Thus ([0, 1] − B) is a countable union of nowhere dense sets, a set of 'the first category' in the terminology introduced by Baire in *his* doctoral thesis *1899a* (see section 3.13). Baire proved that [0, 1] is not of the first category; hence B cannot be of the first category. Like [0, 1], it is of the second category.

4.7. *Conclusion*

By stressing the largeness of the set B, Schönflies seemed to be suggesting that B should not be regarded as negligible in measure, that a definition which implied such a conclusion was inappropriate. Others undoubtedly shared his sentiments. Indeed, we have seen that the idea that a dense set could have zero measure was contrary to the approach to the measure of sets adopted by Harnack, Cantor and many other mathematicians, and championed by Schönflies. Lebesgue's work really settled the issue over the most appropriate definition of measure, for he showed that a Borel-type measure is necessary—a necessary evil, perhaps, but nonetheless necessary. That is, the definition of the integral which accompanies Lebesgue's generalisation of Jordan's theory of measure (as explained at the beginning of section 4.5) is free from most of the defects of the Riemann integral, including those discussed in section 4.4. Thus if a uniformly bounded sequence of Lebesgue-integrable functions, $f_n(x)$, converges to a function $f(x)$ for each x in $[a, b]$, then f is Lebesgue-integrable and

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \int_a^b f_n(x)\,dx. \qquad (4.7.1)$$

And if a function $f(x)$ has a bounded derivative $f'(x)$ on $[a, b]$, then f' is always integrable in Lebesgue's sense and

$$\int_a^b f'(x)\,dx = f(b) - f(a). \qquad (4.7.2)$$

Lebesgue's signal achievement was the discovery that his generalisation of the integral possesses these and many other remarkable properties.[1] By creating his theory of integration Lebesgue had in effect confirmed Fourier's naïve belief that 'arbitrary functions' are not beyond the purview of mathematical analysis.

[1] For a more detailed historical analysis of Lebesgue's contributions see Hawkins *1970a*, chs. 5 and 6. An excellent exposition of Lebesgue's theory is given in Royden *1968a*, chs. 3–5.

Chapter 5

The Development of Cantorian Set Theory

Joseph W. Dauben

5.1. *Introduction*

This chapter explores the early development of set theory, in particular the contributions of the German mathematician Georg Cantor (1845–1918). Though he was joined by mathematicians in the 19th century like Riemann, Hankel, Harnack and du Bois Reymond (among others) in exploring the properties of point sets and their significance for mathematical analysis, Cantor's contributions were in many ways unique. His creation of transfinite numbers was controversial from the beginning, and his professional career was devoted to defending and to promoting his revolutionary work. Perhaps more than most branches of modern mathematics, set theory bears the special stamp of its originator's interests and personality. Thus the historical development of Cantorian set theory demonstrates how the abstract objectivity so often ascribed to scientific theory may be influenced by the character and interests of those who contribute most to its development. This is particularly true of so contentious a subject as the infinite in mathematics, for not only did Cantor have to face strong opposition from mathematicians, but also theologians and philosophers clung to traditions that refused to admit any ground to the actual infinite. In relentlessly supporting the validity of transfinite set theory, he promoted his research until its importance to virtually every branch of mathematics was recognised.

For convenience I shall usually cite particular passages in Cantor's writings from Zermelo's edition of his works (Cantor *Papers* in the bibliography). For more general studies of his life and work, see Fraenkel *1930a*, Meschkowski *1967a*, Grattan-Guinness *1971b* and Dauben *1977a* and *1979a*.

5.2. *The trigonometric background : irrational numbers and derived sets*

Though Cantor's *Dissertation* of 1867, written at the University of Berlin under the auspices of Kummer and Kronecker, was devoted to a difficult problem in number theory (as was his *Habilitationsschrift*, published in 1869), this was not the area which first stimulated his interest in set theory. Having left Berlin early in 1869 to become a *Privatdozent* at the University of Halle, he found that one of his senior colleagues there, Eduard Heine, was working on problems dealing with the theory of trigonometric series (compare sections 3.11 and 3.12). Heine recognised in Cantor a young mathematician of great promise, and encouraged him to take up a very important question in analysis : If a given, arbitrary function could be represented by a trigonometric series, was the representation unique ? Heine had managed to solve a part of the problem in *1870a* by assuming that the given function was almost everywhere continuous, and that the trigonometric series in question was also uniformly convergent almost everywhere. But Cantor was anxious to do away with such restrictions, and to establish the uniqueness theorem in the most general terms possible. (For more details on this work, see Dauben *1971a*.)

This Cantor did, though for his first proof in 1870 he found it necessary to assume that the trigonometric series in question was convergent for all values of x (*Papers*, 80–83). In 1871 he published a short note indicating that it was in fact possible to establish the theorem even if, for certain values of x, either the representation of the function or the convergence of the series could be given up, so long as the total number of such exceptional points remained finite (*Papers*, 84–86). Cantor's greatest achievement (with respect to the uniqueness theorem) came in his *1872a*, when he succeeded in showing that even an infinite number of exceptions might be permitted, so long as they were distributed in a specified way.

Wanting to present this last proof in as simple and as rigorous a way as possible, Cantor found that he had to develop a satisfactory theory of the real numbers in order to deal precisely with the infinite sets of exceptional points which he now had in mind. Criticising earlier approaches for assuming the *existence* of the irrationals as limits of infinite sequences of rationals used to 'define' them, he wanted to present a theory of the irrationals which in no way presupposed their existence. Beginning with the set of all rational numbers A, he introduced sequences of rationals : $a_1, a_2, \ldots, a_n, \ldots$. These sequences were further subject to the condition that, for whatever m, if n was taken large enough, $|a_{n+m}-a_n|<\epsilon$, for any rational number $\epsilon>0$, however small. If the sequence satisfied this condition, Cantor called it a 'fundamental sequence' and said that it had a definite limit b (*Papers*, 93, 186). This was to be taken as a convention to express, not that the sequence $\{a_n\}$ actually had the limit b, or that the number b was presupposed as the limit, but merely that with each such sequence $\{a_n\}$ a number b was associated with it.

Cantor then denoted the collection of numbers b associated with such infinite sequences $\{a_n\}$ by B. Two numbers b and b' defined by two fundamental sequences $\{a_\nu\}$ and $\{a_\nu'\}$ were said to be equal, $b=b'$, if $a_\nu-a_\nu'$ became very small as ν increased without limit. He also noted that by virtue of the fact that any constant sequence $\{a\}$ was a fundamental sequence, then a must be an element of B. Consequently, $A\subset B$, though the converse was clearly false (*Papers*, 93–94).

In an analogous fashion, Cantor considered infinite sequences of elements from B : $b_1, b_2, \ldots, b_n, \ldots$. With each fundamental sequence $\{b_n\}$ there was associated a number c. All such sequences generated from B constituted the domain C. He went on in this way to define higher-order domains from C. Proceeding through λ such constructions, he reached the domain L. Given an element l in this domain L, he called it a number, value or limit (he took these to be the same for his theory of real numbers) of the λ-th kind (*Papers*, 95–96).

Though Cantor had built up the real numbers B from the domain of rationals A, and had then gone on in similar fashion, using infinite sequences to define higher-order domains, he was now faced with the problem of identifying the real numbers so constructed with points of the real line. It was clear that every such point could be associated with one of his real numbers, but it was by no means obvious that to each of his real numbers in B a unique point of the linear continuum must correspond. Therefore, he invoked the axiom : 'To every real number a definite point of the straight line corresponds, whose coordinate is equal to that number' (*Papers*, 97). This identification was to be especially important in terms of Cantor's definition of derived sets of the first and second species, which required the concept of limit-point : 'Given a point set P, if an infinite number of points of the set P lie within every neighbourhood, however small, of a point p, then p is said to be a "limit-point" of the set P' (*Papers*, 98 ; note that p may be a limit-point of P, and yet not belong to the set P itself).

Given any point set P, Cantor noted that every point was either a limit-point of P, or it was not. The set of all limit-points of P was denoted P', and called 'the first derived set' of P. Just as he was able to generate from B an entire system of λ-domains, he did the same with P'. If P' were an infinite point set, then it gave rise to a second derived point set P'', and so on, until after taking successively n such derived sets, it was possible to produce the $(n+1)$-th derived set of P, $P^{(n+1)}$.

The case important for Cantor's extension of his uniqueness theorem was the one in which, after n repetitions, the derived point set $P^{(n)}$ consisted only of a finite number of points, thus making the extension to further (non-empty) derived sets impossible. He designated such sets as point sets of the *first species*, for which the derived set $P^{(n)} = \varnothing$ for some finite value of n. Were $P^{(n)} \neq \varnothing$ for any finite value of n, then P was said to be a point set of the *second species*. It was from these point sets of the second species that he would eventually produce his transfinite numbers, but for his uniqueness theorem of 1872 he was concerned only with point sets of the first species.

Cantor showed that such point sets existed by appealing to a point on the line whose abscissa was determined by a number of the ν-th kind (hence the need for his axiom). Working backwards, he took the sequence of numbers of the $(\nu-1)$-th kind determining ν, then the numbers of the $(\nu-2)$-th kind determining each of these, and so on, until eventually he reached an infinite set of rational numbers in the domain A. By taking the point set corresponding to this set of rationals, he had clearly produced a point set of the ν-th kind. It was then possible to establish the most general of his theorems concerning the uniqueness of representations by means of trigonometric series (*Papers*, 99):

> If an equation is of the form
>
> $$0 = \tfrac{1}{2}d_0 + \sum_{n=1}^{\infty} c_n \sin nx + d_n \cos nx$$
>
> for all values of x with the exception of those which correspond to the points in a [closed] interval $(0 \ldots 2\pi)$ of a given point set P of the ν-th kind, where ν is any whole number, then $d_0 = 0$, $c_n = d_n = 0$.

Cantor's achievement was impressive. Following the success of his application of first species sets to establishing the uniqueness theorem, even for infinite sets of exceptional points, he must have been intrigued by the reasons which might account for the validity of the result. His proof had insisted that the points of exception be distributed in a carefully specified way. Nevertheless, there could be infinitely many such points, which raised the question: How could one characterise the important difference between the rationals and the reals? The rationals were *dense* (between any two rationals there were always infinitely many others), but the set of all rationals was not continuous. It was perhaps natural to suspect that there were more irrationals than rationals, but what did that mean? Try as he might, he could find no reason to establish or to deny the denumerability of the reals.

5.3. *Non-denumerability of the real numbers, and the problem of dimension*

On 29 November 1873, a year after their first meeting in Switzerland, Cantor wrote to his friend Richard Dedekind. In his letter, Cantor posed the problem which his analysis of irrational numbers had directly suggested. Was it possible to correspond uniquely in a one–one fashion the collection of all natural numbers N with the set of all real numbers X of the continuum? He assumed that the answer was 'no'; but he had not been able to find a reason why this should be so, and he hoped that Dedekind might see a simple answer to the dilemma. But Dedekind replied that he could find no reason to prohibit any such correspondence (*Cantor/Dedekind*, 12–20).

Before the year was over, however, Cantor had discovered a valuable key to understanding the nature of continuity, and in 1874 he published an important theorem in Crelle's *Journal*: The set of all real numbers R cannot be corresponded in a unique, one–one fashion with the set of all natural numbers N. In other words, the set R is non-denumerable (*Papers*, 117). The proof ran as follows. Assuming that the real numbers ω were countable, it followed that they could be placed in a one–one correspondence with the natural numbers N:

$$\omega_1, \omega_2, \omega_3, \omega_4, \ldots, \omega_n, \ldots . \qquad (5.3.1)$$

Cantor then claimed that it was possible, given any closed interval $[a, b]$ in R, to find at least one real number $\eta \in R$ such that η failed to be listed as an element of (5.3.1). Assuming $a < b$, he picked the first two numbers from (5.3.1) which fell within the interval $[a, b]$. Denoted by a' and b' respectively, these were used to constitute another interval $[a', b']$. Proceeding analogously, he produced a sequence of nested intervals, reaching $[a^n, b^n]$, where a^n and b^n were the first two numbers from (5.3.1) lying within $[a^{n-1}, b^{n-1}]$. If the number of intervals thus constructed were finite, then at most only one more element from (5.3.1) could lie in $[a^n, b^n]$. It was easy in this case to conclude that a number η could be taken in this interval which was not listed in (5.3.1). Any real number in $[a^n, b^n]$ would suffice, so long as it was not the possible least number indexed in (5.3.1).

On the other hand, if the number of intervals $[a^n, b^n]$ were not finite, Cantor's argument shifted to alternatives in the limit. Since the sequence $a, a', \ldots, a^n, \ldots$ did not increase indefinitely, but was bounded within $[a, b]$, it had to assume an upper limit, which he denoted by a^∞. Similarly, the sequence $b, b', \ldots, b^n, \ldots$ was assigned the lower limit b^∞. Were $a^\infty < b^\infty$, then, as in the finite case, any real number $\eta \in (a^\infty, b^\infty)$ was sufficient to produce the necessary real number not listed in (5.3.1).

However, were $a^\infty = b^\infty$, he reasoned that $\eta = a^\infty = b^\infty$ could not be included as an element of (5.3.1). He designated η as ω_ρ. But ω_ρ, for sufficiently large index n, would be excluded from all intervals nested within $[a^n, b^n]$. Nevertheless, by virtue of the construction that he had given, η had to lie in every interval $[a^n, b^n]$, regardless of index. The contradiction established the proof: R was non-denumerable.

Cantor's proof, coupled with the fact that the set of all algebraic numbers was denumerable, provided an independent corroboration of Liouville's proof *1851a* that there were an infinite number of transcendental numbers in any given interval $[a, b]$ of reals. But this was hardly the most significant part of Cantor's conclusion. As he described it, without particular emphasis (*Papers*, 116):

> This theorem shows why sets of real numbers (for example, the entirety of real numbers $\geqslant 0$ but $\leqslant 1$) cannot be uniquely corresponded with the set of all natural numbers N. Thus I have found the clear difference between a so-called continuum and a set of the nature of the entirety of all algebraic numbers.

But, as he was to discover, the features distinguishing continua from other kinds of sets were not completely described by the fact that they were non-denumerable. Nevertheless, with the idea of denumerability and the existence of non-denumerable sets established, Cantor was now able to make some of his earlier ideas more precise. For example, though he had the basic idea for the transfinite numbers in the sequence of derived sets $P', P'', \ldots, P^{(n)}, \ldots$, the basis for any articulate distinction between $P^{(n)}$ and $P^{(\infty)}$ was lacking. There was no precise basis for defining the first transfinite number ∞ following all finite natural numbers n until it was clear that in fact there were sets much larger than N, sets that could not be counted or enumerated by the indices of natural numbers.

Cantor's next subject of research produced surprising and unexpected results. (They are discussed in more detail in Dauben *1974a*.) Shortly after his discovery that the real numbers were non-denumerable, he must have begun to search for other distinct powers of infinity greater than the power of the real numbers. Early in 1874 he wrote to Dedekind, posing a new but clearly related problem: 'Might it be possible to correspond a surface (a square, perhaps, including its boundaries) with a straight line (perhaps an interval with the inclusion of its endpoints) so that to each point of the surface, one point of the line corresponds, and conversely?' (*Cantor/Dedekind*, 20).

Cantor cautioned that the solution was one of great difficulty, though one might be tempted to say that the answer was clearly 'no', and, even more clearly, that a proof was superfluous. In fact, when he mentioned the same problem to friends while visiting Berlin in the spring of 1874, they were astonished at the seeming ridiculousness of the question (*Cantor/Dedekind*, 21).

More than three years passed before Cantor discovered a way to produce a one–one correspondence between lines and surfaces. Finally, in 1877, he wrote to Dedekind and explained that, contrary to prevailing mathematical opinion, the 'absurd' correspondence between lines and planes was not impossible. The discovery prompted one of his best-known remarks: 'I see it, but I don't believe it!' (*Cantor/Dedekind*, 34).

Although Cantor had originally constructed a one–one correspondence between the points of any ρ-dimensional space and the linear continuum, the basic idea of his proof can be expressed more easily. For the simplest case of the two-dimensional plane and the one-dimensional line, he took any point (x_1, x_2) in the plane and matched it with exactly one y of the line. He did so by considering the infinite decimal expansions

$$x_1 = \alpha_1, \alpha_2, \ldots, \alpha_\nu, \ldots, \tag{5.3.2}$$

$$x_2 = \beta_1, \beta_2, \ldots, \beta_\nu, \ldots. \tag{5.3.3}$$

The corresponding y under his mapping was then determined as follows:

$$y = \alpha_1, \beta_1, \alpha_2, \beta_2, \ldots, \alpha_\nu, \beta_\nu, \ldots. \tag{5.3.4}$$

Unfortunately, there was a difficulty which Dedekind explained in a letter to Cantor of 22 June 1877 (*Cantor/Dedekind*, 27–28). In order to avoid the representation of one and the same value x twice, the assumption had to be added that no representation be allowed which from a certain index on was always zero. Otherwise a number x would have two representations, for example: $x = 0{\cdot}3000\ldots$ and $x = 0{\cdot}2999\ldots$. The only exception to the above restriction would of course be the representation of zero itself. But under these conditions, Cantor's mapping was necessarily incomplete. Any y of the form

$$y = \alpha_1, \beta_1, \alpha_2, \beta_2, \ldots, \alpha_\nu, \beta_\nu, 0, \beta_{\nu+1}, 0, \beta_{\nu+2}, \ldots \tag{5.3.5}$$

was inadmissible under his assumptions, since it would have corresponded to the two points:

$$x_1 = \alpha_1, \alpha_2, \ldots, \alpha_\nu, 0, 0, 0, \ldots, \tag{5.3.6}$$

$$x_2 = \beta_1, \beta_2, \ldots, \beta_\nu, \ldots. \tag{5.3.7}$$

Fortunately, the damage was not irreparable, and he was soon able to find an alternative proof which, though more complex, nevertheless established the general validity of his theorem. It was possible to determine a one–one correspondence which mapped the points of the two-dimensional plane onto the one-dimensional line (*Papers*, 122–125).

Cantor's discovery was so startling because it was completely contrary to what mathematicians had believed for so long. He immediately criticised the work of others, particularly the work of Riemann and Helmholtz, who had assumed that the dimension of a space was uniquely determined by the number of coordinates needed to identify a point in that space. As Cantor had demonstrated, there was no such invariance. But, as Dedekind was quick to point out, Cantor's theorem involved a mapping that was *discontinuous*, and, as everyone had always assumed, dimension was invariant under continuous, one–one correspondences.[1] Above all, the value of Cantor's proof was the justification it gave him for narrowing his study of continuity to the linear continuum of real numbers. His next major publications presented a systematic study of linear point sets, and introduced his transfinite numbers as the keys to producing a general theory of infinite sets.

Though the paper published in 1878 was a triumph for Cantor, it was also the cause of some distress and unpleasantness. In fact, it was the first occasion for open hostilities between Cantor and one of his life-long opponents, his former teacher at the University of Berlin, Leopold Kronecker.

5.4. *First trouble with Kronecker*

The details of Kronecker's programme of arithmeticisation, basing all of mathematics on a finite number of operations involving only the integers, were outlined in his article 'Über den Zahlbegriff' ('On the number-concept': Kronecker *1886a*). Cantor, who had written both his *Dissertation* and *Habilitationsschrift* under Kronecker, could not have been unaware of his extreme position. In the early 1870s Kronecker was building his forces in opposition to such basic concepts as the Bolzano–Weierstrass theorem, upper and lower limits, and the irrational numbers. At one point he had even tried to persuade Heine not to publish his *1872a*, but Heine was not deterred.

As an editor of Crelle's *Journal*, Kronecker was in a position to refuse any article for publication, and by 1879 he was so appalled at the direction Cantor's work was taking that he did just that. Though Cantor had sent his manuscript on the subject of dimension to the editors of Crelle's *Journal* on 12 July 1877, it did not appear immediately.

[1] This assumption was supported shortly after Cantor's startling article by a number of mathematicians who offered proofs in a number of forms and with varying degrees of generality, but the first completely satisfactory argument that dimension was invariant appeared only in Brouwer *1911a*. Cantor's alleged proof of the invariance of dimension is discussed in Dauben *1975a*.

Despite the editors' promise to accept it, and Weierstrass's efforts to promote its appearance, no steps were taken to prepare the paper for press. Cantor, suspecting Kronecker's intervention, became so agitated over the matter that he wrote a bitter letter to Dedekind complaining about the treatment of his work, and raising the possibility of withdrawing it from the journal and asking Vieweg to publish it separately (*Cantor/Dedekind*, 40). Dedekind, acting upon his own experience in such matters, was able to convince Cantor that he should wait. As it turned out, Dedekind was right: Cantor's paper finally appeared in the volume for 1878. But he was so offended by the apparent reluctance of the editors to give his paper speedy notice that he refused to publish again in Crelle's *Journal*.

The delay in publication represented the first major conflict that Cantor was to experience with regard to the acceptance of his work. No longer could he believe that his differences with Kronecker were purely academic. He recognised the extent to which Kronecker would go to prevent the spread of his work. In the years to come, he fought vigorously against all those who refused to allow the completed infinite, the transfinite numbers of the new set theory, into the bounds of accepted mathematics.

5.5. *Descriptive theory of point sets*

Beginning in 1879, Cantor published a series of four papers dealing specifically with the theory of linear point sets. He returned to the concept of derived set that had proven so useful in his research on trigonometric series, and stressed that the analysis of the properties of derived sets would eventually reveal the properties of the continuum.

Some preliminaries were required, starting with this definition: If P lies partially or entirely in the interval $[a, b]$, then it can happen that every interval $[c, d]$ in $[a, b]$ contains points of P. In such a case the set P is said to be *everywhere-dense in the interval* $[a, b]$ (*Papers*, 140–141). Cantor immediately connected the ideas of derived and everywhere-dense sets. A set P was everywhere-dense in an interval $[a, b]$ whenever the first derived set P' of P contained $[a, b]$ itself. Furthermore, everywhere-dense sets were necessarily sets of the second species, while first species sets could never be everywhere-dense.

The second definition is this: Two sets M and N are said to be of the *same power* if to every element of M one element of N corresponds, and conversely, to every element of N one element of M corresponds (*Papers*, 141). Cantor singled out two cases: denumerable sets, whose power was that of the natural numbers N; and continuous, or non-denumerable sets, whose power was that of the real numbers R. Countably infinite sets included the natural numbers, the rational and algebraic numbers. All sets of the first species were also of this first, denumerable, kind. But he noted that the rational and algebraic numbers showed that everywhere-dense sets, and hence sets of the second species, could also be denumerable.

Like the derived sets $P^{(n)}$ of a given set P, the power of P was given intrinsically with the set P. Cantor explained the mathematical importance of the concept of power as follows (*Papers*, 150, 152):

> The *concept of power*, which includes as a special case the concept of whole number, that foundation of the theory of number, and which ought to be considered as the most general genuine origin of sets [Moment bei Mannigfaltigkeiten], is by no means restricted to linear point sets, but can be regarded as an attribute of any *well-defined* collection, whatever may be the character of its elements . . . *Set theory* in the conception used here, if we only consider mathematics for now and forget other applications, includes the areas of arithmetic, function theory and geometry. It contains them in terms of the concept of power and brings them all together in a higher unity. *Discontinuity* and *continuity* are similarly considered from the same point of view and are thus measured with the same measure.

Cantor's next article of 1880 was short. It continued the same brick-laying work of the 1879 article, and sought to reformulate old ideas in the context of linear point sets. It also introduced for the first time his transfinite numbers. Given a set of the second species P, he explained how the first derived set P' of P could then be given the disjoint decomposition

$$P' = \{Q, R\}, \qquad (5.5.1)$$

where Q was the set of all points belonging to first species sets of P', and R was the set of points contained in *every* derived set of P'. He defined this last property (intersection) by

$$R \underset{\mathrm{Df}}{=} \mathscr{D}(P', P'', \ldots). \qquad (5.5.2)$$

Since R was to consist of points belonging to *every* derived set of P', then it was equally true that

$$R = \mathscr{D}(P^{(2)}, P^{(3)}, \ldots), \ldots, \qquad (5.5.3)$$

$$R = \mathscr{D}(P^{(n)}, P^{(n+1)}, \ldots). \qquad (5.5.4)$$

Consequently, he felt justified in defining R, taken from P, as

$$R \underset{\mathrm{Df}}{=} P^{(\infty)}. \qquad (5.5.5)$$

$P^{(\infty)}$ was the derived set of P of order ∞ (*Papers*, 147). Assuming $P^{(\infty)} \neq \emptyset$, he denoted the first derived set of $P^{(\infty)}$ by $P^{(\infty+1)}$, the n-th by $P^{(\infty+n)}$. Continuing in this fashion, it was possible to generate derived sets of the following general form:

$$P^{(n_0\infty^{\nu}+n_1\infty^{\nu-1}+\ldots+n_{\nu})}.$$

By allowing ν to be taken as a variable, one could then produce 'an endless sequence of concepts', as he put it:

$$P^{(n\infty^{\infty})}, P^{(\infty^{\infty}+1)}, P^{(\infty^{\infty}+n)}, P^{(\infty^{n\infty})}, P^{(\infty^{\infty^{n}})}, P^{(\infty^{\infty^{\infty}})}, \ldots \qquad (5.5.6)$$

Cantor summarised the entire procedure in decisive terms: 'we see here a dialectic generation of concepts', he said, 'which always continues further and thus is free of any arbitrariness' (*Papers*, 148). He also took care to add (in a footnote which Zermelo failed to include in Cantor's *Papers*) that he first had the idea of the second species sets and their corresponding transfinite symbols (soon to be named his transfinite ordinal numbers) a decade earlier (Cantor *1880a*, 358). This note was doubtless in response to an accusation made in 1879, in which du Bois Reymond claimed priority in the matter of designating 'everywhere-dense' sets with his own terminology 'pantachisch'.[1] Cantor was anxious to make clear that his work certainly had been done earlier than that of du Bois Reymond. Thus his footnote was intended to underscore the origins of his own work on the subject in his paper on trigonometric series published in 1872. But until he came to terms with the metaphysical nature of the transfinites, he referred to them only as 'infinite symbols' (*Papers*, 160). The derived sets remained the focus of his research for another few years, and the transfinite numbers, the infinite symbols themselves, were taken only for useful tags by means of which derived sets could be distinguished and identified.

The importance of the infinite symbols, however, was demonstrated without delay. The paper of 1879 had left unanswered the question of second species sets and whether they were necessarily everywhere-dense in any given interval. The new infinite symbols made it easy to describe a procedure by which second species sets could be identified

[1] du Bois Reymond had alluded to everywhere-dense sets in 1875; he had named them explicitly in 1879, soon followed with direct reference to Cantor's own paper of 1879. He even suggested in not-so-oblique terms that Cantor was misappropriating ideas not entirely his own. Though the concept of everywhere-dense sets was certainly of long standing, du Bois Reymond clearly felt some claim to priority in designating the property such sets exhibited. He must have hoped his terminology, 'pantachisch', would become standard. Even more clearly, he felt Cantor posed a threat for undisputed recognition. See du Bois Reymond *1880a*, 127–128.

which consisted of a single point (*Papers*, 148). These sets, which Cantor grouped with sets of the first species, were clearly losing candidates in the search for a complete explanation of the nature of the continuum.

In pressing his study of point sets of the second species further, Cantor offered a number of definitions and theorems dealing with various kinds of sets. For example, a set P was said to be *isolated* if it contained none of its limit-points; in other words, when $\mathscr{D}(P, P') = \varnothing$ (*Papers*, 158). Then, given any set P, an isolated set Q resulted by simply removing $\mathscr{D}(P, P')$ from P. Thus

$$Q = P - \mathscr{D}(P, P'), \qquad (5.5.7)$$

and consequently

$$P = Q + \mathscr{D}(P, P'). \qquad (5.5.8)$$

(5.5.8) offered some immediate insights. Clearly, any set P could be considered as a disjoint combination of an isolated set Q and any set that was a divisor of P'. Since $P^{(n+1)} \subseteq P^{(n)}$, it followed that $P^{(n)} - P^{(n+1)}$ was *always* an isolated set. Two decompositions were of special importance, and led to some far-reaching conclusions:

$$P' = (P' - P'') + (P'' - P''') + \ldots + (P^{(n-1)} - P^{(n)}) + P^{(n)}; \qquad (5.5.9)$$

$$P' = (P' - P'') + (P'' - P''') + \ldots + (P^{(n-1)} - P^{(n)}) + (P^{(n)} - P^{(n+1)}) + \ldots + P^{(\infty)}. \qquad (5.5.10)$$

Following five theorems proving the denumerability of certain types of sets, Cantor offered a corollary which dealt with non-denumerable sets: 'If P is a non-denumerable point set, then $P^{(\alpha)}$ is also non-denumerable, whether α is a finite whole number, or if it is one of the infinite symbols' (*Papers*, 160).

The transfinite numbers were still merely symbols, employed only as indices for the sake of adding precision to the distinction between first and second species sets. But it was only a matter of months before Cantor was to alter his goals dramatically, abandoning this older view of the 'infinite symbols' to introduce the new transfinite numbers.

5.6. *The* Grundlagen: *transfinite ordinal numbers, their definitions and laws*

By the end of 1882, Cantor had finished a manuscript: *Grundlagen einer allgemeinen Mannichfaltigkeitslehre* ('Foundations of a general theory of sets': *1883a*), which outlined a defence of his new ideas in theological, philosophical and mathematical terms (*Papers*, 165–208). He was unusually anxious that its publication proceed rapidly. Towards

the end of the year, he wrote nearly every day to Felix Klein, editor of *Mathematische Annalen*, and urged the greatest speed possible.[1] He even visited the press in Leipzig himself, hoping to expedite the appearance of his defence, both mathematical and philosophical, of an entirely new mathematics. The *Grundlagen* established his place as the founder of set theory, and he subsequently overshadowed even his closest rivals like du Bois Reymond and Harnack. It was the beginning of something new, quite startling and profoundly original.

The major achievement of the *Grundlagen* was its presentation of the transfinite numbers as an autonomous and systematic extension of the real numbers. Cantor had reached the point in his research where no progress in set theory, no advances in his study of continuity, were possible without recourse to the transfinite numbers. His own mathematical future hinged in large measure upon the acceptance of the actual infinite by mathematicians.

Cantor admitted that his new ideas might seem risky, but he argued their simplicity and necessity in a straightforward way. With concern for the introduction of ideas previously foreign to mathematics, he suggested (*Papers*, 165) that:

> So daring as this may seem, I can express not only the hope but the firm conviction, that this extension will, in time, have to be regarded as a thoroughly simple, appropriate, and natural one. But I in no way hide from myself the fact that with this undertaking I place myself in a certain opposition to widespread views about the mathematical infinite and to frequently advanced opinions on the nature of number.

Cantor's first concern was to counter mathematicians sympathetic to Kronecker's finitism who might easily have refused to read further than the *Grundlagen*'s first paragraph. He began by explaining the distinction that had long been recognised between the potential and actual infinite. The former was used in mathematics as the very roots of the calculus. It involved essentially the idea of variation, of growing beyond any ascertainable bound, but a state never actually considered as completed or final. He also referred to such infinities as improper infinities. In contrast with these were proper or actual infinities.

The realisation that his transfinite numbers were equally as real mathematically as the finite whole numbers had only recently come to

[1] The correspondence between Cantor and Klein, largely unpublished, is preserved in the archives of the Niedersächsische Staats- und Universitätsbibliothek, Göttingen. In particular, see Cantor's letter to Klein, No. 429, 18 December 1882, and his letter No. 430, 20 December 1882. This correspondence is referred to again in section 5.10 below.

consciousness in Cantor's mind. This recognition represented significant progress: 'I will define the infinite real whole numbers in the following, to which I have been led over the past few years without realising that they were concrete numbers of real meaning' (*Papers*, 166).

In the *Grundlagen* Cantor explained how the sequence of natural numbers 1, 2, 3, . . . had its origin in the repeated addition of *units*. He called this process of defining finite ordinal numbers by the successive addition of units the *first principle of generation*. It was clear that the class of all finite whole numbers (I) had no largest element. Though it was incorrect to speak of a largest element for (I), he believed there was nothing improper in thinking of a new number ω which expressed the natural, regular order of the *entire* set (I). This new number ω, the first transfinite number, was the *first* number following the entire sequence of natural numbers ν. It was then possible to apply the first principle of generation to ω, and to produce additional transfinite ordinal numbers:

$$\omega,\ \omega+1,\ \omega+2,\ \ldots,\ \omega+\nu,\ \ldots . \qquad (5.6.1)$$

Again, since there was no largest element, one could imagine another number representing the entirety, in order, of numbers $\omega+\nu$. Denoting this entirety by 2ω, it was possible to continue further:

$$2\omega,\ 2\omega+1,\ 2\omega+2,\ \ldots,\ 2\omega+\nu,\ \ldots . \qquad (5.6.2)$$

(Later Cantor reversed the order of the terms in ordinal multiplication, so that, for example, '2ω' became '$\omega 2$'; this latter is the modern notation.)

In attempting to characterise this mode of generation, Cantor allowed that ω could be regarded as a limit towards which the natural numbers N increased monotonically but never reached. Lest the analogy seem entirely mistaken, he added that by this he meant only to emphasise the character of ω taken as the first whole number following next after *all* the numbers $n \in N$. The idea of ω as a limit served to satisfy its role as an ordinal, the smallest integer larger than any integer $n \in N$.

This then was the *second principle of generation*. Whenever a sequence of numbers could be considered as limitless in extent, new transfinite numbers could always be generated by positing the existence of some least number larger than any in the given sequence. Cantor expressed the essential feature of this second principle of generation in terms of its logical function (*Papers*, 196):

> I call it the *second principle of generation* of real whole numbers and define them more precisely: if any definite succession of defined whole real numbers exists, for which there is no largest, then a new number is created by means of this second principle of generation which is thought of as the *limit* of those numbers, that is, it is defined as the next number larger than all of them.

By successive application of the two principles it was always possible to produce new numbers, and always in a completely determined succession. In their most general formulation, such numbers could be given in the following form:

$$\nu_0\omega^{\mu}+\nu_1\omega^{\mu-1}+\ldots+\nu_{\mu}.$$

But by proceeding apparently without constraint, there seemed to be no end to the numbers of this second number-class. Were this the case, what distinctions could be drawn between the first and second classes? Cantor was able to add, however, a third principle which he called the *principle of limitation* ('Hemmungsprinzip'), and which was designed to produce natural breaks in the sequence of transfinite numbers. Consequently it was possible to place definite bounds upon the second number-class (II), and to distinguish it from the third and successively higher number-classes, with this definition (*Papers*, 197):

> We define therefore the second number-class (II) as *the collection of all numbers* α *(increasing in definite succession) which can be formed by means of the two principles of generation:*
>
> $$\omega,\ \omega+1,\ \ldots,\ \nu_0\omega^{\mu}+\nu_1\omega^{\mu-1}+\ldots+\nu_{\mu},\ \ldots,\ \omega^{\omega},\ \ldots,\ \alpha,\ \ldots,$$
>
> *with the condition that all numbers preceding* α *(from 1 on) constitute a set of power equivalent to the first number-class (I).*

In the *Grundlagen* he went on to establish that not only were the powers of the two number-classes (I) and (II) distinct, but that in fact the power of the second number-class (II) was the next larger after that of the first number-class (I) (*Papers*, 197–201).

An important advance made possible by the new numbers was the distinction that Cantor made between 'Zahl' and 'Anzahl'. Zahl, or Number, referred to the cardinal sense of the number of objects in a set without regard to the order in which elements occurred; Anzahl, or Numbering, took into consideration the order of elements. The difference was fundamental. For example, all of the following sets have the same cardinal number, they are equal in power, and are all denumerable. Nevertheless, their Numberings, their ordinal numbers, are different:

$$\left.\begin{aligned}
(a_1, a_2, \ldots, a_n, a_{n+1}, \ldots) &= \omega,\\
(a_2, a_3, \ldots, a_{n+1}, a_{n+2}, \ldots, a_1) &= \omega+1,\\
(a_3, a_4, \ldots, a_n, \ldots, a_1, a_2) &= \omega+2,\\
(a_1, a_3, a_5, \ldots;\ a_2, a_4, a_6, \ldots) &= \omega+\omega=2\omega.
\end{aligned}\right\} \qquad (5.6.3)$$

Once his transfinite numbers were defined, Cantor went on to describe their arithmetic and properties such as prime numbers among the transfinites (*Papers*, 201–204). Among the most significant characteristics of the transfinite ordinals was their non-commutativity. In general, $a+b \neq b+a$, nor did $ab=ba$ in all cases. For finite numbers, commutativity of operations was preserved, but not for transfinites. For example:

$$2+\omega=(1, 2, a_1, a_2, \ldots, a_n, \ldots) \neq (a_1, a_2, \ldots, a_n, \ldots, 1, 2)=\omega+2, \quad (5.6.4)$$

$$2\omega=(a_1, a_2, \ldots; b_1, b_2, \ldots) \neq (a_1, b_1, a_2, b_2, \ldots, a_n, b_n, \ldots)=\omega 2. \quad (5.6.5)$$

Two sets were defined to be of the same Numbering (that is, their corresponding ordinal numbers were equal) if they could be corresponded in a one–one fashion such that the order of elements was preserved in each case. In a similar way, the powers of two sets M and N were defined as equivalent if the elements of one set could be corresponded one–one with those of the other.[1]

The newly introduced distinction between Number and Numbering brought new insights to understanding the difference between finite and infinite sets. For finite sets, regardless of ordering, the Numbering of elements was always the same. Infinite sets were much more interesting because of the different Numberings one could find for sets of the same power. The Numbering of sets, therefore, was a concept totally dependent upon the order in which the elements of the set occurred. And there was a correlation between the Number of a set and the Numbering that its elements might produce, depending upon their arrangement: 'Every set of the power of the first class is denumerable by numbers of the second number-class and only by such numbers' (*Papers*, 169).

Though the difference between Number and Numbering was indistinct on finite sets, it helped to explain how the number concept functioned in a double sense, and why there had been confusion for centuries over potential and actual infinities. For finite sets, ordinal and cardinal numbers coincided. But because the two kinds of number were fundamentally different, Cantor could demonstrate the illegitimacy of trying to press properties of finite numbers onto infinite numbers. Once the ordinal/cardinal distinction had been recognised on transfinite

[1] Though Cantor did not identify the powers of infinite sets with transfinite cardinal numbers until after the *Grundlagen* had appeared, he later defined inequalities among cardinal numbers in full detail; see section 5.10 below, and *Papers*, 284–285. The somewhat more involved definitions for inequalities between transfinite ordinal numbers are discussed in section 5.11 below, and *Papers*, 320–325.

numbers, he was able to re-apply the same concepts to finite sets, and in the process find that it was another way to characterise the differences between finite and infinite domains: if a set were finite, then its cardinal and ordinal numbers were the same (*Papers*, 168–169).

5.7. *The continuum hypothesis and the topology of the real line*

One of the major goals of Cantor's transfinite set theory was the answer to a question that seemed quite simple, but one which to this day remains unanswered: What is the power of the continuum? This question, with an answer always believed to be the only possible solution, has come to be known as 'Cantor's continuum hypothesis': The power of the continuum is equivalent to that of the second number-class (II) (*Papers*, 192).

Though Cantor was never able to establish the truth of this conjecture, his *Grundlagen* did manage to make some progress in refining the mathematical description of continuous sets. One such advance was his description of the general conditions necessary and sufficient to constitute a continuum, which involved the idea of perfect sets: A set P is said to be ***perfect*** if it equals its derived set; in other words, if $P=P'$ (*Papers*, 193).

It was clear that continua had to be perfect sets, since the classic example, the set of all real numbers R, clearly equalled the set of its limit-points. But there was a further difficulty: perfect point sets were not necessarily everywhere-dense. As an example, Cantor offered his famous ternary set,[1] the set of all real numbers Z represented by

$$Z=\frac{c_1}{3}+\frac{c_2}{3^2}+\ldots+\frac{c_n}{3^n}+\ldots, \quad c_n=0 \text{ or } 2. \quad (5.7.1)$$

Though he could prove that perfect sets never had the power of the first number-class (I), his ternary set was everywhere-dense in no interval. Consequently, in addition to the fact that continua must be perfect sets, another definition was needed: 'T is a connected point set if, for any two points t and t' and for any arbitrarily small number ϵ, a finite number of point $t_1, t_2, t_3, t_4, \ldots, t_n$ can always be found in T, such that the distances $\overline{tt_1}, \overline{t_1t_2}, \overline{t_2t_3}, \ldots, \overline{t_nt'}$ are all less than ϵ' (*Papers*, 194). Thus the characteristic features of continua were identified: they were both perfect and connected. These, stated Cantor, were the necessary and ***sufficient*** conditions under which a point set could be

[1] Cantor introduced this set in note 11 of the *Grundlagen* (*Papers*, 207). In *Papers*, 235, he proved that it was of measure zero; compare section 4.5.

considered continuous. But there was a serious lacuna in the *Grundlagen* : the question of the power of the continuum was still unanswered. He intimated that he was hopeful a proof would be forthcoming, establishing his conjecture that the power of the continuum was none other than that of the second number-class (II).

The corollaries to such a solution would be numerous. It would immediately follow that all infinite point sets were either of the power of the first or second number-class, something Cantor had long claimed but never proven. It would also establish that the set of all functions of one or more variables represented by infinite series was necessarily equal in power to the second number-class. Likewise, the set of all analytic functions, and that of all functions represented by trigonometric series, would also be shown to be equivalent in power to that of the second number-class (II).

In a continuation of the *Grundlagen* published in 1884 Cantor developed a number of ideas which were related directly to the continuum hypothesis. He began by establishing a number of theorems concerning perfect sets and various kinds of derived sets. Characteristic of these theorems is the following : If P is a point set such that its first derived set P' is of power greater than the first, then ' there are always points which belong to all derived sets $P^{(\alpha)}$, where α is any number of (I) or (II), and the set of all these points, nothing other than the derived set $P^{(\Omega)}$, is always a perfect set ' (*Papers*, 221).

Next, Cantor established a theorem (Theorem E) of interest because it corrected an error in the *Grundlagen* that had been discovered by the Swedish mathematician Ivar Bendixson. Originally, Cantor had claimed that if P' were equal in power to the second number-class, then P' could be uniquely decomposed into two sets, $P'=R\cup S$; the set R he took to be *reducible*, meaning there was always some γ of (I) or (II) such that $R^{(\gamma)}=\emptyset$ (*Papers*, 193 and 222–223). This, in fact, was not true, as Bendixson pointed out.[1] S was a *perfect* set, meaning that for any number γ of (I) or (II), $S^{(\gamma)}=S$. Thus the question arose as to what properties distinguished the denumerable set R from other denumerable sets. This question was answered in one of Bendixson's theorems, as Cantor acknowledged : If R is the set of first power mentioned in Theorem E, then there is always a smallest number α of (I) or (II) such that $\mathcal{D}(R, R^{\alpha})=\emptyset$ (*Papers*, 224).

Cantor then turned his attention to another class of sets closely related to perfect sets : those he termed ' closed sets '. If a set con-

[1] Bendixson's letters to Cantor are kept in the archives of the Institut Mittag-Leffler, Djursholm, Sweden. These archives also contain the correspondence between Cantor and Mittag-Leffler, which is used occasionally in the rest of this chapter and which is mostly unpublished.

tained its first derived set, then it was said to be *closed*, namely :

$$\mathcal{D}(P, P')=P'. \qquad (5.7.2)$$

Every set might be closed by simply adding its first derived set P' : thus $P\cup P'$ was a closed set. Any set P could be decomposed into the sum of two sets, one set Q which was isolated and therefore denumerable, the second set P' which was closed : $P=Q\cup P'$ (*Papers*, 226–227). He went on to discuss the properties of sets dense-in-themselves (for which $\mathcal{D}(P, P')=P$), perfect sets, everywhere-dense sets, and the like, as well as inter-connections between them (*Papers*, 225–229). But he did not come any nearer to answering the question of the power of the continuum itself.

5.8. *Cantor's mental breakdown and non-mathematical interests*

Nothing caused Cantor greater annoyance than did Kronecker and his persistent attacks upon transfinite set theory. Cantor was especially angered by the fact that Kronecker refused to be open, preferring to save his most critical remarks for lectures and informal discussions with students. Kronecker thus carried on his polemic privately, or semi-publically in university, but never openly in print.

In early September 1883, Cantor learned that Kronecker was writing to the French mathematician Hermite and criticising Cantor's work as ' Humbug ' (Cantor to Mittag-Leffler, 5 May 1883). Shortly before Christmas Cantor wrote to Gösta Mittag-Leffler, for a time one of Cantor's closest friends, who was responsible as the founding editor of *Acta mathematica* for having nearly all of Cantor's early work on set theory published in French translation (Cantor *1883b*). Cantor confided that he had written to the Ministry of Education, hoping to annoy Kronecker by applying for a position in Berlin available the following spring. This was a direct expression of Cantor's lifelong belief that he deserved the honour of a position at one of two German universities known for their great mathematicians : either Göttingen or Berlin. But on 30 December 1883 he admitted to Mittag-Leffler that the application in Berlin would come to nothing. He had heard from Weierstrass that the obstacles were largely financial, owing to Kronecker's large salary. The letter to Mittag-Leffler was another occasion for Cantor to re-iterate the bitterness which he felt towards his position at Halle. The entire episode underscored the frustration and hostility which he felt in realizing that there was little he could ever do in the face of powerful opposition to improve his position.

If Kronecker was annoyed at Cantor's move, he returned the challenge masterfully. Early in January 1884 he wrote to Mittag-Leffler asking to publish in the *Acta mathematica* a short paper in which he would show 'that the results of modern function theory and set theory are of no real significance' (see Schönflies *1927a*, 5). At first Cantor was mildly receptive to the idea, believing that the article would at last bring Kronecker's opposition into the open, where it could be directly countered and presumably rejected. But Cantor began to have second thoughts. He feared that Kronecker might reduce his arguments to personal polemics. It all seemed as though Kronecker, by wanting to publish in the *Acta mathematica*, was trying to drive Cantor out of the one journal in which he had found a sympathetic editor, just as Kronecker years earlier had tried to prevent Cantor from publishing any work in Crelle's *Journal*.

Cantor threatened that Mittag-Leffler could expect him to withdraw his support for the journal in the years to follow should any polemical writings appear in the *Acta mathematica* under Kronecker's signature (Schönflies *1927a*, 5). Kronecker apparently never sent anything for the *Acta*, but the threats that Cantor was willing to make even to his friend Mittag-Leffler show how sensitive he could be to the conspiracy he felt was brewing against his work under the auspices of a single man: Leopold Kronecker.

While Cantor was becoming increasingly annoyed by the opposition to his work in Germany, there were other frustrations conspiring to upset his peace of mind in the early part of 1884. The continuum problem seemed as intransigent as ever, though he had reduced it to the problem of showing that perfect sets were equal in power to the second number-class (II).

Then, barely a month later, Cantor experienced his first mental breakdown (see Grattan-Guinness *1971b*, 356; and Peters *1961a*, 15, 27). It came upon him swiftly, unexpectedly, and apparently lasted somewhat more than a month. By the end of June he was sufficiently recovered to write to Mittag-Leffler, but complained that he lacked the energy and interest to return to rigorous mathematical thinking, and was content to take care of trifling administrative matters at the university. He felt capable of little more. But it was significant that he wrote to Mittag-Leffler saying that he was anxious to return to work, and would prefer his research to confining himself to the preparation of his lectures (see Schönflies *1927a*, 9). In fact, as soon as he had recovered sufficient strength, he set off for his favourite vacation resort in the Harz mountains and returned to his analysis of perfect sets. He also undertook the bold step of writing directly to Kronecker, and attempted to put their differences aside (see Meschkowski *1967a*, 237–241).

Less than a week after his letter of reconciliation to Kronecker,

Cantor wrote to Mittag-Leffler announcing at last an extraordinarily simple proof that the continuum was equal in power to the second number-class (II). The proof attempted to show that there were closed sets of the second power. Based upon straightforward decompositions and the fact that every perfect set was of power equal to that of the continuum, he was certain he had triumphed at last. He summarised the heart of his supposed proof in a single sentence: 'You see, therefore, everything comes down to defining a single closed set of the second power. When I have put everything in order, I will send you the details' (Meschkowski *1967a*, 243).

But on 20 October Cantor sent a lengthy letter to Mittag-Leffler announcing the complete failure of the new proof. On 14 November he again wrote saying that he had just found a rigorous proof that the continuum did *not* have the power of the second number-class, or of any number-class. He consoled himself by saying that 'so fatal an error, which one has held for so long, makes it an even greater advance to overcome it' (Schönflies *1927a*, 17). Nevertheless, within twenty-four hours he had decided that his latest proof was wrong, and that the continuum hypothesis was again an open question. It must have been embarrassing for him to have been compelled to reverse himself so often within so short a period of time. Even more discouraging must have been the realisation that the simplicity of the continuum hypothesis concealed difficulties of a high order that he was unable to resolve, despite his best efforts. However, he was not easily discouraged, and his continuing search for new methods and results marked the final and most devastating episode responsible for his disillusionment with mathematics and his discontent with colleagues both in Germany and abroad.

Cantor had developed a theory of simply ordered sets, sets for which given any two distinct elements a and b, either $a<b$ or $b<a$. He was confident that a systematic study of the types of simply ordered sets, in particular the rational and real numbers as given in their natural order, would make new advances possible (for details, see section 5.11 below). Consequently he prepared an article entitled 'Principien einer Theorie der Ordnungstypen' ('Principles of a theory of order-types': *1885a*), which he sent to Mittag-Leffler for the *Acta mathematica*. The manuscript was partly set in type and dated 21 February 1885, but not published until it appeared in Grattan-Guinness *1970b*.

The *Principien* was a remarkable paper. To deal with simply ordered sets like the rationals or reals, Cantor introduced new concepts, including those of coherent and adherent sets and their order-types. Designating the type of simply ordered set represented by the rationals as η, he produced a general theorem establishing the necessary and sufficient conditions for any set to be of type η. He also introduced θ as the simply ordered type represented by the continuum of all real numbers taken in their natural order on [0, 1].

The *Principien*, in addition to such technical advances, contained Cantor's first explicit statement that pure mathematics was nothing other than pure set theory. By this Cantor meant that the principles and results of set theory were so general and so penetrating that all of mathematics could be understood in terms that were essentially set-theoretic in nature (Cantor *1885a*, 84).

Despite the importance of his theory of simply ordered types, to Cantor's dismay Mittag-Leffler wrote on 9 March 1885 to suggest that the *Principien* be withdrawn from press. Mittag-Leffler was convinced that the publication of Cantor's newest research, before he had been able to obtain any positive results from it, would harm his reputation rather than advance it. He added that if Cantor's set theory came into discredit because of the *Principien*, it would take much longer for any of Cantor's work to win general acceptance among mathematicians. He even added that though Cantor's ideas concerning simply ordered sets might never be appreciated in his lifetime, it would perhaps be rediscovered by someone a hundred years later, and then Cantor would receive the credit he deserved! (Grattan-Guinness *1970b*, 102).

In recalling this episode more than a decade later, Cantor confided in Poincaré his real feelings about Mittag-Leffler's request that he not publish the *Principien* in the *Acta mathematica*: ' It was soon clear to me that he was doing this in the interest of his Acta Mathematica ' (*ibid.*, 105). Cantor was deeply hurt by Mittag-Leffler's rejection of his newest research. More than his polemic with Kronecker, more than his nervous breakdown or the trouble he was having in finding a proof that his continuum hypothesis was true, Mittag-Leffler's suggestion that he not print his work in the *Acta mathematica* seemed the cruellest blow of all. Though he never admitted that the incident affected in any way his personal regard and friendship for Mittag-Leffler, he wrote less frequently, and only seldom did he mention matters concerning his research. He felt as though the last mathematician at all sympathetic with his struggle to establish the transfinite numbers had abandoned him. He published only once more in the pages of the *Acta mathematica*, and that was a paper (*1885b*) which had already been accepted.

Instead, Cantor began to concentrate more and more upon problems of philosophy, theology, and the Bacon–Shakespeare controversy (see Grattan-Guinness *1971b*, esp. pp. 363–365; and Meschkowski *1967a*, 172, 264). Isolated and alone in Halle, he began to teach philosophy, and to correspond with theologians who provided a natural outlet for his need to communicate the importance and implications of his work. In turn, his contact with Catholic theologians may have made his own religious sympathies all the stronger. By the early part of 1884 he could write to Mittag-Leffler that he was not the creator of his new work, but merely a reporter (Schönflies *1927a*, 15–16). He was even more direct in a letter written to Hermite during the first month of 1894, in which he claimed that it was God's doing that had led him away from serious mathematics to concerns of theology and philosophy (Meschkowski *1965a*, 514–515):

> But now I thank God, the all-wise and all-good, that He forever denied me the fulfilment of this wish [for a position at university in either Göttingen or Berlin], for He thereby constrained me, through a deeper penetration into theology, to serve Him and His Holy Roman Catholic Church better than I would have been able to with my probably weak mathematical powers through an exclusive occupation with mathematics.

At one stroke Cantor signalled the many disappointments and doubts accumulated over more than two decades. His remarks reflected the frustration that he must have felt at being unable to solve the continuum hypothesis, and the disastrous effects which both the relentless attacks from Kronecker and Mittag-Leffler's response to his work on order-types had occasioned. Realising that no positions were ever going to be offered him in either Göttingen or Berlin, Cantor turned to other interests less demanding and more positively reinforcing. By the end of his life, in the spirit of Pope Leo XIII's encyclical *Aeterni patris*, he saw himself as the servant of God, a messenger or reporter who could use the mathematics he had been given to serve the Roman Catholic Church (see Dauben *1977a*). He firmly believed that ' for the first time Christian philosophy will learn from me the true theory of the infinite ' (Meschkowski *1965a*, 513). Convinced that he had been inspired and helped by God, Cantor was sure that his work was of consequence, despite the failure of mathematicians to understand the importance of his discoveries.

5.9. *Cantor's method of diagonalisation and the concept of coverings*

One of Cantor's most important projects in the late 1880s was the formation of a professional society for the promotion of mathematics in Germany. He viewed the idea as an alternative to the tradition-bound universities and the poorly-organised *Gesellschaft Deutscher Naturforscher und Aerzte* (' Society of German scientists and physicians '). Above all, he felt his own career had been greatly damaged by the premature and prejudiced rejection of his work by the prevailing establishment, led by Kronecker, and he hoped that an independent organisation would provide an open forum, one that would give younger mathematicians encouragement and a fair hearing of new, even radical, ideas.

The new society, the *Deutsche Mathematiker-Vereinigung* (' German Mathematicians' Union '), held its first meeting in 1891 in Halle, and elected Cantor its first president. (On the history of the Union, see especially Gutzmer *1904a* and Gericke *1966a*.) As his own contribution to the Union's first proceedings, he presented a theorem using a powerful new method to establish the existence of non-denumerable sets (*Papers*, 278–280). He thereby reconfirmed his earlier (1874) proof of the non-denumerability of the real numbers (see section 5.3 above), but he was able to go much further. He had made a fundamental advance of major significance for the future of transfinite set theory: ' This proof seems remarkable not only because of its great simplicity, but above all because the principle which it follows can be extended immediately to the general theorem, that the powers of well-defined sets have no maximum, or, what is the same, that every given aggregate L can be replaced by another M which is of greater power than is L ' (*Papers*, 279).

The proofs that Cantor offered in his paper of 1891 hinged on his new method of diagonalisation. Relying upon only two elements, m and w, he considered the collection M of elements $E=(x_1, x_2, \ldots, x_\nu, \ldots)$, where each x_ν was either m or w. As examples, he suggested:

$$E^{\mathrm{I}}=(m, m, m, m, \ldots), \tag{5.9.1}$$

$$E^{\mathrm{II}}=(w, w, w, w, \ldots), \tag{5.9.2}$$

$$E^{\mathrm{III}}=(m, w, m, w, \ldots). \tag{5.9.3}$$

He then claimed that the collection of all such elements M was non-denumerable: ' If $E_1, E_2, E_3, \ldots, E_\nu, \ldots$ is any simply infinite sequence of elements of the set M, then there is always an element E_0 of M which corresponds to no E_ν ' (*Papers*, 278). In his proof he first listed the elements of M, assumed to be denumerable:

$$\left.\begin{array}{l} E_1=(a_{11}, a_{12}, \ldots, a_{1\nu}, \ldots), \\ E_2=(a_{21}, a_{22}, \ldots, a_{2\nu}, \ldots), \\ \ldots\ldots\ldots\ldots \\ E_\mu=(a_{\mu 1}, a_{\mu 2}, \ldots, a_{\mu\nu}, \ldots), \\ \ldots\ldots\ldots\ldots \end{array}\right\} \tag{5.9.4}$$

Each element $a_{\mu,\nu}$ of the array (5.9.4) was taken to be either m or w. He then defined a new sequence $b_1, b_2, \ldots, b_\nu, \ldots$. Again, each

b_ν was either m or w, but determined so that $b_\nu \neq a_{\nu\nu}$. Then $E_0=(b_1, b_2, b_3, \ldots)$ was an element of M, but it was immediately clear that $E_0 \neq E_\nu$ for any value of the index ν. Using only two elements m and w, he had shown that from these alone a new set could be generated, and one of greater power. In fact, the method of diagonalisation provided him with an easy means of showing that the ascending sequence of powers of well-defined sets had no maximum. In other words, given any set L, it was always possible to produce from elements of L another set M which was necessarily of higher power than L itself.

Above all, Cantor had advanced significantly beyond the specific conclusion of his earlier paper of 1874, and had proved the existence of more than just one non-denumerable transfinite power without having to make any reference to irrational numbers or to the limits of infinite sequences. The comprehensiveness of his new method of diagonalisation made the paper of 1891 an important contribution to the development of set theory. As an example, he considered the linear continuum L, the set of all real numbers on $[0, 1]$, and the collection M of single-valued functions $f(x)$ which assumed only the values 0 and 1 for any value of x in $[0, 1]$, and proved that: The set M of single-valued functions $f(x)$ assuming only the values 0 and 1 on the interval $[0, 1]$ is greater in power than the set L of real numbers $x \in [0, 1]$ (*Papers*, 279–280; compare Fraenkel *1953a*, 64).

Similarly, the diagonalisation argument made it possible for Cantor to show that given any set, the set of all its sub-sets was always of a power greater than the parent set itself. Though he apparently did not realise it at the time, the means were now available whereby the continuum hypothesis itself could be given a direct algebraic formulation. But the discovery of how this could be done was not made until July 1895 (see section 5.10 below).

One last feature of Cantor's paper of 1891 deserves notice. Unlike the *Grundlagen*, where powers were never considered as numbers, he had come to see that powers actually represented the sole and necessary generalisation of the concept of cardinal number (*Papers*, 280). Thus regarded as powers, transfinite cardinal numbers should enjoy the same reality and definiteness as did the finite cardinals. The only difference involved the non-commutativity which distinguished transfinite numbers operationally from their finite counterparts. But in emphasising the fact that powers were to be regarded as necessary generalisations of the finite cardinal number concept, he was indicating a significant advance in his conceptualisation and representation of the basic principles of transfinite set theory.

5.10. *The* Beiträge : *transfinite alephs and simply ordered sets*

Cantor's last major publication was his *Beiträge zur Begründung der transfiniten Mengenlehre* (' Contributions to the founding of transfinite set theory '), issued in two parts as *1895a* and *1897a.*[1] In terms of new results, it made few startling advances. Where it was most innovative, it either improved the scope and presentation of procedures like multiplication or exponentiation of transfinite numbers, or it refined more successfully the details of conclusions already obtained. Above all, it introduced for the first time his special notation for the cardinal numbers, the alephs, where in particular $\aleph_0$ expressed the cardinality of denumerably infinite sets and was thus the smallest of the transfinite alephs.

Cantor opened the *Beiträge* with what has become a classic definition : ' By a " set " we understand any collection M of definite, distinct objects m of our perception or of our thought (which will be called the elements of M) into a whole ' (§ 1 ; *Papers*, 282). Powers, which he had come to regard since the *Grundlagen* as cardinal numbers, were defined in terms of the process of abstraction introduced in his *1887a* (*Papers*, 411–412). He now wrote : ' We call " power " or " cardinal number " of M that general concept which, with the help of our active thought-process, arises from the set M, abstracting from the character of its various elements m and from the order in which they occur ' (§ 1 ; *Papers*, 282).

Here Cantor's philosophical idealism was plain. Gottlob Frege later criticised him sharply, however, for depending upon such lax and unrigorous formulations. He particularly disliked Cantor's use of abstraction in defining both ordinal and cardinal numbers. Nevertheless, he believed that Cantor's theory could be salvaged from its uncritical presentation, and once opined that, basically, the results of transfinite set theory were sound, though its foundations required much more careful scrutiny.[2]

Equivalence between sets and their corresponding cardinal numbers was defined no differently in the *Beiträge* than in previous presentations

[1] Part I of the *Beiträge* was translated almost immediately into Italian as Cantor *1895b*. Both parts were translated into French as Cantor *1899a*. English-speaking readers had to wait until 1915 for a translation made by P. E. B. Jourdain (Cantor *1915a*). All references in this chapter to the *Beiträge* are made first to the appropriate section, and then to the corresponding pages of the *Beiträge* in Cantor's *Papers*.

[2] Frege criticized Cantor's idealism even before the *Beiträge* appeared (Frege *1892a*, 270). But he also made it clear that he thought the transfinite set theory could be given a rigorous foundation, one that would make it entirely acceptable mathematically (see Frege *1892a*, 272). For an even more critical evaluation of Cantor's use of abstractions, see Frege's unpublished draft version *1890a* of an earlier review of Cantor's work.

of Cantor's theory. One important question concerned the comparability of cardinals, for which he introduced this definition : Given $a=\overline{\overline{M}}$, $b=\overline{\overline{N}}$ (his way of denoting the double process of abstraction which produced cardinal numbers a and b), then if : (1) there is no proper subset $M' \subset M$ such that $M' \sim N$, and (2) there is a proper subset $N' \subset N$ such that $N' \sim M$, then it is to be said that $a<b$ or $b>a$ (§ 2 ; *Papers*, 284–285).

This was essentially the same definition of the order relation for cardinal numbers that Cantor had given as early as 1887. But he had gone further in 1887 and claimed that if M and N were two non-equivalent sets, then one always had to be equivalent to a proper subset of the other (*Papers*, 413). Moreover, he noted in consequence of his characterisation of order among the powers of sets that whenever two sets M and N could be mapped in a one–one fashion to proper subsets of each other, so that $M \sim N' \subset N$ and $N \sim M' \subset M$, then M and N were necessarily equivalent. This same theorem appeared as Theorem B of the *Beiträge*'s section 2, but it was one which he had never been able to prove directly himself. The theorem was later established independently by Felix Bernstein and E. Schröder, and has subsequently come to bear their names.[1] Cantor noted in the *Beiträge* that the Schröder–Bernstein Theorem, as well as three other theorems dealing with equivalence relations among sets, followed easily from a previously stated but *unproven* theorem, the trichotomy law : If a and b are any two cardinal numbers, then either $a=b$, $a>b$ or $b<a$ (§ 2 ; *Papers*, 285).

Cantor had already shown that, given any two cardinal numbers, only one of these order relations could hold, but he was unable to prove that *exactly* one always had to be true. The pernicious complication concerned the case of two cardinals represented by sets A and B, where A was assumed equivalent to no part of B and B was assumed equivalent to no part of A. He conjectured that this could only happen for finite sets where A and B were equivalent. But he was unable to show that the same could not occur on infinite sets. Consequently, there was no way he could establish the necessary comparability of all cardinal numbers, finite and infinite. The matter was critical, for if they were not comparable, it would be impossible to arrange all cardinal numbers in an ordered sequence, and in turn, it would be impossible to say whether one of two given cardinals was necessarily larger than the other. This was a grave matter for his continuum hypothesis, for if the power of the continuum were a non-comparable cardinal number, then he could never show that it was in fact equivalent to the power of the second number-class, which was a comparable cardinal by virtue

[1] For the papers by Bernstein and Schröder dealing with the equivalence of powers, see Schröder *1898a* and Bernstein *1898a*. See also Zermelo *1901a*, 34–38.

of its definition in terms of a well-ordered set. He assumed all along that every set could be well-ordered, and thus he could overlook a multitude of related difficulties while being able to assert, as a consequence, that all cardinal numbers were comparable. But he never published any proof of these claims, and chose not to raise the subject again in either part I of the *Beiträge* or its successor of 1897.

Like his definitions of order relations, rules for the addition and multiplication of any two cardinal numbers were first given in an earlier paper, Cantor's *Mitteilungen* (*1887a*), but now took this form: The union of two sets M and N, which have no elements in common, is denoted (M, N). The cardinal number of (M, N) depends only upon the cardinal numbers $a=\overline{\overline{M}}$ and $b=\overline{\overline{N}}$. This leads to the definition of the sum of a and b: $a+b=\overline{\overline{(M, N)}}$ (§ 3; *Papers*, 285–286). Since the power of the sets in question was independent of the order of the elements in either set or their union, the addition of cardinal numbers was commutative, and $a+b=b+a$. For multiplication, 'every element m of a set M may be joined with each element n of another set to form a new element (m, n); we denote the set of all these elements (m, n) by $(M . N)$, and call this the "union set of M and N". Thus $(M . N)=\{(m, n)\}$. If $a=\overline{\overline{M}}$ and $b=\overline{\overline{N}}$, then the product ab is defined as $ab=\overline{\overline{(M . N)}}$' (§ 3; *Papers*, 286).

There was a conspicuous difficulty with this definition. Products were defined for cardinal numbers in terms of their corresponding sets. These were combined to form a new set, a product set, but this was done for only two sets at a time. Thus the procedure could not be extended directly to include any more than finitely many products. But in 1895 Cantor explained how it was possible to represent the power of the continuum as an infinite exponentiation of the form $2^{\aleph_0}$. It was consequently of special importance to be able to define transfinite exponentiation. He was able to do so by a significant innovation: the idea of the covering of a set (§ 4; *Papers*, 287):

> By a 'covering of the set N with elements of the set M' or more simply, by a 'covering of N by M', we understand a rule by which, to each element n of N a definite element of M is corresponded, whereby one and the same element of M may be used repeatedly. The element corresponded with n from M is clearly a single-valued function of n and can therefore be designated $f(n)$; it is called the 'covering function of n', and the corresponding covering of N is called $f(N)$.

Drawing upon an example already used in the paper of 1891 to show that the set of single-valued functions was at least equivalent to the power of the continuum, Cantor suggested a covering by two elements of M, m_0 and m_1. If n_0 represented a particular element from N, then a covering function could be given such that $f(n_0)=m_0$, and $f(n)=m_1$ for all other values of n in N excepting the value n_0.

The set of all such covering functions served as the basis for defining the exponentiation of cardinal numbers. Specifically, the set of all different coverings of a set N by the set M produced a set with elements $f(N)$, which Cantor denoted $(N|M)$. This he called 'the covering set', and $(N|M)=\{f(N)\}$. Since the definition depended only upon the cardinal numbers $a=\overline{\overline{M}}$ and $b=\overline{\overline{N}}$, then the cardinal number of $(N|M)$ served to define the exponentiation:

$$a^b \underset{\mathrm{Df}}{=} \overline{\overline{(N|M)}}. \tag{5.10.1}$$

Cantor was very much pleased with the new results he could suddenly obtain with merely 'a few strokes of the pen' (*Papers*, 289), and he sent off word of his discovery on 19 July 1895 directly to Felix Klein, editor of *Mathematische Annalen*. By then the first part of the *Beiträge* was already in press, but he was determined to insert the new definition of exponentiation, and to explain its ramifications in a freshly written fourth section. The language of his letter to Klein appeared nearly word-for-word, without change, in the *Beiträge*, which still bore the date March 1895 though he had not by then included the very important conclusions made possible by the definition and rules of exponentiation for cardinal numbers. As he told Klein on 19 July 1895, 'from the following example one can see how fertile the simple formulae extended from the powers are'. The example offered was none other than an exact algebraic determination of the power of the continuum, expressed in terms of two other known cardinals of lesser degree. This was something that he had never managed to do previously.

Cantor recognised that the power of the linear continuum, denoted by $\mathfrak{o}$, could be represented as well by the set of all representations:

$$x=\frac{f(1)}{2}+\frac{f(2)}{2^2}+\ldots+\frac{f(\nu)}{2^\nu}+\ldots, \tag{5.10.2}$$

where $f(\nu)=0$ or 1; x represented numbers of the continuum $[0, 1]$ as given in the binary system (§ 4; *Papers*, 288–289). Though the numbers $x=(2\nu+1)/2^\mu<1$ were represented twice, they were the only numbers in $[0, 1]$ that failed to have a unique representation, and they could be discounted in the question of cardinality, as he showed, since they were only countably infinite in number. By denoting the set of elements with a double representation $\{s_\nu\}$, then

$$2^{\aleph_0}=\overline{\overline{(\{s_\nu\}, X)}}, \tag{5.10.3}$$

X representing the entire set of x given in (5.10.2). Removing from X any countable set $\{t_\nu\}$ and denoting the remainder by X_1, then

$$X = (\{t_\nu\}, X_1) = (\{t_{2\nu-1}\}, \{t_{2\nu}\}, X_1), \quad (5.10.4)$$

$$(\{s_\nu\}, X) = (\{s_\nu\}, \{t_\nu\}, X_1), \quad (5.10.5)$$

and since

$$\{t_{2\nu-1}\} \sim \{s_\nu\}, \{t_{2\nu}\} \sim \{t_\nu\}, X_1 \sim X_1, \quad (5.10.6)$$

then

$$X \sim (\{s_\nu\}, X), \text{ and thus } 2^{\aleph_0} = \overline{\overline{X}} = \mathfrak{o}. \quad (5.10.7)$$

For the first time, algebraically, Cantor had a firm grasp of what the power of the continuum must be. Even before he had introduced his new symbol '$\aleph_0$' for the smallest transfinite cardinal number, he was using it to show the light he hoped the exponentiation $2^{\aleph_0}$ would shed on his long-standing promise to establish the truth of his conjecture.

5.11. *Simply ordered sets and the continuum*

Cantor realised that, in order to describe completely the structure of the continuum, more refined means were necessary than appeals to well-ordered sets and their ordinal and cardinal numbers. Since no well-ordered set possessed the more interesting and essential properties of the continuum, in particular the property of everywhere-denseness, he turned to the study of simply ordered sets in order to advance his study of continuity. A set S for Cantor was *simply ordered* if there is some rule by which all its elements are ordered such that given any two, one can always be said to precede the other. Thus for any two elements m_1 and m_2 of a simply ordered set, either $m_1 \prec m_2$ or $m_1 \succ m_2$. If $m_1 \prec m_2$, $m_2 \prec m_3$, then it is always true that $m_1 \prec m_3$ (§ 7; *Papers*, 296). Different sets could be arranged with different ordinal properties. For example, the rationals in their natural order on the real line were everywhere-dense, though they could be arranged to form a denumerable sequence.

Cantor designated the order-type of any given set M by $\overline{M}$: 'Every ordered set M has a definite "order-type", or more briefly a definite "type", which we denote by $\overline{M}$; by this we understand the general concept which arises from M if we only abstract from the character of its elements, but retain the order in which the elements occur' (§ 7; *Papers*, 297). Two simply ordered sets were said to be *similar*, expressed $M \sim N$, if and only if there was a one–one correspondence between the two preserving the order in which elements occurred in both sets. Since two simply ordered sets could have the same order-type if and only if they were similar, it followed that $\overline{M} = \overline{N}$ and $M \sim N$, each implying the other. Moreover, since sets of equal type were always of equal cardinality, $M \sim N$ always implied that $\overline{\overline{M}} = \overline{\overline{N}}$, though the converse was clearly not true in general.

Cantor denoted ordinal numbers by lower case letters of the Greek alphabet. For example, the well-ordered set of natural numbers 1, 2, 3, ... was denoted as type ω. The simply ordered set of rational numbers in their natural order on [0, 1] was denoted by the order-type η, and the simply ordered set of real numbers on [0, 1] in their natural order by θ.[1]

In reversing the order of elements,[2] a new set was produced which Cantor denoted $^*M$, with order-type $^*\alpha$, assuming that $\alpha = \overline{M}$. For finite types it was always true that $^*\alpha = \alpha$. To make possible the combination of various order-types, arithmetic operations were also introduced. These followed the definitions given earlier in the *Grundlagen*, and included the familiar caveats concerning the non-commutative character of transfinite operations whenever order-types were concerned.

The set R of rational numbers within (0, 1) was a particularly fascinating one as a simply ordered type. Though it was countable, and therefore of cardinality $\aleph_0$, it produced distinctly different order-types depending upon how the precedence of elements might be taken. In characterising the properties peculiar to R, Cantor noted that it was denumerable. that there was neither a lowest nor a highest element, and that between any two elements of R there were infinitely many others belonging to the set. Thus R was everywhere-dense. These properties, he claimed, were the necessary and sufficient properties determining the order-type η, exemplified by the rationals (§ 9; *Papers*, 304).

Characterising the ordinal properties of continuous sets, however, posed a special problem. The special character of limit-points, a familiar component of continuous sets, also had to be translated into the language of order-types. This Cantor did, first by defining fundamental sequences, and then by introducing the concept of limit elements (§ 10; *Papers*, 307–308). Drawing upon the elements of his theory of real numbers, he reminded his readers that every fundamental sequence $\{x_\nu\}$ in the linear continuum X had a limit element x_0 in X, and conversely, every element of X was the limit element of some fundamental

[1] Cantor's first substantial effort to study the types of simply ordered sets was the *Principien (1885a)*. Cantor discussed the simply ordered type ω in the *Beiträge*, § 9 (*Papers*, 303–307), and the type θ in § 11 (*Papers*, 310–311).

[2] Cantor first introduced the concept of inverse order-types in the *Principien*, and later devoted part of section 7 of the *Beiträge* to the same material (*Papers*, 299).

sequence in X. Moreover, as the set of all real numbers in [0, 1], X contained as a subset the set of all rational numbers R of type η, and thus it was true that between any two elements x_0 and x_1 of X there were an infinite number of additional elements of X.

Collecting all of these properties together, Cantor claimed that quite apart from the specific example of X, such properties were both sufficient and necessary to characterise the ordinal type θ of any linearly continuous domain. He formulated the entire matter as follows (§ 11 ; *Papers*, 310):

> If a set M is so constituted that
> 1) It is 'perfect',
> 2) That there is a set S contained in M with cardinality $\bar{\bar{S}}=\aleph_0$, which is so related to M that between any two elements m_0 and m_1 of M elements of S occur,
> then [the order of M is of type θ,] $\bar{M}=\theta$.

This theorem brought the first part of Cantor's *Beiträge* to a close. It was two years before the sequel to Part I appeared in 1897. Sometime in 1895 he discovered the first of the paradoxes of set theory, specifically those involving the largest ordinal and cardinal numbers (see section 6.6). He had probably come upon them in the course of trying to establish his comparability theorem for transfinite cardinal numbers, and in the attempt to deal with the related questions of whether every transfinite power was necessarily an aleph, and whether every set could in fact be well-ordered. These problems may account for his delay in forwarding the second half of the *Beiträge* to Felix Klein for printing. They may also explain why Part I made few references to well-ordered sets and went no further than to introduce the first transfinite aleph, $\aleph_0$.

Perhaps Cantor believed he might soon be able to resolve the complications which the paradoxes of set theory seemed to raise. It may be that he even hoped to apply the new ideas introduced in Part I of the *Beiträge* to produce a proof that every set could be well-ordered. In turn this would settle both the theorem concerning comparability of the powers of all transfinite sets, as well as the question of whether every cardinal number was also an aleph. He may even have thought that the new algebraic formulations of the relations between cardinal numbers he had most recently discovered would enable him to solve the continuum hypothesis itself and to prove conclusively that $2^{\aleph_0}=\aleph_1$. But this was not to be.

5.12. *Well-ordered sets and ordinal numbers*

Part II of the *Beiträge* presented the bulk of Cantor's important theory of the transfinite ordinal and cardinal numbers. These had not featured with any prominence at all in Part I, but they now appeared in a detailed study which would carry his readers beyond $\aleph_0$ to the first of the non-denumerable, transfinite alephs.

The first step was to define the concept of well-ordered set. This was essential, for Cantor had already shown that sets could be ordered in many different ways, with diverse properties. But in arithmetic, as he had indicated as early as the *Grundlagen*, well-ordered sets were intrinsic to the process of counting by the successive addition of units. Moreover, he found in the concept of well-ordered sets the rigorous foundation for his transfinite numbers, something that his earlier 'principles of generation' had failed to offer in the *Grundlagen* (§ 12 ; *Papers*, 312):

> A simply ordered set F is said to be well-ordered, if its elements f *increase* from a lowest f_1 *in a definite succession*, so that the following two conditions are fulfilled:
> I. '*There is in F a smallest element f_1*'.
> II. '*If F' is any subset of F and if F contains one or more elements larger than all elements of F', then there exists in F an element f' which follows next after the entirety F', so that there is no element of F which falls between F' and f'*'.

Before Cantor could advance adequate definitions for his transfinite ordinal and cardinal numbers, he had to introduce the concept of sections or segments ('Abschnitte') of well-ordered sets: 'If f is any element of a well-ordered set F different from its first element f_1, then we call the set A of all elements of F which $\prec f$ a "section of F", and in fact the section of F determined by f. We call the set R of all other elements of F *including* f the "remainder of F" ...' (§ 13 ; *Papers*, 314). For example, given the well-ordered set $F=(a_1, a_2, \ldots ; b_1, b_2, \ldots ; c_1, c_2, c_3)$, a_3 determined the segment (a_1, a_2) and the remainder $(a_3, \ldots ; b_1, \ldots ; c_1, c_2, c_3)$; b_1 determined the segment $(a_1, \ldots)$ and the remainder $(b_1, \ldots ; c_1, c_2, c_3)$; c_2 determined the segment $(a_1, \ldots ; b_1, \ldots ; c_1)$ and the remainder (c_2, c_3).

Inequalities were introduced between segments as follows. Given two segments A and A' determined respectively by two elements f and f' of F, A' was said to be a segment of A if $f' \prec f$. In such cases, A' was said to be the lesser segment, written $A' < A$. Clearly, for every segment A of F, $A < F$. In terms of such segments, Cantor was able to state clearly the relations that were possible under similar correspondences between any two well-ordered sets F and G in the following form: If F and G are any two well-ordered sets, then either (1) F and G are similar to each other, (2) there is a definite segment B_1 of G which is similar to F, or (3) there is a definite segment A_1 of F which is similar to G, and each of these three cases excludes the possibility of the other two (§ 13; *Papers*, 319).

The results of this theorem were translated directly into important conclusions concerning the order of any two ordinal numbers in general, once these had been defined as follows: 'Every simply ordered set M has a definite *order-type* $\overline{M}$; it is the general concept which arises from M if the character [but not the order] . . . is abstracted from the elements of M. . . . We call the order-type of a well-ordered set F the "ordinal number" corresponding to it' (§ 14; *Papers*, 320–321).

Order-types, defined as the concept obtained from a well-ordered set by abstracting all individual properties of the elements while retaining their order, were represented as follows: $\alpha = \overline{F}$, $\beta = \overline{G}$. Given any two such sets F and G such that $\overline{F} = \alpha$ and $\overline{G} = \beta$, then the theorem above (concerning the relations possible between segments of similar sets) insured that only three mutually exclusive possibilities could occur. Either: (1) $F \sim G$, in which case $\alpha = \beta$; (2) G contained a definite segment B_1 such that $F \sim B_1$, then $\alpha < \beta$; or (3) There was a definite segment A_1 of F such that $G \sim A_1$: then $\alpha > \beta$.

Moreover, because of the comparability of segments, it followed immediately that if α and β were any two ordinal numbers, then exactly one of three possibilities was necessarily true: either $\alpha < \beta$, $\alpha = \beta$ or $\alpha > \beta$. The nature of segments also ensured that such relations were transitive: among any three ordinal numbers, if $\alpha < \beta$ and $\beta < \gamma$, then $\alpha < \gamma$, and it followed that the set of all ordinal numbers, taken in their order of magnitude, constituted a simply ordered set. Later Cantor showed that the set of all ordinal numbers was actually well-ordered, since every subset of the set of all ordinals had a least element and every element of the set had a definite and unique successor.

Cantor used the first number-class (I) of finite ordinals ν to determine the first transfinite cardinal number $\aleph_0$. However, to introduce the second transfinite cardinal $\aleph_1$, it was similarly necessary to establish securely the succession of transfinite ordinal numbers of the second number class. First he gave this definition: '*The second number-class* $Z(\aleph_0)$ *is the entirety* $\{\alpha\}$ *of all order-types* α *of well-ordered sets of the cardinality* $\aleph_0$' (§ 15; *Papers*, 325). He went on to prove that $Z(\aleph_0)$ was well-ordered, from which he could then define its cardinal number $\aleph_1$, and establish that $\aleph_0 < \aleph_1$.

Cantor ended the *Beiträge*, not with discussion of the cardinality

of the continuum and the solution of his continuum hypothesis that $2^{\aleph_0} = \aleph_1$, but instead with a detailed study of the arithmetic character of the numbers of the second number-class. At first he paid special attention to the numbers of $Z(\aleph_0)$ which could be expressed as polynomials in ω. In fact, such numbers could always be uniquely expressed in the form:

$$\phi = \omega^{\mu}\nu_0 + \omega^{\mu-1}\nu_1 + \ldots + \nu_{\mu}. \qquad (5.12.1)$$

Later Cantor was able to generalise such representations for ordinals of the second number-class without restricting the degree μ to finite values. No transfinite ordinals of the form ω^{ω}, for example, could be included rigorously in his transfinite number theory until he had established a satisfactory means of introducing the product of transfinitely many ordinal numbers. To do so, he invented the process of transfinite induction which was similar, but by no means the same as, the familiar mathematical (or complete) induction on well-ordered sets of type ω (§ 18; *Papers*, 336–339).

In the final section of the *Beiträge*, Cantor investigated the properties of a special kind of number, the 'epsilon numbers' ϵ of $Z(\aleph_0)$ for which $\omega^{\epsilon} = \epsilon$. These were central to his introduction of transfinite cardinals, since these were determined by those 'initial elements' of each number-class which could not be reached or produced by any arithmetic or exponential combination of elements preceding them. Initial elements, like ω and Ω, were transfinite numbers which were preceded by no numbers of equal power. Moreover, to every transfinite power there was only one such number, and every such 'first number' was necessarily an epsilon number. Thus there was a basic connection between the succession of transfinite alephs and the epsilon numbers which he introduced at the very end of the *Beiträge* (§ 20; *Papers*, 347–351).

Cantor's presentation of the principles of transfinite set theory in the *Beiträge* was elegant, but ultimately disappointing. One might have thought that, at long last, having given the extensive and rigorous foundations for the transfinite ordinal numbers of the second number-class, he would then have gone on to discuss the higher cardinal numbers in some detail. In particular, one might have expected him to fulfil his promise made in Part I to establish not only the entire succession of transfinite cardinal numbers $\aleph_0, \aleph_1, \ldots, \aleph_\nu, \ldots$, but to prove as well the existence of $\aleph_\omega$, and to show that in fact there was no end to the ever-increasing sequence of transfinite alephs. But instead, the final sections of the *Beiträge* were devoted to an analysis of the number-theoretic properties of transfinite ordinals. The entire manner of his handling of the transfinite cardinals in the *Beiträge* was fundamentally unsatisfying because it seemed so anti-climactic.

By the time that Cantor came to write the *Beiträge*, the continuum hypothesis seemed as elusive as ever, despite a tantalising hope that coverings, which led to the formulation $2^{\aleph_0} = \aleph_1$, might provide the key for which he had searched so long. But by 1897 he had discovered the paradoxes of set theory, he had failed to establish directly the comparability of all cardinal numbers, and he had not managed to find any proof that every set could be well-ordered (for details on all these matters, see sections 6.6 and 6.9). These obstacles seemed to leave no alternative: rather than produce the complete and absolutely certain solutions to the outstanding problems that his set theory had raised, he was forced to accept something less. Instead, he sought to present the elements and internal workings of his theory of transfinite sets as rigorously as his research completed since the appearance of the *Grundlagen* would allow. Abstract, and independent of point sets and physical examples, the *Beiträge* represented his last effort to present mathematicians with the basic features of his transfinite set theory. He hoped that at last the theory would speak for itself, and that its utility and interest would be acknowledged accordingly.

5.13. *Cantor's formalism and his rejection of infinitesimals*

Cantor always insisted that his transfinite numbers arose naturally and necessarily from the elements of *sets*, and thus he was convinced that his characterisation of the infinite was the *only* characterisation possible. This attitude was reflected in his eagerness to represent transfinite set theory as absolutely certain, complete, open neither to variant opinions nor to opposing interpretations. In this connection the work of the Italian mathematician G. Veronese was particularly unwelcome, because it advanced a theory of the infinite very different from Cantor's in a number of fundamental respects. Cantor devoted some of his most vituperative correspondence, as well as a portion of the *Beiträge*, to attacking what he described at one point at the 'infinitesimal Cholera bacillus of mathematics', which had spread from Germany through the work of Thomae, du Bois Reymond and Stolz, to infect Italian mathematics. Mostly at issue, but not exclusively, was the question of infinitesimals. (For more details, see Dauben *1977b*.)

Very early in his career Cantor had denied any role to infinitesimals in determining the nature of continuity, and by 1886 he had devised a proof that the existence of such entities was in his view impossible (*Papers*, 407–409). Thus any attempt to urge their legitimacy could be interpreted as a direct challenge to one of the most basic principles of Cantorian set theory, since it was in terms of the character of his transfinite numbers that he had argued the impossibility of infinitesimals. Moreover, any acceptance of infinitesimals necessarily meant that his own theory of number was incomplete. Thus to accept the work of Thomae, du Bois Reymond, Stolz and Veronese was to deny the perfection of Cantor's own creation. Understandably, Cantor launched a thorough campaign to discredit Veronese's work in every way possible. Veronese had just published a German translation of his *Fondamenti di geometria*, and Cantor felt it was timely to warn everyone of its manifold errors.[1] But the very nature of Cantor's own view of his set theory determined from the start the attack he would make, and the fate of its outcome.

Cantor underscored in the *Beiträge* that his concept of 'ordinal type', together with that of 'cardinal number' or 'power', included *everything* capable of being numbered that was thinkable (§ 7; *Papers*, 300). To him this meant that no further generalisations were conceivable. Moreover, there was nothing in the least way arbitrary about his definitions of number; cardinals and order-types were perfectly natural extensions of the number concept. Equally free from arbitrariness was his condition for the equality of two order-types, which was given in terms of their similarity. This condition followed with absolute necessity from the concept of order-type, and hence permitted no alteration. He claimed that Veronese's failure to understand this absolute character of the transfinite numbers was the major source of error in his misguided attempt to establish a different sort of infinite number in his *Fondamenti di geometria*. He could reject Veronese's definition of the equality of his 'numbers of ordered groups' by pointing out that it was viciously circular and therefore meaningless. To employ the concept 'not equal' in a definition of equality *presupposed* that one already knew what was meant by 'equality'. Thus the *petitio principii* rendered Veronese's entire approach suspect and mathematically unsound (§ 7; *Papers*, 301).

But there were more reasons for rejecting Veronese's theory of number, reasons which concerned the impossibility of results which he had obtained. Examining both Veronese's transfinite numbers and his infinitesimals, Cantor easily spotted 'erroneous' conclusions. Not surprisingly, his criticisms were based on the incompatibility of Veronese's conclusions with his own.

[1] Veronese *1891a*; the German translation by A. Schepp appeared three years later (Veronese *1894a*). Cantor told Killing in a letter of 3 June 1895, that it was appropriate to warn everyone of Veronese's errors; see page 156 in the second of his letter-books. (These were volumes in which he would draft his letters before mailing a final copy. Three survive, and all are now part of the Cantor *Nachlass*, recently given to the Göttingen Academy of Sciences. In the following text, *Letter-book* I refers to the book used by him between 1884 and 1888, and *Letter-book* II denotes the one used between 1890 and 1895. For further details, see Grattan-Guinness *1971b*, 348–349.)

Concerning Veronese's infinitely large numbers, Cantor once commented that as soon as he saw the equation $2 \cdot \infty = \infty \cdot 2$, he knew that the entire theory which Veronese had developed was false (*Letter-book* II, 165). Assuming the absolute character of his own theory of transfinite numbers, Cantor concluded that any theory of the infinite would have to be comparable with his; one major requirement was the non-commutativity of arithmetic operations for infinite ordinal numbers. Since Veronese's numbers clearly violated such necessary laws, they were inadmissable. To Cantor's way of thinking, it was as simple as that.

Infinitesimals, at the opposite extreme, were equally unwelcome, and were high on Cantor's list of mathematical 'ghosts and chimareas' (*Letter-book* II, 30, 138). His proof of the self-contradictory nature of the idea of infinitesimals was based upon a property he regarded as common to all finite numbers, expressed by the Axiom of Archimedes: If a and b are any two positive numbers, then there exists a positive integer n such that $na > b$ (*Papers*, 408–409). Basically, he refused to regard this as an axiom at all, but argued that it followed directly from the concept of linear number (*Letter-book* I, 96; II, 16, 137; *Papers*, 409). Numbers were *linear* if finitely or infinitely many of them could be added producing yet another linear magnitude. But Cantor assumed all along what Hilbert in his *Grundlagen der Geometrie* (*1899a*) was to call the Axiom of Continuity, and consequently all of his assertions followed directly. In particular, since the Axiom of Continuity and the Axiom of Archimedes implied each other, one could be derived from the other, and from Cantor's view there was no difficulty in asserting that the Archimedean 'Axiom' could be proven. In terms of his assumptions, it was as provable as the non-existence of infinitesimals. Moreover, had he agreed that the Archimedean property of the real numbers was merely axiomatic, then there was no reason to prevent the development of number systems by merely denying the axiom, so long as consistency was still preserved. But to have allowed this would have left him open to the challenge that, if infinitesimals could be produced without contradiction, then his own theory of number would have been contravened.

On the other hand, were the Axiom of Archimedes not an 'Axiom' at all, but a theorem which could be proven from other accepted principles, then Cantor could rest assured that it was impossible simply to deny the proposition and produce a consistent theory of infinitesimals. He was so persuasive, in fact, with his disavowal of infinitesimals that he was able to convince Peano, who wrote an article *1892a* on the subject in his *Rivista di matematica*. Bertrand Russell went even further than Peano, and argued in *The principles of mathematics* that mathematicians, completely understanding the nature of real numbers, could safely conclude that the non-existence of infinitesimals was firmly established. He was wise to add, however, that if it were ever possible to speak of infinitesimal numbers, it would have to be in some radically new sense (Russell *1903a*, 334–337).

Finally, there was an additional argument, one that Cantor found equally persuasive in rejecting the attempts Thomae, du Bois Reymond, Stolz and Veronese had made to develop logically sound theories of the infinitesimal. Once more his reasons were based on his view of set as the ultimate origin of any concept of number. In writing to Veronese on the subject, Cantor accused the infinitesimalists of talking nonsense, since in the realm of the possible there were no infinitely small entities. He stressed that his transfinite numbers were linked with *real ideas* produced directly from *sets*, and he challenged Veronese to show any *real ideas* corresponding to the supposed infinitesimals (*Letter-book* II, 15). Until Veronese could do so, Cantor insisted that deviation from the 'Axiom' of Archimedes, which he took as proven, was an error of the greatest seriousness.

5.14. *Conclusion*

Following publication of the *Beiträge* and its translation almost immediately into Italian and French, Cantor's ideas became widely known and were circulated among mathematicians of the highest rank the world over. The value of his transfinite set theory was recognised almost immediately, and soon his ideas were fuelling heated polemics between widely divided camps of mathematical opinion. Though he never seemed able to avoid controversy over the nature of his work, he was, after 1895, increasingly defended by younger and more energetic mathematicians. No longer was he left to face the opposition alone. Though Kronecker had died in 1891, he was replaced in the phalanx of dissenters by mathematicians like Poincaré; but Cantor could begin to count an ever-increasing, always impressive array of those ready to join in support of set theory. For him, the crusade was nearly over, and though the difficulties were by no means satisfactorily resolved, it was widely recognised at last that Cantor had contributed something of lasting significance to the world.

From the Calculus to Set Theory, 1630–1910

An Introductory History

Edited and with an introduction by

I. Grattan-Guinness

with chapters by

H. J. M. Bos
R. Bunn
J. W. Dauben
I. Grattan-Guinness
T. W. Hawkins
K. Møller Pedersen

Princeton University Press
Princeton and Oxford