9 Partial Differentiation

9.1 Multi-variable Functions

Recall that a function is a map from one set to another, $f : A \to B$. So far, we have been interested in functions from and to some subset of the real numbers, $f : \mathbb{R} \to \mathbb{R}$. However, we can consider much more general functions than this. For example, the function $f : \mathbb{R}^2 \to \mathbb{R}, f(x, y) = x^2 + y$ maps **two** real inputs to a unique output.

To see how this function behaves, consider firstly the function $f_0 : \mathbb{R} \to \mathbb{R}, f_0(x) = x^2$. We know this function well and can plot it on a graph. Now consider the similar function $f_1(x) = x^2 + 1$. This is just a translation of f_0. Similarly, we can define $f_k(x) = f_0(x) + k$ for any value of k and have an intuitive understanding of this function. We can also draw plots of any of these functions for arbitrary values of k, as in Fig. 1.

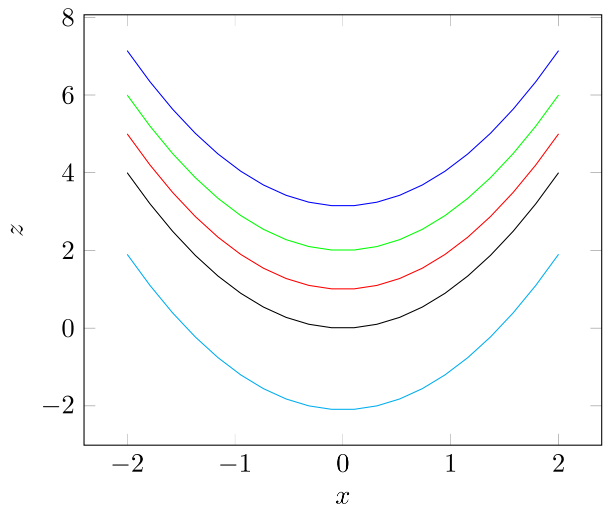


Figure 1: Plots of $z = x^2 + k$ for $k = 0$ (black), $k = 1$ (red), $k = 2$ (green), $k = \pi$ (blue) and $k = -2.1$ (cyan). These values of k are not important and were chosen to emphasise that such curves exist for **any** value of k.

Rather than colour-coding the graphs, as in Fig. 1, we could depict the value of k on a separate y axis, as in Fig. 2.

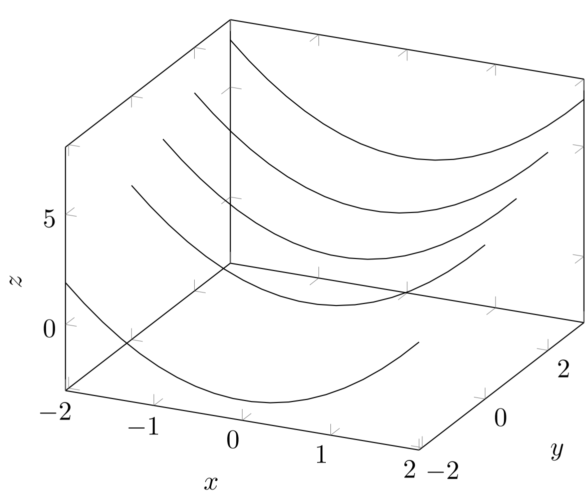


Figure 2: Plots of $z = x^2 + k$ for $k = 0, k = 1, k = 2, k = \pi, k = -2.1$ with k now shown by position on the y axis.

Since there are an infinite number of such curves, if we plot them all we build a two-dimensional surface as in Fig. 3, and we have a visual representation of our function $f(x, y) = x^2 + y$ via the 3D graph $z = f(x, y)$.

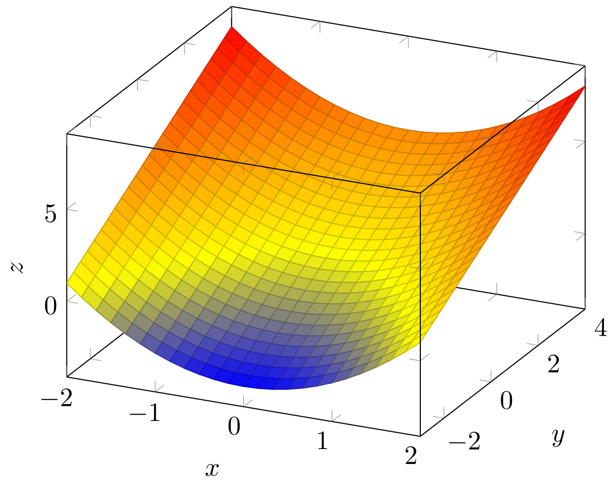


Figure 3: Plot of $z = x^2 + y$.

We approached this by considering quadratics in x with different values of y. This corresponds to looking at cross-sections of the full surface for constant values of y. Fig. 4 shows several cross-sections of the full surface that give back ordinary one-dimensional functions of the form $f_k(x) = x^2 + k$ for some constant value of k.

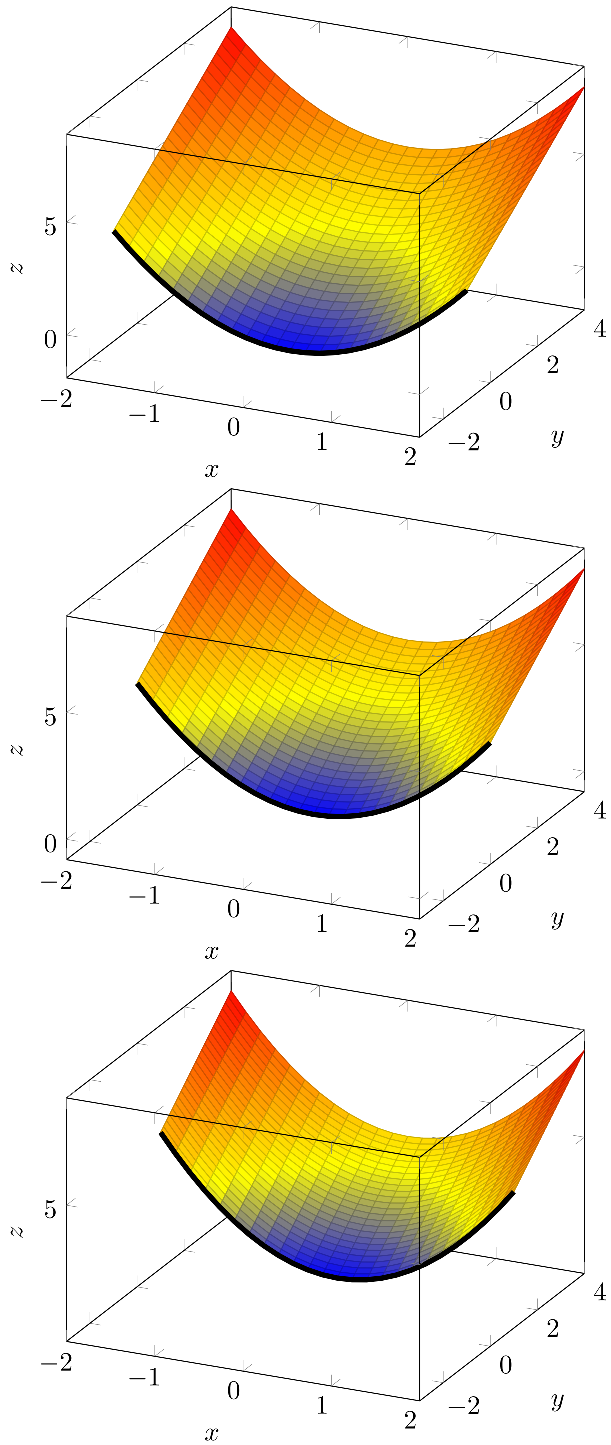


Figure 4: Cross-sections of the function $f(x, y) = x^2 + y$ for $y = -1, y = 0, y = 1$.

For any of these functions, we could differentiate and find the slope of the cross-section:

$$\frac{\mathrm{d}f_k}{\mathrm{d}x} = \frac{\mathrm{d}(x^2 + k)}{\mathrm{d}x} = 2x\,. \tag{1}$$

Note that we **chose** to build up the full surface as the set of functions f_k. We could equally have considered $g_0(y) = y$, $g_1(y) = 1^2 + y$, $g_\pi(y) = \pi^2 + y$, etc. and built the surface the other way round. More generally, we can define an infinite set of functions $g_m(y) = m^2 + y$ for some constant m. Taking cross-sections through the surface parallel to the y axis will give these functions instead as shown in Fig. 5.

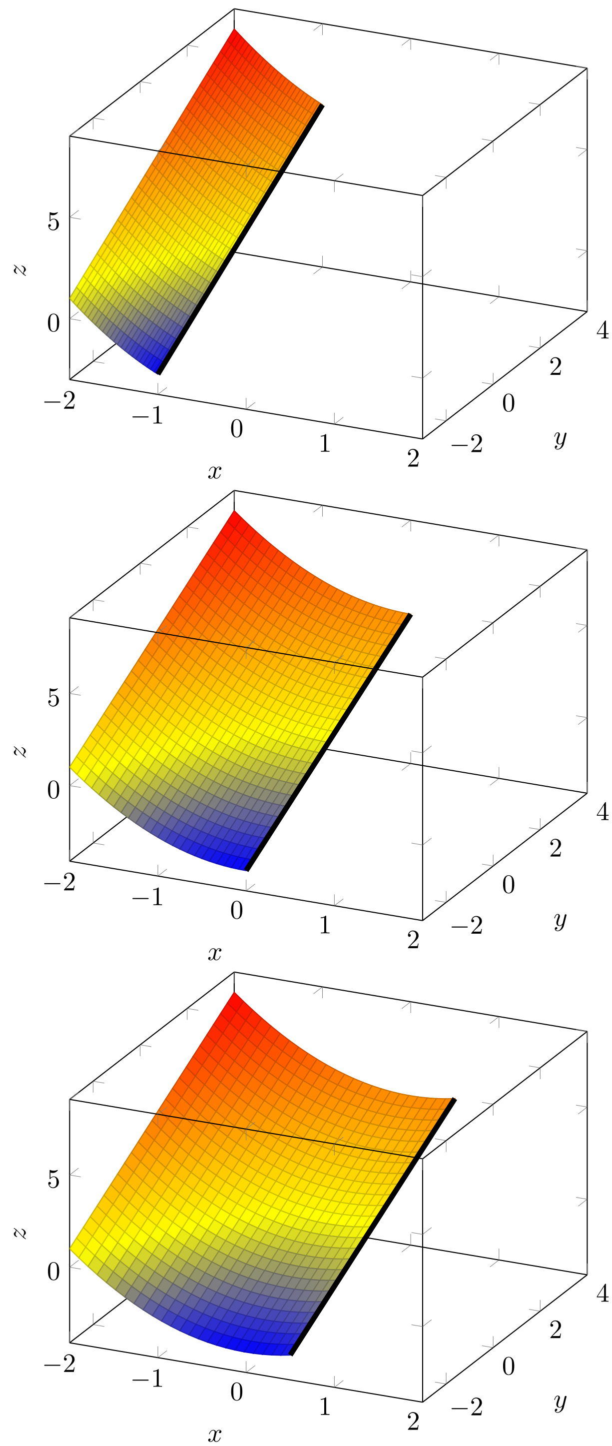


Figure 5: Cross-section of $f(x, y) = x^2 + y$ at $x = -1$, $x = 0$ and $x = 0.5$.

Again, we can differentiate any of these functions to find the slope of the cross-section in this direction:

$$\frac{\mathrm{d}g_m}{\mathrm{d}y} = \frac{\mathrm{d}(m^2 + y)}{\mathrm{d}y} = 1\,. \tag{2}$$

9.2 Partial Differentiation

The functions that we have considered so far have had only a trivial relationship between x and y. Consider the more interesting function $f(x, y) = yx^2$. In the same way as before, we can build up the graph of this function by considering its cross-sections in both x and y. The graph of $z = f(x, y)$ is given in Fig. 6.

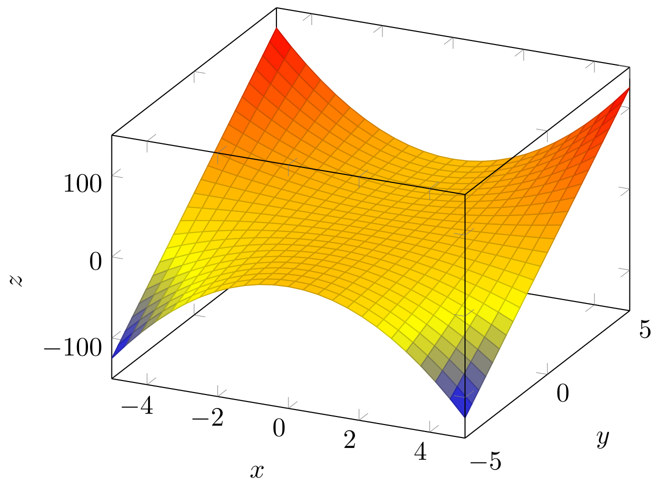


Figure 6: Graph of $z = yx^2$.

The cross-sections in this case are more interesting as we can see in Fig. 7.

As before, we can think of these cross-sections as one-dimensional functions and differentiate them to find the slope. This time, however, it is immediately clear from the graph that the derivative function now **depends on** y. In particular, when $y = 0$, the slope of the cross-section is identically 0. In general, we have

$$\frac{\mathrm{d}f_k}{\mathrm{d}x} = \frac{\mathrm{d}(kx^2)}{\mathrm{d}x} = 2kx\,. \tag{3}$$

We want a general approach to this idea of finding the slope of a cross-section without having to plot the graph or build up an infinite set of 1D functions. This is where the partial derivative comes in. The partial derivative of a function $f(x, y)$ with respect to x is the derivative with respect to x of the cross-section at some specified value of y. That is, we consider the slope of the function as we vary x while keeping y constant. To distinguish this from the ordinary derivative, we use a curly “d”. The definition of the partial derivative is:

$$\begin{aligned}\frac{\partial f(x,y)}{\partial x} &= \lim_{\delta x\to 0}\frac{f(x+\delta x, y) - f(x,y)}{\delta x} \text{ and}\\ \frac{\partial f(x,y)}{\partial y} &= \lim_{\delta y\to 0}\frac{f(x, y+\delta y) - f(x,y)}{\delta y}\,.\end{aligned} \tag{4}$$

While this is the formal definition, in practice, we simply differentiate as normal, using all of the familiar standard cases, but treating one of the variables as a constant.

9.3 Example

Find the partial derivatives with respect to x and y of $f(x, y) = x^2 \sin(y) + 3x$ and $g(x, y) = x^y$.

9.4 Second Derivatives

Just as in the case of 1D functions, we can differentiate more than once. If we consider the cross-section of a function for constant y as a 1D function, we can differentiate that curve twice with

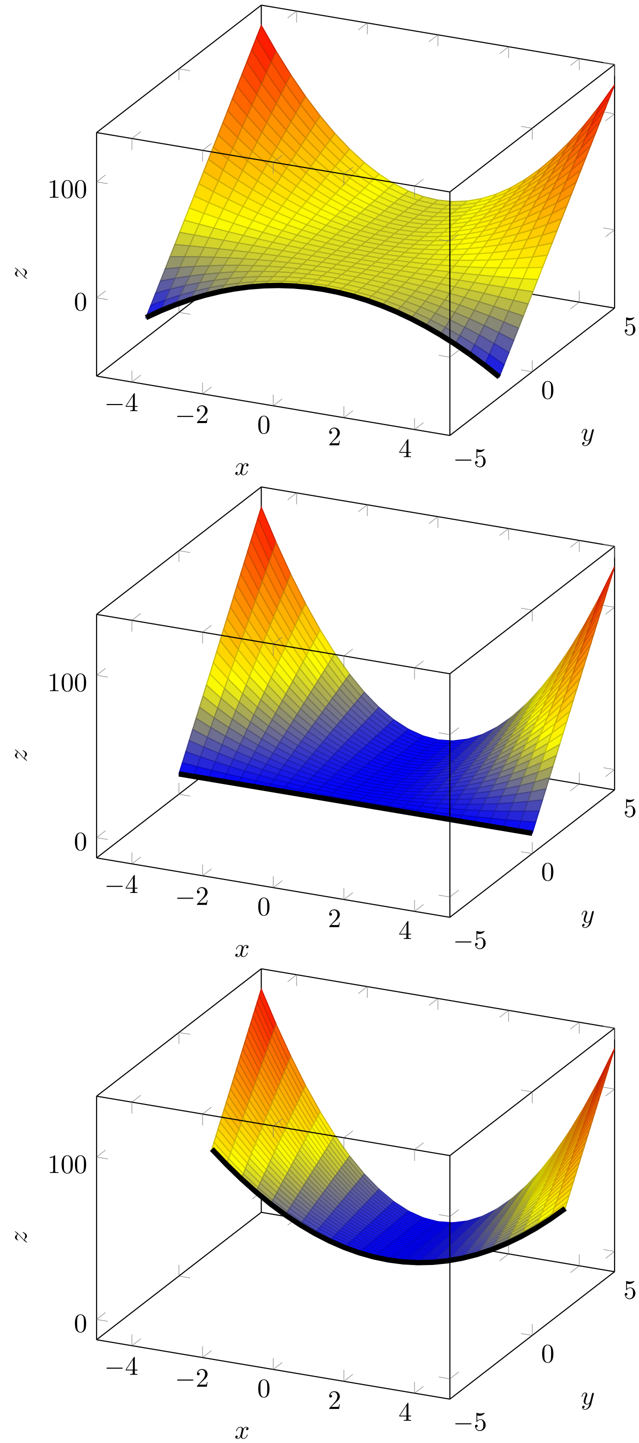
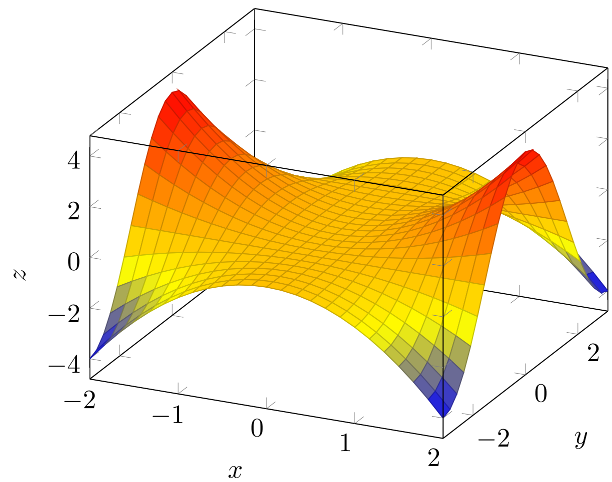


Figure 7: Cross-sections of $z = yx^2$ for $y = -2, y = 0, y = 2$.

respect to x, and vice versa. This will tell us useful information about the curvature of the cross-section in each direction. We can denote these second derivatives as

$$\frac{\partial^2 f(x,y)}{\partial x^2} \text{ and } \frac{\partial^2 f(x,y)}{\partial y^2}\,. \tag{5}$$

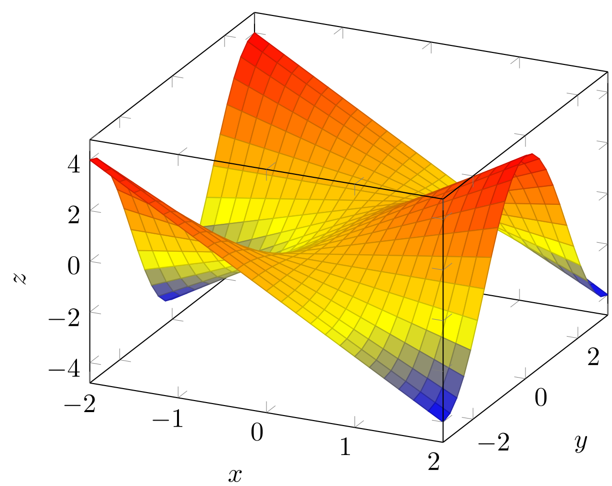
There are also mixed partial derivatives that are a little trickier to visualise. To see what they represent, notice that the partial derivative of a 2D function is itself a 2D function, and consider the graphs of both a function $f(x, y)$ and its partial derivative $\partial f(x, y)/\partial x$. For example, the graph of $f(x, y) = x^2 \cos(y)$ looks like



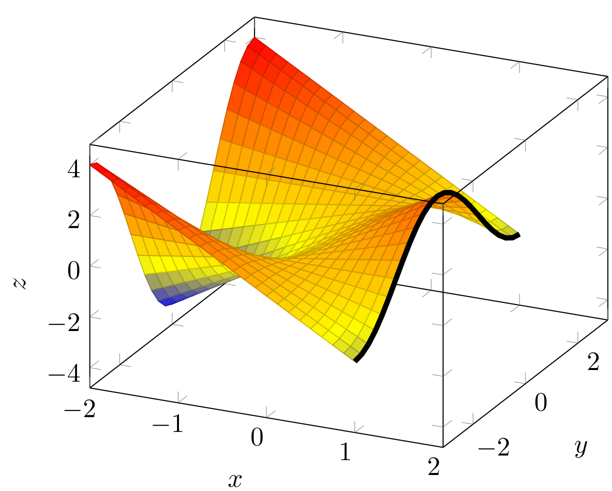
while its derivative

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

has a graph that looks like



for which we can then look at a y cross-section



and find the slope in the y direction. The first derivative represents the rate at which the function varies with x, while the mixed second derivative represents the rate at which the first derivative varies with y. This is denoted

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

Notice that the order of the variables in the denominator matters in principle. However, in practice, we will **usually** find that

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

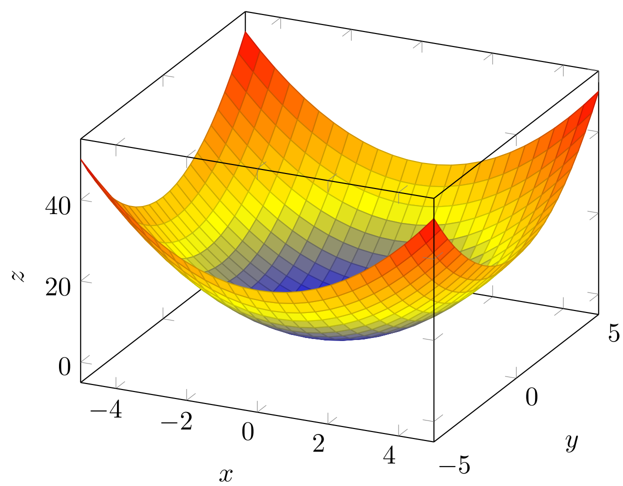
if the function is reasonably well behaved.

9.5 Example

Show that the mixed second partial derivatives of $f(x, y) = x^2 \cos(y)$ are equal.

9.6 Stationary Points

Consider again the cross-section of the function $f(x, y) = x^2 + y$ for $y = 0$ shown in Fig. 4. There is a minimum in the cross-section at $x = 0$. However, when we look at the full function, it is clear that there is no minimum – if a ball were dropped on the graph, it would roll along the y axis. On the other hand, the graph $z = x^2 + y^2$ shown below clearly has a minimum at $(x, y) = (0, 0)$:



It is not enough for one of the partial derivatives to be 0 for a point to be stationary – we require both to be 0 **at the same point**.

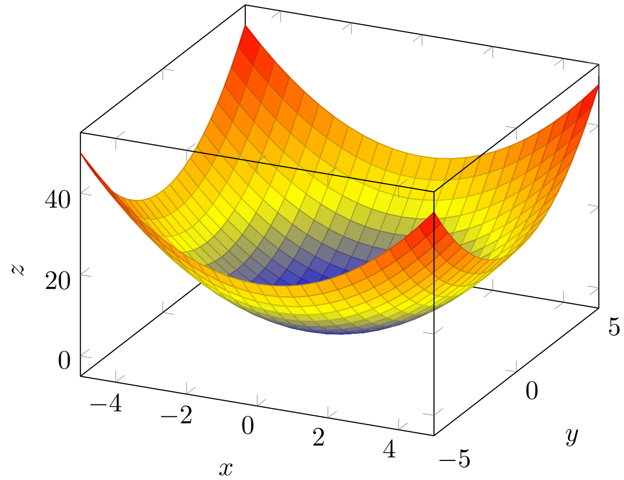
9.7 Example

Find the stationary points of $f(x, y) = -x^2 + xy + y^2 + 6y$.

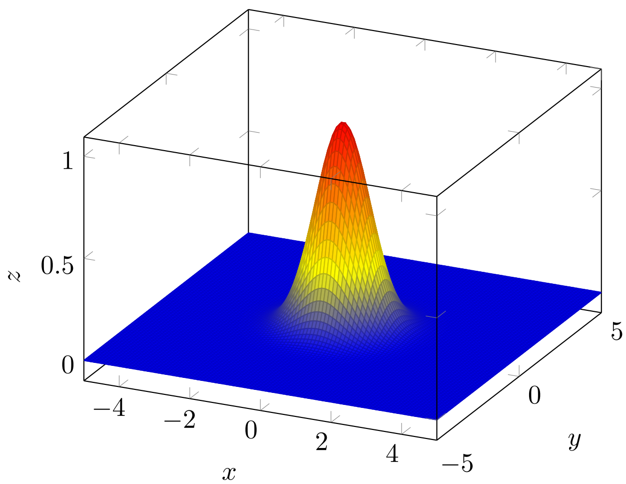
9.8 The Nature of Stationary Points

For multivariable functions, we have various types of stationary point to consider. For example:

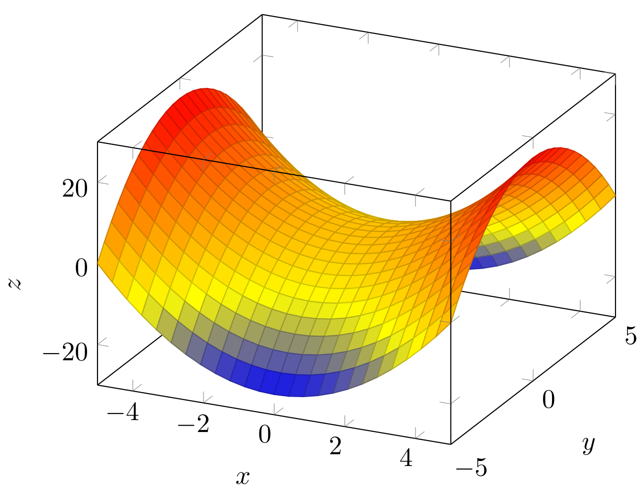
- $z = x^2 + y^2$ has a minimum at $(x, y) = (0, 0)$:



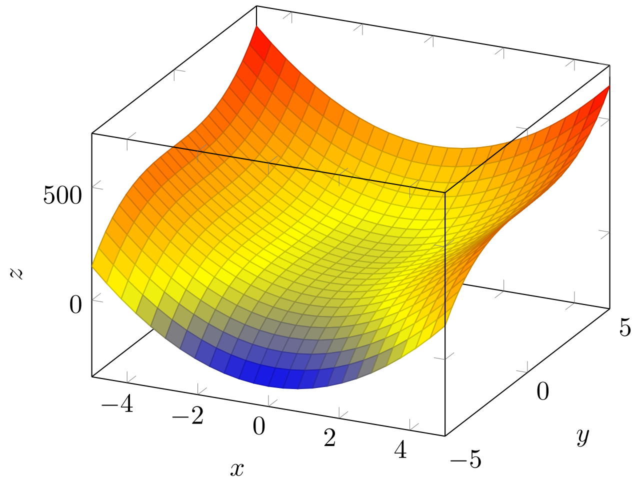
- $z = e^{-(x^2+y^2)}$ has a maximum at $(x, y) = (0, 0)$:



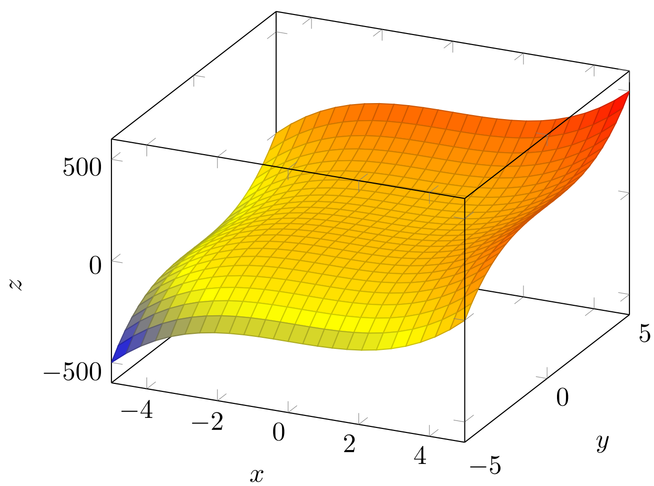
- $z = x^2 - y^2$ has a saddle point at $(x, y) = (0, 0)$:



- $z = 16x^2 + 2y^3$ has a saddle point at $(x, y) = (0, 0)$:



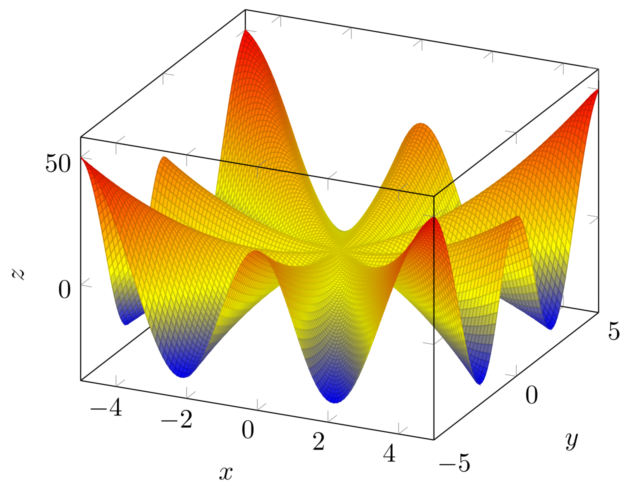
- $z = 2x^3 + 2y^3$ has a saddle point at $(x, y) = (0, 0)$:



Notice that the nature of the saddle point is quite different in these last three examples. In the first, there is a minimum in x but a maximum in y. In the second, there is a minimum in x and a 1D saddle point in y. In the third, there are 1D saddle points in both x and y. All of these are classed as saddle points. In general, a saddle point is any stationary point that is not an extremum.

9.9 Determining the Nature of Stationary Points

At first glance, it may appear that the nature of a stationary point depends only on the values of $\partial^2 f/\partial x^2$ and $\partial^2 f/\partial y^2$, since a positive value in both "must" be a minimum. However, a counter-example will show why this logic fails: the graph



has a saddle point at the origin. However, notice that cross-sections along both the x and y axes have minima at the origin. So it is not enough to consider the second derivatives in only two directions – we somehow need to consider the second derivative in **all** directions. We will not prove it here, but the correct quantity to calculate is

$$D = \frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 , \tag{6}$$

which is often denoted

$$D = f_{xx} f_{yy} - f_{xy}^2 , \tag{7}$$

where the subscripts are a shorthand for "the partial derivative with respect to...".

It can be shown that if $D < 0$ at a stationary point, then it is a saddle point. If $D > 0$, then it is necessary to consider f_{xx} and f_{yy} separately. In this case, if both f_{xx} and f_{yy} are positive, the point is a maximum. If both are negative, it is a minimum. If they have opposite signs, it is a saddle point.

This is summarised in the flow chart in Fig. 8.

9.10 The Total Derivative

Consider a small change in the arguments of a two-variable function, $x \mapsto x + \delta x$, $y \mapsto y + \delta y$. This induces a small change in the function,

$$\begin{aligned}
\delta f(x,y) &= f(x+\delta x, y+\delta y) - f(x,y) \\
&= f(x+\delta x, y+\delta y) - f(x, y+\delta y) + f(x, y+\delta y) - f(x,y) \\
&= \frac{f(x+\delta x, y+\delta y) - f(x, y+\delta y)}{\delta x}\delta x + \frac{f(x, y+\delta y) - f(x,y)}{\delta y}\delta y \\
&\approx \frac{f(x+\delta x, y) - f(x,y)}{\delta x}\delta x + \frac{f(x, y+\delta y) - f(x,y)}{\delta y}\delta y \\
&= \frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial y}\delta y \, .
\end{aligned} \tag{8}$$

This approximation gets better as $\delta x \to 0$, so the total differential is

$$\mathrm{d}f(x,y) = \frac{\partial f}{\partial x}\mathrm{d}x + \frac{\partial f}{\partial y}\mathrm{d}y \, . \tag{9}$$

If the value of f is ultimately dependent on a single parameter, t, then we can define the total derivative of f in terms of the ratio of small changes in f to small changes in t, $\frac{\mathrm{d}f}{\mathrm{d}t}$.

Suppose now that f is determined by t only indirectly through its dependence on x and y, which in turn depend on t. For example, if the length x and width y of a sheet of metal depend

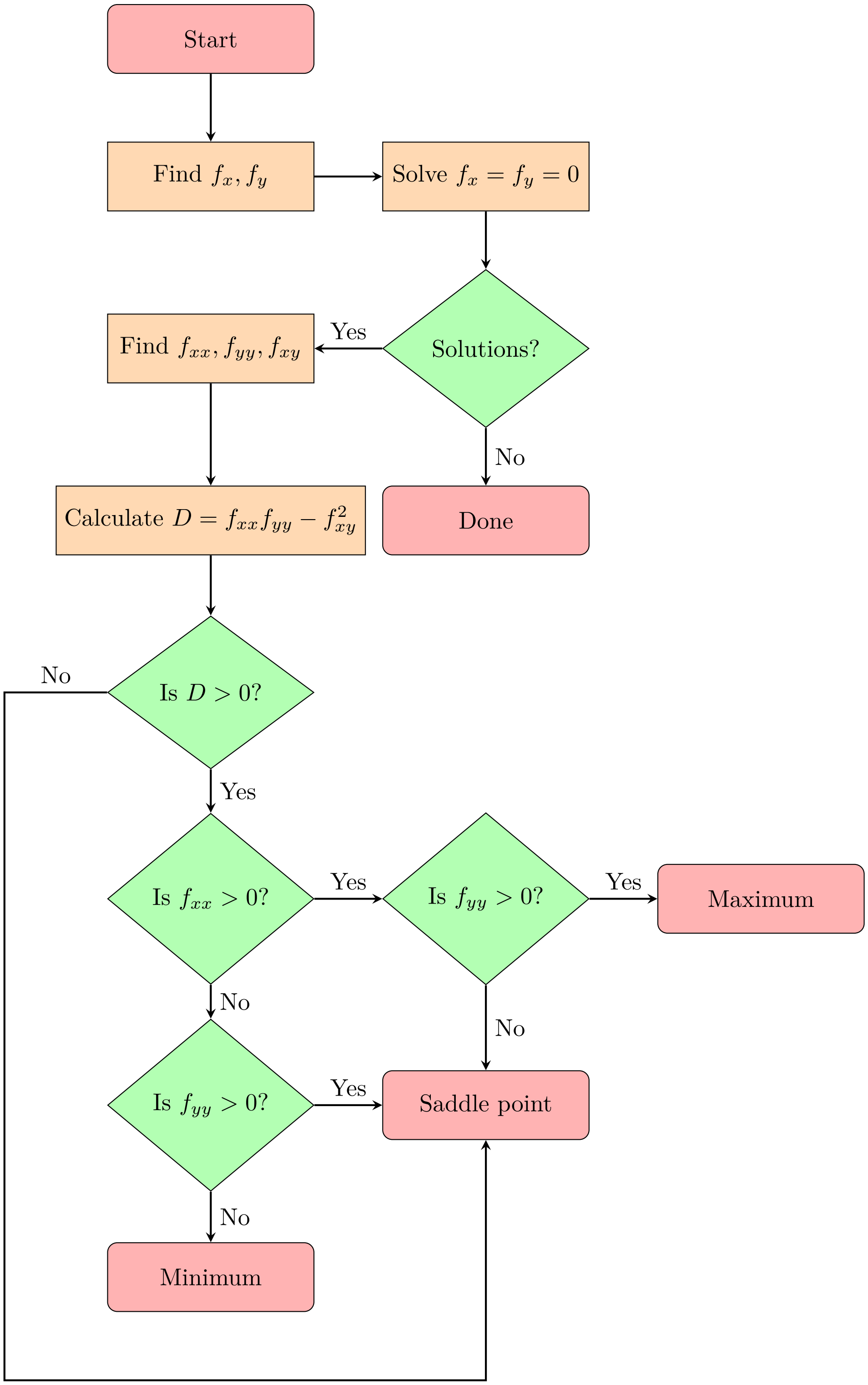


Figure 8: Stationary Points of 2D Functions Flow Chart

on the temperature t, then the area of the sheet, f, depends on temperature via x and y. In such a situation, an infinitesimal change in t, $\mathrm{d}t$, induces changes $\mathrm{d}x$ and $\mathrm{d}y$, and therefore induces a change $\mathrm{d}f$ in f. The total derivative of f with respect to t is then

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial f}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t}\,, \tag{10}$$

which is the multivariate generalisation of the chain rule.

9.11 Example

A good example of this may help clarify the situation. Consider a closed gas-filled container with a piston at one end so that the volume can be controlled. Suppose the container is also in thermal contact with a heat sink so that we can control its temperature. This means that we can indirectly control the pressure in the container through the gas law, $pV = nRT$. A small change in the temperature, $\mathrm{d}T$, or the volume, $\mathrm{d}V$, will induce a corresponding small change in the pressure, $\mathrm{d}p$, with

$$\mathrm{d}p = \frac{\partial p}{\partial V}\mathrm{d}V + \frac{\partial p}{\partial T}\mathrm{d}T\,.$$

Now suppose that both the volume and the temperature are controlled by an automated process so that they follow well-defined functions of time:

$$V = V(t) \text{ and } T = T(t)\,.$$

Then the derivatives $\mathrm{d}V/\mathrm{d}t$ and $\mathrm{d}T/\mathrm{d}t$ represent how quickly these values change over time. But this also means that the pressure is ultimately a function of time, $p = p(t)$, even though we are not directly controlling it. In this case, the rate of change of pressure will be

$$\frac{\mathrm{d}p}{\mathrm{d}t} = \frac{\partial p}{\partial V}\frac{\mathrm{d}V}{\mathrm{d}t} + \frac{\partial p}{\partial T}\frac{\mathrm{d}T}{\mathrm{d}t}\,.$$

9.12 Inter-dependent Functions

Returning to the example of the gas law, notice that we can rearrange the law to suit different situations. If we control the temperature and volume, then it makes sense to treat pressure as a function of these. But in other situations, we may want to treat temperature as a function of pressure and volume, or volume as a function of temperature and pressure. More generally, if we have three variables uniquely connected by some relationship, then we can choose which one we treat as the dependent variable:

$$x = x(y, z) \text{ or } y = y(z, x) \text{ or } z = z(x, y)\,.$$

Using the total derivative, we can deduce some useful relationships between the partial derivatives of these variables. To emphasise which variables are being held constant, we will denote the constant variables as subscripts, so

$$\left(\frac{\partial x}{\partial y}\right)_z$$

means "the rate of change of x when y changes and z is held constant".

Considering each of these variables as the dependent variable gives the following expressions for their differentials:

$$\begin{aligned}
\mathrm{d}x &= \left(\frac{\partial x}{\partial y}\right)_z \mathrm{d}y + \left(\frac{\partial x}{\partial z}\right)_y \mathrm{d}z\,, \\
\mathrm{d}y &= \left(\frac{\partial y}{\partial z}\right)_x \mathrm{d}z + \left(\frac{\partial y}{\partial x}\right)_z \mathrm{d}x\,, \\
\mathrm{d}z &= \left(\frac{\partial z}{\partial x}\right)_y \mathrm{d}x + \left(\frac{\partial z}{\partial y}\right)_x \mathrm{d}y\,.
\end{aligned}$$

We can now substitute the expression for $\mathrm{d}y$ into the one for $\mathrm{d}x$ to get

$$\begin{aligned}
\mathrm{d}x &= \left(\frac{\partial x}{\partial y}\right)_z \left[\left(\frac{\partial y}{\partial z}\right)_x \mathrm{d}z + \left(\frac{\partial y}{\partial x}\right)_z \mathrm{d}x\right] + \left(\frac{\partial x}{\partial z}\right)_y \mathrm{d}z \\
0 &= \left[\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial x}\right)_z - 1\right] \mathrm{d}x + \left[\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x + \left(\frac{\partial x}{\partial z}\right)_y\right] \mathrm{d}z\,.
\end{aligned}$$

Now choosing x and z as the independent variables means that we can vary $\mathrm{d}x$ without affecting $\mathrm{d}z$, and vice versa. This means the two terms of the last line in Eq. 9.12 must both be 0 independently. The first of these gives the relation

$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial x}\right)_z = 1\,.$$

Notice that we arbitrarily chose the dependent variables to arrive at this relationship, so we must have similar relationships for other pairs of variables as well. In particular, we could have shown that

$$\left(\frac{\partial x}{\partial z}\right)_y \left(\frac{\partial z}{\partial x}\right)_y = 1\,.$$

Substituting this into the second term of Eq. 9.12, we also find

$$\begin{aligned}
\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x + \left(\frac{\partial x}{\partial z}\right)_y &= 0 \\
\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x &= -\left(\frac{\partial x}{\partial z}\right)_y \\
\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y &= -1\,.
\end{aligned}$$

This relationship proves to be particularly useful in thermodynamics.