(ii) They are distinctively of the twentieth century (or in the case of set theory, post-1870), having developed virtually from scratch in recent times.

(iii) Because of this they are accessible to anyone with a smattering of mathematics—certainly more accessible than any other modern topic of comparable importance.

(iv) They throw completely new light on the question "What is mathematics?" In fact, they arise from the first serious attempts to answer this question.

23.2 Sets

Sets became established in mathematics in the late nineteenth century as a result of attempts to answer certain questions about the real numbers. First, what *is* a real number? Several equivalent answers were given around 1870, all involving infinite sets or sequences. The simplest was that of Dedekind (1872), who defined a real number to be a partition (or "cut") of the rational numbers into two sets, L and U, such that each member of L is less than all members of U. If one has a preconceived notion of real number, such as a point x on a line, then L and U are uniquely determined by x as the sets of rational points to left and right of it, respectively. Thus if x is preconceived, then L and U are no more than auxiliary concepts that enable x to be handled in terms of rationals, as Eudoxus did (Section 4.2). Dedekind's breakthrough was to realize that no preconceived x was necessary: x could be *defined* as the pair (L,U). Thus the concept of sets of rationals was a basis for the concept of real number.

Dedekind cuts gave a precise model for the continuous number line $\mathbb{R}$, since they filled all the gaps in the rationals. Indeed, wherever there is a gap in the rationals, the real number that fills it is essentially the gap itself: the pair of sets L,U to left and right of it. Other formulations of this *completeness* property of $\mathbb{R}$ are also easy consequences of Dedekind's definition. For example, each bounded set of reals (L_i,U_i) has a *least upper bound* (L,U). L is simply the union of the sets L_i.

Dedekind seemed to have settled the ancient problem of explaining the continuous in terms of the discrete, but in penetrating as far as he did, he also uncovered deeper problems. The central problem is that the completeness of $\mathbb{R}$ entails its *uncountability*, a phenomenon discovered by Cantor

(1874). The *countable* sets are those that can be put in one-to-one correspondence with $\mathbb{N} = \{0,1,2,\dots\}$, and they include the set of rationals and the set of algebraic numbers, as Cantor also discovered. But if $\mathbb{R}$ is countable, this means that all reals can be included in a sequence $x_0, x_1, x_2, \dots$. Cantor (1874) showed that this is impossible by selecting from each sequence $\{x_m\}$ of distinct reals a subsequence $a_0, b_0, a_1, b_1, a_2, b_2, \dots$, such that

$$a_0 < a_1 < a_2 < \cdots < b_2 < b_1 < b_0$$

and with each x_m *outside* one of the nested intervals $(a_0, b_0) \supset (a_1, b_1) \supset (a_2, b_2) \supset \cdots$. It follows that any common element of all the (a_n, b_n) is a real $x \neq$ each x_m. A common element obviously exists if the sequence of intervals is finite, and if the sequence is infinite, it exists by completeness, as the least upper bound of the a_n. The common element x is a "gap" in the given sequence $\{x_m\}$.

The uncountability of $\mathbb{R}$ has been a great challenge to set theorists and logicians ever since its discovery. The most successful response to this challenge has been the theory of *ordinal numbers*, which grew out of Cantor's (1872) investigation of Fourier series (see Section 13.4). The existence of a Fourier series for a function f depends largely on the structure of the set of discontinuities of f, and thus leads to the problem of analyzing the complexity of point sets. Cantor measured complexity by the number of iterations of the prime operation ($'$) of taking the limit points of a set. For example, if $S = \{0, 1/2, 3/4, 7/8, \dots, 1\}$, then the prime operation can be applied once, and $S' = \{1\}$. It can happen that S' itself has limit points, so that S'' also exists. In fact, one can find a set S for which $S', S'', \dots, S^{(n)}, \dots$ exist for all finite n, so one can envisage iterating the prime operation an infinite number of times. In the case where all the $S^{(n)}$ exist, Cantor (1880) took their intersection, thereby defining

$$S^{\infty} = \bigcap\nolimits_{n=1,2,3,\dots} S^{(n)}.$$

He viewed ∞ as the first infinite ordinal number. To avoid confusion with higher infinite numbers soon to appear, I shall use the modern notation ω for the first infinite ordinal.

Having made the leap to ω, it is easy to go further. $\left(S^{(\omega)}\right)' = S^{(\omega+1)}$, $\left(S^{(\omega+1)}\right)' = S^{(\omega+2)}$, ..., and the intersection of this new infinite sequence is $S^{\omega\cdot 2}$, where $\omega \cdot 2$ is the first infinite number after ω, $\omega+1$, $\omega+2$, ... After $\omega \cdot 2$, one has

$$\omega\cdot 2+1,\ \omega\cdot 2+2,\ \dots,\ \omega\cdot 3, \dots,\ \omega\cdot 4,\ \dots,\ \dots,\ \omega\cdot\omega,\ \dots.$$

All these can actually be realized as numbers of iterations of the prime operation on sets of reals. We can also investigate the ordinal numbers independently of this realization, as an extension of the concept of natural number.

Cantor (1883) viewed the ordinals as being generated by two operations:

(i) Successor, which for each ordinal α gives the next ordinal, $\alpha+1$.

(ii) Least upper bound, which for each set $\{\alpha_i\}$ of ordinals gives the least ordinal $\geq$ each α_i.

The most elegant formalization of these notions was given by von Neumann (1923). The empty set $\emptyset$ (not considered by Cantor) is taken to be the ordinal 0, the successor of α is $\alpha \cup \{\alpha\}$, and the least upper bound of $\{\alpha_i\}$ is simply the union of the α_i. Thus

$$\begin{aligned} 0 &= \emptyset, \\ 1 &= \{0\}, \\ 2 &= \{0,1\}, \\ &\cdots \\ \omega &= \{0,1,2,\ldots,n,\ldots\}, \\ \omega+1 &= \{0,1,2,\ldots,n,\ldots,\omega\}, \end{aligned}$$

and so on. The natural ordering of the ordinals is then given by set membership, $\in$, and in particular the members of an ordinal α are all ordinals smaller than α.

Cantor's principle (ii) generates ordinals of breathtaking size, since it gives the power to transcend any set of ordinals already defined. In particular, an ordinal of uncountable size is on the horizon as soon as one thinks of the concept of countable ordinal, as Cantor did (1883). He defined an ordinal α to be countable (or, as he later put it, of *cardinality* or cardinal number $\aleph_0$) if α could be put in one-to-one correspondence with $\mathbb{N}$. For example,

$$\omega \cdot 2 = \{0,1,2\ldots,\omega,\omega+1,\omega+2,\ldots\}$$

is countable because of its obvious correspondence with

$$\mathbb{N} = \{0,2,4,\ldots,1,3,5,\ldots\}.$$

The least upper bound of the countable ordinals is the least *uncountable* ordinal, ω_1. Sets in one-to-one correspondence with ω_1 are of the next cardinality, $\aleph_1$. Ordinals of cardinality $\aleph_1$ have a least upper bound ω_2 of cardinality $\aleph_2$, and so on.

Having found this orderly way of generating successive uncountable cardinals, Cantor reconsidered the uncountable set $\mathbb{R}$. Although no method of generating members of $\mathbb{R}$ in the manner of ordinals was apparent, Cantor conjectured that the cardinality of $\mathbb{R}$ was $\aleph_1$. This conjecture has since become known as the *continuum hypothesis*. By 1900 it was recognized as the outstanding open problem of set theory, and Hilbert (1900a) made it number one on the famous list of problems he presented to the mathematical community. There have been two outstanding results on the continuum problem since 1900, but they seem to make it less likely that we will ever know whether the continuum hypothesis is correct. Gödel (1938) showed that the continuum hypothesis is *consistent* with standard axioms for set theory, but Cohen (1963) showed that its negation is also consistent. Thus the continuum hypothesis is independent of standard set theory, in the same way that the parallel postulate is independent of Euclid's other postulates. Whether this means that the notion of "set" is open to different natural interpretations, like the notion of "straight line," is not yet clear.

EXERCISES

Cantor's 1874 proof of the uncountability of $\mathbb{R}$ is based on the following construction. Given a sequence $x_0, x_1, x_2, \ldots$ of distinct reals, he found a "gap" in them by picking out $a_0, b_0, a_1, b_1, \ldots$ as follows.

$$\begin{aligned}
a_0 &= x_0, \\
b_0 &= \text{first } x_m \text{ with } a_0 < x_m, \\
a_1 &= \text{first } x_m \text{ after } b_0 \text{ with } a_0 < x_m < b_0, \\
b_1 &= \text{first } x_m \text{ after } a_1 \text{ with } a_1 < x_m < b_0, \\
a_2 &= \text{first } x_m \text{ after } b_1 \text{ with } a_1 < x_m < b_1. \\
&\vdots
\end{aligned}$$

23.2.1 Explain why the sequence $a_0, b_0, a_1, b_1, a_2, b_2, \ldots$ has the "gap" property described above: each x_m is outside one of the nested intervals $(a_0, b_0) \supset (a_1, b_1) \supset (a_2, b_2) \supset \cdots$.

We now explore how far we can enlarge the set of natural numbers and still have a countable set.