however, concluded in Section 21 that this phenomenon was indeed intrinsic to infinite sets (it in fact leads to one possible definition of infinite set) and a more elaborate quantitative comparison of infinite sets is not precluded. In fact, a good part of the book elaborates on the apparently paradoxical relationship of infinite sets with proper subsets, and shows that it must be accorded a central role in the whole theory. Just such a comparison of sizes of infinite sets was undertaken by Georg Cantor, with considerable mathematical rigor.

Exercise 2.11: In Example 1 of §20, Bolzano shows that the set of real numbers y between 0 and 12 can be put in one-to-one correspondence with a proper subset of itself, namely the numbers x between 0 and 5, by means of the equation

$$5y = 12x.$$

Verify that $25y = 12x^2$ and $5y = 12(5 - x)$ will also give such one-to-one correspondences. Can you give additional examples?

Exercise 2.12: Show that the points of a 1 inch by 1 inch square can be put into one-to-one correspondence with the points of a 2 inch by 2 inch square and with the points of a 2 inch by 1 inch rectangle.

Exercise 2.13: Give additional examples of sets that may be put in one-to-one correspondence with proper subsets of themselves.

2.3 Cantor's Infinite Numbers

On the occasion of Georg Cantor's confirmation, his father wrote a letter to him that Cantor would always remember. It foreshadowed much that was to happen to Cantor in later years.

> No one knows beforehand into what unbelievably difficult conditions and occupational circumstances he will fall by chance, against what unforeseen and unforeseeable calamities and difficulties he will have to fight in the various situations of life.
>
> How often the most promising individuals are defeated after a tenuous, weak resistance in their first serious struggle following their entry into practical affairs. Their courage broken, they atrophy completely thereafter, and even in the best case they will still be nothing more than a so-called ruined genius!...
>
> But they lacked that steady heart, upon which everything depends! Now, my dear son! Believe me, your sincerest, truest and most experienced friend—this sure heart, which must live in us, is: a truly religious spirit....

PHOTO 2.4. Cantor.

> But in order to prevent as well all those other hardships and difficulties which inevitably rise against us through jealousy and slander of open or secret enemies in our eager aspiration for success in the activity of our own specialty or business; in order to combat these with success one needs above all to acquire and to appropriate the greatest amount possible of the most basic, diverse technical knowledge and skills. Nowadays these are an absolute necessity if the industrious and ambitious man does not want to see himself pushed aside by his enemies and forced to stand in the second or third rank [36, pp. 274 ff.].

By the time Cantor encountered his own great "unforeseen calamities and difficulties" he had acquired that religious spirit and steady heart that his father believed to be so indispensable.

Having been irresistably drawn to mathematics, Cantor excelled in it as a student, without neglecting diversity in learning, as his father had admonished him. In 1869, at age 22, three years after he completed his dissertation at the University of Berlin, he joined the University of Halle as a lecturer. Through his colleague Eduard Heine (1821–1881), Cantor became interested in analysis, and soon drew the attention of important researchers, most notably Leopold Kronecker (1823–1891), a professor at the University of Berlin and a very influential figure in German mathematics, both mathematically as well as politically. Kronecker held very strong views about the admissibility of certain mathematical techniques and approaches. To him is attributed the statement "God Himself made the whole numbers—everything else is the work of men." As a consequence, he disliked the style of analysis done by Weierstrass, involving a concept of the real numbers that was not based on whole numbers, and which was not constructive to a high degree. As soon as Cantor began to pursue his set-theoretic research, he lost the support of Kronecker, who quickly became one of his strongest critics. Cantor's desire to obtain a position in the mathematically much more stimulating environment of Berlin never materialized, and Kronecker's opposition had much to do with it.

The series of six articles on infinite linear sets that Cantor published between 1878 and 1884 contained in essence the whole of his set-theoretic discoveries. His creative period was close to an end. Shortly thereafter, in the spring of 1884, he suffered his first mental breakdown. But by the fall he was back working on the Continuum Hypothesis. An important supporter of Cantor in those years was the Swedish mathematician Gösta Mittag-Leffler (1846–1927). A student of Weierstrass, Mittag-Leffler became professor at the University of Stockholm in 1881. Shortly afterwards he founded the new mathematics journal *Acta Mathematica*, which quickly became very influential, with articles by the most eminent researchers. His relationship with Cantor was close, and Mittag-Leffler offered to publish Cantor's papers when other journals were rather reluctant to do so. In 1885, Cantor was about to publish two short articles in the *Acta Mathematica* containing some new ideas on ordinal numbers. While the first article was being typeset, Cantor received the following letter from Mittag-Leffler, requesting that Cantor withdraw the article:

> I am convinced that the publication of your new work, before you have been able to explain new positive results, will greatly damage your reputation among mathematicians. I know very well that basically this is all the same to you. But if your theory is once discredited in this way, it will be a long time before it will again command the attention of the mathematical world. It may well be that you and your theory will never be given the justice you deserve in your lifetime. Then the theory will be rediscovered in a hundred years or so by someone else, and then it will subsequently be found out that you already had it all. Then, at least, you will be given justice. But in this way [by publishing the article], you will exercise no significant influence, which you naturally desire as does everyone who carries out scientific research [36, p. 138].

The effect on Cantor was devastating. Disillusioned and without friends in the mathematical community, he decided by the end of 1885 to abandon mathematics. Instead, he turned to philosophy, theology, and history. To his surprise, he found more support for his work among theologians than he had ever received from mathematicians. Soon, however, he was back at work on an extension and generalization of his earlier work. As his last major set-theoretic work, he published a grand summary of his set theory between 1895 and 1897, his *Beiträge zur Begründung der Transfiniten Mengenlehre* (Contributions to the Founding of Transfinite Set Theory) [28, 282–351]. The following excerpt from this work concerns the arithmetic of cardinal numbers. The English translation is taken from [26], pp. 85–97, 103–108.

In the wake of this publication, his work was finally recognized for its full significance, and while controversy continued to surround set theory, Cantor was not alone anymore to defend it. Set theory was here to stay. His mathematical career had reached its zenith, and he was able to see that he had indeed created something of lasting significance.

Georg Cantor, from *Contributions to the Founding of the Theory of Transfinite Numbers*

§1
The Conception of Power or Cardinal Number

By an "aggregate"[4] we are to understand any collection into a whole M of definite and separate objects m of our intuition or our thought. These objects are called the "elements" of M.

In signs we express this thus:

(1) $$M = \{m\}.$$

We denote the uniting of many aggregates $M, N, P, \ldots$, which have no common elements, into a single aggregate by

(2) $$(M, N, P, \ldots).$$

The elements of this aggregate are, therefore, the elements of M, of N, or $P, \ldots$, taken together.

We will call by the name "part" or "partial aggregate" of an aggregate M any other aggregate M_1 whose elements are also elements of M.

If M_2 is a part of M_1 and M_1 is a part of M, then M_2 is part of M.

Every aggregate M has a definite "power," which we will also call its "cardinal number."

We will call by the name "power" or "cardinal number" of M the general concept which, by means of our active faculty of thought, arises from the

[4] The modern terminology is "set."

aggregate M when we make abstraction of the nature of its various elements m and of the order in which they are given.

We denote the result of this double act of abstraction, the cardinal number or power of M, by

$$\overline{\overline{M}}. \tag{3}$$

Since every single element m, if we abstract from its nature, becomes a "unit," the cardinal number $\overline{\overline{M}}$ is a definite aggregate composed of units, and this number has existence in our mind as an intellectual image or projection of the given aggregate M.

We say that two aggregates M and N are "equivalent," in signs

$$M \sim N \text{ or } N \sim M, \tag{4}$$

if it is possible to put them, by some law, in such a relation to one another that to every element of each one of them corresponds one and only one element of the other. To every part M_1 of M there corresponds, then, a definite equivalent part N_1 of N, and inversely.

If we have such a law of co-ordination of two equivalent aggregates, then, apart from the case when each of them consists only of one element, we can modify this law in many ways. We can, for instance, always take care that to a special element m_0 of M a special element n_0 of N corresponds. For if, according to the original law, the elements m_0 and n_0 do not correspond to one another, but to the element m_0 of M the element n_1 of N corresponds, and to the element n_0 of N the element m_1 of M corresponds, we take the modified law according to which m_0 corresponds to n_0 and m_1 to n_1 and for the other elements the original law remains unaltered. By this means the end is attained.

Every aggregate is equivalent to itself:

$$M \sim M. \tag{5}$$

If two aggregates are equivalent to a third, they are equivalent to one another; that is to say:

$$\text{from } M \sim P \text{ and } N \sim P \text{ follows } M \sim N. \tag{6}$$

Of fundamental importance is the theorem that two aggregates M and N have the same cardinal number if, and only if, they are equivalent: thus,

$$\text{from } M \sim N \text{ we get } \overline{\overline{M}} = \overline{\overline{N}}, \tag{7}$$

and

$$\text{from } \overline{\overline{M}} = \overline{\overline{N}}, \text{ we get } M \sim N. \tag{8}$$

Thus the equivalence of aggregates forms the necessary and sufficient condition for the equality of their cardinal numbers.

In fact, according to the above definition of power, the cardinal number $\overline{\overline{M}}$ remains unaltered if in the place of each of one or many or even all elements

m of M other things are substituted. If, now, $M \sim N$, there is a law of co-ordination by means of which M and N are uniquely and reciprocally referred to one another; and by it to the element m of M corresponds the element n of N. Then we can imagine, in the place of every element m of M, the corresponding element n of N substituted, and, in this way, M transforms into N without alteration of cardinal number. Consequently

$$\overline{\overline{M}} = \overline{\overline{N}}.$$

The converse of the theorem results from the remark that between the elements of M and the different units of its cardinal number $\overline{\overline{M}}$ a reciprocally univocal (or bi-univocal) relation of correspondence subsists.[5] For, as we saw, $\overline{\overline{M}}$ grows, so to speak, out of M in such a way that from every element m of M a special unit of $\overline{\overline{M}}$ arises. Thus we can say that

$$M \sim \overline{\overline{M}}. \tag{9}$$

In the same way $N \sim \overline{\overline{N}}$. If then $\overline{\overline{M}} = \overline{\overline{N}}$, we have, by (6), $M \sim N$.

We will mention the following theorem, which results immediately from the conception of equivalence. If $M, N, P, \dots$ are aggregates which have no common elements, $M', N', P', \dots$ are also aggregates with the same property, and if

$$M \sim M',\ N \sim N',\ P \sim P',\ \dots\ ,$$

then we always have

$$(M, N, P, \dots) \sim (M', N', P', \dots).$$

Early on in this section Cantor defines the concept of *cardinal number* of an aggregate (or set, in modern parlance) which, in his words, is "the general concept which, by means of our active faculty of thought, arises from the aggregate M when we make abstraction of the nature of its various elements m and of the order in which they are given." Cantor then observes that the result $\overline{\overline{M}}$ of these abstractions is again a set. Thus Cantor employs a process of abstraction to define a concept based on a common feature shared by a collection of objects. For instance, the concept "blue" is abstracted from that quality which all blue objects have in common, and "three" arises from the quality that all sets with three elements have in common. In particular, Cantor notes that not only does the abstraction of the quality of having three elements have nothing to do with the particular nature of the elements, it is also independent of the order in which they are given. However, the reader presumably is struck by the seeming circularity of this definition.

[5] One-to-one correspondence.

The confusion stems from the subtle fact that the first occurrence of "three" in the above sentence is a noun to be defined through the adjective "three." A slightly different formulation of the same idea is to say that the noun "three" denotes that property which all sets have in common that can be put in one-to-one correspondence with the set $\{a, b, c\}$. (In fact, in §5, Cantor gives roughly this definition for the finite cardinal number 3.) This second definition seems at first to avoid the problem of circularity, though it may seem a bit arbitrary. What is special about the set $\{a, b, c\}$? Why not use $\{X, Y, Z\}$? For that matter, why not use the set $\{x, y\}$? The obvious answers are that there is nothing special about $\{a, b, c\}$; $\{X, Y, Z\}$ would do just as well, since it also has three elements; but $\{x, y\}$ does not work, since it has only two elements. We have not escaped circularity after all. Both attempts to define "three" suffer from the same essential problem: they both rely on the idea that we can recognize whether or not a given set has three elements *before* we have a definition of three. The puzzle as to how we can recognize that some particular concrete object has an abstract property (i.e., that $\{a, b, c\}$ has three elements or that the Cookie Monster is blue) without having a definition of the property (three or blue) dates back at least as far as Plato and has an extensive literature.

A further logical problem with this way of defining cardinal numbers (or anything, for that matter) is that one might define a property and give it a name without being assured that it really exists, or that a unique object of this kind exists. (Bertrand Russell argues in [149, pp. 114–15] that definition by abstraction is never valid.) Later mathematicians and logicians objected to Cantor's approach and searched for a definition that would incorporate the existence of the thing so defined by giving the definition in terms of things already known to exist. Russell and, independently, Gottlob Frege around the turn of the century proposed to define the cardinal number of a set as the collection of all sets that could be put in one-to-one correspondence with it, thereby specifying a unique object. However, this definition also leads to serious logical problems, since any such collection is subject to Russell's paradox mentioned in the introduction, that is, it is too big to be a set and thereby lies outside the realm of a theory of sets free of paradoxes. This difficulty in turn led to a further refinement of the idea of cardinal number, through the use of ordinal numbers, proposed by John von Neumann in 1928 [125, p. 265], and has become the modern standard.

§2
"Greater" and "Less" with Powers

If for two aggregates M and N with the cardinal numbers $\mathfrak{a} = \overline{\overline{M}}$ and $\mathfrak{b} = \overline{\overline{N}}$, both the conditions:

(a) There is no part of M which is equivalent to N,
(b) There is a part N_1 of N, such that $N_1 \sim M$,

are fulfilled, it is obvious that these conditions still hold if in them M and N are replaced by two equivalent aggregates M' and N'. Thus they express a definite relation of the cardinal numbers $\mathfrak{a}$ and $\mathfrak{b}$ to one another.

Further, the equivalence of M and N, and thus the equality of $\mathfrak{a}$ and $\mathfrak{b}$, is excluded; for if we had $M \sim N$, we would have, because $N_1 \sim M$, the equivalence $N_1 \sim N$, and then, because $M \sim N$, there would exist a part M_1 of M such that $M_1 \sim M$, and therefore we should have $M_1 \sim N$; and this contradicts the condition (a).

Thirdly, the relation of $\mathfrak{a}$ to $\mathfrak{b}$ is such that it makes impossible the same relation of $\mathfrak{b}$ to $\mathfrak{a}$; for if in (a) and (b) the parts played by M and N are interchanged, two conditions arise which are contradictory to the former ones.

We express the relation of $\mathfrak{a}$ to $\mathfrak{b}$ characterized by (a) and (b) by saying: $\mathfrak{a}$ is "less" than $\mathfrak{b}$ or $\mathfrak{b}$ is "greater" than $\mathfrak{a}$; in signs

$$\mathfrak{a} < \mathfrak{b} \text{ or } \mathfrak{b} > \mathfrak{a}. \tag{1}$$

We can easily prove that,

$$\text{if } \mathfrak{a} < \mathfrak{b} \text{ and } \mathfrak{b} < \mathfrak{c}, \text{ then we always have } \mathfrak{a} < \mathfrak{c}. \tag{2}$$

Similarly, from the definition, it follows at once that, if P_1 is part of an aggregate P, from $\mathfrak{a} < \overline{\overline{P_1}}$ follows $\mathfrak{a} < \overline{\overline{P}}$ and from $\overline{\overline{P}} < \mathfrak{b}$ follows $\overline{\overline{P}}_1 < \mathfrak{b}$.

We have seen that, of the three relations

$$\mathfrak{a} = \mathfrak{b},\ \mathfrak{a} < \mathfrak{b},\ \mathfrak{b} < \mathfrak{a},$$

each one excludes the two others. On the other hand, the theorem that, with any two cardinal numbers $\mathfrak{a}$ and $\mathfrak{b}$, one of those three relations must necessarily be realized, is by no means self-evident and can hardly be proved at this stage.

Not until later, when we shall have gained a survey over the ascending sequence of the transfinite cardinal numbers and an insight into their connexion, will result the truth of the theorem:

A. If $\mathfrak{a}$ and $\mathfrak{b}$ are any two cardinal numbers, then either $\mathfrak{a} = \mathfrak{b}$ or $\mathfrak{a} < \mathfrak{b}$ or $\mathfrak{a} > \mathfrak{b}$.

Here, Cantor states the trichotomy principle, comments that this assertion is by no means self-evident, and promises to provide a proof later. No use of the assertion is made in the text at hand. Later Cantor claimed in a letter to Dedekind [28, p. 447] to have a proof that relied on the Well-Ordering Theorem.

§3
The Addition and Multiplication of Powers

The union of two aggregates M and N which have no common elements was denoted in §1, (2), by (M, N). We call it the "union-aggregate of M and N."

If M' and N' are two other aggregates without common elements, and if $M \sim M'$ and $N \sim N'$, we saw that we have

$$(M, N) \sim (M', N').$$

Hence the cardinal number of (M, N) only depends upon the cardinal numbers $\overline{\overline{M}} = \mathfrak{a}$ and $\overline{\overline{N}} = \mathfrak{b}$.

This leads to the definition of the sum of $\mathfrak{a}$ and $\mathfrak{b}$. We put

$$\mathfrak{a} + \mathfrak{b} = \overline{\overline{(M, N)}}. \tag{1}$$

Since in the conception of power, we abstract from the order of the elements, we conclude at once that

$$\mathfrak{a} + \mathfrak{b} = \mathfrak{b} + \mathfrak{a}; \tag{2}$$

and, for any three cardinal numbers $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$, we have

$$\mathfrak{a} + (\mathfrak{b} + \mathfrak{c}) = (\mathfrak{a} + \mathfrak{b}) + \mathfrak{c}. \tag{3}$$

We now come to multiplication. Any element m of an aggregate M can be thought to be bound up with any element n of another aggregate N so as to form a new element (m, n); we denote by $(M.N)$ the aggregate of all these bindings (m, n), and call it the "aggregate of bindings of M and N."[6] Thus

$$(M.N) = \{(m, n)\}. \tag{4}$$

We see that the power of $(M.N)$ only depends on the powers $\overline{\overline{M}} = \mathfrak{a}$ and $\overline{\overline{N}} = \mathfrak{b}$; for, if we replace the aggregates M and N by the aggregates

$$M' = \{m'\} \text{ and } N' = \{n'\}$$

respectively equivalent to them, and consider m, m' and n, n' as corresponding elements, then the aggregate

$$(M'.N') = \{(m', n')\}$$

is brought into a reciprocal and univocal correspondence with $(M.N)$ by regarding (m, n) and (m', n') as corresponding elements. Thus

$$(M'.N') \sim (M.N). \tag{5}$$

We now define the product $\mathfrak{a}.\mathfrak{b}$ by the equation

$$\mathfrak{a}.\mathfrak{b} = \overline{\overline{(M.N)}}. \tag{6}$$

---

[6] This is just the Cartesian product of M and N.

An aggregate with the cardinal number $\mathfrak{a}.\mathfrak{b}$ may also be made up out of two aggregates M and N with the cardinal numbers $\mathfrak{a}$ and $\mathfrak{b}$ according to the following rule: We start from the aggregate N and replace in it every element n by an aggregate $M_n \sim M$; if, then, we collect the elements of all these aggregates M_n to a whole S, we see that

(7) $$S \sim (M.N),$$

and consequently

$$\overline{\overline{S}} = \mathfrak{a}.\mathfrak{b}.$$

For, if, with any given law of correspondence of the two equivalent aggregates M and M_n, we denote by m the element of M which corresponds to the element m_n of M_n, we have

(8) $$S = \{m_n\};$$

thus the aggregates S and $(M.N)$ can be referred reciprocally and univocally to one another by regarding m_n and (m, n) as corresponding elements.

From our definitions result readily the theorems:

(9) $$\mathfrak{a}.\mathfrak{b} = \mathfrak{b}.\mathfrak{a},$$

(10) $$\mathfrak{a}.(\mathfrak{b}.\mathfrak{c}) = (\mathfrak{a}.\mathfrak{b}).\mathfrak{c},$$

(11) $$\mathfrak{a}.(\mathfrak{b} + \mathfrak{c}) = \mathfrak{a}\mathfrak{b} + \mathfrak{a}\mathfrak{c};$$

because:

$$(M.N) \sim (N.M),$$
$$(M.(N.P)) \sim ((M.N).P),$$
$$(M.(N,P)) \sim ((M.N),(M.P)).$$

Addition and multiplication of powers are subject, therefore, to the commutative, associative, and distributive laws.

§4
The Exponentiation of Powers

By a "covering of the aggregate N with elements of the aggregate M," or, more simply, by a "covering of N with M," we understand a law by which with every element n of N a definite element of M is bound up, where one and the same element of M can come repeatedly into application. The element of M bound up with n is, in a way, a one-valued function of n, and may be denoted by $f(n)$; it is called a "covering function of n." The corresponding covering of N will be called $f(N)$.

Two coverings $f_1(N)$ and $f_2(N)$ are said to be equal if, and only if, for all elements n of N the equation

(1) $$f_1(n) = f_2(n)$$

is fulfilled, so that if this equation does not subsist for even a single element $n = n_0$, $f_1(N)$ and $f_2(N)$ are characterized as different coverings of N. For example, if m_0 is a particular element of M, we may fix that, for all n's

$$f(n) = m_0;$$

this law constitutes a particular covering of N with M. Another kind of covering results if m_0 and m_1 are two different particular elements of M and n_0 a particular element of N, from fixing that

$$\begin{aligned} f(n_0) &= m_0 \\ f(n) &= m_1, \end{aligned}$$

for all n's which are different from n_0.

The totality of different coverings of N with M forms a definite aggregate with the elements $f(N)$; we call it the "covering-aggregate of N with M" and denote it by $(N|M)$. Thus:

$$(N|M) = \{f(N)\}. \tag{2}$$

If $M \sim M'$ and $N \sim N'$, we easily find that

$$(N|M) \sim (N'|M'). \tag{3}$$

Thus the cardinal number of $(N|M)$ depends only on the cardinal numbers $\overline{\overline{M}} = \mathfrak{a}$ and $\overline{\overline{N}} = \mathfrak{b}$; it serves us for the definition of $\mathfrak{a}^{\mathfrak{b}}$:

$$\mathfrak{a}^{\mathfrak{b}} = \overline{\overline{(N|M)}}. \tag{4}$$

For any three aggregates, M, N, P, we easily prove the theorems:

$$((N|M).(P|M)) \sim ((N, P)|M), \tag{5}$$

$$((P|M).(P|N)) \sim (P|(M.N)), \tag{6}$$

$$(P|(N|M)) \sim ((P.N)|M), \tag{7}$$

from which, if we put $\overline{\overline{P}} = \mathfrak{c}$, we have, by (4) and by paying attention to §3, the theorems for any three cardinal numbers, $\mathfrak{a}, \mathfrak{b}$, and $\mathfrak{c}$:

$$\mathfrak{a}^{\mathfrak{b}}.\mathfrak{a}^{\mathfrak{c}} = \mathfrak{a}^{\mathfrak{b}+\mathfrak{c}}, \tag{8}$$

$$\mathfrak{a}^{\mathfrak{c}}.\mathfrak{b}^{\mathfrak{c}} = (\mathfrak{a}.\mathfrak{b})^{\mathfrak{c}}, \tag{9}$$

$$(\mathfrak{a}^{\mathfrak{b}})^{\mathfrak{c}} = \mathfrak{a}^{\mathfrak{b}.\mathfrak{c}}. \tag{10}$$

We see how pregnant and far-reaching these simple formulæ extended to powers are by the following example. If we denote the power of the linear continuum X (that is, the totality X of real numbers x such that $x \geq 0$ and ≤ 1) by $\mathfrak{c}$, we easily see that it may be represented by, amongst others, the formula:

$$\mathfrak{c} = 2^{\aleph_0}, \tag{11}$$

where §6 gives the meaning of $\aleph_0$. In fact, by (4), $2^{\aleph_0}$ is the power of all representations

$$(12) \qquad x = \frac{f(1)}{2} + \frac{f(2)}{2^2} + \cdots + \frac{f(\nu)}{2^\nu} + \cdots \quad \text{(where } f(\nu) = 0 \text{ or } 1)$$

of the numbers x in the binary system. If we pay attention to the fact that every number x is only represented once, with the exception of the numbers $x = \frac{2\nu+1}{2^\mu} < 1$, which are represented twice over, we have, if we denote the "enumerable" totality of the latter by $\{s_\nu\}$,

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

If we take away from X any "enumerable" aggregate $\{t_\nu\}$ and denote the remainder by X_1, we have:

$$\begin{aligned} X &= (\{t_\nu\}, X_1) = (\{t_{2\nu-1}\}, \{t_{2\nu}\}, X_1), \\ (\{s_\nu\}, X) &= (\{s_\nu\}, \{t_\nu\}, X_1), \\ \{t_{2\nu-1}\} &\sim \{s_\nu\}, \quad \{t_{2\nu}\} \sim \{t_\nu\}, \quad X_1 \sim X_1; \end{aligned}$$

so

$$X \sim (\{s_\nu\}, X),$$

and thus (§1)

$$2^{\aleph_0} = \overline{\overline{X}} = \mathfrak{c}.$$

From (11) follows by squaring (by §6, (6))

$$\mathfrak{c}.\mathfrak{c} = 2^{\aleph_0}.2^{\aleph_0} = 2^{\aleph_0+\aleph_0} = 2^{\aleph_0} = \mathfrak{c},$$

and hence, by continued multiplication by $\mathfrak{c}$,

$$(13) \qquad \mathfrak{c}^\nu = \mathfrak{c},$$

where ν is any finite cardinal number.

If we raise both sides of (11) to the power $\aleph_0$ we get

$$\mathfrak{c}^{\aleph_0} = (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0.\aleph_0}.$$

But since, by §6, (8), $\aleph_0.\aleph_0 = \aleph_0$, we have

$$(14) \qquad \mathfrak{c}^{\aleph_0} = \mathfrak{c},$$

The formulæ (13) and (14) mean that both the ν-dimensional and the $\aleph_0$-dimensional continuum have the power of the one-dimensional continuum. Thus the whole contents of my paper in Crelle's *Journal*, vol. lxxxiv, 1878, are derived purely algebraically with these few strokes of the pen from the fundamental formulæ of the calculation with cardinal numbers.

Cantor's idea to define exponentiation of cardinals by using the set of all functions from one set to another was a stroke of genius. Why this is a plausible generalization of exponentiation for finite numbers will be explored in the exercises.

The power of this idea can be seen in the example of the interval X of real numbers from 0 to 1. The numbers $f(\nu)$ in equation (12) are simply the zeros and ones in the binary expansion of the number x. Thus, the number x corresponds to a sequence of zeros and ones, in other words, a function f from the natural numbers $\{1, 2, \ldots, \nu, \ldots\}$ to the two-element set $\{0, 1\}$, and therefore the set of all such functions has, by Cantor's definition, the cardinality $2^{\aleph_0}$.

Cantor carefully notes that certain real numbers have two different representations in terms of sequences of zeros and ones. For instance, the number $\frac{1}{4}$ has the two binary representations $0.010000\ldots$ and $0.00111\ldots$. This set of real numbers with more than one binary expansion, he claims, is enumerable, in other words it is equivalent to the set of natural numbers. (The reader may show this as a challenging exercise or read Section 7 of Cantor's work.) A clever correspondence argument then shows that the double representations are not numerous enough to prevent X from having cardinality $2^{\aleph_0}$.

§6
The Smallest Transfinite Cardinal Number Aleph-Zero

Aggregates with finite cardinal numbers are called "finite aggregates," all others we will call "transfinite aggregates" and their cardinal numbers "transfinite cardinal numbers."

The first example of a transfinite aggregate is given by the totality of finite cardinal numbers ν; we call its cardinal number (§1) "Aleph-zero" and denote it by $\aleph_0$; thus we define

(1) $$\aleph_0 = \overline{\overline{\nu}}.$$

That $\aleph_0$ is a *transfinite* number, that is to say, is not equal to any finite number μ, follows from the simple fact that, if to the aggregate $\{\nu\}$ is added a new element e_0, the union-aggregate $(\{\nu\}, e_0)$ is equivalent to the original aggregate $\{\nu\}$. For we can think of this reciprocally univocal correspondence between them: to the element e_0 of the first corresponds the element 1 of the second, and to the element ν of the first corresponds the element $\nu + 1$ of the other. By §3 we thus have

(2) $$\aleph_0 + 1 = \aleph_0.$$

But we showed in §5 that $\mu + 1$ is always different from μ, and therefore $\aleph_0$ is not equal to any finite number μ.

The number $\aleph_0$ is greater than any finite number μ:

(3) $$\aleph_0 > \mu.$$

This follows, if we pay attention to §3, from the three facts that $\mu = \overline{\overline{(1, 2, 3, \ldots, \mu)}}$, that no part of the aggregate $(1, 2, 3, \ldots, \mu)$ is equivalent to the aggregate $\{\mu\}$, and that $(1, 2, 3, \ldots, \mu)$ is itself a part of $\{\nu\}$.

On the other hand, $\aleph_0$ is the least transfinite cardinal number. If $\mathfrak{a}$ is any transfinite cardinal number different from $\aleph_0$, then

(4) $$\aleph_0 < \mathfrak{a}.$$

This rests on the following theorems:

A. Every transfinite aggregate T has parts with the cardinal number $\aleph_0$.

Proof. If, by any rule, we have taken away a finite number of elements $t_1, t_2, \ldots, t_{\nu-1}$, there always remains the possibility of taking away a further element t_ν. The aggregate $\{t_\nu\}$, where ν denotes any finite cardinal number, is a part of T with the cardinal number $\aleph_0$, because $\{t_\nu\} \sim \{\nu\}$ (§1).

B. If S is a transfinite aggregate with the cardinal number $\aleph_0$, and S_1 is any transfinite part of S, then $\overline{\overline{S_1}} = \aleph_0$.

Proof. We have supposed that $S \sim \{\nu\}$. Choose a definite law of correspondence between these two aggregates, and, with this law, denote by s_ν that element of S which corresponds to the element ν of $\{\nu\}$, so that

$$S = \{s_\nu\}.$$

The part S_1 of S consists of certain elements s_κ of S, and the totality of numbers κ forms a transfinite part K of the aggregate $\{\nu\}$. By Theorem G of §5 the aggregate K can be brought into the form of a series

$$K = \{\kappa_\nu\},$$

where

$$\kappa_\nu < \kappa_{\nu+1};$$

consequently we have

$$S_1 = \{s_{\kappa_\nu}\}.$$

Hence it follows that $S_1 \sim S$, and therefore $\overline{\overline{S_1}} = \aleph_0$.

From A and B the formula (4) results, if we have regard to §2.

From (2) we conclude, by adding 1 to both sides,

$$\aleph_0 + 2 = \aleph_0 + 1 = \aleph_0,$$

and, by repeating this

(5) $$\aleph_0 + \nu = \aleph_0.$$

We also have

$$\aleph_0 . \aleph_0 = \aleph_0.$$

Proof. By (6) of §3, $\aleph_0 . \aleph_0$ is the cardinal number of the aggregate of bindings

$$\{(\mu, \nu)\},$$

where μ and ν are any finite cardinal numbers which are independent of one another. If also λ represents any finite cardinal number, so that $\{\lambda\}, \{\mu\}$, and $\{\nu\}$ are only different notations for the same aggregate of all finite numbers, we have to show that

$$\{(\mu, \nu)\} \sim \{\lambda\}.$$

Let us denote $\mu+\nu$ by ρ; then ρ takes all the numerical values $2, 3, 4, \ldots$, and there are in all $\rho - 1$ elements (μ, ν) for which $\mu + \nu = \rho$, namely:

$$(1, \rho - 1), (2, \rho - 2), \ldots , (\rho - 1, 1).$$

In this sequence imagine first the element $(1, 1)$, for which $\rho = 2$, put, then the two elements for which $\rho = 3$, then the three elements for which $\rho = 4$, and so on. Thus we get all the elements (μ, ν) in a simple series:

$$(1, 1); (1, 2), (2, 1); (1, 3), (2, 2), (3, 1); (1, 4), (2, 3), \ldots ,$$

and here, as we easily see, the element (μ, ν) comes at the λth place, where

$$\lambda = \mu + \frac{(\mu + \nu - 1)(\mu + \nu - 2)}{2}. \tag{9}$$

The variable λ takes every numerical value $1, 2, 3, \ldots$, once. Consequently, by means of (9), a reciprocally univocal relation subsists between the aggregates $\{\lambda\}$ and $\{(\mu, \nu)\}$.

Recall from Section 4 that the real numbers between zero and one have cardinality $2^{\aleph_0}$. In the 1891 paper *Über eine Elementare Frage der Mannigfaltigkeitslehre* (On an Elementary Question in the Theory of Sets) [28], from which we include an excerpt below [28, pp. 278 f.], Cantor shows that if $\mathfrak{a}$ is a cardinal number, then $2^{\mathfrak{a}} > \mathfrak{a}$. In particular, $2^{\aleph_0} > \aleph_0$. This means that the set of real numbers has larger cardinality than the natural (and rational) numbers.

Georg Cantor, from

On an Elementary Question in the Theory of Sets

Namely, if m and w are any two distinct characters, we form a collection M of elements

$$E = (x_1, x_2, \ldots , x_\nu, \ldots)$$

which depends on infinitely many coordinates $x_1, x_2, \ldots , x_\nu, \ldots$, each of which is either m or w. Let M be the set of all elements E.

Amongst the elements of M are for example the following three

$$
\begin{aligned}
E^{\mathsf{I}} &= (m, m, m, m, \dots)\,, \\
E^{\mathsf{II}} &= (w, w, w, w, \dots)\,, \\
E^{\mathsf{III}} &= (m, w, m, w, \dots)\,.
\end{aligned}
$$

I now claim that such a manifold[7] M does not have the power of the series $1, 2, \dots, \nu, \dots$.

This follows from the following theorem:

"If $E_1, E_2, \dots, E_\nu, \dots$ is any simply infinite sequence of elements of the manifold M, then there is always an element E_0 of M which does not agree with any E_ν."

To prove this let

$$
\begin{aligned}
E_1 &= (a_{11}, a_{12}, \dots, a_{1\nu}, \dots)\,, \\
E_2 &= (a_{21}, a_{22}, \dots, a_{2\nu}, \dots)\,, \\
&\dots \\
E_\mu &= (a_{\mu 1}, a_{\mu 2}, \dots, a_{\mu\nu}, \dots)\,. \\
&\dots
\end{aligned}
$$

Here each $a_{\mu\nu}$ is a definite m or w. A sequence $b_1, b_2, \dots, b_\nu$ will now be so defined, that each b_ν is also equal to m or w and *different* from $a_{\nu\nu}$.

So if $a_{\nu\nu} = m$, then $b_\nu = w$, and if $a_{\nu\nu} = w$, then $b_\nu = m$.

If we then consider the element

$$E_0 = (b_1, b_2, b_3, \dots)$$

of M, then one sees immediately that the equation

$$E_0 = E_\mu$$

cannot be satisfied for any integer value of μ, since otherwise for that value of μ and all integer values of ν

$$b_\nu = a_{\mu\nu}$$

and also in particular

$$b_\mu = a_{\mu\mu},$$

which is excluded by the definition of b_μ. It follows immediately from this theorem that the totality of all elements of M cannot be put in the form of a series: $E_1, E_2, \dots, E_\nu, \dots$, since otherwise we would have the contradiction that a thing E_0 would be an element of M and also not an element of M.

[7] Set.

This proof appears remarkable not only due to its great simplicity, but in particular for the reason that the principle employed in it can be directly extended to the general theorem, that the powers of well-defined point sets have no maximum, or what amounts to the same, that to every given point-set L can be associated another one M which has a higher power than L.

Exercise 2.14: For $M = \{1, 2, 3, 4, 5\}$ and $N = \{a, b, c, d, e, f, g, h\}$, what are $\overline{\overline{M}}$ and $\overline{\overline{N}}$? Show that $\overline{\overline{M}} < \overline{\overline{N}}$ using the definition given in §2.

Exercise 2.15: Verify the distributive law for powers; that is, prove (11) in §3.

Exercise 2.16: If A, B, C are sets, show that $A.(B.C) \neq (A.B).C$, but they are equivalent.

Exercise 2.17: Which real numbers between 0 and 1 have more than one binary expansion? How many different binary expansions for a number are possible?

Exercise 2.18: Why are covering functions so named?

Exercise 2.19: What are 0^0, 0^1, and 1^0?

Exercise 2.20: Prove (3) in §4.

Exercise 2.21:

1. Let M and N be aggregrates such that $\overline{\overline{M}} = 3$ and $\overline{\overline{N}} = 2$. Show that $3^2 = \overline{\overline{(N \mid M)}}$.
2. If $\overline{\overline{M}} = m$ and $\overline{\overline{N}} = n$ where m and n are positive integers, show that $m^n = \overline{\overline{(N \mid M)}}$.
3. Suppose $\overline{\overline{M}} = m$ where m is a positive integer, and let $\mathcal{P}(M)$ denote the power set of M (i.e., the set of all subsets of M). Show that

$$\overline{\overline{\mathcal{P}(M)}} = \overline{\overline{(M \mid 2)}} = 2^m.$$

Exercise 2.22: Prove (8), (9), and (10) in §4.

Exercise 2.23: Show that $\aleph_0 + \aleph_0 = \aleph_0$.

Exercise 2.24: Verify formula (9) of §6 by representing Cantor's series of ordered pairs (or "bindings" as he calls them)

$$(1, 1); (1, 2), (2, 1); (1, 3), \ldots$$

as points in the plane and indicating their order of precedence.

Exercise 2.25: Generalize Cantor's argument in the last source to prove that $2^{\mathfrak{m}} > \mathfrak{m}$ for any cardinal number $\mathfrak{m}$.

Exercise 2.26: List an infinite sequence of infinite cardinal numbers.

Exercise 2.27: Show that Frege's proposal for a definition of cardinal numbers leads to a paradox. Recall that he proposes to define the number 1, for instance, as the equivalence class of all sets with one element, that is,

$$1 = \{\{\mathbf{a}\} \mid \mathbf{a} \text{ is a set}\}.$$

Exercise 2.28: Read Cantor's §7 on ordinal types.

2.4 Zermelo's Axiomatization

Ernst Zermelo's work is situated at two opposite ends of the mathematical spectrum. He wrote his dissertation in 1899 at the University of Göttingen on the very applied topic "Hydrodynamic Investigations of Currents on the Surface of a Sphere." Five years later he proved his Well-Ordering Theorem, far removed from any mathematics that seemed applicable at the time. Almost all of his publications fall into either the area of physics and the calculus of variations, or are related to set theory. A notable combination of applied and pure mathematics can be found in a paper he wrote on applications of set theory to chess. He is best known through his axiomatization of Cantorian set theory, which represents his most important contribution.

Following are excerpts (pp. 261–267) from Zermelo's 1908 paper *Untersuchungen über die Grundlagen der Mengenlehre, I* (Investigations on the

PHOTO 2.5. Zermelo.