Multivariate Calculus

Lachlan Dufort-Kennett

March 30, 2025

Contents

Preface

Chapter 1

Multivariate Functions

1.1 Introduction

In single-variable calculus and real analysis, we study functions of the form $f : \mathbb{R} \to \mathbb{R}$. These are functions which take a single variable x and send it to a real number. We explored this theory and defined concepts such as limits, continuity, derivatives, integrals, and power series. In multivariate calculus, we use what we have learned to generalise these ideas to functions of the form $f : \mathbb{R}^n \to \mathbb{R}$, which send an **ordered n-tuple** $\boldsymbol{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ to a real number $f(\boldsymbol{x}) \in \mathbb{R}$.

We will mostly be interested in *bivariate* and *trivariate* functions, that is functions from subsets of $\mathbb{R}^2$ and $\mathbb{R}^3$ to $\mathbb{R}$. This is both because they are the simplest case of multivariate functions and also because they have the most familiar applications in the real world. Examples of these are equations for 3D surfaces, complex analysis, or many areas of physics such as classical mechanics.

1.2 Functions of Two Variables

Definition 1.1. A function of two variables takes an ordered pair (x, y) and returns a real number $f(x, y)$.

$$f : \mathbb{R}^2 \to \mathbb{R} \quad \text{or} \quad (x, y) \mapsto f(x, y). \tag{1.1}$$

We can equivalently consider the ordered pair (x, y) as a 2D vector $\boldsymbol{x}$, it makes no difference to the mathematics.

How do we visualise these functions? These functions define **surfaces** in 3D space with cartesian coordinates $(x, y, f(x, y))$. Another way we can represent them is with a contour plot (contours are all the values of x and y which give a specific value of f).

Example 1.1. Consider the function $f(x, y) = (1 - x^2)(1 + y^2)$.

There are many examples of surfaces which are commonly encountered in geometry. We will go through some of these now.

Example 1.2. A **plane** is defined by the function $f(x, y) = c - ax - by$. With $z = f(x, y)$, we get the equation

$$ax + by + z = c, \tag{1.2}$$

which is a plane with normal vector $(a, b, 1)$.

Example 1.3. Consider the function $f(x, y) = \sqrt{R^2 - x^2 - y^2}$. With $z = f(x, y)$, this gives

$$x^2 + y^2 + z^2 = R^2, \tag{1.3}$$

which is the equation for a **sphere** of radius R centered at the origin. Since f is a real-valued function, it is only defined for $R^2 - x^2 - y^2 \geq 0$, so the domain of f is

$$x^2 + y^2 \leq R, \tag{1.4}$$

and the range of f is

$$0 \leq z \leq R. \tag{1.5}$$

Example 1.4. Consider the function $f(x, y) = z = \sqrt{\frac{x^2}{2} + y^2}$. In the z-x plane, we get the equation $z = \pm\frac{x}{\sqrt{2}}$. In the z-y plane, we get $z = \pm y$. These are both straight lines. Now consider a contour for some $z = z_0$:

$$\frac{x^2}{2} + y^2 = z_0^2. \tag{1.6}$$

This is an ellipse. Therefore, this function represents an **elliptical cone**.

Example 1.5. Now consider $f(x, y) = x^2 + y^2$. In this case, we get parabolas in the x-z and y-z planes and circles parallel to the x-y plane. This is called a **elliptic paraboloid**.

Example 1.6. Finally, consider $f(x, y) = x^2 - y^2$. This is a positive parabola in one direction but a negative parabola in the other. In the $z = 0$ plane we have $x = \pm y$, which are straight lines. This is called a **hyperbolic paraboloid**.

1.3 Limits and Continuity of Multivariate Functions

Recall that for a limit to exist at a point for a single-variable function, it must be defined and have the same value when approaching from the positive and negative direction. With more than one variable, we now have *infinitely many* directions to approach a point from. Thus, for a limit to exist at a point it has to have the same value when approaching from **every direction** and also be **path independent**. If two paths do not give the same value for the limit, then the limit does not exist.

Example 1.7. Does the function $f(x,y) = \frac{x^2-y^2}{x^2+y^2}$ have a limit as $(x,y) \to (0,0)$? Consider the limit along the x-axis ($y = 0$):

$$\lim_{x\to 0} f(x,0) = \lim_{x\to 0} \frac{x^2}{x^2} = 1. \tag{1.7}$$

Now look at the limit along the y-axis ($x = 0$):

$$\lim_{y\to 0} f(0,y) = lim_{y\to 0} \frac{-y^2}{y^2} = -1 \tag{1.8}$$

So, since the limits along the different paths have different values, the limit does not exist.

Since there is no ordering on tuples like there is for real numbers, it is impossible to prove the existence of limits in general like we would in 1D. Why can't we simply use the absolute value $|(x,y)| = \sqrt{x^2+y^2}$? Because this corresponds to a straight line path towards the limit. One of the best strategies is to convert to **polar coordinates**, basically using trigonometry to relabeli every point in 2D space using the distance from the origin and the angle from the x-axis (like the polar form of complex numbers).

$$x = r\cos\theta \qquad r = \sqrt{x^2+y^2} \tag{1.9}$$

$$y = r\sin\theta \qquad \theta = \tan^{-1}\left(\frac{y}{x}\right). \tag{1.10}$$

Example 1.8. Find the limit

$$\lim_{(x,y)\to(0,0)} \frac{x^2}{\sqrt{x^2+y^2}}. \tag{1.11}$$

Using polar coordinates, we rewrite the function inside the limit as

$$\frac{x^2}{\sqrt{x^2+y^2}} = \frac{r^2\cos^2\theta}{r} = r\cos^2\theta. \tag{1.12}$$

Now we can rewrite the limit as

$$\lim_{(x,y)\to(0,0)} \frac{x^2}{\sqrt{x^2+y^2}} = \lim_{r\to 0} r\cos^2\theta = 0. \tag{1.13}$$

Since the last limit does not depend on θ, it is path-independent and so the limit is well-defined.

The property of continuity is defined completely analogously to how it is defined in 1D.

Definition 1.2. A function $f(x,y)$ is **continuous** at a point (x_0,y_0) if

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y) = f(x_0,y_0). \tag{1.14}$$

Example 1.9. If we define the function in the last exercise as

$$f(x,y) = \begin{cases} \frac{x^2}{\sqrt{x^2+y^2}} & (x,y) \neq (0,0) \\ 0 & x = y = 0, \end{cases} \tag{1.15}$$

then f is a continuous function.

1.4 Partial Derivatives

In 1D we only had one direction to define a slope, but in more than one dimension we now have infinitely many. We can define the rate of change *with respect to* one of the variables by defining the derivative in the same way we would have in 1D but keeping all of the other variables constant.

Definition 1.3. Consider a function of two variables $f(x,y)$. The **partial derivative** with respect to x at a point (x_0, y_0) is defined as the limit

$$\left(\frac{\partial f}{\partial x}\right)_y \bigg|_{(x_0,y_0)} = f_x(x_0,y_0) = \lim_{h\to 0} \frac{f(x_0+h,y_0) - f(x_0,y_0)}{h}, \tag{1.16}$$

if it exists. Likewise, the partial derivative with respect to y is defined as

$$\left(\frac{\partial f}{\partial y}\right)_y \bigg|_{(x_0,y_0)} = f_y(x_0,y_0) = \lim_{h\to 0} \frac{f(x_0,y_0+h) - f(x_0,y_0)}{h}. \tag{1.17}$$

Most of the time the parentheses in the fractional notation are removed because they are there to show explicitly what variables are being kept constant. In cases where there are variables which depend on each other such as thermodynamics or statistical mechanics, this can greatly increase clarity, but for simple calculus with independent variables it is not necessary.

Note that the partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are still functions of two variables. If they are both *continuous*, then $f(x,y)$ is differentiable.

1.5 Higher Order Derivatives

Just like in 1D we define higher order partial derivatives recursively as partial derivatives of partial derivatives. We can also obtain **mixed partial derivative** by changing the variable we differentiate with respect to. For example, the four possible second derivatives of a function of two variables are

$$\frac{\partial^2 f}{\partial x^2} = f_{xx} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right), \quad \frac{\partial^2 f}{\partial y \partial x} = f_{xy} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) \tag{1.18}$$

$$\frac{\partial^2 f}{\partial y^2} = f_{yy} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right), \quad \frac{\partial^2 f}{\partial x \partial y} = f_{yx} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right). \tag{1.19}$$

Under certain conditions (which are quite general for the functions that we will be studying), it can be shown that the two mixed derivatives are the same

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}, \quad f_{xy} = f_{yx}. \tag{1.20}$$

I.e., the order of differentiation can be swapped.

1.6 The Chain Rule

Suppose we have two one-variable functions; $y(x)$ and $x(t)$. Then using the 1D chain rule we can define

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\mathrm{d}y}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}t}. \tag{1.21}$$

Suppose now we have a two-variable function $f(x, y)$ where the variables x and y are themselves functions of another variable $x(t)$, $y(t)$. How can we find $\frac{\mathrm{d}f}{\mathrm{d}t}$? We could substitute in for t but this could be very complicated.

If t changes by a small amount Δt, x will change by a small amount Δx.

$$x(t + \Delta t) = x(t) + \Delta x \implies \frac{\Delta x}{\Delta t} = \frac{x(t + \Delta t) - x(t)}{\Delta t}. \tag{1.22}$$

This becomes $\frac{\mathrm{d}x}{\mathrm{d}t}$ in the limit $\Delta t \to 0$. The same thing happens to y. Since both x and y change when t changes, $f(x, y)$ will change by a value Δf given by

$$\Delta f = f(x + \Delta x, y + \Delta y) - f(x, y) \tag{1.23}$$

$$= \underbrace{f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y)}_{\times \frac{\Delta x}{\Delta x}} + \underbrace{f(x, y + \Delta y) - f(x, y)}_{\times \frac{\Delta y}{\Delta y}} \tag{1.24}$$

$$\begin{aligned} \implies \frac{\Delta f}{\Delta t} &= \frac{f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y)}{\Delta x}\left(\frac{\Delta x}{\Delta t}\right) \\ &+ \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}\left(\frac{\Delta y}{\Delta t}\right). \end{aligned} \tag{1.25}$$

Now we take the limit as $\Delta t \to 0$ and (if the limit exists) we get

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \left(\frac{\partial f}{\partial x}\right)_y \frac{\mathrm{d}x}{\mathrm{d}t} + \left(\frac{\partial f}{\partial y}\right)_x \frac{\mathrm{d}y}{\mathrm{d}t} \tag{1.26}$$

Note that this is an full derivative since both x and y are functions of t.

Example 1.10. Consider the surface $z(x, y) = x^2 + y^2$ where $x = t^2$, $y = \sin t$. Then

$$\left(\frac{\partial z}{\partial x}\right)_y = 2x, \quad \left(\frac{\partial z}{\partial y}\right)_x = 2y, \quad \frac{\mathrm{d}x}{\mathrm{d}t} = 2t, \quad \frac{\mathrm{d}y}{\mathrm{d}t} = \cos t, \tag{1.27}$$

so

$$\frac{\mathrm{d}z}{\mathrm{d}t} = (2x)(2t) + (2y)(\cos t) \tag{1.28}$$
$$= 4xt + 2y\cos t \tag{1.29}$$
$$= 4t^3 + 2\sin t\cos t. \tag{1.30}$$

Example 1.11. Consider a two-variable function $z(x, y)$, where both x and y are themselves functions of two variables given by

$$x(s,t) = s + t, \quad y(s,t) = s - t. \tag{1.31}$$

Then we can calculate the first derivatives of x and y as

$$\frac{\partial x}{\partial s} = 1, \quad \frac{\partial x}{\partial t} = 1, \quad \frac{\partial y}{\partial s} = 1, \quad \frac{\partial y}{\partial t} = -1, \tag{1.32}$$

and therefore we can find the partial derivatives of z with respect to s and t using the chain rule:

$$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s} = \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y} \tag{1.33}$$
$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t} = \frac{\partial z}{\partial x} - \frac{\partial z}{\partial y}. \tag{1.34}$$

Similarly, we can calculate the second derivatives by using the chain rule again.

$$\frac{\partial^2 z}{\partial s^2} = \frac{\partial}{\partial s}\left(\frac{\partial z}{\partial s}\right) = \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial s}\right)\frac{\partial x}{\partial s} + \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial s}\right)\frac{\partial y}{\partial s} \tag{1.35}$$
$$= \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x} + \frac{\partial z}{\partial y}\right) + \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x} + \frac{\partial z}{\partial y}\right) \tag{1.36}$$
$$= \frac{\partial^2 z}{\partial x^2} + 2\frac{\partial^2 z}{\partial x \partial y} + \frac{\partial^2 z}{\partial y^2}, \tag{1.37}$$
$$\frac{\partial^2 z}{\partial t^2} = \frac{\partial}{\partial t}\left(\frac{\partial z}{\partial t}\right) = \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial t}\right)\frac{\partial x}{\partial t} + \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial t}\right)\frac{\partial y}{\partial t} \tag{1.38}$$
$$= \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x} - \frac{\partial z}{\partial y}\right) - \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x} - \frac{\partial z}{\partial y}\right) \tag{1.39}$$
$$= \frac{\partial^2 z}{\partial x^2} - 2\frac{\partial^2 z}{\partial x \partial y} + \frac{\partial^2 z}{\partial y^2}. \tag{1.40}$$

This links to the wave equation because if we set $\frac{\partial^2 z}{\partial t^2} = \frac{\partial^2 z}{\partial s^2}$, we conclude that $\frac{\partial^2 z}{\partial x \partial y} = 0$, which implies that $z = f(x) + g(y) = f(s+t) + g(s-t)$ is the general solution.

1.7 Implicit Differentiation

Implicit differentiation may be the most common concept in multivariate calculus to have been encountered before studying the subject.

Suppose we have two single variable functions $x(t)$ and $y(t)$ which represent a path in 2D space and a two-variable function $z = f(x, y)$ representing the height of a surface. Then $z(x(t), y(t))$ represents the variation in height of the surface along the path. If we choose this path to be a contour (where z remains constant) then we have

$$\frac{\partial z}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial z}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} = 0. \tag{1.41}$$

So by using the 1D chain rule, we get

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}x} = -\frac{\left(\frac{\partial z}{\partial x}\right)_y}{\left(\frac{\partial z}{\partial y}\right)_x}. \tag{1.42}$$

This allows us to calculate $\frac{\mathrm{d}y}{\mathrm{d}x}$ for an implicitly defined function.

Example 1.12. Consider the implicitly defined function

$$y^4 + 3y - 4x^3 - 5x - 1 = 0. \tag{1.43}$$

If we consider this function as a contour of a two-variable function $z(x, y) = 0$, then we can use the formula above to find

$$\frac{\partial z}{\partial x} = -12x^2 - 5, \quad \frac{\partial z}{\partial y} = 4y^3 + 3 \tag{1.44}$$

$$\implies \frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{-12x^2 - 5}{4y^3 + 3} = \frac{12x^2 + 5}{4y^3 + 3}. \tag{1.45}$$

This technique readily generalises to n dimensions for however many variables an implicit function may have. For example, suppose we have an implicit function of three variables which defines a surface $w(x, y, z) = 0$ Due to the constraint that $w = 0$, we only really have 2 degrees of freedom, which we will denote t and s. Using the chain rule, we get

$$\frac{\mathrm{d}w}{\mathrm{d}t} = \frac{\partial w}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial w}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} + \frac{\partial w}{\partial z}\frac{\mathrm{d}z}{\mathrm{d}t} = 0. \tag{1.46}$$

Suppose we want to find $\frac{\partial z}{\partial x}$. Without loss of generality, we can define $x = t$, $y = s$, $z = z(t, s)$ (choosing x and y determines the value of z even if we can't solve for it analytically). Then since $\frac{\partial x}{\partial t} = 1$ and $\frac{\partial y}{\partial t} = 0$, we get

$$\frac{\partial w}{\partial t} = \frac{\partial w}{\partial x} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial x} = 0 \tag{1.47}$$

$$\implies \frac{\partial z}{\partial x} = -\frac{\frac{\partial w}{\partial x}}{\frac{\partial w}{\partial z}}. \tag{1.48}$$

$$\frac{\partial^2 f}{\partial s^2} = \frac{\partial}{\partial s}\left(\frac{\partial f}{\partial s}\right) = \frac{\partial}{\partial s}\left(\frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s}\right) \tag{1.49}$$

$$= \frac{\partial}{\partial s}\left(\frac{\partial f}{\partial x}\right)\frac{\partial x}{\partial s} + \frac{\partial f}{\partial x}\frac{\partial^2 x}{\partial s^2} + \frac{\partial}{\partial s}\left(\frac{\partial f}{\partial y}\right)\frac{\partial y}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial^2 y}{\partial s^2} \tag{1.50}$$

$$= \left\{\frac{\partial^2 f}{\partial x^2}\frac{\partial x}{\partial s} + \frac{\partial^2 f}{\partial x \partial y}\frac{\partial y}{\partial s}\right\}\frac{\partial x}{\partial s} + \left\{\frac{\partial^2 f}{\partial y^2}\frac{\partial y}{\partial s} + \frac{\partial f}{\partial x \partial y}\frac{\partial x}{\partial s}\right\}\frac{\partial y}{\partial s} + \frac{\partial f}{\partial x}\frac{\partial^2 x}{\partial s^2} + \frac{\partial f}{\partial y}\frac{\partial^2 y}{\partial s^2} \tag{1.51}$$

$$= \frac{\partial^2 f}{\partial x^2}\left(\frac{\partial x}{\partial s}\right)^2 + 2\frac{\partial^2 f}{\partial x \partial y}\frac{\partial x}{\partial s}\frac{\partial y}{\partial s} + \frac{\partial^2 f}{\partial y^2}\left(\frac{\partial y}{\partial s}\right)^2 + \frac{\partial f}{\partial x}\frac{\partial^2 x}{\partial s^2}\frac{\partial f}{\partial x}\frac{\partial^2 y}{\partial s^2}. \tag{1.52}$$

1.8 Change of Variables

Say we have a bivariate function $f(x, y)$ where $x = x(s, t)$ and $y = y(s, t)$. Here, we could express f purely in terms of s and t by substituting in for x and y, i.e. $f(x(s, t), y(s, t)) = f(s, t)$.

Example 1.13. Let $f(x, y) = xy$, $x = s^2 t$, $y = s \cos t$. Then

$$f(s, t) = (s^2 t)(s \cos t) = s^3 t \cos t. \tag{1.53}$$

If we want to evaluate partial derivatives, then we can use the chain rule using the fomulae in exercise 1.11.

A common change of variables in two and three dimensions is to change into *polar* coordinates.

Definition 1.4. In two dimensions, **plane polar coordinates** are defined as

$$x = r\cos\theta \qquad r = \sqrt{x^2 + y^2} \tag{1.54}$$

$$y = r\sin\theta \qquad \theta = \tan^{-1}\left(\frac{y}{x}\right). \tag{1.55}$$

Now any derivatives of a function with respect to cartesian coordinates can be transformed to derivatives with respect to polar coordinates using the formulae from before.

$$\frac{\partial x}{\partial r} = \cos\theta, \quad \frac{\partial x}{\partial \theta} = -r\sin\theta \tag{1.56}$$

$$\frac{\partial y}{\partial r} = \sin\theta, \quad \frac{\partial y}{\partial \theta} = r\cos\theta \tag{1.57}$$

$$\implies \frac{\partial f}{\partial r} = \cos\theta\frac{\partial f}{\partial x} + \sin\theta\frac{\partial z}{\partial y} \tag{1.58}$$

$$\frac{\partial f}{\partial \theta} = r\cos\theta\frac{\partial f}{\partial r} - r\sin\theta\frac{\partial f}{\partial x}. \tag{1.59}$$

The second derivatives can be written as

$$\frac{\partial^2 f}{\partial r^2} = \cos^2\theta\frac{\partial^2 f}{\partial x^2} + \sin^2\theta f y + 2\cos\theta\sin\theta\frac{\partial^2 f}{\partial x \partial y} \tag{1.60}$$

$$\frac{\partial^2 f}{\partial \theta^2} = -r\frac{\partial f}{\partial r} + r^2\left\{\sin^2\theta\frac{\partial^2 f}{\partial x^2} + \cos^2\theta\frac{\partial^2 f}{\partial y^2} - 2\cos\theta\sin\theta\frac{\partial^2 f}{\partial x \partial y}\right\}. \tag{1.61}$$

Example 1.14. Rewrite Laplace's equation, given in 2D cartesian coordinates as

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0, \tag{1.62}$$

in plane polar coordinates.

$$\frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2} = 0 \tag{1.63}$$

These formulae generalise to any number of variables. Suppose we have a function of n variables $f(x_1, x_2, \ldots, x_n)$ where each variable x_i is itself a function of m variables $x_i(t_1, t_2, \ldots, t_m)$. Then the partial derivative of f with respect to t_j is defined as

$$\frac{\partial f}{\partial t_j} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial t_j} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial t_j} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial t_j} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\frac{\partial x_i}{\partial t_j}. \tag{1.64}$$

Some examples in 3 dimensions include **cylindrical polar coordinates**

$$x = \rho\cos\phi \qquad \rho = \sqrt{x^2 + y^2} \tag{1.65}$$

$$y = \rho\sin\phi \qquad \phi = \tan^{-1}\left(\frac{y}{x}\right) \tag{1.66}$$

$$z = z \tag{1.67}$$

and **spherical polar coordinates**

$$x = r\sin\theta\cos\phi \qquad r = \sqrt{x^2 + y^2 + z^2} \tag{1.68}$$

$$y = r\sin\theta\sin\phi \qquad \theta = \tan^{-1}\left(\frac{\sqrt{x^2 + y^2}}{z}\right) \tag{1.69}$$

$$z = r\cos\theta \qquad \phi = \tan^{-1}\left(\frac{y}{x}\right). \tag{1.70}$$

1.9 Directional Derivative

We have seen that for a function of n variables, we can define partial derivatives in each of the n directions given by the coordinate axes. However, what if we want to define a rate of change along one of the infinitely many other directions not aligned with any coordinate axis? Consider a function in 3D $f(x, y, z)$ and let c_1, c_2 be two real numbers such that $f(x, y, z) = c_1$ and $f(x, y, z) = c_2$ define two surfaces $\mathcal{S}_1$ and $\mathcal{S}_2$. Imagine we have two points P and Q on $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively. How does the function f vary along the line joining P and Q? Let $P = (x_1, y_1, z_1)$, and let $\hat{\boldsymbol{u}} = l\hat{\boldsymbol{\imath}} + m\hat{\boldsymbol{\jmath}} + n\hat{\boldsymbol{k}}$ be a unit vector in the direction of $\overrightarrow{PQ}$. Then we can parameterise the points on the line through P and Q by some parameter t as

$$x = x_1 + tl, \quad y = y_1 + tm, \quad z = z_1 + tn. \tag{1.71}$$

Now we can calculate the (full) derivative of f with respect to s using the chain rule:

$$
\begin{aligned}
\frac{\mathrm{d}f}{\mathrm{d}t} &= \frac{\partial f}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial f}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} + \frac{\partial f}{\partial z}\frac{\mathrm{d}z}{\mathrm{d}t} && (1.72)\\
&= \frac{\partial f}{\partial x}l + \frac{\partial f}{\partial y}m + \frac{\partial f}{\partial z}n && (1.73)\\
&= \nabla f \cdot \hat{\boldsymbol{u}}, && (1.74)
\end{aligned}
$$

where we defined $\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right)$. We call $\nabla f \cdot \hat{\boldsymbol{u}}$ the **directional derivative** of f in the direction $\hat{\boldsymbol{u}}$ and most of the time denote it $\nabla_{\hat{\boldsymbol{u}}} f$. ∇f is called the **gradient** of f. It is a vector-valued derivative which defines a kind of "slope field" that gives the directional derivative in some direction when dotted with a unit vector. As we can see from the definition, finding the dot product of ∇f with any of the cartesian unit vectors gives us the partial derivatives we were working with before.

Example 1.15. Condsider the function $f(x, y, z) = xy^2z^3$. Find the directional derivative in the direction of the vector $\boldsymbol{u} = 2\hat{\boldsymbol{\imath}} + 6\hat{\boldsymbol{\jmath}} + 3\hat{\boldsymbol{k}}$ at the point $(1, 1, 1)$.
Start by finding the gradient of f:

$$
\begin{aligned}
\nabla f &= \frac{\partial f}{\partial x}\hat{\boldsymbol{\imath}} + \frac{\partial f}{\partial y}\hat{\boldsymbol{\jmath}} + \frac{\partial f}{\partial z}\hat{\boldsymbol{k}} && (1.75)\\
&= y^2z^3\hat{\boldsymbol{\imath}} + 2xyz^3\hat{\boldsymbol{\jmath}} + 3xy^2z^2\hat{\boldsymbol{k}}. && (1.76)
\end{aligned}
$$

Therefore, at the point $(1, 1, 1)$ the gradient has the value $\hat{\boldsymbol{\imath}} + 2\hat{\boldsymbol{\jmath}} + 3\hat{\boldsymbol{k}}$. Hence

$$
\begin{aligned}
\nabla_{\hat{\boldsymbol{u}}} f|_{(1,1,1)} &= (\hat{\boldsymbol{\imath}} + 2\hat{\boldsymbol{\jmath}} + 3\hat{\boldsymbol{k}}) \cdot \left(\frac{2\hat{\boldsymbol{\imath}} + 6\hat{\boldsymbol{\jmath}} + 3\hat{\boldsymbol{k}}}{\sqrt{2^2 + 6^2 + 3^2}}\right) && (1.77)\\
&= (\hat{\boldsymbol{\imath}} + 2\hat{\boldsymbol{\jmath}} + 3\hat{\boldsymbol{k}}) \cdot \left(\frac{2}{7}\hat{\boldsymbol{\imath}} + \frac{6}{7}\hat{\boldsymbol{\jmath}} + \frac{3}{7}\hat{\boldsymbol{k}}\right) && (1.78)\\
&= \frac{23}{7}. && (1.79)
\end{aligned}
$$

1.10 Tangent Planes

Let $\boldsymbol{t}$ be a tangent vector to the surface $f(x, y, z) = c$ at a point P. Then the directional derivative of f in the direction of $\boldsymbol{t}$ has to be 0.

$$
\nabla_{\hat{\boldsymbol{t}}} f|_P = \nabla f|_P \cdot \hat{\boldsymbol{t}} = 0 \tag{1.80}
$$

This implies that the gradient ∇f is perpendicular to $\boldsymbol{t}$ at P, i.e. it is *normal* to the surface. Thus, the gradient is *always* perpendicular to level sets.

Example 1.16. Consider the surface $f(x, y, z) = x - y^2 + xz = -1$. Find a vector normal to the surface at $P = (1, 2, 2)$.

Calculate the gradient and evaluate it at P:

$$\nabla f = \frac{\partial f}{\partial x}\hat{\boldsymbol{\imath}} + \frac{\partial f}{\partial y}\hat{\boldsymbol{\jmath}} + \frac{\partial f}{\partial z}\hat{\boldsymbol{k}} = (1+z)\hat{\boldsymbol{\imath}} - 2y\hat{\boldsymbol{\jmath}} + x\hat{\boldsymbol{k}} \tag{1.81}$$

$$\implies \nabla f|_P = 3\hat{\boldsymbol{\imath}} - 4\hat{\boldsymbol{\jmath}} + \hat{\boldsymbol{k}}. \tag{1.82}$$

Note that we for any vector $\boldsymbol{u}$ we have

$$\nabla f|_P \cdot \hat{\boldsymbol{u}} = |\nabla f|_P||\hat{\boldsymbol{u}}|\cos\theta = |\nabla f|_P|\cos\theta, \tag{1.83}$$

where θ is the angle between the two vectors. This shows that the maximum rate of change of f is in the direction of ∇f (perpendicular to the level sets). We can also get tangent vectors by setting the above equation to 0 as we saw before. It may seem like there is only one equation for 3 variables, but in fact there is only 1 degree of freedom because we are looking for a vector in a tangent plane.

For a surface $f(x, y, z) = c$ and a point $P = (x_0, y_0, z_0)$ on the surface, the tangent plane at P is given by the equation

$$(x - x_0)\left.\frac{\partial f}{\partial x}\right|_P + (y - y_0)\left.\frac{\partial f}{\partial y}\right|_P + (z - z_0)\left.\frac{\partial f}{\partial z}\right|_P = 0 \tag{1.84}$$

Example 1.17. Consider the surface $f(x, y, z) = xy^2 + x^2 z = 7$ and the point $P = (1, 2, 3)$.

$$\nabla f = (y^2 + 2xz)\hat{\boldsymbol{\imath}} + 2xy\hat{\boldsymbol{\jmath}} + x^2\hat{\boldsymbol{k}} \tag{1.85}$$

$$\implies \nabla f|_P = 10\hat{\boldsymbol{\imath}} + 4\hat{\boldsymbol{\jmath}} + \hat{\boldsymbol{k}}. \tag{1.86}$$

Therefore the equation for the tangent plane at P is given by

$$10(x - 1) + 4(y - 2) + (z - 1) = 0 \tag{1.87}$$

$$\implies 10x + 4y + z = 21. \tag{1.88}$$

Chapter 2

Optimisation

2.1 Taylor Series of Multivariate Functions

First we recap some facts about Taylor series for univariate functions. If a single-variable function f is **analytic**, then it can be represented around a point x_0 by its Taylor series:

$$f(x) = \sum_{n=0}^{\infty} \frac{(x-x_0)^n}{n!} \frac{\mathrm{d}^n f}{\mathrm{d}x^n} = f(x_0) + (x-x_0) \left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{x_0} + \frac{(x-x_0)^2}{2} \left.\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}\right|_{x_0} + \dots. \tag{2.1}$$

This power series is valid for any value of x within the **radius of convergence**, given by

$$|x - x_0| < \lim_{n\to\infty} \left| \frac{\frac{1}{n!}\frac{\mathrm{d}^n f}{\mathrm{d}x^n}}{\frac{1}{(n+1)!}\frac{\mathrm{d}^{n+1} f}{\mathrm{d}x^{n+1}}} \right| = \lim_{n\to\infty} \left| (n+1) \frac{\frac{\mathrm{d}^n f}{\mathrm{d}x^n}}{\frac{\mathrm{d}^{n+1} f}{\mathrm{d}x^{n+1}}} \right| = R. \tag{2.2}$$

Recall that we can differentiate a power series without changing the radius of convergence. If we truncate the power series after a finite numbers of terms $n = k$, then there is an error between the value of the power series and the actual function which is given by

$$R_{k+1} = \frac{(x-x_0)^{k+1}}{(k+1)!} \left.\frac{\mathrm{d}^{k+1} f}{\mathrm{d}x^{k+1}}\right|_c, \tag{2.3}$$

where c is a real number between x and x_0. This is Taylor's theorem and it allows us to obtain bounds on the size of the error.

Example 2.1. What is the maximum possible error for the degree 4 Taylor polynomial for the function e^x expanded about $x_0 = 0$ and used to approximate the function at $x = 0.2$?

To examine a function at a point outside the radius of convergence, we have to change the expansion point x_0. We can do this by substituting.

Example 2.2. The function $f(u) = \frac{1}{1-u}$ has a power series representation about the point $u_0 = 0$ which has a radius of convergence $R = 1$.

$$\frac{1}{1+u} = 1 - u + u^2 - u^3 + \dots \quad -1 < u < 1. \tag{2.4}$$

Suppose we want a series representation for this function which convergences for $-1 < x < 3$. We make the substitution $u = \frac{x-1}{2}$ (equation of a straight line)

$$\frac{1}{1+\frac{x-1}{2}} = 1 - \left(\frac{x-1}{2}\right) + \left(\frac{x-1}{2}\right)^2 - \left(\frac{x-1}{2}\right)^3 + \dots \quad -1 < \frac{x-1}{2} < 1. \tag{2.5}$$

Rearranging this equation gives

$$\frac{1}{1+x} = \frac{1}{2}\left[1 - \left(\frac{x-1}{2}\right) + \left(\frac{x-1}{2}\right)^2 - \left(\frac{x-1}{2}\right)^3 + \dots\right]. \tag{2.6}$$

Lets now try to find a Taylor series for a function of two variables. A taylor polynomial is the a polynomial approximation for a function $f(x, y)$ about a point (x_0, y_0). So we want to capture how f *varies* in a local area around (x_0, y_0). Consider a small line segment given by $x = x_0 + ht$, $y = y_0 + kt$. Then the first and second derivatives with respect to the parameter t is given by

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial f}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} = h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y} \tag{2.7}$$

$$\frac{\mathrm{d}^2 f}{\mathrm{d}t^2} = \frac{\mathrm{d}}{\mathrm{d}t}\left(h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y}\right) = h\left(\frac{\partial^2 f}{\partial x^2}h + \frac{\partial^2 f}{\partial x \partial y}k\right) + k\left(\frac{\partial^2 f}{\partial x \partial y}h + \frac{\partial^2 f}{\partial y^2}k\right) \tag{2.8}$$

$$= h^2\frac{\partial^2 f}{\partial x^2} + 2kh\frac{\partial^2 f}{\partial x \partial y} + k^2\frac{\partial^2 f}{\partial y^2}. \tag{2.9}$$

Now we can express $f(x(t), y(t)) = z(t)$ and expand $z(t)$ in a Taylor series about $t = 0$:

$$z(t) = \underbrace{z(0)}_{f(x_0,y_0)} + t\underbrace{\left.\frac{\mathrm{d}z}{\mathrm{d}t}\right|_{t=0}}_{f'(x_0,y_0)} + \frac{t^2}{2}\underbrace{\left.\frac{\mathrm{d}^2 z}{\mathrm{d}t^2}\right|_{t=0}}_{f''(x_0,y_0)} + \dots. \tag{2.10}$$

Note that $ht = x - x_0$, $kt = y - y_0$, so

$$t\left.\frac{\mathrm{d}z}{\mathrm{d}t}\right|_{t=0} = (x - x_0)\left.\frac{\partial f}{\partial x}\right|_{(x_0,y_0)} + (y - y_0)\left.\frac{\partial f}{\partial y}\right|_{(x_0,y_0)} \tag{2.11}$$

$$\frac{t^2}{2}\left.\frac{\mathrm{d}^2 z}{\mathrm{d}t^2}\right|_{t=0} = \frac{(x - x_0)^2}{2}\left.\frac{\partial^2 f}{\partial x^2}\right|_{(x_0,y_0)} + \frac{(y - y_0)^2}{2}\left.\frac{\partial^2 f}{\partial y^2}\right|_{(x_0,y_0)} \tag{2.12}$$

$$+ (x - x_0)(y - y_0)\left.\frac{\partial^2 f}{\partial x \partial y}\right|_{(x_0,y_0)}. \tag{2.13}$$

Hence our Taylor expansion for a function of two variables about a point (x_0, y_0) is

$$f(x, y) = f(x_0, y_0) + (x - x_0)\left.\frac{\partial f}{\partial x}\right|_{(x_0,y_0)} + (y - y_0)\left.\frac{\partial f}{\partial y}\right|_{(x_0,y_0)} + \frac{(x - x_0)^2}{2}\left.\frac{\partial^2 f}{\partial x^2}\right|_{(x_0,y_0)} \tag{2.14}$$

$$+ \frac{(y - y_0)^2}{2}\left.\frac{\partial^2 f}{\partial y^2}\right|_{(x_0,y_0)} + (x - x_0)(y - y_0)\left.\frac{\partial^2 f}{\partial x \partial y}\right|_{(x_0,y_0)} \tag{2.15}$$

Example 2.3. Find a second-order Taylor approximation for the function $f(x, y) = \sin xe^y$ about the point $(0, 0)$.

$$
\begin{aligned}
f(0,0) &= 0 \quad f_{xx}(0,0) = 0 &(2.16)\\
f_x(0,0) &= 1 \quad f_{yy}(0,0) = 0 &(2.17)\\
f_y(0,0) &= 0 \quad f_{xy}(0,0) = 1. &(2.18)
\end{aligned}
$$

Hence by the formula above the Taylor expansion is

$$f(x, y) = x + xy + \dots \tag{2.19}$$

2.2 Stationary Points

For functions of one variable the possible types of stationary points are minima, maxima, and inflection points. For multivariate functions, we can have all of these as well as **saddle points**, which are like a minimum in one direction but a maximum in another. To investigate the nature of stationary points for a 1D function, we look at the second derivative. Consider the Taylor series of a 1D function:

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + \frac{(x - x_0)^2}{2} f''(x_0) + \dots \tag{2.20}$$

If we set $f'(x_0) = 0$ (stationary point), then we have

$$f(x) - f(x_0) = \underbrace{\frac{(x - x_0)^2}{2}}_{\text{alway positive}} f''(x_0) + \dots. \tag{2.21}$$

So if $f''(x_0) > 0$, then $f(x) - f(x_0) > 0$ and $f(x_0)$ is a **local minimum**. Similarly if $f''(x_0) < 0$, then $f(x_0)$ is a **local maximum**. If $f''(x_0) = 0$ but $f'''(x_0) \neq 0$, then

$$f(x) - f(x_0) = f'''(x_0) \underbrace{\frac{(x - x_0)^3}{6}}_{\text{changes sign depending on } x} + \dots. \tag{2.22}$$

This leads to an inflection point.

We can use the same technique to investigate stationary points of two-variable functions. Using equation 2.14, if (x_0, y_0) is a stationary point then $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$ and

$$
\begin{aligned}
f(x, y) - f(x_0, y_0) &= \frac{(x - x_0)^2}{2} f_{xx}(x_0, y_0) + \frac{(y - y_0)^2}{2} f_{yy}(x_0, y_0) &(2.23)\\
&+ (x - x_0)(y - y_0) f_{xy}(x_0, y_0) + \dots. &(2.24)
\end{aligned}
$$

First consider the case where $f_{xx}(x_0, y_0) = f_{yy}(x_0, y_0) = 0$. Then we have

$$f(x, y) - f(x_0, y_0) = (x - x_0)(y - y_0) f_{xy}(x_0, y_0) + \dots. \tag{2.25}$$

The sign varies since $x - x_0$ and $y - y_0$ can be positive or negative. Therefore this is a **saddle point**.

Now consider the case where $f_{xx}(x_0, y_0) \neq 0$. In this case we can write

$$\Delta f = \frac{1}{f_{xx}}\left(\frac{\Delta x^2}{2} f_{xx}^2 + \frac{\Delta y^2}{2} f_{yy} f_{xx} + \Delta x \Delta y f_{xy} f_{xx}\right) + \dots \tag{2.26}$$

$$= \frac{1}{f_{xx}}\left(\frac{\Delta x^2}{2} f_{xx}^2 + \Delta x \Delta y f_{xy} f_{xx} + \Delta y^2 f_{xy}^2 - \Delta y^2 f_{xy}^2 + \Delta y^2 f_{yy} f_{xx}\right) + \dots \tag{2.27}$$

$$= \frac{1}{f_{xx}}\left\{\left(\frac{\Delta x}{\sqrt{2}} f_{xx} + \frac{\Delta y}{\sqrt{2}} f_{yy}\right)^2 + \frac{\Delta y^2}{2}(f_{xx} f_{yy} - f_{xy}^2)\right\} + \dots . \tag{2.28}$$

The quantity in the first set of parentheses is always positive, whereas the quantity in the second set, often called the **discriminant** $\Delta = f_{xx} f_{yy} - fxy^2$ can be positive or negative. If $\Delta > 0$ and $f_{xx} > 0$, then the whole expression is > 0 and we have a **local minimum**. If $\Delta > 0$ but $f_{xx} < 0$, then $\Delta f < 0$ and we have a **local maximum**. Finally, if $\Delta < 0$ then the sign can change (why?) and we have a **saddle point**. The exact same analysis can be applied to the case $f_{yy} \neq 0$.

Example 2.4. Consider the function $f(x, y) = x^3 - 6xy + y^3$. Stationary points occur when

$$f_x = 3x^2 - 6y = 0 \tag{2.29}$$

$$f_y = 3y^2 - 6x = 0. \tag{2.30}$$

These equations are satisfied at the points $(0, 0)$ and $(2, 2)$. Let's calculate the discriminant and find the nature of the stationary points.

$$\Delta = f_{xx} f_{yy} - f_{xy}^2 = (6x)(6y) - (-6)^2 = 36xy - 36. \tag{2.31}$$

Hence at $(0, 0)$, $\Delta = -36$ so we have a saddle point. At $(2, 2)$, $\Delta = 108$ and $f_{xx} = 24$ so we have a local minimum.

Example 2.5. Consider $f(x, y) = x^4 + y^4 - 2(x - y)^2$. We have stationary points when

$$f_x = 4x^3 - 4(x - y) = 0 \tag{2.32}$$

$$f_y = 4y^3 + 4(x - y) = 0. \tag{2.33}$$

Adding the two equations gives $x = -y$, then using one of the equations gives the solutions $(0, 0)$, $(\sqrt{2}, -\sqrt{2})$ and $(-\sqrt{2}, \sqrt{2})$.

$$\Delta = f_{xx} f_{yy} - f_{xy}^2 = (12x^2 - 4)(12y^2 - 4) - (4)^2 = 144x^2y^2 - 48x^2 - 48y^2. \tag{2.34}$$

At $(\sqrt{2}, -\sqrt{2})$, $\Delta = 384 > 0$ and $f_{xx} = 20 > 0$ so this is a local minimum. By the same argument, there is also a local minimum at $(-\sqrt{2}, \sqrt{2})$. At $(0, 0)$, $\Delta = 0$ so we can't say anything about the nature of the stationary point using the discriminant. Take $x = y$, then $f(x, y = x) = 2x^4 \geq 0$. But now consider small x and y such that $(x - y)^2 > x^4 + y^4$, then $f(x, y) < 0$. So since the sign can change, $(0, 0)$ must be a saddle point.

2.3 Optimisation Subject to Constraints

Suppose we want to find the minimum or maximum of a functions along a certain path. This is equivalent to optimising the function with a **constraint** on the variables. There are many ways of solving these optimisation problems, but two of the most common are the **substitution method** and the **method of Lagrange multipliers**. The substitution method works well for simple problems where we can separate the variables in the constraint. For more complex problems, the method of Lagrange multipliers is very powerful.

Let's look at the substitution method first. Suppose we have a bivariate function $f(x, y)$ and a constraint between the two variables $y = y(x)$. Then we can substitution $f(x, y = y(x))$ which defines another function $g(x)$ that we can differentiate to get the extremum.

Example 2.6. Minimise $f(x, y) = x^2 + y^2$ subject to the constraint $y = 1 - x$. We define the composite function $g(x)$ as

$$g(x) = f(x, y = 1 - x) = x^2 + (1 - x)^2 = 2x^2 - 2x + 1. \tag{2.35}$$

We can now find the derivative and set it equal to 0:

$$g'(x) = 4x - 2 = 0. \tag{2.36}$$

So the function $f(x, y)$ is minimised along the path at $x = \frac{1}{2}$, $y = 1 - \frac{1}{2} = \frac{1}{2}$.

Example 2.7. Minimise $f(x, y, z) = x^2 + 2y^2 + z$ subject to the constraint $x + y^2 - z = 0$.

$$g(x, y) = f(x, y, z = x + y^2) = x^2 + x + 3y^2. \tag{2.37}$$

$$\frac{\partial g}{\partial x} = 2x + 1 = 0, \quad \frac{\partial g}{\partial y} = 6y = 0, \tag{2.38}$$

so $x = -\frac{1}{2}$ and $y = 0$. It can be shown using the second derivatives that this is a minimum. Hence the minimum of $f(x, y, z)$ subject to the constraint is $f(-\frac{1}{2}, 0, -\frac{1}{2}) = -\frac{1}{4}$.

2.4 Method of Lagrange Multipliers

If we can't solve an optimisation problem using the substitution method then the method of Lagrange multipliers is the next best thing. Suppose we have a function of three variables $f(x, y, z)$ and a constraint which we write in the general form $g(x, y, z) = 0$. At an stationary point on the constraint surface, the directional derivative of f in a direction tangent to the surface will be 0. Since $g(x, y, z) = 0$ defines a level set, the gradient of g is normal to the surface everywhere. Thus, at a stationary point on the constraint surface, the gradient of f is *parallel* to the gradient of g.

$$\exists \lambda \in \mathbb{R} \text{ such that } \nabla f = \lambda \nabla g. \tag{2.39}$$

We call λ the **Lagrange multiplier**.

In order to optimise a function subject to a constraint using this fact, we construct the function

$$F(x,y,z,\lambda) = f(x,y,z) - \lambda g(x,y,z). \tag{2.40}$$

Then setting all the first derivatives equal to 0 gives

$$\frac{\partial F}{\partial x} = \frac{\partial f}{\partial x} - \lambda\frac{\partial g}{\partial x} = 0 \implies \frac{\partial f}{\partial x} = \lambda\frac{\partial g}{\partial x} \tag{2.41}$$
$$\frac{\partial F}{\partial y} = \frac{\partial f}{\partial y} - \lambda\frac{\partial g}{\partial y} = 0 \implies \frac{\partial f}{\partial y} = \lambda\frac{\partial g}{\partial y} \tag{2.42}$$
$$\frac{\partial F}{\partial z} = \frac{\partial f}{\partial z} - \lambda\frac{\partial g}{\partial z} = 0 \implies \frac{\partial f}{\partial z} = \lambda\frac{\partial g}{\partial z} \tag{2.43}$$
$$\frac{\partial F}{\partial \lambda} = -g(x,y,z) = 0. \tag{2.44}$$

Note that the first three equations are exactly the condition $\nabla f = \lambda \nabla g$.

Example 2.8. Minimise $f(x,y) = x^2+y^2$ subject to the constraint $g(x,y) = x+y-1 = 0$. This is the same problem from before which we can solve using the substitution method, but let's now solve it with the method of Lagrange multipliers. First we construct the function $F(x,y,\lambda)$, find the derivatives and set them equal to 0.

$$F(x,y,\lambda) = f(x,y) - \lambda g(x,y) = x^2 - y^2 - \lambda(x+y-1). \tag{2.45}$$

$$\frac{\partial F}{\partial x} = 2x - \lambda = 0 \tag{2.46}$$
$$\frac{\partial F}{\partial y} = 2y - \lambda = 0 \tag{2.47}$$
$$\frac{\partial F}{\partial \lambda} = x + y - 1 = 0. \tag{2.48}$$

Now adding the first two equations and substituting the last gives

$$2x + 2y - 2\lambda = 0 \tag{2.49}$$
$$x + y = \lambda \tag{2.50}$$
$$\implies \lambda = 1. \tag{2.51}$$

So using the first two equations gives $x = \frac{1}{2}$ and $y = \frac{1}{2}$ like we found before.

Example 2.9. Find and categorise the stationary points of the function $f(x,y,z) = x+y+z$ on the surface of a sphere centered at the origin with radius 5. This constraint takes the form $g(x,y,z) = x^2+y^2+z^2-5^2 = 0$. The function $F(x,y,z,\lambda)$ and its derivatives (set equal to 0) are:

$$F(x,y,z,\lambda) = x + y + z - \lambda(x^2 + y^2 + z^2 - 25) \tag{2.52}$$
$$= \frac{\partial F}{\partial x} = 1 - 2x\lambda = 0 \tag{2.53}$$
$$= \frac{\partial F}{\partial y} = 1 - 2y\lambda = 0 \tag{2.54}$$
$$= \frac{\partial F}{\partial z} = 1 - 2z\lambda = 0 \tag{2.55}$$
$$= \frac{\partial F}{\partial \lambda} = x^2 + y^2 + z^2 - 25 = 0. \tag{2.56}$$

Thus, we have $x = y = z = \frac{1}{2\lambda}$ by the first three derivatives and $\frac{3}{4\lambda^2} = 25$ by the last. From this we we obtain $\lambda = \pm\frac{\sqrt{3}}{10}$ which gives $x = y = z = \pm\frac{5}{\sqrt{3}}$. We therefore have two stationary points, $(\frac{5}{\sqrt{3}}, \frac{5}{\sqrt{3}}, \frac{5}{\sqrt{3}})$ and $(-\frac{5}{\sqrt{3}}, -\frac{5}{\sqrt{3}}, -\frac{5}{\sqrt{3}})$. To find the nature of the stationary points, we expand about the stationary points (we will use the constraint to ensure we remain on the sphere):

$$x = \frac{5}{\sqrt{3}} + \varepsilon \tag{2.57}$$
$$y = \frac{5}{\sqrt{3}} + \delta \tag{2.58}$$
$$z = \sqrt{25 - \left(\frac{5}{\sqrt{3}} + \varepsilon\right)^2 - \left(\frac{5}{\sqrt{3}} + \delta\right)^2} \tag{2.59}$$
$$= \frac{5}{\sqrt{3}}\sqrt{1 - \frac{2\sqrt{3}}{5}(\varepsilon + \delta) - \frac{2}{25}(\varepsilon^2 + \delta^2)} \tag{2.60}$$
$$\approx \frac{5}{\sqrt{3}}(1 - \frac{\sqrt{3}}{5}(\varepsilon + \delta) - \frac{3}{50}(\varepsilon^2 + \delta^2) - \frac{\sqrt{3}}{20}(\varepsilon + \delta)^2) \tag{2.61}$$
$$\tag{2.62}$$

Where in the last line we have used a Taylor series expansion. Hence

$$f(x,y,z) = x + y + z \tag{2.63}$$
$$= \frac{5}{\sqrt{3}} + \varepsilon + \frac{5}{\sqrt{3}} + \delta + \frac{5}{\sqrt{3}} - (\varepsilon + \delta) - \frac{15}{50\sqrt{3}}(\varepsilon^2 + \delta^2) - \frac{1}{4}(\varepsilon + \delta)^2 \tag{2.64}$$
$$= 5\sqrt{3} - \text{something positive (and small?)} \tag{2.65}$$

This is > 0, so we have a local maximum.

Example 2.10. Consider an ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. What is the largest rectangle that can fit inside the ellipse? To turn this into an optimisation problem that we can solve, consider the top right corner of the ellipse (x, y). Due to symmetry, the area of a rectangle with a corner at (x, y) is $4xy$, and this is the function that we want to *minimise* under the constraint that x and y lie on the ellipse. Using the method of Lagrange multipliers:

$$F(x, y, \lambda) = 4xy - \lambda\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1\right) \tag{2.66}$$

$$\frac{\partial F}{\partial x} = 4y - \frac{2x\lambda}{a^2} = 0 \tag{2.67}$$

$$\frac{\partial F}{\partial y} = 4x - \frac{2y\lambda}{b^2} = 0 \tag{2.68}$$

$$\frac{\partial F}{\partial \lambda} = \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0. \tag{2.69}$$

From the x and y derivatives, we get $x = \frac{y\lambda}{2b^2}$ and $y = \frac{x\lambda}{2a^2}$. Combining these equations with the last one, we find $\lambda = 2ab$. Hence by back-substitution we get $x = \frac{a}{\sqrt{2}}$ and $y = \frac{b}{\sqrt{2}}$, which implies $f(x, y) = 2ab$.

Chapter 3

Integration Techniques

3.1 Rational Functions and Square Roots of Rational Functions

The first thing to look for is if the denominator can be factorised. In this case, the first action should be to split the fraction up into partial fractions. If $f(x) = \frac{p(x)}{q(x)}$ with $\deg(p) < \deg(q)$ and $q(x) = (x - a_1)(x - a_2)\dots(x - a_n)$, then

$$f(x) = \frac{A_1}{x - a_1} + \frac{A_2}{x - a_2} + \dots + \frac{A_n}{x - a_n}. \tag{3.1}$$

Example 3.1.

$$\frac{1}{x^2 - 1} = \frac{A}{x - 1} + \frac{B}{x + 1} = \frac{1}{2(x - 1)} + \frac{1}{2(x + 1)}. \tag{3.2}$$

Example 3.2.

$$\frac{3x + 10}{(x - 2)(x^2 + 4)} = \frac{A}{x - 2} + \frac{Bx + C}{x^2 + 4} \tag{3.3}$$

$$\int \frac{3x + 10}{(x - 2)(x^2 + 4)}\, dx = \int \left(\frac{2}{x - 2} + \frac{-2x - 1}{x^2 + 4} \right) dx \tag{3.4}$$

$$= \int \left(\frac{2}{x - 2} - \frac{2x}{x^2 + 4} - \frac{1}{x^2 + 4} \right) dx \tag{3.5}$$

$$= 2\ln|x - 2| - \ln(x^2 + 4) - \frac{1}{2}\tan^{-1}\left(\frac{x}{2}\right) + c. \tag{3.6}$$

3.2 Integration by Substitution

If $f(x) = h(g(x))g'(x)$ where $h(x)$ is a function we know the integral of, we can use the technique of substitution. Note that

$$\frac{\mathrm{d}}{\mathrm{d}x}(H(g(x))) = h(g(x))g'(x), \tag{3.7}$$

therefore

$$\int f(x)\mathrm{d}x = H(g(x)) + c. \tag{3.8}$$

To solve, make the substitution $u = g(x)$ to use this formula.

Expression inside Integral	Suggested Substitution
$\sqrt{a^2 - x^2}$	$x = a\sin\theta$ or $a\cos\theta$
$a^2 + x^2$	$x = a\tan\theta$
$\sqrt{x^2 - a^2}$	$x = a\cosh\theta$
$\sqrt{x^2 + a^2}$	$x = a\sinh\theta$
$a^2 - x^2$	$x = a\tan\theta$

3.3 Derivatives of Trigonometric Functions

Some fractions integrate to inverse trigonometric formulae.

Example 3.3. Note that $\sqrt{1-x^2}$ is the equation for the unit circle.

$$\int \frac{1}{\sqrt{1-x^2}}\,dx = \sin^{-1} x + c$$

$$\int \frac{1}{1+x^2}\,dx = \tan^{-1} x + c$$

$$\int \frac{f'(x)}{\sqrt{a^2 - [f(x)]^2}}\,d = \sin^{-1}\frac{f(x)}{a} + c$$

$$\int \frac{f'(x)}{a^2 + [f(x)]^2}\,dx = \frac{1}{a}\tan^{-1}\frac{f(x)}{a} + c$$

Some other fractions integrate to logs.

Example 3.4. Possible substitution: let $x = \sqrt{k}\tan\theta$.

$$\int \frac{1}{\sqrt{x^2 \pm k}}\,dx = \ln|x + \sqrt{x^2 \pm k}| + c.$$

3.4 Solving Integrals Using Recurrence Relations

The following technique is useful for evaluating integrals of functions raised to an arbitrary integer power. Consider the integral of $\ln x$. This can be found using integration by parts:

$$\int \ln x\mathrm{d}x = x\ln x - x + c. \tag{3.9}$$

Similarly, the integral of $(\ln x)^2$ is

$$\int (\ln x)^2 \mathrm{d}x = x(\ln x)^2 - 2\int \ln x \mathrm{d}x. \tag{3.10}$$

Hence the integral of $(\ln x)^n$ is given by

$$I_n = \int (\ln x)^n \mathrm{d}x = x(\ln x)^n - n\int (\ln x)^{n-1} \mathrm{d}x = x(\ln x)^n - nI_{n-1}. \tag{3.11}$$

Example 3.5. Let $I_n = \int x^n e^x \mathrm{d}x$. Integrating by parts, we get

$$I_n = x^n e^x - \int n x^{n-1} e^x \mathrm{d}x = x^n e^x - nI_{n-1}. \tag{3.12}$$

Example 3.6. Let $I_n = \int \sin^n x \mathrm{d}x$.

Sometimes it is possible to use this technique without using integration by parts.

Example 3.7. Consider $I_n = \int \tan^n x \mathrm{d}x$.

3.5 Standard Antiderivatives

Here are some standard antiderivatives.

$$\int e^{ax} \mathrm{d}x = \frac{1}{a} e^{ax} + c \tag{3.13}$$

$$\int \sin(ax) \mathrm{d}x = -\frac{1}{a} \cos(ax) + c \tag{3.14}$$

$$\int \cos(ax) \mathrm{d}x = \frac{1}{a} \sin(ax) + c \tag{3.15}$$

$$\int \tan(ax) \mathrm{d}x = -\frac{1}{a} \ln|\cos(ax)| + c \tag{3.16}$$

$$\int \sec^2(ax) \mathrm{d}x = \frac{1}{a} \tan(ax) + c \tag{3.17}$$

$$\int \frac{1}{\sqrt{a^2 - x^2}} dx = \sin^{-1}\left(\frac{x}{|a|}\right) + c \tag{3.18}$$

$$\int \frac{-1}{\sqrt{1 - x^2}} dx = \cos^{-1}(x) + c \tag{3.19}$$

$$\int \frac{1}{a^2 + x^2} dx = \frac{1}{a} \tan^{-1}\left(\frac{x}{a}\right) + c \tag{3.20}$$

$$\int \frac{1}{\sqrt{a^2 + x^2}} dx = \sinh^{-1}\left(\frac{x}{|a|}\right) + c \tag{3.21}$$

$$\int \frac{1}{\sqrt{x^2 - 1}} dx = \cosh^{-1}(x) + c \tag{3.22}$$

$$\int \frac{1}{1 - x^2} dx = \tanh^{-1}(x) + c. \tag{3.23}$$

Chapter 4

Double Integrals

4.1 Multivariate Integration

When integrating over more that one variable, we integrate over an **area element** $\mathrm{d}A$ or **volume element** $\mathrm{d}V$ as opposed to a **line element** $\mathrm{d}x$ in 1D. Just as an integral of a 1D function represents the *area* under a curve, an integral of a 2D function represents the *volume* under a surface. We can also find the mean value of a function over a region by

$$\bar{f}(x,y) = \frac{\iint_A f(x,y)\mathrm{d}A}{\iint_A \mathrm{d}A}. \tag{4.1}$$

4.2 Integration over Rectangular Regions

A rectangular region is bounded by four sets of cartesian coordinates. The area element is given by $\mathrm{d}A = \mathrm{d}x\mathrm{d}y$.

$$I = \int_{y=b}^{y=d} \int_{x=a}^{x=c} f(x,y)\mathrm{d}x\mathrm{d}y. \tag{4.2}$$

We can evaluate this integral by choosing one of the 1D integrals to evaluate first while keeping the other variable constant, just like we do when calculating partial derivatives. If we can find $F(x,y)$ satisfying $\frac{\partial F}{\partial x} = f(x,y)$, then $g(y) = F(c,y) - F(a,y) = F(x,y) + c(y)$.

$$\int_a^c f(x,y)\mathrm{d}x = g(y) \tag{4.3}$$

$$\implies I = \int_b^d g(y)\mathrm{d}y. \tag{4.4}$$

This is possible due to **Fubini's theorem**

Example 4.1. Calculate the integral of $f(x,y) = \sqrt{xy}$ over a rectangular region with opposite corners at $(1,4)$ and $(4,9)$.

$$I = \int_4^9 \int_1^4 \sqrt{xy}\mathrm{d}x\mathrm{d}y = \int_4^9 \frac{2}{3}\sqrt{x^3y}\bigg|_1^4 \mathrm{d}y = \int_4^9 \frac{14}{3}\sqrt{y}\mathrm{d}y = \frac{37}{9}y^{\frac{3}{2}}\bigg|_4^9 = \frac{532}{9}. \tag{4.5}$$

Example 4.2. Calculate the integral of $f(x,y) = \frac{y}{(x+y^2)^2}$ over a rectangular region with opposite corners at $(1,4)$ and $(4,9)$.

$$I = \int_1^2 \int_0^1 \frac{y}{(x+y^2)^2} \mathrm{d}x\mathrm{d}y = \int_1^2 \left. \frac{-y}{x+y^2} \right|_0^1 \mathrm{d}y = \ln y - \frac{1}{2}\ln(1+y^2)\Big|_1^2 = \frac{1}{2}\ln\frac{8}{5}. \quad (4.6)$$

This second example is a non-seperable function, which influences which order it is easier to do the integrals in.

4.3 Regions Bounded by Lines and Curves

To integrate over a non-rectangular region in cartesian coordinates, the order in which we integrate matters. This is because the limits of the inner integral will *depend* on the value of the outer variable. For example, say we want to calculate the area of a triangular region. Suppose the triangle has corners at the origin, $(1,0)$ and $(0,1)$. The slope therefore is given by the equation $x+y=1$. It doesn't change the difficulty which variable we choose to integrate over first here so let's pick x as the inner integral. The outer limits for y are 0 to 1. For a fixed value of y, x goes from 0 to $1-y$.

$$\implies A = \int_0^1 \int_0^{1-y} \mathrm{d}x\mathrm{d}y = \int_0^1 (1-y)\mathrm{d}y = y - \frac{1}{2}y^2\Big|_0^1 = \frac{1}{2}. \quad (4.7)$$

Example 4.3. Now consider integrating the function $f(x,y) = \frac{1-y}{1-x}$ over the region above. It will be easier if we switch the order of integration here, so we integrate over y first:

$$\int_0^1 \int_0^{1-x} \frac{1-y}{1-x} \mathrm{d}y\mathrm{d}x = \int_0^1 \frac{1}{1-x}\left[y - \frac{1}{2}y^2\right]_0^{1-x} \mathrm{d}x \quad (4.8)$$

$$= \int_0^1 \frac{1}{2}(1+x)\mathrm{d}x = \frac{1}{2}\left(x + \frac{1}{2}x^2\right)\Big|_0^1 = \frac{3}{4}. \quad (4.9)$$

Example 4.4. Find the integral of the function $f(x,y) = e^{-y^2}$ over a triangular region bounded by the lines $x=0$, $y=x$ and $y=1$. This function has no antiderivative, so we need to do the x integral first. y goes from 0 to 1 and x goes from 0 to y for a fixed value of y.

$$\int_0^1 \int_0^y e^{-y^2} \mathrm{d}x\mathrm{d}y = \int_0^1 xe^{-y^2}\Big|_0^y \mathrm{d}y \quad (4.10)$$

$$= \int_0^1 ye^{-y^2}\mathrm{d}y = \frac{1}{2}\left(1 - \frac{1}{e}\right). \quad (4.11)$$

Example 4.5. Find the mean value of x and y over a region bounded by $y=x$ and $y=4x(1-x)$. To avoid having to split the region into two, make x the outer variable. Hence, x goes from 0 to $\frac{3}{4}$ (show this) and y goes from x to $4x(1-x)$. We have three integrals to evaluate: 1, x, and y over the region.

$$\int_0^{\frac{3}{4}}\int_x^{4x(1-x)} \mathrm{d}y\mathrm{d}x = \int_0^{\frac{3}{4}}(3x-4x^2)\mathrm{d}x = \left[\frac{3}{2}x^2-\frac{4}{3}x^3\right]_0^{\frac{3}{4}} = \frac{9}{32} \tag{4.12}$$

$$\int_0^{\frac{3}{4}}\int_x^{4x(1-x)} x\mathrm{d}y\mathrm{d}x = \int_0^{\frac{3}{4}}(3x^2-4x^3)\mathrm{d}x = \left[x^3-x^4\right]_0^{\frac{3}{4}} = \frac{27}{256} \tag{4.13}$$

$$\int_0^{\frac{3}{4}}\int_x^{4x(1-x)} y\mathrm{d}y\mathrm{d}x = \int_0^{\frac{3}{4}}\left(8x^4-16x^3+\frac{15}{2}x^2\right)\mathrm{d}x = \left[\frac{8}{5}x^5-4x^4+\frac{15}{6}x^3\right]_0^{\frac{3}{4}} \tag{4.14}$$

$$= \frac{27}{160}. \tag{4.15}$$

So using the formula from before, we have

$$\bar{x} = \frac{37/256}{9/32} = \frac{3}{8}, \quad \bar{y} = \frac{27/160}{9/32} = \frac{3}{5}. \tag{4.16}$$

4.4 Integration using Plane Polar Coordinates

Just like in 1D we can make a substitution to simplify the problem, in more than one dimension we can change the coordinate system to make things easier. One of the most common changes in coordinates is into polar coordinates. In order to change, we need to figure out what the area element, which is given by $\mathrm{d}A = \mathrm{d}x\mathrm{d}y$ in cartesian coordinates, is represented by in polar coordinates.

Consider an annulus of thickness $\mathrm{d}r$ and consider a section (is this the right word?) of the annulus with thickness $\mathrm{d}\theta$. The area of the annulus is given by the difference in the area of two circles:

$$\pi[(r+\mathrm{d}r)^2 - r^2] = \pi[2r\mathrm{d}r + \mathrm{d}r^2] = 2\pi r\mathrm{d}r \text{ as } \mathrm{d}r \to 0. \tag{4.17}$$

The fraction of a circle between θ and $\theta + \mathrm{d}\theta$ is given by $\frac{\mathrm{d}\theta}{2\pi}$, so the area of the infinitesimal area is

$$\mathrm{d}A = \frac{\mathrm{d}\theta}{2\pi}2\pi r\mathrm{d}r = r\mathrm{d}r\mathrm{d}\theta. \tag{4.18}$$

The more general approach to calculating the new area element in another coordinate system is to use the **Jacobian**. For a transformation between two cordinates $(x, y) \to (u, v)$, the area element $\mathrm{d}A$ is given by

$$\mathrm{d}A = J\mathrm{d}u\mathrm{d}v \tag{4.19}$$

, where

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \left|\frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}\right|. \tag{4.20}$$

Hence in polar coordinates $x = r\cos\theta$, $y = r\sin\theta$, we have

$$\mathrm{d}A = \left|\frac{\partial x}{\partial r}\frac{\partial y}{\partial \theta} - \frac{\partial x}{\partial \theta}\frac{\partial y}{\partial r}\right|\mathrm{d}r\mathrm{d}\theta = \left|r\cos^2\theta - (-r\sin^2\theta)\right|\mathrm{d}r\mathrm{d}\theta = r\mathrm{d}r\mathrm{d}\theta. \tag{4.21}$$

Example 4.6. Find the integral of $\sqrt{x^2+y^2}$ over the sector between $\theta = \frac{\pi}{6}$ and $\theta = \frac{\pi}{3}$ for $r \leq a$. This would be a very tricky integral to evaluate in cartesian coordinates but in polar coordinates the region is extremely simple. The limits of the integrals don't depend on r and θ, r goes from 0 to a and θ goes from $\frac{\pi}{6}$ to $\frac{\pi}{3}$. The function also has a very simple form in polar coordinates $\sqrt{x^2+y^2} = r$.

$$\int_{\frac{\pi}{6}}^{\frac{\pi}{3}} \int_0^a r^2 \mathrm{d}r\mathrm{d}\theta = \frac{\pi}{6} \int_0^a r^2 \mathrm{d}r = \frac{\pi a^3}{18}. \tag{4.22}$$

Example 4.7. Calculate the integral of x^2 over a circular disc of radius a. Here, θ goes from 0 to 2π and r goes from 0 to a.

$$\int_0^{2\pi} \int_0^a r^2 \cos^2\theta r \mathrm{d}r\mathrm{d}\theta = \int_0^{2\pi} \frac{a^4}{4} \cos^2\theta \mathrm{d}\theta = \frac{\pi}{4} a^4. \tag{4.23}$$

Example 4.8. Find the area bounded by two spirals given by $r_1(\theta) = a\theta^{\frac{1}{2}}$ and $r_2(\theta) = b\theta^{\frac{1}{2}}$, where $b > a > 0$. This is no longer a circularly-symmetric area region, so the order in which we integrate matters because the inner and outer limits will depend on each other. If we choose r as the inner variable, then the calculation is easier because r_1 and r_2 are given purely in terms of θ and so they will be the inner limits for fixed θ.

$$A = \int_0^{2\pi} \int_{r_1(\theta)}^{r_2(\theta)} r \mathrm{d}r\mathrm{d}\theta = \int_0^{2\pi} \frac{1}{2} r^2 \Big|_{r_1(\theta)}^{r_2(\theta)} \mathrm{d}\theta \tag{4.24}$$

$$= \frac{1}{2} \int_0^{2\pi} (b^2\theta - a^2\theta)\mathrm{d}\theta \tag{4.25}$$

$$= \frac{1}{4}(b^2 - a^2)\theta^2 \Big|_0^{2\pi} = \frac{1}{4}(b^2-a^2)(4\pi^2) = \pi^2(b^2-a^2). \tag{4.26}$$

Example 4.9. Calculate the area of an ellipse. The equation of an ellipse is $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. To do this, we will first simplify the equation using the change of variables $u = \frac{x}{a}$, $v = \frac{y}{b}$. Then the equation of an ellipse becomes $u^2 + v^2 = 1$, which is the equation of a circle. Note that the area element is $\mathrm{d}A = ab\mathrm{d}u\mathrm{d}v$. Now our integration region is a circle so we can use polar coordinates, $u = r\cos\theta$, $v = r\sin\theta$, and the area element is given by $\mathrm{d}A = abr\mathrm{d}r\mathrm{d}\theta$. Hence,

$$A = \int_0^{2\pi} \int_0^1 abr\mathrm{d}r\mathrm{d}\theta = 2\pi ab \frac{1}{2} r^2 \Big|_0^1 = \pi ab. \tag{4.27}$$

Example 4.10. Consider the integral of the function $(x^2 + y^2)$ over the ellipse from the last example. Using the same coordinate transformation from last time, $x^2 + y^2 =$

$a^2r^2\cos^2\theta + b^2r^2\sin^2\theta$. The limits are the same as in the last example. Hence,

$$\int_0^{2\pi}\int_0^1 r^3(a^2\cos^2\theta + b^2\sin^2\theta)ab\mathrm{d}r\mathrm{d}\theta = \int_0^{2\pi}\frac{ab}{4}(a^2\cos^2\theta + b^2\sin^2\theta)\mathrm{d}\theta \tag{4.28}$$

$$= \frac{\pi ab}{4}(a^2+b^2). \tag{4.29}$$

Chapter 5

Triple Integrals

5.1 Integration over Regions Bounded by Planes

Just as a double integral of a function over a region of 2D space represents the volume under a surface, the triple integral of a function over a region of 3D space represents the *hypervolume* under a volume. To evaluate a triple integral, we extend the methods that we have developed for evaluating double integrals. Instead of integrating over an infinitesimal area element, we integrate over a **volume element**. In cartesian coordinates, this has the form $\mathrm{d}V = \mathrm{d}x\mathrm{d}y\mathrm{d}z$.

When deciding the limits of integration, we make the same considerations that we do for double integrals. The limits of the outer integral must be fixed, the limits of the middle integral can depend on the outer variable but *not* on the inner variable. The limits of the inner variable can depend on both the middle and outer variables.

Example 5.1. Find the coordinates of the centre of the tetrahedron bounded by the x-y, x-z, y-z planes and the plane $x + y + z = 1$. Just like in 2D, the centre of the region, which is simply the average value of the coordinates, is given by

$$\bar{x} = \frac{\iiint_V x \mathrm{d}V}{\iiint_V \mathrm{d}V}, \ldots \tag{5.1}$$

We only need to find the average x position in this case since the process will be symmetric for y and z. Choosing z as the outer variable, the limits are 0 and 1. Now consider a cross section of the region V for a constant z. x and y are related by

$$x + y = 1 - z, \tag{5.2}$$

so if we choose y as the middle variable, then the limits are 0 and $1 - z$. Then similarly, the limits for x are 0 and $1-z-y$. Thus, we have

$$\int_0^1 \int_0^{1-z} \int_0^{1-z-y} x \mathrm{d}x \mathrm{d}y \mathrm{d}z = \int_0^1 \int_0^{1-z} \frac{1}{2}(1-z-y)^2 \mathrm{d}y \mathrm{d}z \tag{5.3}$$

$$= \int_0^1 -\frac{1}{6}(1-z-y)^3 \Big|_0^{1-z} \mathrm{d}z \tag{5.4}$$

$$= \int_0^1 \frac{1}{6}(1-z)^3 \mathrm{d}z = -\frac{1}{24}(1-z)^4 \Big|_0^1 = \frac{1}{24}. \tag{5.5}$$

The volume of the tetrahedral region is given by

$$\int_0^1 \int_0^{1-z} \int_0^{1-z-y} \mathrm{d}x \mathrm{d}y \mathrm{d}z = \int_0^1 \int_0^{1-z} (1-z-y) \mathrm{d}y \mathrm{d}z \tag{5.6}$$

$$= \int_0^1 -\frac{1}{2}(1-z-y)^2 \Big|_0^{1-z} \mathrm{d}z \tag{5.7}$$

$$= \int_0^1 \frac{1}{2}(1-z)^2 \mathrm{d}z = -\frac{1}{6}(1-z)^3 \Big|_0^1 = \frac{1}{6}. \tag{5.8}$$

Hence

$$\bar{x} = \bar{y} = \bar{z} = \frac{1/24}{1/6} = \frac{1}{4}, \tag{5.9}$$

so the centre of the tetrahedron is the point $\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$.

5.2 Integrals over Volumes Bounded by Curved Surfaces

Just like in 2D, when we have a complicated volume to integrate over making a change of variables can sometimes simplify the problem greatly. If a volume is symmetric about an axis, then we can write $\mathrm{d}V = \mathrm{d}A\mathrm{d}z$, where z points along the axis of symmetry and $\mathrm{d}A$ is the differential cross-sectional area.

Example 5.2. Find the integral of z over an inverted cone of height h and width $2a$. The cone has a straight edge. The distance between the z axis and the surface of the cone is given by

$$R_{\text{edge}}(z) = a - \frac{a}{h}z = a\left(1 - \frac{z}{h}\right). \tag{5.10}$$

Hence if we take z as the outer variable, then the integral can be written as

$$\iiint_V z \mathrm{d}V = \int_0^h z \iint_{A(z)} \mathrm{d}A(z) \mathrm{d}z, \tag{5.11}$$

where the inner double integral is simply the area of the 2D region $A(z)$, which is

$\pi R_{\text{edge}}(z)^2$. Hence

$$\int_0^h z \iint_{A(z)} \mathrm{d}A(z)\mathrm{d}z = \pi a^2 \int_0^h z\left(1-\frac{z}{h}\right)^2 \mathrm{d}z \tag{5.12}$$

$$= \pi a^2 h^2 \int_0^1 u(1-u)^2 \mathrm{d}u \tag{5.13}$$

$$= \pi a^2 h^2 \left[\frac{1}{2}a^2 - \frac{2}{3}u^3 + \frac{1}{4}u^4\right]_0^1 = \frac{\pi a^2 h^2}{12}. \tag{5.14}$$

This method applies in general if we have a function and a volume which are both symmetric about an axis. Such volume are bounded by surfaces called **surfaces of revolution**. In this case,

$$I = \pi \int_{z_1}^{z_2} R_{\text{edge}}(z)^2 f(z)\mathrm{d}z. \tag{5.15}$$

Example 5.3. Find the integral of the decaying exponential e^{-z} over the volume bounded by the paraboloid $az = x^2 + y^2$ with $a > 0$ and $0 \le z \le b$. Note that in plane polar coordinates we have $R^2 = x^2 + y^2$, so the radius of the volume is given by $R_{\text{edge}}(z) = \sqrt{az}$. Hence by using the general formula, we get

$$\iiint_V e^{-z}\mathrm{d}z = \pi a \int_0^b z e^{-z}\mathrm{d}z \tag{5.16}$$

$$= \pi a \left(-ze^{-z}\big|_0^b + \int_0^b e^{-z}\mathrm{d}z\right) \tag{5.17}$$

$$= \pi a\left[-be^{-b} - e^{-b} + 1\right] = \pi a[1-(1+b)e^{-b}]. \tag{5.18}$$

Common surfaces of revolution have radii given by

$$\text{Cylinder: } R = a \tag{5.19}$$

$$\text{Cone: } R = az \tag{5.20}$$

$$\text{Paraboloid: } R = a\sqrt{z} \tag{5.21}$$

$$\text{Sphere: } R = \sqrt{a^2 - z^2} \tag{5.22}$$

$$\text{Spheroid: } R = \sqrt{a^2 - b^2 z^2} \tag{5.23}$$

$$\text{Hyperboloid of one sheet: } R = \sqrt{a^2 + z^2} \tag{5.24}$$

$$\text{Hyperboloid of two sheets: } R = \sqrt{z^2 - a^2} \quad |z| \ge a. \tag{5.25}$$

What about general functions over volumes bounded by surfaces of revolution? In this case we make a change of variables to cylindrical coordinates with the z axis aligned along the axis of revolution. Choosing z as the outer variable, the cross-section at fixed z is a circle which allows us to use plane polar coordinates.

Example 5.4. Calculate the integral of x^2 over the inverted cone from before. Since a cone is bounded by a surface of revolution, we switch to cylindrical coordinates and choose z as the outer variable.

$$I = \iiint_V x^2 \mathrm{d}V = \int_0^h \iint_{A(z)} x^2 \mathrm{d}A(z). \tag{5.26}$$

For the inner double integral, the area element is $R\mathrm{d}R\mathrm{d}\phi$ and the limits are $0 \leq R \leq a\left(1-\frac{z}{h}\right)$ and $0 \leq \phi \leq 2\pi$. Hence,

$$I = \int_0^h \int_0^{2\pi} \int_0^{a\left(1-\frac{z}{h}\right)} R^3 \cos^2\phi \mathrm{d}R\mathrm{d}\phi \tag{5.27}$$

$$= \int_0^h \int_0^{2\pi} \frac{1}{4}\cos^2\phi a^4 \left(1-\frac{z}{h}\right)^4 \tag{5.28}$$

$$= \int_0^h \frac{\pi a^4}{4}\left(1-\frac{z}{h}\right)^4 \mathrm{d}z \tag{5.29}$$

$$= \frac{\pi a^4 h}{4}\int_0^1 (1-u)^4 \mathrm{d}u = \frac{\pi a^4 h}{4}\left[-\frac{1}{5}(1-u)^5\right]_0^1 = \frac{1}{20}\pi a^4 h. \tag{5.30}$$

Example 5.5. Find the integral of xyz over a volume bounded by the $x-y$, $x-z$, $y-z$ planes and a sphere of radius a. The equation for the sphere is $x^2+y^2+z^2=a^2$. Note that $x^2+y^2=R^2$, hence $R_{\text{edge}}=\sqrt{a^2-z^2}$. Doing this integral in cylindrical coordinates, the limits for ϕ are 0 and $\frac{\pi}{2}$.

$$\iiint_V xyz\mathrm{d}V = \int_0^a z\int_0^{\frac{\pi}{2}}\int_0^{\sqrt{a^2-z^2}} R^3\cos\phi\sin\phi \mathrm{d}R\mathrm{d}\phi\mathrm{d}z \tag{5.31}$$

$$= \int_0^a z\left(\frac{1}{2}\right)\left(\frac{1}{4}\left(\sqrt{a^2-z^2}\right)^4\right)\mathrm{d}z \tag{5.32}$$

$$= \frac{1}{8}\int_0^a z(a^2-z^2)^2\mathrm{d}z \tag{5.33}$$

$$= \frac{1}{8}\left[\frac{1}{2}a^4z^2-\frac{1}{2}a^2z^4+\frac{1}{6}z^6\right]_0^a = \frac{1}{8}\left(\frac{1}{6}a^6\right) = \frac{1}{48}a^6. \tag{5.34}$$

5.3 Integration over Volumes Bounded by Portions of a Sphere

If the region of integration is a sphere (or part of one), then changing to spherical coordinates will simplify the problem as the limits of integration will be constant. The volume element in spherical coordinates is given by $\mathrm{d}V = \mathrm{d}r \cdot r\mathrm{d}\theta \cdot R\mathrm{d}\phi = r^2\sin\theta\mathrm{d}r\mathrm{d}\theta\mathrm{d}\phi$. Another case where spherical coordinates are useful is the volume of integration is the entirety of 3D space and the function to be integrated over depends only on the distance from a point (it has spherical symmetry).

Example 5.6. Find the integral of the function $e^{-\frac{r}{h}}$ over $\mathbb{R}^3$.

$$\iiint_{\mathbb{R}^3} e^{-\frac{r}{h}} \mathrm{d}V = \int_0^{2\pi} \int_0^{\pi} \int_0^{\infty} e^{-\frac{r}{h}} r^2 \sin\theta \mathrm{d}r \mathrm{d}\theta \mathrm{d}\phi \tag{5.35}$$

$$= 4\pi \int_0^{\infty} r^2 e^{-\frac{r}{h}} \mathrm{d}r \tag{5.36}$$

$$= 4\pi \left[(-hr^2 - 2h^2 r - 2h^3) e^{-\frac{r}{h}}\right]_0^{\infty} = 8\pi h^3. \tag{5.37}$$

Bibliography

[1] Aidan Naughton, Jean Reinaud, and David Dritschel. Lecture notes for MT2503 Multivariate calculus. University of St. Andrews, 2020.

[2] H. M. Schey. *Div, Grad, Curl, and All That: An Informal Text on Vector Calculus.* W.W. Norton & Company, 2005.