# 1

# Prelude to Modern Analysis

## 1.1 Introduction

The primary purpose of this chapter is to review a number of topics from analysis, and some from algebra, that will be called upon in the following chapters. These are topics of a classical nature, such as appear in books on advanced calculus and linear algebra. For our treatment of modern analysis, we can distinguish four fundamental notions which will be particularly stressed in this chapter. These are

(a) set theory, of an elementary nature;
(b) the concept of a function;
(c) convergence of sequences; and
(d) some theory of vector spaces.

On a number of occasions in this chapter, we will also take the time to discuss the relationship of modern analysis to classical analysis. We begin this now, assuming some knowledge of the points (a) to (d) just mentioned.

Modern analysis is not a new brand of mathematics that replaces the old brand. It is totally dependent on the time-honoured concepts of classical analysis, although in parts it can be given without reference to the specifics of classical analysis. For example, whereas classical analysis is largely concerned with functions of a real or complex variable, modern analysis is concerned with functions whose domains and ranges are far more general than just sets of real or complex numbers. In fact, these functions can have domains and ranges which are themselves sets of functions. A function of this more general type will be called an operator or mapping. Importantly, very often any set will do as the domain of a mapping, with no specific reference to the nature of its elements.

This illustrates how modern analysis generalises the ideas of classical analysis. At the same time, in many ways modern analysis simplifies classical analysis because it uses a basic notation which is not cluttered with the symbolism that characterises many topics of a classical nature. Through this, the unifying aspect of modern analysis appears because when the symbolism of those classical topics is removed a surprising similarity becomes apparent in the treatments formerly thought to be peculiar to those topics.

Here is an example:

$$\int_a^b k(s,t)x(t)\,dt = f(s), \quad a \leqslant s \leqslant b,$$

is an *integral equation*; $f$ and $k$ are continuous functions and we want to solve this to find the continuous function $x$. The left-hand side shows that we have operated on the function $x$ to give the function $f$, on the right. We can write the whole thing as

$$Kx = f,$$

where $K$ is an operator of the type we just mentioned. Now the essence of the problem is clear. It has the same form as a matrix equation $A\mathbf{x} = \mathbf{b}$, for which the solution (sometimes) is $\mathbf{x} = A^{-1}\mathbf{b}$. In the same way, we would like the solution of the integral equation to be given simply as $x = K^{-1}f$. The two problems, stripped of their classical symbolism, appear to be two aspects of a more general study.

The process can be reversed, showing the strong applicability of modern analysis: when the symbolism of a particular branch of classical analysis is restored to results often obtained only because of the manipulative ease of the simplified notation, there arise results not formerly obtained in the earlier theory. In other cases, this procedure gives rise to results in one field which had not been recognised as essentially the same as well-known results in another field. The notations of the two branches had fully disguised the similarity of the results.

When this occurs, it can only be because there is some underlying structure which makes the two (or more) branches of classical analysis appear just as examples of some work in modern analysis. The basic entities in these branches, when extracted, are apparently combined together in a precisely corresponding manner in the several branches. This takes us back to our first point of the generalising nature of modern analysis and of the benefit of working with quite arbitrary sets. To combine the elements of these sets together requires some basic ground rules and this is why, very often and predominantly in this book, the sets are assumed to be vector spaces: simply because vector spaces are sets with certain rules attached allowing their elements to be combined together in a particular fashion.

We have indicated the relevance of set theory, functions and vector spaces in our work. The other point, of the four given above, is the springboard that takes us from algebra into analysis. In this book, we use in a very direct fashion the notion of a convergent sequence to generate virtually every result.

We might mention now, since we have been comparing classical and modern analysis, that another area of study, called functional analysis, may today be taken as identical with modern analysis. A functional is a mapping whose range is a set of real or complex numbers and functional analysis had a fairly specific meaning (the analysis of functionals) when the term was first introduced early in the 20th century. Other writers may make technical distinctions between the two terms but we will not.

In the review which follows, it is the aim at least to mention all topics required for an understanding of the subsequent chapters. Some topics, notably those connected with the points (a) to (d) above, are discussed in considerable detail, while others might receive little more than a definition and a few relevant properties.

## 1.2 Sets and numbers

A *set* is a concept so basic to modern mathematics that it is not possible to give it a precise definition without going deeply into the study of mathematical logic. Commonly, a set is described as any collection of objects but no attempt is made to say what a 'collection' is or what an 'object' is. We are forced in books of this type to accept sets as fundamental entities and to rely on an intuitive feeling for what a set is.

The objects that together make up a particular set are called *elements* or *members* of that set. The list of possible sets is as long as the imagination is vivid, or even longer (we are hardly being precise here) since, importantly, the elements of a set may themselves be sets.

Later in this chapter we will be looking with some detail into the properties of certain sets of numbers. We are going to rely on the reader's experience with numbers and not spend a great deal of time on the development of the real number system. In particular, we assume familiarity with

(a) the *integers*, or whole numbers, such as $-79$, $-3$, $0$, $12$, $4{,}063{,}180$;
(b) the *rational numbers*, such as $-\frac{5}{3}$, $\frac{11}{17}$, which are numbers expressible as a ratio of integers (the integers themselves also being examples);
(c) those numbers which are not rational, known as *irrational numbers*, such as $\sqrt{2}$, $\sqrt[3]{15}$, $\pi$;
(d) the *real numbers*, which are numbers that are either rational or irrational;
(e) the ordering of the real numbers, using the inequality signs $<$ and $>$ (and the use of the signs $\leqslant$ and $\geqslant$);
(f) the representation of the real numbers as points along a line; and
(g) the fact, in (f), that the real numbers fill the line, leaving no holes: to every point on the line there corresponds a real number.

The final point is a crucial one and may not appear to be so familiar. On reflection however, it will be seen to accord with experience, even when expressed in such a vague way. This is a crude formulation of what is known as the *completeness* of the real number system, and will be referred to again in some detail subsequently.

By way of review, we remark that we assume the ability to perform simple manipulations with inequalities. In particular, the following should be known. If $a$ and $b$ are real numbers and $a < b$, then

$$-a > -b;$$
$$\frac{1}{a} > \frac{1}{b}, \text{ if also } a > 0 \text{ or } b < 0;$$
$$\sqrt{a} < \sqrt{b}, \text{ if also } a \geqslant 0.$$

With regard to the third property, we stress that the use of the radical sign ($\sqrt{\phantom{x}}$) always implies that the nonnegative root is to be taken. Bearing this comment in mind, we may define the *absolute value* $|a|$ of any real number $a$ by

$$|a| = \sqrt{a^2}.$$

More commonly, and equivalently of course, we say that $|a|$ is $a$ whenever $a > 0$ and $|a|$ is $-a$ whenever $a < 0$, while $|0| = 0$. For any real numbers $a$ and $b$, we have

$$|a+b| \leqslant |a| + |b|, \qquad |ab| = |a|\,|b|.$$

These may be proved by considering the various combinations of positive and negative values for $a$ and $b$.

We also assume a knowledge of *complex numbers*: numbers of the form $a + ib$ where $a$ and $b$ are real numbers and $i$ is an imaginary unit, satisfying $i^2 = -1$.

This is a good place to review a number of definitions and properties connected with complex numbers. If $z = a+ib$ is a complex number, then we call the numbers $a$, $b$, $a - ib$ and $\sqrt{a^2 + b^2}$ the *real part*, *imaginary part*, *conjugate* and *modulus*, respectively, of $z$, and denote these by $\operatorname{Re} z$, $\operatorname{Im} z$, $\overline{z}$ and $|z|$, respectively. The following are some of the simple properties of complex numbers that we use. If $z$, $z_1$ and $z_2$ are complex numbers, then

$$\begin{gathered}
\overline{\overline{z}} = z, \\
\overline{z_1 + z_2} = \overline{z}_1 + \overline{z}_2, \\
\overline{z_1 z_2} = \overline{z}_1\, \overline{z}_2, \\
|\operatorname{Re} z| \leqslant |z|, \quad |\operatorname{Im} z| \leqslant |z|, \\
z\overline{z} = |z|^2, \\
|z_1 + z_2| \leqslant |z_1| + |z_2|, \\
|z_1 z_2| = |z_1|\,|z_2|.
\end{gathered}$$

It is essential to remember that, although $z$ is a complex number, the numbers $\operatorname{Re} z$, $\operatorname{Im} z$ and $|z|$ are real. The final two properties in the above list are important generalisations of the corresponding properties just given for real numbers. They can be generalised further, in the natural way, to the sum or product of three or four or more complex numbers.

Real numbers, complex numbers, and other sets of numbers, all occur so frequently in our work that it is worth using special symbols to denote them.

**Definition 1.2.1** The following symbols denote the stated sets:

$\mathbf{N}$, the set of all positive integers;
$\mathbf{Z}$, the set of all integers (positive, negative and zero);
$\mathbf{Q}$, the set of all rational numbers;
$\mathbf{R}$, the set of all real numbers;
$\mathbf{R}_+$, the set of all nonnegative real numbers;
$\mathbf{C}$, the set of all complex numbers.

Other sets will generally be denoted by ordinary capital letters and their elements by lower case letters; the same letter will not always refer to the same set or element. To indicate that an object $x$ is an element of a set $X$, we will write $x \in X$; if $x$ is not an element of $X$, we will write $x \notin X$. For example, $\sqrt{2} \in \mathbf{R}$ but $\sqrt{2} \notin \mathbf{Z}$. A statement such as $x, y \in X$ will be used as an abbreviation for the two statements $x \in X$ and $y \in X$. To show the elements of a set we always enclose them in braces and give either a complete listing (for example, $\{1, 2, 3\}$ is the set consisting of the integers 1, 2 and 3), or an indication of a pattern (for example, $\{1, 2, 3, \dots\}$ is the set $\mathbf{N}$), or a description of a rule of formation following a colon (for example, $\{x : x \in \mathbf{R},\ x \geqslant 0\}$ is the set $\mathbf{R}_+$). Sometimes we use an abbreviated notation (for example, $\{n : n = 2m,\ m \in \mathbf{N}\}$ and $\{2n : n \in \mathbf{N}\}$ both denote the set of all even positive integers).

An important aspect in the understanding of sets is that the order in which their elements are listed is irrelevant. For example, $\{1, 2, 3\}$, $\{3, 1, 2\}$, $\{2, 1, 3\}$ are different ways of writing the same set. However, on many occasions we need to be able to specify the first position, the second position, and so on, and for this we need a new notion. We speak of *ordered pairs* of two elements, *ordered triples* of three elements, and, generally, *ordered $n$-tuples* of $n$ elements with this property that each requires for its full determination a list of its elements and the order in which they are to be listed. The elements, in their right order, are enclosed in parentheses (rather than braces, as for sets). For example, $(1, 2, 3)$, $(3, 1, 2)$, $(2, 1, 3)$ are different ordered triples. This is not an unfamiliar notion. In ordinary three-dimensional coordinate geometry, the coordinates of a point provide an example of an ordered triple: the three ordered triples just given would refer to three different points in space.

We give now a number of definitions which help us describe various manipulations to be performed with sets.

**Definition 1.2.2**

(a) A set $S$ is called a *subset* of a set $X$, and this is denoted by $S \subseteq X$ or $X \supseteq S$, if every element of $S$ is also an element of $X$.

(b) Two sets $X$ and $Y$ are called *equal*, and this is denoted by $X = Y$, if each is a subset of the other; that is, if both $X \subseteq Y$ and $Y \subseteq X$. Otherwise, we write $X \neq Y$.

(c) A set which is a subset of any other set is called a *null set* or *empty set*, and is denoted by $\varnothing$.

(d) A set $S$ is called a *proper* subset of a set $X$ if $S \subseteq X$, but $S \neq X$.

(e) The *union* of two sets $X$ and $Y$, denoted by $X \cup Y$, is the set of elements belonging to at least one of $X$ and $Y$; that is,

$$X \cup Y = \{x : x \in X \text{ or } x \in Y \text{ (or both)}\}.$$

(f) The *intersection* of two sets $X$ and $Y$, denoted by $X \cap Y$, is the set of elements belonging to both $X$ and $Y$; that is,

$$X \cap Y = \{x : x \in X \text{ and } x \in Y\}.$$

(g) The *cartesian product* of two sets $X$ and $Y$, denoted by $X \times Y$, is the set of all ordered pairs, the first elements of which belong to $X$ and the second elements to $Y$; that is,

$$X \times Y = \{(x, y) : x \in X,\ y \in Y\}.$$

(h) The *complement* of a set $X$, denoted by $\sim X$, is the set of elements that do not belong to $X$; that is, $\sim X = \{x : x \notin X\}$. The complement of $X$ relative to a set $Y$ is the set $Y \cap \sim X$; this is denoted by $Y \backslash X$.

For some simple examples illustrating parts of this definition, we let $X = \{1, 3, 5\}$ and $Y = \{1, 4\}$. Then

$$X \cup Y = \{1, 3, 4, 5\}, \qquad X \cap Y = \{1\},$$
$$X \times Y = \{(1, 1), (1, 4), (3, 1), (3, 4), (5, 1), (5, 4)\},$$
$$Y \times X = \{(1, 1), (1, 3), (1, 5), (4, 1), (4, 3), (4, 5)\}.$$

We see that in general $X \times Y \neq Y \times X$. The set $Y \backslash X$ is the set of elements of $Y$ that do not belong to $X$, so here $Y \backslash X = \{4\}$.

The definitions of union, intersection and cartesian product of sets can be extended to more than two sets. Suppose we have $n$ sets $X_1$, $X_2$, $\ldots$, $X_n$. Their union, intersection and cartesian product are defined as

$$\begin{aligned}
X_1 \cup X_2 \cup \cdots \cup X_n &= \bigcup_{k=1}^{n} X_k \\
&= \{x : x \in X_k \text{ for at least one } k = 1,\ 2,\ \ldots,\ n\}, \\
X_1 \cap X_2 \cap \cdots \cap X_n &= \bigcap_{k=1}^{n} X_k \\
&= \{x : x \in X_k \text{ for all } k = 1,\ 2,\ \ldots,\ n\}, \\
X_1 \times X_2 \times \cdots \times X_n &= \prod_{k=1}^{n} X_k \\
&= \{(x_1, x_2, \ldots, x_n) : x_k \in X_k \text{ for } k = 1, 2, \ldots, n\},
\end{aligned}$$

respectively (the cartesian product being a set of ordered $n$-tuples). The notations in the middle are similar to the familiar sigma notation for addition, where we write

$$x_1 + x_2 + \cdots + x_n = \sum_{k=1}^{n} x_k,$$

when $x_1$, $x_2$, ..., $x_n$ are numbers.

For cartesian products only, there is a further simplification of notation when all the sets are equal. If $X_1 = X_2 = \cdots = X_n = X$, then in place of $\prod_{k=1}^{n} X_k$ or $\prod_{k=1}^{n} X$ we write simply $X^n$, as suggested by the $\times$ notation, but note that there is no suggestion of multiplication: $X^n$ is a set of $n$-tuples. In particular, it is common to write $\mathbf{R}^n$ for the set of all $n$-tuples of real numbers and $\mathbf{C}^n$ for the set of all $n$-tuples of complex numbers.

It is necessary to make some comments regarding the definition of an empty set in Definition 1.2.2(c). These are gathered together as a theorem.

**Theorem 1.2.3**

(a) *All empty sets are equal.*
(b) *The empty set has no elements.*
(c) *The only set with no elements is the empty set.*

To prove (a), we suppose that $\varnothing_1$ and $\varnothing_2$ are any two empty sets. Since an empty set is a subset of any other set, we must have both $\varnothing_1 \subseteq \varnothing_2$ and $\varnothing_2 \subseteq \varnothing_1$. By the definition of equality of sets, it follows that $\varnothing_1 = \varnothing_2$. This proves (a) and justifies our speaking of 'the' empty set in the remainder of the theorem. We prove (b) by contradiction. Suppose $x \in \varnothing$. Since for any set $X$ we have $\varnothing \subseteq X$ and $\varnothing \subseteq \sim X$, we must have both $x \in X$ and $x \in \sim X$. This surely contradicts the existence of $x$, proving (b). Finally, we prove (c), again by contradiction. Suppose $X$ is a set with no elements and suppose $X \neq \varnothing$. Since $\varnothing \subseteq X$, this means that $X$ is not a subset of $\varnothing$. Then there must be an element of $X$ which is not in $\varnothing$. But $X$ has no elements so this is the contradiction we need. □

All this must seem a bit peculiar if it has not been met before. In defence, it may be pointed out that sets were only introduced intuitively in the first place and that the inclusion in the concept of 'a set with no elements' is a necessary addition (possibly beyond intuition) to provide consistency elsewhere. For example, if two sets $X$ and $Y$ have no elements in common and we wish to speak of their intersection, we can now happily say $X \cap Y = \varnothing$. (Two such sets are called *disjoint*.)

Manipulations with sets often make use of the following basic results.

**Theorem 1.2.4** *Let $X$, $Y$ and $Z$ be sets. Then*

(a) $\sim(\sim X) = X$,
(b) $X \cup Y = Y \cup X$ *and* $X \cap Y = Y \cap X$ (*commutative rules*),
(c) $X \cup (Y \cup Z) = (X \cup Y) \cup Z$ *and* $X \cap (Y \cap Z) = (X \cap Y) \cap Z$ (*associative rules*),
(d) $X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z)$ *and* $X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$ (*distributive rules*).

We will prove only the second distributive rule. To show that two sets are equal we must make use of the definition of equality in Definition 1.2.2(b).

First, suppose $x \in X \cap (Y \cup Z)$. Then $x \in X$ and $x \in Y \cup Z$. That is, $x$ is a member of $X$ and of either $Y$ or $Z$ (or both). If $x \in Y$ then $x \in X \cap Y$; if $x \in Z$ then $x \in X \cap Z$. At least one of these must be true, so $x \in (X \cap Y) \cup (X \cap Z)$. This proves that $X \cap (Y \cup Z) \subseteq (X \cap Y) \cup (X \cap Z)$. Next, suppose $x \in (X \cap Y) \cup (X \cap Z)$. Then $x \in X \cap Y$ or $x \in X \cap Z$ (or both). In both cases, $x \in X \cap (Y \cup Z)$ since in both cases $x \in X$, and $Y \subseteq Y \cup Z$ and $Z \subseteq Y \cup Z$. Thus $X \cap (Y \cup Z) \supseteq (X \cap Y) \cup (X \cap Z)$. Then it follows that $X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$, completing this part of the proof. □

The following theorem gives two of the more important relationships between sets.

**Theorem 1.2.5 (De Morgan's Laws)** *Let $X$, $Y$ and $Z$ be sets. Then*

$$Z \backslash (X \cap Y) = Z \backslash X \cup Z \backslash Y \quad \textit{and} \quad Z \backslash (X \cup Y) = Z \backslash X \cap Z \backslash Y.$$

There is a simpler form of de Morgan's laws for ordinary complements:

$$\sim(X \cap Y) = \sim X \cup \sim Y \quad \text{and} \quad \sim(X \cup Y) = \sim X \cap \sim Y.$$

To prove the first of these, suppose $x \in \sim(X \cap Y)$. Then $x \notin X \cap Y$ so either $x \notin X$ or $x \notin Y$. That is, $x \in \sim X$ or $x \in \sim Y$, so $x \in \sim X \cup \sim Y$. This proves that $\sim(X \cap Y) \subseteq \sim X \cup \sim Y$. Suppose next that $x \in \sim X \cup \sim Y$. If $x \in \sim X$ then $x \notin X$ so $x \notin X \cap Y$, since $X \cap Y \subseteq X$. That is, $x \in \sim(X \cap Y)$. The same is true if $x \in \sim Y$. Thus $\sim X \cup \sim Y \subseteq \sim(X \cap Y)$, so we have proved that $\sim(X \cap Y) = \sim X \cup \sim Y$.

We can use this, the definition of relative complement, and a distributive rule from Theorem 1.2.4 to prove the first result of Theorem 1.2.5:

$$\begin{aligned} Z\backslash(X \cap Y) &= Z \cap {\sim}(X \cap Y) = Z \cap ({\sim}X \cup {\sim}Y) \\ &= (Z \cap {\sim}X) \cup (Z \cap {\sim}Y) = Z\backslash X \cup Z\backslash Y. \end{aligned}$$

The second of de Morgan's laws is proved similarly. □

**Review exercises 1.2**

(1) Let $a$ and $b$ be real numbers. Show that

  (a) $||a| - |b|| \leqslant |a - b|$,
  (b) $|a - b| < \epsilon$ if and only if $b - \epsilon < a < b + \epsilon$,
  (c) if $a < b + \epsilon$ for every $\epsilon > 0$ then $a \leqslant b$.

(2) Suppose $A \cup B = X$. Show that $X \times Y = (A \times Y) \cup (B \times Y)$, for any set $Y$.

(3) For any sets $A$ and $B$, show that

  (a) $A\backslash B = A$ if and only if $A \cap B = \varnothing$,
  (b) $A\backslash B = \varnothing$ if and only if $A \subseteq B$.

## 1.3 Functions or mappings

We indicated in Section 1.1 how fundamental the concept of a function is in modern analysis. (It is equally important in classical analysis but may be given a restricted meaning there, as we remark below.) A function is often described as a rule which associates with an element in one set a unique element in another set; we will give a definition which avoids the undefined term 'rule'. In this definition we will include all associated terms and notations that will be required. Examples and general discussion will follow.

**Definition 1.3.1** Let $X$ and $Y$ be any two nonempty sets (which may be equal).

(a) A *function* $f$ from $X$ into $Y$ is a subset of $X \times Y$ with the property that for each $x \in X$ there is precisely one element $(x, y)$ in the subset $f$. We write $f \colon X \to Y$ to indicate that $f$ is a function from $X$ into $Y$.

(b) The set $X$ is called the *domain* of the function $f \colon X \to Y$.