1

THE COMPLETENESS OF THE REALS

We should make it clear from the beginning that we are not going to construct the real numbers from the rational numbers. On the contrary, we shall assume that there is a set **R** of real numbers in which the familiar 'algebraic' properties of addition, subtraction, multiplication, division and ordering are known to hold.† As usual, we write **Z** for the subset of **R** consisting of the *integers* 0, 1, -1, 2, -2, ... and **Q** for the set of all *rational numbers* m/n where m, n are integers and $n \neq 0$. On one small point we are unorthodox (if it is ever unorthodox to follow Bourbaki): we say that a real number x is *positive* if $x \geqslant 0$ and x is *strictly positive* if $x > 0$.

We shall also make use of the 'geometric' representation of the real numbers as points on a line. Let $|x| = x$ if $x \geqslant 0$, $|x| = -x$ if $x < 0$ and call $|x-y|$ the *distance* between the points x, y. Conventionally an *origin* is marked on the line to represent the number zero and a point distinct from the origin is marked to represent the number one. The representation is unique once these two points are given. When only one such line is under discussion it is natural to draw it across the page parallel to the lines of type and it is usual to have the point 1

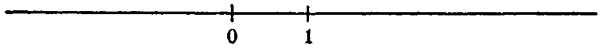

Fig. 1

on the right of the origin (Fig. 1). We refer to this line as the *real line* and denote it by the same symbol **R**. In a sense, no results are based on this representation, because one can always return to the *set* **R** and its *elements* x, but there are many advantages in having two languages for real numbers, not least the relief from monotony. The more pictorial language is a help to the intuition, especially in arguments involving the ordering of the real numbers, and it also suggests some other rather vivid terms. For example, if $a \leqslant b$, the points x satisfying $a \leqslant x \leqslant b$ form the *closed interval* $[a, b]$, the points x satisfying $a < x < b$ form the *open interval* (a, b), and the points x satisfying $a \leqslant x < b$, $a < x \leqslant b$, respectively, form the *half-open intervals* $[a, b)$, $(a, b]$. We refer to these

† A list and brief discussion of axioms for the real numbers will be given in § 1.3.

as the *bounded intervals* on **R** and give them the *length* $b-a$. We shall not follow the common practice of augmenting the real line with two 'points at infinity' $-\infty$ and ∞. Nevertheless it is often convenient to write $[a, \infty)$, (a, ∞), $(-\infty, a]$, $(-\infty, a)$ for the sets defined by $x \geqslant a$, $x > a$, $x \leqslant a$, $x < a$, respectively, and $(-\infty, \infty)$ for the whole real line **R**; we refer to these as *unbounded intervals*.

There is just one other property that we shall need, viz. the *completeness* of the real numbers. This property is essential to any thoroughgoing discussion of the 'analytic' ideas of limit and convergence, and these in turn are essential to any theory of integration. Our present chapter gives an elementary introduction to the completeness of **R** in terms of increasing sequences. We hope that this will help to motivate our construction of the Lebesgue integral in Chapter 3, but any reader who is already familiar with these ideas will probably be content to move on at once to Chapter 2.

1.1 The Axiom of Completeness

We shall assume a rudimentary knowledge of convergence. A *sequence* of real numbers is a function which associates a real number s_n with each integer $n \geqslant 1$. The number s_n, which is the *value* of the sequence for the integer n, is often called the n-th *term* of the sequence. It is usual to write the sequence as

$$s_1, s_2, s_3, \ldots$$

or simply $\{s_n\}$ for short.

In the construction of the Lebesgue integral we shall be dealing almost exclusively with *increasing* sequences and *decreasing* sequences of real numbers, i.e. those which satisfy

$$s_n \leqslant s_{n+1} \quad \text{for all} \quad n,$$
$$s_n \geqslant s_{n+1} \quad \text{for all} \quad n,$$

respectively. A sequence which is either increasing or decreasing will be called *monotone*. *Strictly increasing* and *strictly decreasing* sequences are defined with the strict inequalities.

The question of convergence for monotone sequences is quite simple, for here our geometric picture of points on the real line and our intuition make it abundantly clear what is going to happen. Suppose that

$$s_1 \leqslant s_2 \leqslant s_3 \leqslant \ldots,$$

then as n increases the points s_n on the real line never move to the left. They either move indefinitely far to the right, in which case we say that

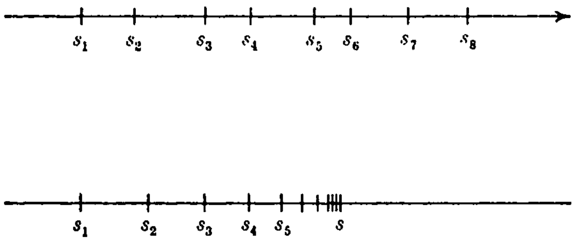

Fig. 2

$\{s_n\}$ *diverges* (to infinity), or they get nearer and nearer to some fixed point s, in which case we say that $\{s_n\}$ *converges* (to s) (see Fig. 2).

This property suggested by our intuition is exactly what we mean by the completeness of the real numbers. But we must now try to define more carefully what is meant by the last rather vague paragraph. As a first step we introduce the idea of boundedness: the sequence $\{s_n\}$ is *bounded above* if there is a real number K satisfying $s_n \leqslant K$ for all n; in this case K is called an *upper bound* of the sequence. A *divergent* increasing sequence of real numbers may now be defined as an increasing sequence which has no upper bound – more colloquially, a sequence which increases without bound. For example, the increasing sequence 1, 2, 3, ... is divergent as there is no real number K which satisfies $n \leqslant K$ for all integers n.†

In exactly the same way we say that a sequence $\{s_n\}$ of real numbers is *bounded below* if there is a real number L satisfying $s_n \geqslant L$ for all n, and we refer to L as a *lower bound* of the sequence. A decreasing sequence is *divergent* if it has no lower bound.

Let $\{s_n\}$ be a sequence of real numbers and s a real number. Recall that $\{s_n\}$ *converges* to (the *limit*) s if, given any $\epsilon > 0$, there is an integer N such that $|s_n - s| < \epsilon$ for all $n \geqslant N$.

We have now accumulated enough language to state the

Axiom of Completeness for the Real Numbers. *A sequence of real numbers which is increasing and bounded above converges to a real number.*

It is clear that a decreasing sequence $\{t_n\}$ bounded below by the real number L corresponds to an increasing sequence $\{-t_n\}$ bounded above by the real number $-L$. It therefore follows from the Axiom of Completeness that a *sequence of real numbers which is decreasing and bounded below converges to a real number.*

† This statement is often called the Axiom of Archimedes. In fact it can be deduced from our Axiom of Completeness: see Ex. 1.

A sequence like $\{1-1/n\}$ does not show the full power of the Axiom of Completeness because it is clear from first principles that this sequence converges to the limit 1. The main point of the Axiom is that it guarantees the *existence* of a limit for an increasing sequence provided we can prove that it is bounded above. The familiar sequence for which

$$s_n = \left(1+\frac{1}{n}\right)^n$$

is a good illustration of this point. Let $n, r \geqslant 2$. The $(r+1)$-th term in the Binomial expansion of s_n is

$$\left(1-\frac{1}{n}\right)\dots\left(1-\frac{r-1}{n}\right)\bigg/ r!$$

which increases strictly with n. Each term in the expansion of s_{n+1} is not less than the corresponding term in the expansion of s_n and the former expansion contains one extra positive term; thus $s_{n+1} > s_n$. The displayed term is less than $1/r!$ Thus

$$s_n < 1+1+1/2!+\dots+1/n! \leqslant 1+1+\tfrac{1}{2}+\dots+\tfrac{1}{2}^{n-1} < 3.$$

The Axiom of Completeness now assures us that $\{(1+1/n)^n\}$ converges to a real number between 2 and 3 (the base of natural logarithms) which we denote by e.

The above definition of convergence is not restricted to increasing or decreasing sequences of real numbers. It is convenient to gather under the heading of Proposition 1 a few of the most elementary facts about convergent sequences. As a shorthand for the statement 'the sequence $\{s_n\}$ converges to the limit s' we often write

$$s_n \to s \quad \text{or} \quad s = \lim s_n.$$

Proposition 1. *Suppose that $s_n \to s$ and $t_n \to t$. Then*

(i) $s_n + t_n \to s+t$;
(ii) $s_n - t_n \to s-t$;
(iii) $s_n t_n \to st$;
(iv) *if* $t_n \neq 0$ *for all* n *and* $t \neq 0$, *then* $s_n/t_n \to s/t$;
(v) *if* $s_n \geqslant t_n$ *for all* n, *then* $s \geqslant t$.

If a sequence $\{a_n\}$ of real numbers is given, we refer to the formal expression† Σa_n as a *series* (of real numbers) whose n-th *term* is a_n.

† The reader who is unhappy about the words 'formal expression' may prefer to define a series of real numbers as a sequence $\{a_n\}$ of real numbers together with the operation of addition: thus a series is an ordered pair $(\{a_n\}, +)$ which we at once replace by the shorter notation Σa_n.

Furthermore, if the sequence $\{s_n\}$ of partial sums

$$s_n = a_1 + a_2 + \dots + a_n$$

converges to a limit s, we say that *the series* Σa_n *converges to the sum* s; in keeping with historical custom and usage we shall also allow Σa_n or $a_1 + a_2 + \dots$ to denote the sum s in this case.

It is clear that we may recapture the sequence $\{a_n\}$ from the sequence $\{s_n\}$ merely by noting that $a_1 = s_1$ and $a_n = s_n - s_{n-1}$ for $n \geqslant 2$. Thus the theory of sequences and the theory of series (of real numbers) are practically synonymous; it is a matter of convenience which language we adopt in a given problem.

There is one further notion of convergence which is most aptly expressed in terms of series: we say that Σa_n is *absolutely convergent* if the series $\Sigma|a_n|$ is convergent. For this terminology to make sense we must prove that an absolutely convergent series is convergent. In view of Proposition 1 the following result tells us rather more.

Proposition 2. *A series of real numbers is absolutely convergent if and only if it can be expressed as the difference of two convergent series of positive real numbers.*

Proof. (i) Let $a_n = b_n - c_n$ where b_n, c_n are positive and $\Sigma b_n, \Sigma c_n$ are convergent. Then

$$|a_1| + \dots + |a_n| \leqslant b_1 + \dots + b_n + c_1 + \dots + c_n$$

is bounded and so $\Sigma|a_n|$ is convergent by the Axiom of Completeness.

(ii) Suppose that $\Sigma|a_n|$ is convergent. Write

$$\begin{aligned} a_n^+ &= a_n \quad \text{if} \quad a_n \geqslant 0, \\ a_n^+ &= 0 \quad \text{if} \quad a_n < 0; \\ a_n^- &= -a_n \quad \text{if} \quad a_n \leqslant 0, \\ a_n^- &= 0 \quad \text{if} \quad a_n > 0. \end{aligned}$$

Then
$$a_n = a_n^+ - a_n^-$$
and
$$|a_n| = a_n^+ + a_n^-.$$

The partial sums $|a_1| + \dots + |a_n|$ are bounded; *a fortiori* the partial sums $a_1^+ + \dots + a_n^+$, $a_1^- + \dots + a_n^-$ are bounded, and so by the Axiom of Completeness the series Σa_n^+, Σa_n^- are convergent.

A familiar decimal 'expansion' such as 4·283... is shorthand for the sum of the series
$$4 + 2/10 + 8/10^2 + 3/10^3 + \dots.$$

We are assured of the convergence of any such expansion because the partial sums are increasing and bounded above. In this particular case the partial sums

$$4,\ 4{\cdot}2,\ 4{\cdot}28,\ 4{\cdot}283, \ldots$$

are bounded above by 5. The Axiom of Completeness is simply stating what most of us have taken for granted every time we have written an infinite decimal expansion.

Another important fact emerges here. If an arbitrary real number k is given, we may construct a decimal expansion whose sum is k. We may as well assume that k is positive; otherwise we can find a decimal expansion for $-k$. It is helpful to refer to Fig. 3: if the points representing all the integers n are marked on the real line then the point k lies in exactly one of the half-open intervals $[n, n+1)$.† Let this interval be denoted by $[a_0, a_0+1)$ so that a_0 is the largest integer $\leqslant k$. If the interval $[a_0, a_0+1)$ is divided into ten intervals of type $[\ ,\ )$ and of equal length, then the point k lies in exactly one of these intervals. We may denote the left-hand end point of this interval by

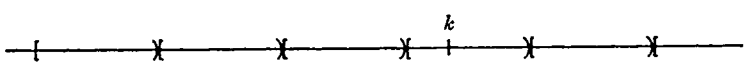

Fig. 3

$a_0 + a_1/10$, where $0 \leqslant a_1 \leqslant 9$, i.e. in decimal notation, $a_0{\cdot}a_1$ is the largest (integer/10) $\leqslant k$. By further subdividing into ten half-open intervals of equal length we find $a_0{\cdot}a_1a_2$ which is the largest (integer/10^2) $\leqslant k$; and so on inductively. By our construction $a_0{\cdot}a_1a_2 \ldots a_n$ differs from k by less than $1/10^n$ and so the decimal expansion converges to k. In other words

$$k = a_0{\cdot}a_1a_2\ldots.$$

The fact that any real number has a decimal expansion is of great practical value in all kinds of applied mathematics. It is also of profound theoretical importance because it means that one can find rational numbers (even of the type $m/10^n$) arbitrarily close to any given real number. It might be thought that one could therefore dispense with real numbers and deal with rational numbers exclusively. This would be disastrous in many ways. It was a source of anguish to the Pythagorean School in the 6th century B.C. Their idea of number in arithmetic was limited to integers and rationals; in this context they were aware that there is no (rational) number x which satisfies the simple

† We accept this as one of the elementary 'algebraic' properties of ordering on the real line; but see Ex. 5.

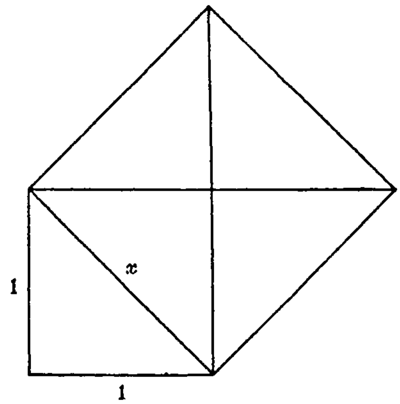

Fig. 4

equation $x^2 = 2$. In the realm of geometry, however, the most famous theorem of Pythagoras, applied to the diagonals of a unit square, provides a length x whose square is 2 (Fig. 4). The existence of such irrational or 'unutterable' numbers, as they called them, was viewed with such concern that the members were sworn to secrecy and forbidden to mention them to any outsider. Their dilemma was essentially resolved two centuries later by Eudoxus, but more than two thousand years elapsed before the mathematical world as a whole was prepared to accept the notion of the 'continuum' or, in other words, the completeness of the reals. (Using the Axiom of Completeness we can easily prove the existence of a positive *real* number whose square is 2: see Ex. 3.) In modern terms, this property of completeness is essential to the theory of limits and convergence; without it the differential calculus could not even make a beginning.

Exercises

1. Deduce the Axiom of Archimedes (footnote p. 3) from the Axiom of Completeness.

2. If $s_n \to s$ and $t_n \to t$ show that

$$\max\{s_n, t_n\} \to \max\{s, t\} \quad \text{and} \quad \min\{s_n, t_n\} \to \min\{s, t\}.$$

(Notation: if $a \geqslant b$, $\max\{a,b\} = a$, $\min\{a,b\} = b$; if $a < b$, $\max\{a,b\} = b$, $\min\{a,b\} = a$.)

3. For $n \geqslant 1$ let s_n be the largest (integer/10^{n-1}) whose square is less than 2 (e.g. $s_4 = 1{\cdot}414$). Show that $\{s_n\}$ is increasing and $\{s_n + 1/10^{n-1}\}$ is decreasing. Hence show that there is a positive real number whose square is 2.

4. If two infinite decimal expansion are different in form but converge to the same real number, show that one ends in a string of 0's and the other ends in a string of 9's (e.g. $1{\cdot}000\ldots$ and $0{\cdot}999\ldots$).

5. *The Well-Ordering Axiom* for the positive integers states that any non-empty set of positive integers has a smallest member. You are given a real number k; use the Well-Ordering Axiom and the Axiom of Archimedes (Ex. 1) to show that there is a largest integer less than or equal to k.

1.2 Infima and Suprema

Let us consider an increasing sequence $\{s_n\}$ which converges to a limit s. The diagram on the real line (Fig. 5) makes it intuitively clear that

$$s_n \leqslant s \quad \text{for all} \quad n. \tag{i}$$

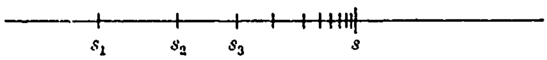

Fig. 5

We can prove this quite easily as follows. If (i) were not true, then for some N, $s_N > s$ and so all subsequent s_n, which of course satisfy $s_n \geqslant s_N$, could not get any closer to s than s_N (Fig. 6). This contradicts the definition of convergence.

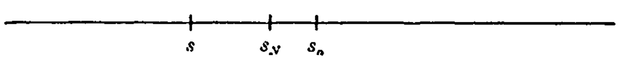

Fig. 6

Suppose now that we are given an increasing sequence $\{s_n\}$ bounded above by K. According to our Axiom of Completeness we are assured of the existence of a limit s. The diagram on the real line (Fig. 7) again makes it clear that the limit s satisfies

$$s \leqslant K. \tag{ii}$$

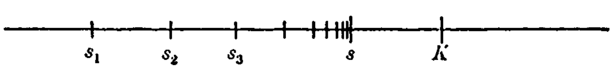

Fig. 7

The proof is quite brief. If we had $s > K$, then for s_n near enough to s we should have $s_n > K$ which would contradict the given condition $s_n \leqslant K$ for all n (see Fig. 8).

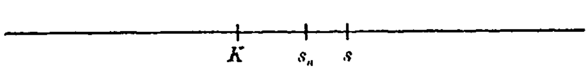

Fig. 8

Condition (i) means that s itself is an upper bound for the sequence $\{s_n\}$. On the other hand, condition (ii) tells us that no upper bound K

can be smaller than s. We combine these two conditions quite simply by saying that s is the *least upper bound* of the sequence $\{s_n\}$. The following theorem summarises our knowledge so far of the behaviour of increasing sequences.

Theorem 1. *An increasing sequence $\{s_n\}$ of real numbers is convergent if, and only if, it is bounded above; in this case the limit of the sequence is the same as its least upper bound.*

In a much more general way, if we have an arbitrary non-empty set X of real numbers, a real number K such that $x \leqslant K$ for all x in X will be called an *upper bound* for the set X. If such a K exists we say that X is bounded above (by K). In fact K may not exist; as we have mentioned before, the set of integers $\{1, 2, 3, \ldots\}$ has no upper bound. If an upper bound does exist the question arises: is there a *least* upper bound? In other words, is there an upper bound B such that no number smaller than B is an upper bound? Clearly a least upper bound, assuming that it exists, is *unique*, for if there were two such least upper bounds, neither would allow the other to be smaller.

It is also clear what we mean by a *lower bound* for a non-empty set X and a *greatest lower bound*. If we write $-X$ for the set of all $-x$, where x runs through all the numbers in X, then we may use the 'reflection' $x \to -x$ which reverses the order on the real line to transfer results from X to $-X$, and vice versa. It is clear that B is the least upper bound of X if and only if $-B$ is the greatest lower bound of $-X$.

As a convenient abbreviation we shall often refer to the least upper bound of X as the *supremum* of X, the greatest lower bound of X as the *infimum* of X, and write these as sup X, inf X, for short.

Theorem 2. (i) *A non-empty set of real numbers that is bounded above has a least upper bound* (*supremum*).

(ii) *A non-empty set of real numbers that is bounded below has a greatest lower bound* (*infimum*).

Proof. In view of the above remarks it is enough to prove (ii). Let X be a set of real numbers containing the number x and assume that X is bounded below by L, say. Of course, L may not be an integer, but there is an integer l in the interval $(L-1, L]$† which is also a lower bound for X; there are only finitely many integers in $[l, x]$† so there is a largest integer a_0 which is lower bound for X.† Among the numbers

† We accept these as reasonable statements about the real line; but see Ex. 3 and the footnote on p. 3.

$a_0+r/10$ ($r = 0, 1, \ldots, 9$) there is a largest $a_0+a_1/10$ which is a lower bound for X; in other words there is a largest (integer/10) $a_0 \cdot a_1$ which is a lower bound for X. Then there is a largest (integer/100) $a_0 \cdot a_1 a_2$ which is a lower bound for X; and so on. The expansion $a_0 \cdot a_1 a_2 \ldots$ converges to a real number s because of the Axiom of Completeness. It only remains to show that s is the greatest lower bound of X.

Consider an arbitrary element x of X. Every term of the sequence

$$a_0, \quad a_0 \cdot a_1, \quad a_0 \cdot a_1 a_2, \ldots$$

is a lower bound for X and so is less than or equal to x. We have seen above that this implies that $s \leqslant x$. Since x was arbitrary, this means that s is a lower bound for X.

Assume that $k > s$. As $1/10^n \to 0$ we may find an integer N for which $1/10^N < k-s$. In other words

$$k > s+1/10^N$$

and so certainly
$$k > a_0 \cdot a_1 a_2 \ldots a_N + 1/10^N.$$

But this last number is too large to be a lower bound (by our construction of the sequence of decimals) and so k cannot be a lower bound. This shows finally that s is the greatest lower bound of the set X, as required.

Many mathematicians prefer to state Theorem 2 as an axiom and to deduce our Axiom of Completeness from it. (This approach is discussed very briefly in § 1.3.) Our most important reason for favouring the statement in terms of monotone sequences is that these sequences feature in our definition and main theorem for the Lebesgue integral. We also believe that many students are more confident with the idea of a bounded increasing sequence than they are with the idea of a least upper bound.

Exercises

1. If the non-empty set X of real numbers is bounded above show that $\sup X$ is an element of X if and only if X possesses a largest member.

2. Let X, Y be non-empty sets of real numbers which are bounded above. Show that
$$\sup (X \cup Y) = \max\{\sup X, \sup Y\}.$$
(For the definition of max see Ex. 1.1.2.)

Let Z be the set of all $x+y$ for x in X, y in Y. Show that
$$\sup Z = \sup X + \sup Y.$$

3. Let x be an element of X and suppose that X is bounded below by L. Use the Axiom of Archimedes (footnote on p. 3) to show that there is an