numbers are to provide an arithmetic model for the points in a line in the manner which is implicit in coordinate geometry with its tacit assumption that every point on the x- and y-axes is indexed by a number, and (b) that the geometrical interpretation supports a realist conception of limits of infinite sequences, via the picture of points as limits of sequences of approximation to them, limits which are therefore given independently of the sequences of which they are the limits. It is this which tends to underwrite the conception of infinite convergent sequences of rational numbers as actual, completed, infinities rather than as potential infinities.

Tiles, Mary
The Philosophy of Set Theory
Mineola: Dover 2004

TILES

5

Cantor's Transfinite Paradise

It was Cantor's work which gave sense to the question 'How many points are there in a line?', a question which previously lacked any precise sense. To avoid prejudging the question of whether Cantor should be seen as an inventor or a discoverer, the notion of 'sense' here can be treated as relating to the mathematician's understanding of the question. This was certainly changed by Cantor's work, whether we think that this was a change in the concepts involved or a more adequate grasping of unchanging concepts. But by looking at the way in which the question of the number of points on a line comes to have a sense we may be able to shed an indirect light on the discovery/invention issue as well as on the question of exactly what sense the question has been given.

Before Cantor developed his theory of transfinite numbers, the natural, and the only available answer, to the question was 'Infinitely many', and this was a way of saying that there is no number of points in a line, they are without number. This answer is not devoid of content for it indicates that given any finite number of points in a line there will always be more. In other words, the indeterminacy surrounding the totality of points in a line is of a specific kind and is unlike that surrounding the totality of angels that can dance on the head of a pin. One can at least start counting the points in a line, but cannot see any way to stop. So one knows that certain numerical answers are wrong, and to the extent that such answers are ruled out, the question has some sense; it is not wholly meaningless.

Thus we see that the first prerequisite for giving any further sense to the question is to have some more determinate conception of the totality of points in a line. In addition the concept of number needs to be extended or modified in such a way that even collections which cannot be fully counted may none the less be supposed to contain a determinate number of elements – may be assigned a number. The first of these, as was seen in chapter 4, is fulfilled by the development of the theory of real numbers. The second, which concerns us here, is effected by analyses of the concept of number which link numbering to the establishment of one–one correspondences rather than specifically to counting (counting is just one way of setting up a one–one correspondence between natural numbers and members of the set being counted).

1 Sets and Cardinal Numbers

The basic idea behind set theoretic analyses of the notion of positive whole number (natural number) is that it is to sets, or to collections of things, that numbers are assigned. A flock of sheep is counted and the number reached is the number of sheep in the flock. But it is also possible to compare sets in respect of the number of things they contain without actually counting them. One may, for example, establish that there are just as many cups as saucers on a tray by checking to see that there are no saucerless cups and no saucers without cups on them, without actually counting to find out how many of either there are. One can be in a position to say that there must be the same number of each without being able to say what that number is by establishing that there is a one–one correspondence between the set of cups on the tray and the set of saucers on the tray. Counting can then be seen as establishing the existence of a one–one correspondence between a finite set of objects and a subset of the natural numbers. If this correspondence can be thought to exist independently of anyone setting it up, then any finite set must contain a determinate number of objects, whether it has been counted or not. For any two finite sets it is obvious that there can be a one–one correspondence between them only if they contain the same number of elements, and it is tempting to generalize this to all sets, whether finite or infinite. This is precisely the step which Cantor took, but it is not entirely straightforward.

Suppose that we were to define 'number' by stipulating that a collection A has the same number of elements as another collection

B if (and only if) there is a one–one correspondence between them (a correlation which assigns to each member of A exactly one member of B and to each member of B exactly one member of A). Since this does not make any mention of whether the collections in question are finite or infinite it would seem to legitimize an extension of the notion of number into the infinite, enabling one to think of an infinite collection as containing a determinate number of things.

But the matter is not quite so simple. For as was mentioned in chapter 3, infinite sets are characterized by the seemingly paradoxical property of being such that they can be put into one–one correspondence with proper parts of themselves. This is paradoxical because it seems intuitively obvious that there must always be more elements in any whole than in some proper part of it.

The set of natural numbers $N = \{0, 1, 2, 3, \ldots n \ldots\}$ can be put in one–one correspondence with the set which consists only of the even numbers $E = \{0, 2, 4, 6, \ldots 2n, \ldots\}$ by pairing each n in N with $2n$, in E. Because E is infinite its supply of elements will never run out even though one would instinctively want to say that there must be twice as many elements in N as in E. And Dedekind defined infinite sets by reference to this characteristic: a set A is *infinite* if, and only if, there is a one–one correspondence between A and a set X which is a proper subset of A.

This means that if one were to say that two infinite sets contain the same number of elements when there is a one–one correspondence between them, and if one remains convinced that the size of any whole must always be greater than that of any of its proper parts, then the number of elements in an infinite set cannot be thought to be a measure of its size. For example, N and E will have the same 'number' of elements even though there are infinitely many numbers in N which are not in E, so that in this sense N is 'bigger than' E. This suggests that the elements of an infinite set are without number not just because they cannot be exhaustively counted but also because the notion of number, as a measure of size, can get no grip here. All infinite sets seem to come out as being of the same 'size' if one–one correspondence is taken as indicating sameness of size for sets. Indeed, if all infinite sets could be put into one–one correspondence with each other, one would be justified in treating the classification 'infinite' as an undifferentiated refusal of numerability. But given Cantor's discovery that there are infinite sets which cannot be put into one–one correspondence with each other, this conclusion is less compelling.

His proof, that for any set A (whether finite or infinite) there can be no one–one correspondence between A and the set of all subsets of A (the power set of A, denoted by $P(A)$), is important because it entails that there can be no one–one correspondence between the natural numbers and the real numbers. It immediately follows that the set N of natural numbers cannot be put into one–one correspondence with its power set $P(N)$. Since (a) each subset of the natural numbers can be uniquely correlated with an infinite sequence of zeros and ones, (b) each such sequence can be read as a binary decimal representation of a real number in the interval (0, 1) and thus as representing a point on the unit line, and (c) the real numbers in (0, 1) index all the points on whatever is chosen as the unit line, this means that the points on a line cannot be put into one–one correspondence with the natural numbers.

Cantor interpreted this impossibility of one–one correspondence as meaning that there must be 'more' points on a line than there are natural numbers (since there are clearly at least as many points in a line as there are natural numbers). More generally he interpreted it as licensing an attempt to extend the notion of number into the infinite. On this basis it became necessary to recognize a division between those sets which are denumerable, i.e. which can be put in one–one correspondence with the natural numbers, and those which, like the set of points in a line, are non-denumerable, i.e. for which no such correspondence exists.

Following Cantor in taking the first steps in this direction slightly more formally:

Definition A *set*, or *aggregate*, is any collection into a whole M of definite and separate objects m of our intuition or thought.

Assumption Every set, or aggregate, has a determinate 'power' or 'cardinal number'.

Definition Two sets M and N have the same power, or cardinal number ($C(N) = C(M)$) if, and only if, there is a one–one correspondence between them.

Cantor's way of introducing this definition is to say:

> We will call by the name 'power' or 'cardinal number' of M the general concept which, by means of our active faculty of thought, arises from the aggregate M when we make abstraction from the nature of its various elements m and of the order in which they are given. (Cantor, 1955, p. 86)

If this is interpreted as an attempt to define a concept by reference to a mental act, something performed privately by each individual for himself, it is hardly a rigorous or adequate definition. But it can also be treated as a commentary on the precise condition under which two sets are to be said to have the same power. Whatever 'power' or 'cardinal number' is, it is a property of a set which does not depend on the specific nature of the elements it contains nor on the order in which the elements are given because neither of these are relevant to determining whether two sets are the same in this respect.

Likewise, the 'definition' of 'set' is less a definition than an attempt at explication of something which is being given the status of a primitive, undefined, term. For example Hausdorff introduces the term 'set' as follows:

> A set is formed by the grouping together of single objects into a whole. A set is a plurality thought of as a unit. (1957, p. 11)

What is implicit in both these explications is the thought that a set is a determinate collection of objects (a whole given after its parts) whose identity is entirely dependent on its members (the objects collected) and not on any method by which they may have been grouped or collected. Cantor, in particular, wanted to treat all sets as far as possible by analogy with finite sets. A finite set can be specified simply by listing its members, in the form $\{a, b, c, d\}$ where a, b, c, d need have nothing in common (a wasp, a London bus, Mount Everest, Mrs Thatcher). An infinite set might, by analogy, be thought to be specifiable by an infinite list. It might be just a contingent human limitation not to be able to think of an infinite set as a unit without going via some common characteristic of its elements, or some principle for selecting or generating them.

Here then, although classes, treated extensionally, would be counted as sets, sets are not restricted to being classes, i.e. are not restricted to being collections of objects which are the extensions of terms.

It is important to the development of set theoretic analyses of the natural numbers that sets be determinate collections of objects. Sets are to be just the sort of collections whose members can, or could in principle, be counted and thus assigned a number. But it is not easy to make the conditions of numerability explicit. To say that two sets have the same power if, and only if, there is a one–one correspondence between them does not yet entitle one to call these powers 'cardinal numbers' (where the sense of 'cardinal number' is derived from finite sets and natural numbers – the sense in which we may think of such sets as containing a determinate (finite) number of objects even though they have not been counted). At the very least we need to be able to say when the power, or cardinal number, of one set is greater or less than that of another.

Definition Given any two sets A and B

$C(A) \leqslant C(B)$ iff there is a subset B° of B such that $C(B^\circ) = C(A)$. *i.e.*
$C(A) \leqslant C(B)$ iff $(\exists B^\circ)(B^\circ \subseteq B \;\&\; C(B^\circ) = C(A))$.
$C(A) < C(B)$ iff $C(A) \leqslant C(B)$ and $C(A) \neq C(B)$.

But to know that this defines even a partial order relation it is necessary to know that

$$C(A) \leqslant C(B) \;\&\; C(B) \leqslant C(A) \Rightarrow C(A) = C(B)$$

i.e. that if there is a subset B° of B such that there is a one–one correspondence between A and B° and there is a subset A° of A such that there is a one–one correspondence between B and A°, then there is a one–one correspondence between A and B. For finite sets this is obvious, but not for infinite sets. The proof that it holds for infinite sets is known as the Schröder–Bernstein theorem (for a proof see, for example, Rotman and Kneebone, 1966, p. 49). But even this result, does not give all that is necessary for powers to

look like cardinal numbers. Given any two sets A and B there are four possibilities:

1 There is $B^\circ \subseteq B$ such that $C(A) = C(B^\circ)$, and there is $A^\circ \subseteq A$ such that $C(A^\circ) = C(B)$.
2 There is $B^\circ \subseteq B$ such that $C(A) = C(B^\circ)$, but there is no $A^\circ \subseteq A$ such that $C(A^\circ) = C(B)$.
3 There is no $B^\circ \subseteq B$ such that $C(A) = C(B^\circ)$, but there is $A^\circ \subseteq A$ such that $C(A^\circ) = C(B)$.
4 There is no $B^\circ \subseteq B$ such that $C(A) = C(B^\circ)$, and there is no $A^\circ \subseteq A$ such that $C(A^\circ) = C(B)$.

We then have:

1 implies $C(A) = C(B)$
2 implies $C(A) < C(B)$
3 implies $C(B) < C(A)$
4 implies A and B are incomparable in respect of cardinality.

It can readily be shown that if either A or B or both are finite then case 4 will not arise, but the proof that 4 can never occur when both A and B are infinite requires a further assumption about sets – the assumption that every set can be well-ordered (ordered in such a way that each of its non-empty subsets has a least element).

Unable to prove the comparability of all sets in respect of cardinality, Cantor adopted it as an assumption. With this assumption it is possible to operate with powers in such a way that they do indeed begin to behave like cardinal numbers.

Definitions Let A and B be sets. Let $C(A) = a$, $C(B) = b$, then

1 If A and B are disjoint, $a + b = C(A \cup B)$ where $A \cup B = \{x: x \in A \text{ or } x \in B\}$.
2 $a \cdot b = C(A \times B)$ where $A \times B = \{\langle x, y\rangle: x \in A \;\&\; y \in B\}$.
3 $a^b = C(A^B)$ where $A^B = \{f: f \text{ is a function from } B \text{ to } A\}$.

It can then be proved that powers, or cardinalities, behave very much as numbers should, and that in the case of finite sets we get all the results we should expect. If $A = \{a, b\}$ and $B = \{k, m, n\}$ and we put $C(A) = 2$, $C(B) = 3$ we find that

$$
\begin{aligned}
2 + 3 &= C(\{a, b, k, m, n\}) = 5 \\
2 \cdot 3 &= C(\{\langle a, k\rangle, \langle a, m\rangle, \langle a, n\rangle, \langle b, k\rangle, \langle b, m\rangle, \langle b, n\rangle\}) = 6 \\
2^3 &= C(\{\{\langle k, a\rangle, \langle m, a\rangle, \langle n, a\rangle\}, \{\langle k, b\rangle, \langle m, b\rangle, \langle n, b\rangle\}, \\
&\quad \{\langle k, a\rangle, \langle m, a\rangle, \langle n, b\rangle\}, \{\langle k, b\rangle, \langle m, b\rangle, \langle n, a\rangle\}, \\
&\quad \{\langle k, a\rangle, \langle m, b\rangle, \langle n, b\rangle\}, \{\langle k, b\rangle, \langle m, a\rangle, \langle n, a\rangle\}, \\
&\quad \{\langle k, a\rangle, \langle m, b\rangle, \langle n, a\rangle\}, \{\langle k, b\rangle, \langle m, a\rangle, \langle n, b\rangle\}\}) = 8
\end{aligned}
$$

But, as might be expected, given that infinite sets can be put in one–one correspondence with proper subsets of themselves, infinite cardinalities do not behave quite like finite ones and their 'arithmetic' may seem a bit surprising. For example, if $\aleph_0 = C(N)$ where N is the set of natural numbers, then for any finite n

$$
\begin{aligned}
n\aleph_0 &= \aleph_0 + \aleph_0 + \ldots (n \text{ times}) = \aleph_0 \\
\aleph_0^n &= \aleph_0 \cdot \aleph_0 \cdot \ldots (n \text{ times } n) = \aleph_0
\end{aligned}
$$

Both the even numbers, E, and the odd numbers, O, can be put in one–one correspondence with the whole natural number sequence, i.e. $C(E) = C(O) = C(N) = \aleph_0$. And since $N = E \cup O$ and E and O are disjoint, $C(E) + C(O) = C(N)$, i.e. $\aleph_0 + \aleph_0 = \aleph_0$.

What this arithmetic of cardinal numbers does give is a way of expressing the relationship between the cardinality of a given set A and its power set P(A) (the set of all subsets of A). For there is a one–one correspondence between subsets of A and the set of all functions f from A to a two element set, such as $\{0, 1\}$, where each subset is considered as the set of those elements of A for which f takes the value 1.

$$\text{so} \quad C(\mathrm{P}(A)) = C(2^A) = 2^{C(A)}$$

Since it has already been established that the cardinal number of the points on a (unit) line is the same as that of the real numbers in the interval (0, 1), and that this in turn is the same as the cardinal number of the power set of the natural numbers P(N), we can now put a 'number' on the points in a (unit) line, namely $C(\mathrm{P}(N)) = 2^{\aleph_0}$.

But if, by analogy with $2^3 = 8$, we ask for an 'evaluation' of $2^{\aleph_0}$ we find that we do not have the means of supplying an answer. In particular Cantor thought that there are no infinite cardinal numbers in between $\aleph_0$ and $2^{\aleph_0}$, i.e. that $2^{\aleph_0}$ is the next infinite cardinal number after $\aleph_0$, but was unable to prove it. This is what has become known as Cantor's continuum hypothesis.

The situation so far is that it has just been assumed (for want of the means of providing a proof), that given any two sets A and B, either $C(A) = C(B)$, or $C(A) < C(B)$ or $C(B) < C(A)$. In other words it has been assumed that cardinalities, or cardinal numbers, can be arranged in a single linear order. But just making that assumption does not tell us anything about the nature of the cardinal number 'sequence', about how to establish where any given cardinality lies in it, or even whether it is correct to talk about there being a *next* cardinal number after $\aleph_0$. Our assumption does not rule out the possibility that infinite cardinalities might, like the rational numbers, be densely ordered. If that were the case, there would always be another cardinal number between any two given cardinalities and given any cardinal number there would be no 'next' one.

Once we get into the domain of infinite cardinalities the only procedure we so far have for reliably generating higher cardinalities is exponentiation – repeatedly taking the cardinal number of the power set of a given set. So we can form a series of sets

$$N, \mathrm{P}(N), \mathrm{P}(\mathrm{P}(N)), \mathrm{P}(\mathrm{P}(\mathrm{P}(N))) \ldots \text{ with cardinal numbers}$$

$$\aleph_0, 2^{\aleph_0}, 2^{\aleph_0}, 2^{2^{2^{\aleph_0}}} \ldots$$

In the finite case, exponentiation does not take us from one cardinal number to the next, there are lots of numbers in between 2 and 2^3, and even more between $2^3 = 8$, and 2^8. But we also know that infinite cardinalities do not behave in the same way as finite ones, so the question of whether there are or are not any infinite numbers in between those in the above series is an open question. The whole problem is that the theory of cardinality on its own has so effectively severed the connection between number and measure of size that it gives rise to no numerical scale by reference to which infinite sets might be 'measured'. The idea of a scale is one which involves an order; in demanding a scale we are asking for an ordering of cardinalities. But the infinite cardinalities were introduced by disregarding the associations of number with counting and with ordering, so it is perhaps not surprising that this way of introducing transfinite numbers yields 'numbers' which, even if ordered in reality, as the assumption of comparability asserts, cannot be put in order or named in order by us.

2 Transfinite Ordinal Numbers

However, as was seen in chapter 4, it was not with cardinal notions that Cantor first started to extend the notion of number into the transfinite (Cantor, 1883). His initial extension was of the natural number *sequence* into the transfinite, using numbers as a measure of the number of times an operation has been repeated. He first introduced his transfinite ordinal numbers as numbers which are *generated* in a sequence and thus as an extension of the natural number sequence, which is generated in counting by the principle of adding one to the previous number. Thus his first principle of generation for ordinal numbers is as follows.

First Principle of Generation The addition of a unit to a number which has alredy been formed.

Used on its own this principle just gives us the ordinary natural numbers, or numbers belonging to what Cantor calls the first number class.

First Number Class (I) $= 0, 1, 2, 3, 4, \ldots$

His second principle of generation is one which allows for the formation of the first infinite ordinal numbers as limit numbers. We imagine that the set of natural numbers can be run through in order and, assuming they constitute an actually infinite set, that there must be an infinite bound to the numbers required. The second principle of generation allows for the formation of a 'number', ω, to stand for the first number which is greater than all the finite numbers. ω is thought of as a limit which the sequence $0, 1, 2, 3, \ldots$ approaches but never attains.

Second Principle of Generation If there is defined any definite succession of real integers of which there is no greatest, a new number is created, which is defined as the next greatest to them all.

Once one of these new, infinite ordinal numbers has been introduced, the first principle of generation will apply to it so that we get a new sequence $\omega + 1$, $\omega + 2$, $\omega + 3$, ... This, being a succession of integers with no greatest member, is a sequence to which the second principle can be reapplied so that we get another infinite ordinal number $\omega + \omega$ or $\omega \cdot 2$. The infinite ordinal numbers generated by repeated applications of these first two principles alone form what Cantor called the second number class.

Second Number Class (II) $= \omega, \omega + 1, \ldots, \omega + n, \ldots, \omega \cdot 2, (\omega \cdot 2) + 1, \ldots \omega \cdot 3, \ldots, \omega \cdot \omega, \ldots$

All the numbers in the second number class can, however, be thought of as the numbers obtained by introducing a more or less complicated order on the sequence of natural numbers. Any sequence which, like $0, 1, 2, 3, \ldots$, is a linear sequence with no last member and which involves only one infinite sequence will be called a *simply infinite sequence*. The sequence formed by running through all the natural numbers from 2 on and then tacking 0 and 1 on at the end after all the rest, is not a simply infinite sequence but one whose ordinal number is $\omega + 2$, i.e. a simply infinite sequence followed by a two element sequence $(2, 3, 4 \ldots 0, 1)$. Similarly the sequence formed by running through all the even numbers and then all the odd numbers, $(0, 2, 4, \ldots 1, 3, 5, \ldots)$, has the ordinal number $\omega + \omega$, i.e. it is one simply infinite sequence followed by another. The sequence formed by running through all the numbers divisible by 2, followed by all those divisible by 3, followed by all those divisible by 5, and so on for all the prime numbers.

$$(2, 4, 6, \ldots 3, 6, 9, \ldots 5, 10, 15, \ldots \quad \ldots \quad \ldots)$$

is a simply infinite sequence of simply infinite sequences, whose ordinal number is $\omega \cdot \omega$, since there is no greatest prime number. This means that although we have generated a lot of infinite ordinal numbers (numbers which depend on the order in which a set is given) they are all such that they are ordinal numbers of sets which can be put in one–one correspondence with the natural numbers (denumerable sets), and indeed are all ordinal numbers which can be assigned to the set of natural numbers when it is listed (or 'counted') in something other than its natural order.

So what we have is a proliferation of infinite ordinal numbers which all apply to sets having the same cardinality, $\aleph_0$. These first two principles on their own do not generate any ordinal number which could be the number of points in a line, since this set has a cardinality greater than $\aleph_0$. Thus Cantor introduces a third principle of generation, which he also called the principle of limitation, or the principle of interruption.

Third Principle – Principle of Limitation All the numbers formed next after ω should be such that the aggregate of numbers preceding each one should have the same power (or cardinality) as the first number class. These numbers constitute the second number class.

The idea behind this principle is to delimit a totality of ordinal numbers produced by the first two principles (the second number class) in such a way that the second principle can then be applied to give a new number (ω_1) which is defined as the next number greater than all of the numbers belonging to the second number class. The first two principles can then be reapplied to further extend the ordinal sequence. A more general form of the principle of limitation allows this process to go on indefinitely by sectioning the numbers generated into number classes to which the second principle can then be applied.

General Principle of Limitation All the numbers formed next after ω_α should be such that the aggregate of numbers preceding each one should have the same power (or cardinality) as the $(\alpha + 1)$th number class. These numbers then form the $(\alpha + 2)$th number class.

Cantor proved that the second number class cannot be put into one–one correspondence with the first and that there can be no set with a cardinality in between those of the two number classes. More generally he proved that the cardinality of the $(\alpha + 1)$th number class will be greater than that of the αth number class and that there

can be no set with a cardinal number in between these. Thus the cardinal number of the second number class is the next after $\aleph_0$, and is labelled $\aleph_1$. Any set whose ordinal number is ω_1 or more will also have a cardinality greater than $\aleph_0$, i.e. will be a non-denumerable set.

3 Ordinal Numbers and Cantor's Continuum Hypothesis

Our question about the number of points in a line thus now receives a further sense. We can now ask whether its cardinality is the next after $\aleph_0$ by asking whether $2^{\aleph_0} = \aleph_1$. The more precise form of Cantor's continuum hypothesis asserts that this is the case. We can also ask whether it is possible to assign the points in a line, in their natural order, an ordinal number. If this were possible then it would be relatively easy to answer the question about the cardinality of the continuum because we would merely have to know which number class the relevant ordinal number falls into. The problem is that the points in a line in their natural order cannot be assigned an ordinal number, even one which involves going through infinitely many infinite sequences.

To see this it is necessary to examine more closely what is involved in assigning to a set an ordinal number, making explicit some of the assumptions which have been made. The ordinal number sequence was extended into the infinite by imitating as closely as possible the ordinary natural number sequence. Similarly the application of infinite ordinal numbers to sets is to imitate as closely as possible the application of the finite, natural numbers when they are used in counting to give the number of elements in a finite collection.

When the elements of a finite set are counted this can be seen as setting up a one–one correspondence between these elements and the first part of the natural number sequence given in order. When the elements of the set are exhausted the last number used gives the number of elements in the set. This process of counting involves selecting the elements counted in a particular order; when a set is counted it is counted in a particular order. When the set to be 'counted' is infinite the process of counting does not stop so one cannot assign a number on the basis of saying that it is the last one to be used. But one can continue to use the idea of putting the elements of the set in an order which matches the order of the first part of the ordinal number sequence up to some specific ordinal number. This was what was done with the examples of the different orderings (ways of counting) the natural number sequence, and is the idea used in extending the ordinary arithmetic operations to the ordinal numbers. This is based on the idea that the addition of ordinal numbers involves placing two ordered sequences end to end to obtain a new, extended ordered sequence. Thus if $A = \langle a, b \rangle$ and $B = \langle c, d, e \rangle$ are ordered sets, the ordered union $A + B = \langle a, b, c, d, e \rangle$, whereas the ordered union $B + A \langle c, d, e, a, b \rangle$. This means that the addition of ordinal numbers will not, in the infinite case, be commutative. Multiplication of ordinal numbers is defined as repeated addition. Where infinite sequences are involved the ordinary laws of arithmetic are not all obeyed. For example:

$$1 \cdot \omega = 1 + 1 + 1 \ldots = \omega$$
$$2 \cdot \omega = 2 + 2 + 2 + \ldots = (1 + 1) + (1 + 1) + (1 + 1) = \ldots = \omega$$
$$\omega \cdot 2 = \omega + \omega = (1 + 1 + 1 + \ldots) + (1 + 1 + 1 + \ldots)$$
$$\omega \cdot (\omega + 1) = (\omega + \omega + \omega + \ldots) + \omega = (\omega \cdot \omega) + \omega$$
$$(\omega + 1) \cdot \omega = (\omega + 1) + (\omega + 1) + (\omega + 1) + \ldots = \omega \cdot \omega$$

For a set to be assigned an ordinal number it must be possible to order it in the same kind of way that the sequence of ordinal numbers is ordered. The ordinal number sequence is constructed by starting a sequence, letting it run on infinitely and then taking the next number after all of those and starting again. There are lots of bits of the sequence which have no greatest member (there is no greatest natural number, for example) but every bit has a least member (it always starts somewhere). So if a set is to have an ordinal number it must be possible to arrange its elements in a linear order which is such that every non-empty subset has a least element, i.e. to impose a well-ordering on it.

But the points in a line (or the real numbers in the interval (0, 1)) given in their natural order are not well-ordered. There are indefinitely many subsections of the line which have no first point. For example the set of points corresponding to numbers greater than $\frac{1}{2}$ and less than 1. Since there is no real number immediately after $\frac{1}{2}$ (between any given number and $\frac{1}{2}$ there will always be another number) there is no least element to this set. So if the points in a line were to be assigned an ordinal number it would have to be possible to impose on them an order which is different from their natural order and which is a well-ordering. This can be done for the rational numbers, for they too are not well-ordered by the natural order; there is no least rational number greater than $\frac{1}{2}$ and less than one. But it is possible to write out all positive rational numbers (with some repetitions) using two dimensions:

	1	2	3	4	5	6	...
1	1/1	1/2	1/3	1/4	1/5	1/6	...
2	2/1	2/2	2/3	2/4	2/5	2/6	...
3	3/1	3/2	3/3	3/4	3/5	3/6	...
4	4/1	4/2	4/3	4/4	4/5	4/6	...
5	5/1	5/2	5/3	5/4	5/5	5/6	...
6	6/1	6/2	6/3	6/4	6/5	6/6	...
.	.	.	.	.	.	.	
.	.	.	.	.	.	.	
.	.	.	.	.	.	.	

The numbers in the array can be listed by following the arrows, giving the sequence:

1/1,	1/2,	2/1,	3/1,	2/2,	1/3,	1/4,	2/3,	3/2,	4/1,	5/1,	...
1	2	3	4	5	6	7	8	9	10	11	...

And this gives not just a well-ordering but also a one–one correspondence with the natural numbers in their natural order. Each rational number can be expressed in a form x/y, where x and y are relatively prime. The numbers of this form constitute an infinite subset of that listed (infinite because $n/1$ is included for each natural number n). The listing thus effects a one–one correspondence between the positive rational numbers and a subset of the natural numbers in their natural order. Thus the rational numbers, ordered in this way, have ordinal number ω and are shown to be denumerable, i.e. to have cardinality $\aleph_0$.

However, it is not possible to produce an ordering on the real numbers in the same sort of way, for if it were, there would be a one–one correspondence between them and the natural numbers.

One can use a square array to show that no enumeration of the real numbers in (0, 1) can be complete; there must always be a number which has been missed out. Suppose that the real numbers are given in terms of their binary decimal expansion, then when we list them we get an array:

	r_1	r_2	r_3	r_4	...
1	0	1	1	0	...
2	0	0	1	1	...
3	1	1	1	0	...
4	1	0	1	1	...
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.

Define a new number r given by reading down the diagonal of the array and interchanging zeros and ones, i.e. given by the decimal expansion .1100.... This number differs from all those listed because, for each n, it differs in the nth decimal place from r_n. This means that there are no immediate grounds for supposing either that there is a well-ordering of the points in a line or that there is not.

The extension of the ordinal number sequence into the transfinite gives a way of generating not only infinite ordinal numbers, but also a sequence of infinite cardinal numbers, the numbers of the successive number classes, which is such that each one is the 'next greatest' after the one which it follows. So the ordinal number sequence also provides a scale of cardinal numbers on which one might hope to locate $2^{\aleph_0}$. But the problem, as Cantor saw, was that the two routes to cardinal numbers are largely independent. There is no guarantee that every cardinality (the power of every set) will have a representative amongst the cardinalities of the number classes (classes of ordinal numbers) generated by Cantor's three principles. Without such a guarantee one has no reason to suppose that there must be a definite answer to 'Where on the sequence of alephs indexing the number classes, does $2^{\aleph_0}$ lie?' In proposing that $2^{\aleph_0} = \aleph_1$ Cantor also assumed that every set, not just the set of points on a line, can be well-ordered, and it was this assumption that entitled him to claim that every set, including the set of points on a line, has at least one ordinal number. Any set which has an ordinal number will have a cardinal number which is amongst the alephs. If it is not the case that every set can be well-ordered, then the cardinalities of those sets which cannot be so ordered will not be represented by cardinal numbers arising from the construction of the ordinal number sequence.

But if the ordinal number sequence is just a construct, something Cantor brought into existence by defining certain symbols, principles for generating more of them and rules for manipulating them, there would be an essential asymmetry between the status of infinite cardinal and infinite ordinal numbers. This asymmetry would not justify Cantor's persistent attempts to prove his continuum hypothesis. For he believed not only that it was true but that, as a mathematical truth, it should also be provable.

4 Order Types

Cardinal numbers were introduced in terms of one–one correspondences between independently given sets. The continuum hypothesis seeks to equate the cardinality of one such independently given set with that of a set of ordinal numbers, a constructed set. But if the ordinal numbers are purely mental constructs there are no grounds for supposing that this equation can be *proved* as a mathematical theorem, any more than one would suppose that the number of planets can be mathematically proved to be nine. Only those properties of ordinal numbers which follow from the way in which they have been constructed could be expected to be provable, not anything about their relation to independently given sets. This raises the philosophically crucial question of whether it makes sense to suppose that there are mathematical truths which are not provable. Could the continuum hypothesis be true (or false) but not provably one or the other? If it were true but not provable (Cantor believed it to be true but could not prove it) what sort of grounds, if any, could we have for thinking it to be true? These are questions whose discussion is to be postponed until chapter 9, but which can never be very far away when exploring a newly created branch of mathematics (the study of transfinite numbers, the discipline of transfinite set theory, was created by Cantor and his successors, whether the object of their study was created by them or not).

But although Cantor first introduced the transfinite ordinals via principles of construction, he did not regard these numbers as *mere* products of mental construction. Their justification to the title 'number' requires more than generation in a sequence and more than definition of 'arithmetic' operations as purely symbolic manipulations. They must function something like numbers in that they can be applied and seen as giving a certain kind of information about the sets to which they are applied. In other words they have to be shown to have a use. This requires that the arithmetic operations be interpretable as operations applied to well-ordered sets. To make the parallel with cardinal numbers closer Cantor gives a definition, in terms of one–one correspondence, of what it is for two sets to have the same ordinal number.

Definition Two well-ordered sets A, ordered by $\leqslant_1$, and B ordered by $\leqslant_2$, have the same order type (or ordinal number) ($O(A, \leqslant_1) = O(B, \leqslant_2)$) if, and only if, there is a one–one, order preserving correspondence c between A and B, i.e. c is such that for every x, y in A, $x \leqslant_1 y$ iff $c(x) \leqslant_2 c(y)$.

Definition The segment A_a of an element a of a well-ordered set $\langle A, \leqslant\rangle$ is the subset of A which consists of those elements of A which precede a, i.e.

$$A_a = \{x : x \in A \ \& \ x < a\}$$

It can then be proved that a well-ordered set does not have the same order type as any of its segments, although in the case of an infinite set it is still possible for the whole set to have the same order type as one of its proper subsets (one which is not a segment). For example, take the set N of natural numbers given in their natural order. Any segment of N will be a finite set, the set of numbers less than n for some natural number n. Since every finite set has a greatest member and N does not, none of these segments is of the same order type as N. But the set of even numbers does have the same order type as N, since the mapping $f(n) = 2n$ from N to E is one–one and order preserving.

Definition If $\langle A, \leqslant_1\rangle$ and $\langle B, \leqslant_2\rangle$ are well-ordered sets then $O(A, \leqslant_1) < 0(B, \leqslant_2)$ if, and only if, A has the same order type as some segment of B, i.e. if, and only if, there is an element b of B, such that $O(A, \leqslant_1) = O(B_b, \leqslant_2)$

The comparability of well-ordered sets can then be proved, i.e. given any two well-ordered sets $\langle A, \leqslant_1\rangle$ and $\langle B, \leqslant_2\rangle$ either

$O(A, \leqslant_1) = O(B, \leqslant_2)$, or $O(A, \leqslant_1) < O(B, \leqslant_2)$, or $O(B, \leqslant_2) < O(A, \leqslant_2)$.

The ordinal numbers are themselves well-ordered and, in the finite case, the set of numbers which are less than n, $\{0, 1, 2, \ldots n-1\}$, itself contains n members. This suggests using the ordinal number sequence as a standard well-ordered set, a scale against which all others can be compared, and thereby assigned an ordinal number.

Definition A set M, well-ordered by $\leqslant$, has ordinal number α if, and only if, M ordered by $\leqslant$ has the same order type as the set of ordinal numbers less than α under their natural ordering.

Addition and multiplication of ordinal numbers can now be associated with operations on well-ordered sets.

Definition If $O(A, \leqslant_1) = \alpha$ and $O(B, \leqslant_2) = \beta$ and A and B are disjoint then $\alpha + \beta = O(\langle A \cup B\rangle)$, where $\langle A \cup B\rangle$ is the ordered union of A and B, i.e. $A \cup B$ ordered by the relation $\leqslant$ defined as follows: $x \leqslant y$ if, and only if (a) $x, y \in A \ \& \ x \leqslant_1 y$, or (b) $x, y \in B \ \& \ x \leqslant_2 y$, or (c) $x \in A \ \& \ y \in B$.

In other words $\alpha + \beta$ is the ordinal number of the well-ordered set which results from first running through all of A and following this by all of B in their given orders, and this definition holds whether A and B are sets of ordinal numbers or of objects of other kinds. Multiplication was defined in terms of repeated addition: $\alpha \cdot \beta = \alpha + \alpha + \alpha \ \ldots \ \beta$ times. Thus $\alpha \cdot \beta$ will be the ordinal number of the ordered union of a whole sequence, whose ordinal number is β, of disjoint sets, each with ordinal number α.

It is now also possible to give an alternative definition of cardinal number.

Definition The cardinal number C(X) of a set X is the least ordinal number α such that there is a one–one correspondence between X and $\{x: x$ is an ordinal number and $x < \alpha\}$.

This definition can replace the preceding one, on the assumption that every set can be well-ordered. If this assumption fails then either it has to be allowed that there are some sets which lack cardinal numbers or it has to be allowed that there are cardinal numbers which cannot be identified with ordinal numbers.

In this way the theory of infinite ordinal and cardinal numbers can be integrated, but this is done in such a way that the basis for further investigation lies in a study of sets. For questions about what numbers there are, and what their relations are, have been made dependent on questions concerning what sets exist and what are the relations between them. Even the generation of the ordinal number sequence is dependent on the power of the principle of limitation to mark off the ordinal numbers into sets, or classes, to which the second principle of generation can then be applied. So the question of what ordinal numbers there are depends, for its answer, on an answer to the question 'What classes of ordinal numbers are there?'

5 Set Theoretic Paradoxes

Moreover, this becomes a pressing question in the light of what has become known as the Burali–Forti paradox. If one supposes that there are no limitations on the formation of classes of ordinal numbers, then it must be the case that there is a class consisting of all the ordinal numbers. But if these numbers can be 'limited' to form a class, it is certainly a well-ordered class and so must itself have an ordinal number. The second principle of generation would assign it a new number, that which is the next greatest after all the ordinals But there can be no ordinal number greater than all the ordinals. Yet the ordinal number, Ω, of the set of all ordinals cannot be an ordinal number belonging to that set, for then it would be a well-ordered set which has the same ordinal number as a segment of itself, $\{x: x$ is an ordinal number & $x < \Omega\}$. But it can be proved that no well-ordered set can have the same ordinal number as a proper segment of itself. So if the theory of ordinal numbers is to be consistent, clearly the totality of ordinal numbers cannot be allowed to form a set or class to which the second principle of generation can be applied or to which the notion of ordinal number can be applied. Some more precise delimitation of the permissible 'limitations' of the ordinal number sequence is required. The second principle of generation is what allows entry into the transfinite domain; without it there would be no infinite ordinal numbers. But the vagueness inherent in it concerns what is to count as a defined definite succession of real integers; clearly not every candidate for being a definite succession of ordinal numbers will do.

The problem is not limited to the ordinal numbers, however. A very similar situation, exhibited by Cantor's paradox, occurs in the case of cardinal numbers. Cantor proved that for any set A, the cardinal number of P(A) is strictly greater than that of A, $C(P(A)) > C(A)$. This entails that there can be no greatest cardinal number, for given a set of no matter what cardinality, its power set will have a greater cardinality. But consider now the set U consisting of all objects (including classes). This is the greatest possible set, since it includes all other sets as subsets. Moreover for any two sets A and B, if A is a subset of B, $C(A) \leqslant C(B)$. So that every set must have a cardinal number less than or equal to that of U. Hence $C(U)$ is the greatest cardinal number, contradicting our previous conclusion. So again consistency would seem to require that the set U should not be treated as a totality to which the notion of number can be applied.

These paradoxes clearly pose a threat to the whole theory of transfinite numbers. At the very least their claim to mathematical legitimacy requires that use of these notions should not lead to contradictions. Even the most ardent proponent of the view that mathematics is a free creative activity recognizes the consistency constraint – the mathematician can create what realms he wants provided they are free from contradiction. In a pure dream world contradictions don't matter, but if mathematicians are to go in for giving proofs, and if their theories are to be used, then contradictions must be avoided. However, the appearance of contradictions in a new theory does not condemn it at once. It is quite possible that it has been incorrectly formulated, or that some inappropriate assumption has been made. In any new development, whether of a game, a piece of legislation, a computer program or a radar system, the prototype will need improving, modifying and generally tidying up before it will function properly.

Cantor's reaction to the paradoxes was to introduce a distinction between the ordinary infinite, which is, in his view, a proper domain of mathematical study and subject to the numerical methods proposed in his transfinite arithmetic, and the absolute infinite which is beyond all numbering, measuring and human reasoning. He thus effectively extends the notion of number so that it can now be applied to totalities which previously had to be regarded as being without number, but even so there still have to be some totalities which are without number. The members of any absolutely infinite collection are without either ordinal or cardinal number. But this still leaves a problem which is how to tell whether a given collection is absolutely infinite or just ordinarily infinite.

That the fundamental difficulty here lies not so much with the notion of number as with the notion of set to which it has been inseparably linked in Cantor's theory, is suggested by Russell's paradox which does not involve numbers at all, only sets. This paradox arises by modifying the proof (given on pp. 63–4) that there can be no one–one correspondence between a set A and its power set $P(A)$. Since sets (power sets in particular) can contain other sets as members, it seems sensible to ask whether a given set belongs to itself. For example the set of all sets containing more than two elements must belong to itself, whereas the set of sets containing less than two elements does not. Consider then the set R of all those sets which do not belong to themselves. R must either belong to itself or not. If R belongs to itself, then it satisfies its own defining conditions, i.e. it does not belong to itself. So R cannot belong to itself, but then it does satisfy its own defining condition and does belong to itself. So we are caught in a contradiction either way. A consistent set theory cannot then allow R to exist, for it is a contradictory set. But this raises the difficult question of the grounds on which R is to be excluded. It will not be enough simply to legislate R out of existence since there might well be other problematic sets waiting to surprise us. To be sure of getting a consistent set theory, something without which the theory of transfinite numbers can never be assured of consistency, there have to be general principles governing set existence. These principles

have either to be laid down or discovered by appeal to some more general considerations.

There are two slightly divergent concerns here. There is the concern of the mathematician to get a working theory, one about which he can be reasonably confident, and there is that of the philosopher worried about the nature and status of this mathematical activity. He wants to know not merely what principles will produce a workable set theory, but what sort of justification, if any, these principles can have, whether they are ultimate first principles or whether they can be justified by appeal to some more basic principles. A justification from principles known to be reliable would constitute a guarantee of consistency, but in the absence of any justification he will want to know what possible assurance there can be of consistency.

The mathematical approach to systematizing and rigorizing a body of unsystematized procedures is that established with Euclid, namely, axiomatization. Producing a minimal list of axioms or postulates about what exists and about what operations can be performed and then proceeding systematically to show that most of the previously used results, together with many others, can now be proved from these axioms. It is this approach which gives the question concerning the number of points on a line its present meaning and so this is what will be sketched in the next chapter. The more philosophical concerns, which have been raised here will be postponed until we have a clearer view of the mathematical situation.