partial F}{\partial x} = \frac{\partial u}{\partial x}\frac{df}{du}.$$

   Second, if $F(x) = f(u(x), v(x))$ then

$$\frac{dF}{dx} = \frac{du}{dx}\frac{\partial f}{\partial u} + \frac{dv}{dx}\frac{\partial f}{\partial v}.$$

   Notice that the partial derivatives in the formula become ordinary derivatives wherever the function being differentiated is a function of only one variable.

Example 1.11 Let $w = u^2 + v^2$ where $u = \sin\theta$ and $v = \cos\phi$. Use the chain rule to calculate w_θ and w_ϕ in terms of θ and ϕ.

Solution :

Answer: $w_\theta = \sin 2\theta$ and $w_\phi = -\sin 2\phi$. □

Example 1.12 Let z be a function of x and y, where y is a function of x. Hence z may also be regarded both as a function two variables (x and y) and of one variable (x).

Express $\dfrac{dz}{dx}$ in terms of $\dfrac{\partial z}{\partial x}$ and $\dfrac{\partial z}{\partial y}$.

Solution :

Answer: $\frac{dz}{dx} = \frac{\partial z}{\partial x} + \frac{dy}{dx}\frac{\partial z}{\partial y}$. ☐

Remark We can also use the chain rule to calculate higher order derivatives. This is a bit more complicated than calculating first order derivatives. This topic is considered later in connection with *partial differential equations.*

Example 1.13 Suppose f is a function of u and v, where $u = x^2y$ and $v = x + y$. So f is ultimately a function of x and y. Find $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ and $\frac{\partial^2 f}{\partial x^2}$ in terms of partial derivatives with respect to u and v.

Solution :

Answer: $\frac{\partial f}{dx} = 2xy\frac{\partial f}{\partial u} + \frac{\partial f}{\partial v}$, $\frac{\partial f}{dy} = x^2\frac{\partial f}{\partial u} + \frac{\partial f}{\partial v}$ and $\frac{\partial^2 f}{\partial x^2} = 2y\frac{\partial f}{du} + 4x^2y^2\frac{\partial^2 f}{\partial u^2} + 4xy\frac{\partial^2 f}{\partial u \partial v} + \frac{\partial^2 f}{\partial v}$. ☐