CAMBRIDGE

Cambridge University Press
978-1-108-98691-5 — Set Theory
John P. Burgess
Excerpt
More Information

1 Historical Roots

Although in retrospect others (Bernard Bolzano, Richard Dedekind) can be viewed as precursors, set theory was largely the creation of a single individual, Georg Cantor, beginning in the 1870s, and his key work (Cantor, 1915) remains highly readable to this day. He launched the field with two results on questions with ancient roots.

1.1 Strings to Ordinals

Pythagoreans noted that if the lengths of otherwise similar strings are in the ratio 2:1, the shorter sounds an octave higher. Why? Because it vibrates twice as quickly. In modern mathematical language, if the graph of the displacement of the center of the string with time approximates $y = \cos x$ for the longer, it will approximate $y = \cos 2x$ for the shorter. No real string vibrates so simply, and a better approximation for the long string would be $y = a_1 \cos + a_2 \cos 2x$, with the amplitude a_1 of the "fundamental" much larger than the amplitude a_2 of the "overtone." By the eighteenth century, workers in analysis, the branch of mathematics beginning with calculus, were dealing with infinite trigonometric series:

$$y = (a_1 \cos x + b_1 \sin x) + (a_2 \cos 2x + b_2 \sin 2x) + (a_3 \cos 3x + b_3 \sin 3x) + \ldots$$

The "vibrating string controversy" engaging Leonhard Euler and others concerned how wide a class of functions can be represented in this form. The dispute exposed, beyond endemic deficiencies of rigor in the treatment of infinite series, lack of a common understanding about what is meant by a *function*. The ensuing nineteenth-century rigorization of analysis, besides banning any literal infinities or infinitesimals, explaining contexts containing the symbol ∞ without assuming it to denote anything in isolation, fixed on the maximally general notion of function, under which *any* correlation between inputs and outputs counts, as long as there is one and only one output per input. Improved rigor eventually led to consensus about the existence of trigonometric series representations.

But with existence there come uniqueness questions. Could a function have *two different* representations? Does the constant function zero have any other than the trivial one with $a_n = b_n = 0$ for all n? Bernhard Riemann showed it does not if the sequence converges for all x. But what if one allows an exceptional point for which convergence is not assumed? Enter Cantor. It turns out that even then triviality holds (and, as a conclusion, we get what we did not assume as a premise, convergence even at the exceptional point). Indeed, one can allow two or any finite number of exceptional points. One can even allow infinitely many as long as they are all *isolated* from one another,

CAMBRIDGE

Cambridge University Press
978-1-108-98691-5 — Set Theory
John P. Burgess
Excerpt
More Information

meaning that for each exceptional x there is a positive ε with no *other* exceptional points between $x - \varepsilon$ and $x + \varepsilon$. One can even allow a *doubly* exceptional point, not isolated from other exceptional points. Indeed, one can allow two or any finite number. One can even allow infinitely many as long as they are isolated from one another. One can even allow a *triply* exceptional point. And so on. And as one goes on, it becomes natural to switch from speaking in the plural of the exceptional points to speaking in the singular of the *set* E of which they are *elements*. What it means to treat E as a single item is to think of operations being applicable to it. The relevant operation on sets Cantor called *derivation*, discarding isolated points. Let E_0 be E itself, and let E_{n+1} be the derived set of E_n. Reimann's result was that uniqueness holds if $E_0 = \varnothing$, the empty set, with no elements. Cantor's results were that uniqueness holds if any of $E_1,\ E_2,\ E_3, \ldots$ is empty. Moreover, if we let E_ω be the intersection of the E_n, the set of x belonging to all of them, uniqueness still holds if $E_\omega = \varnothing$. Moreover, the results continue, with sets indexed by:

$$\omega + 1, \omega + 2, \omega + 3, \ldots \omega + \omega = \omega \cdot 2, \omega \cdot 3, \omega \cdot 4, \ldots \omega \cdot \omega = \omega^2, \omega^3, \omega^4, \ldots \omega^\omega$$

and more. Here are Cantor's *transfinite ordinal numbers*, and, as the notation suggests, he introduced an arithmetic for them, with addition, multiplication, and exponentiation.

1.2 Quadrature to Cardinals

Euclid shows many geometrical figures can be constructed with straightedge and compass, indicating the steps involved and proving they lead to the desired result. Thus one can *duplicate the square*, or construct, given the side of a square, the side of a square of twice the area, just by taking the diagonal of the original square. To show a construction *not* possible is more difficult, and requires an analysis available only with the modern coordinate methods, which transform geometric into algebraic problems. Thus *duplicating the cube*, constructing, given the side of a cube, the side of a cube of twice the volume, turns out equivalent to obtaining a key number, $\sqrt[3]{2}$, from rational numbers by addition, subtraction, multiplication, division, and extraction of square roots. And this was proved impossible in the 1830s, disposing of an ancient problem. For *quadrature of the circle*, constructing for a given circle a square of equal area, the key number is π. Now, although $\sqrt[3]{2}$ is not obtainable in the way indicated, it is at least an *algebraic* number in the sense of a solution to a polynomial equation:

$$a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 = 0$$

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-108-98691-5 — Set Theory
John P. Burgess
Excerpt
More Information

with rational coefficients a_i, namely, $x^3–2 = 0$. It was conjectured, however, that π is not even algebraic in this sense. Joseph Liouville showed nonalgebraic or *transcendental* numbers exist. Then *e*, the basis of the natural logarithms, was shown to be one by Charles Hermite, and, finally, π by Ferdinand von Lindemann. Between these last two, Cantor showed that the vast majority of real numbers are transcendental.

Since the sets of algebraics and transcendentals are infinite, to say one has more elements than the other requires a definition of when the *transfinite cardinal*, or number of elements of one infinite set, *A*, is equal or unequal to that of another, *B*. Cantor took as his standard of equality the existence of a *bijection* between *A* and *B*, a relation under which each element of *A* is associated with exactly one element of *B*, and vice versa. In the case of the set *N* of natural numbers, the existence of a bijection with a set *B* means that the elements of *B* can be *enumerated* or listed in a sequence indexed by 0, 1, 2, . . ., as in Table 1. An infinite set whose elements can be so enumerated is called *denumerable*, while a set that is *either* denumerable *or* finite is called *countable*.

The number of elements of a denumerable set Cantor called $\aleph_0$ (pronounced “aleph nought”). What the table shows is that signed integers and positive rationals both have cardinal or size $\aleph_0$; so do the signed rationals. Nowadays, a finite sequence of keystrokes is transmitted electronically as a sequence of zeros and ones, the binary numeral for some natural number that may be considered a code for the sequence. This makes the set of such sequences denumerable, in order of increasing code number. Then, since a polynomial equation of degree *n* has at most *n* solutions, each algebraic number can be denoted by an expression such as “the second smallest solution to $2x^3 - 9x^2 - 6x + 3 = 0$” and given a code number accordingly. But their denumerability was established in correspondence between Dedekind and Cantor long before the digital age began.

By contrast, Cantor showed that the whole set *R* of real numbers (and hence the set of transcendentals, left over when we remove the algebraics) is *not* denumerable. No countable set can contain even just those whose decimal

Table 1 Denumerable sets

Set	Enumeration									
Natural numbers	0	1	2	3	4	5	6	7	8	. . .
Integers	0	1	−1	2	−2	3	−3	4	−4	. . .
Positive rationals	1/1	1/2	2/1	1/3	2/3	3/2	3/1	1/4	3/4	. . .

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-108-98691-5 — Set Theory
John P. Burgess
Excerpt
More Information

Table 2 The diagonal argument

Index	Zero-one sequence									
0	0*	0	0	0	0	0	0	0	0	. . .
1	1	1*	1	1	1	1	1	1	1	. . .
2	0	1	0*	1	0	1	0	1	0	. . .
3	1	0	1	0*	1	0	1	0	1	. . .
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .

expansion involves only 0s and 1s; or what is the same, all infinite zero-one sequences; or what is the same, all sets of natural numbers, each such being representable by the zero-one sequence with one in the *n*th place if and only if *n* is in the set. This he established by his famous *diagonal argument*. Suppose we have an enumeration of some set *S* of infinite zero-one sequences, as in Table 2. Go down the diagonal, marked with asterisks. Take in order for each *n* the digit appearing in the *n*th place in the *n*th row of the table. This gives 0100 . . . . Now swap the zeros and the ones. This gives 1011 . . ., a sequence that does not belong to the denumerable set *S*, since it differs in the *n*th place from the *n*th sequence. Cantor called the cardinal of the real numbers or points of the line ***c***. Analogously to the results in Table 1 in this discussion, he showed that the positive real numbers, or even just those in a finite interval, also have cardinal ***c***, as do pairs of real numbers, or equivalently complex numbers. He also introduced an arithmetic, with addition, multiplication, and exponentiation, for his cardinals.

Cantor's audacious introduction of ω and ℵ when mathematicians had just finished explaining away ∞ provoked a reaction. But Cantor's theory won acceptance among leaders in the rising generation fairly quickly (as examples they put forth, such as the one-, two-, and three-dimensional *Cantor set*, *Sierpinski carpet*, and *Menger sponge*, whose images appear all over the Internet today, captured the imagination of amateurs). The leading mathematician David Hilbert insisted: "No one shall expel us from the paradise Cantor created for us."

2 The Notion of Set

Many objections turned on certain *paradoxes*. Cantor, unlike his contemporary Gottlob Frege, never made the assumptions that led to these paradoxes, but he did not make clear enough what assumptions he *was* making. His successors had to be more clear and explicit. Explicit axiomatization began in the first decade of the twentieth century with Ernst Zermelo (1908/1967). His system,

© in this web service Cambridge University Press www.cambridge.org

with additions and amendments, mainly by Abraham Fraenkel (1922/1967), remains that accepted today, when it is recognized that the paradoxes result mostly from confusing the notion of set behind the axioms of *Zermelo–Fraenkel set theory with Choice* (ZFC) with other ideas.

2.1 Collections

The expression “a multiplicity of objects” begins singular but ends plural, and may be understood as referring either to a *plurality*, a many, or to a *universal*, a one as opposed to a many. Universals include *properties*, which are *intensional*, meaning that two may be different even while having exactly the same instances, as with the stock example *being a coin in my pocket* and *being a penny in my pocket*, which are distinct properties even if I have no coins in my pocket but pennies. They also include *aggregates* completely determined by their components. One kind, topic of a theory called *mereology*, is a *fusion* of a plurality of component parts into a single whole, in a way that permits different pluralities to have the same fusion, as do the eight ranks and the eight files of a chessboard, the fusion being the selfsame chessboard in either case. By contrast we have *collections*, in which many are gathered into a one without losing track of which many they were.

The notion of collection in Frege (1893) was that of an *extension*. Here we start with all objects, and take what he called a *concept* (associated with a predicate), and divide objects into those that fall under the concept (satisfy the predicate) and those that do not. The collection of those that do is the extension of the concept, so that the extensions of two concepts are the same if and only if the concepts are *coextensive*, having exactly the same things falling under them. Graphically, we may represent the unbounded range of all objects with which we start as an unbounded blank page, and represent the extension as given by a dividing line or curve separating objects inside from objects outside, as in Figure 1. But for Frege, the extension is itself an object: If represented by a dot, that dot must fall on the page on one side or the other of the division – but which? That is the question indicated by the question marks in the figure.

Bertrand Russell raised an embarrassing issue about the extension R of the concept: it *is an extension that as an object is outside, not inside, itself*. In the case of the universal extension, V, the extension of *is self-identical*, V is inside itself since *everything* is inside V. In the case of the empty extension $\varnothing$, the extension of *nonself-identical*, $\varnothing$ is outside itself since *nothing* is inside $\varnothing$. Hence $\varnothing$ is inside, and V is outside, the Russell extension R. But just as the statement *this very statement is false* seems to be true if it is false and false if it is

CAMBRIDGE
Cambridge University Press
978-1-108-98691-5 — Set Theory
John P. Burgess
Excerpt
More Information

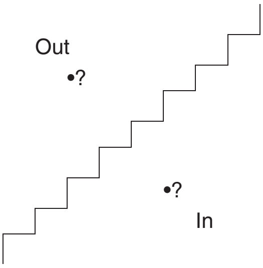

Figure 1 An extension

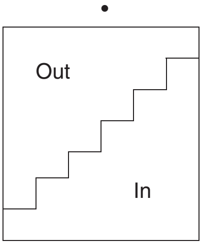

Figure 2 An ensemble

true, so *R* seems to be inside itself if outside itself, and outside if inside. This is the *Russell paradox* as Russell (1902) put it to Frege.

Contrasting with this inconsistent "top down" notion of extension is the "bottom up" notion of an *ensemble*. Here we start with a given "universe of discourse," which might be represented by a box, and a predicate will, like a curve in a Venn diagram, mark off the ensemble of things in the universe that do satisfy it from things in the universe that do not. The ensemble does *not*, however, itself belong to the universe. A dot representing it would lie outside the box, as in Figure 2. Implicit here is the possibility of *iteration*. We can add a new box atop the original, to accommodate all the dots representing ensembles of things in the lower box, and then more. But there are two ways to implement this idea.

On the *layered* approach of the *theory of types*, deriving from Russell (1908) by way of Frank Ramsey (1925), we have a hierarchy with *individuals* at the bottom type zero, collections called *classes* of type zero items at type one, classes of type one items at type two, and so on. Even if we assume *no* items at type zero, there will be one item at type one, the empty class $\varnothing_1$ of type zero items, and then two items at type two, the empty class $\varnothing_2$ of type one items, and the singleton class $\{\varnothing_1\}_2$ of the one item at type one. At type three, there will be four items, as in Table 3. With one item at type zero, there will be two at type

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-108-98691-5 — Set Theory
John P. Burgess
Excerpt
More Information

Table 3 The layered hierarchy

…	
4	Sixteen Items
3	$\varnothing_3$, $\{\varnothing_2\}_3$, $\{\{\varnothing_1\}_2\}_3$, $\{\varnothing_2,\{\varnothing_1\}_2\}_3$
2	$\varnothing_2$, $\{\varnothing_1\}_2$
1	$\varnothing_1$
0	No Items

one, then four, then sixteen. But with only finitely many individuals, there will only ever be only finitely many items of any one type. For mathematical purposes, Russell assumed infinitely many individuals.

2.2 Sets

By contrast, we have the *cumulative* approach, where successive boxes are nested, like Chinese boxes or Russian dolls, each higher one adding a new level of collections called *sets*. In box zero are individuals or *Urelemente*; at level one, sets whose elements are individuals; in box one, individuals and level-one sets; at level two, any new sets whose elements come from box one; in box two, box-one and level-two items; and so on.

In ZFC, we consider only *pure* sets, without individuals. There then will be no items at level zero, one item, the empty set $\varnothing$, at level one, in box one. As for level two, from the one item in box one can be formed two sets: the empty set $\varnothing$ and its singleton $\{\varnothing\}$, but the former we already have, so only the latter is new. In box three will be four items, two new at level three. In box four will be sixteen items, twelve new at level four. And so on, as in Table 4.

After all finite levels, we may recognize a box ω containing everything of finite level but nothing new, and then form a level $\omega + 1$ for sets whose elements come from level ω, meaning from any finite level, but do not themselves appear at any such level, containing as they do sets of arbitrarily high finite level. We can then continue through the transfinite ordinals. Zermelo at first claimed for his axioms only that they permitted none of the known deductions of contradictions, and seemed adequate to develop Cantor's set theory (as they are with Fraenkel's friendly amendments). Only later (as in Zermelo, 1930) did something like the picture in the table emerge.

The ideal of rigor is that one should list in advance all *primitives*, notions assumed meaningful without definition, and *postulates* or axioms, results assumed true without demonstration, and given these principles all further

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-108-98691-5 — Set Theory
John P. Burgess
Excerpt
More Information

Table 4 The cumulative hierarchy

…	…
ω + 1	{Ø, {Ø}, {{Ø}}, {{{Ø}}}, …} and Many Other New Items
ω	No New Items
…	…
4	Twelve New Items
3	{{Ø}}, {Ø,{Ø}}
2	{Ø}
1	Ø
0	No Items

Table 5 Primitive logical notions

Symbol	Operation	Reading
¬	Negation	"not"
∧	Conjunction	"and"
∨	Disjunction	"or"
∀	Universal quantification	"for all"
∃	Existential quantification	"for some" or "there exists"

notions or results should be logically derived, by definition or deduction. In set theory, there is just one primitive, written with a stylized epsilon symbol, $x \in y$, read "x is an element of y" or "x is in y" or "y contains x." All other notions must be defined in terms of this and the logical notion of identity using the logical operators in Table 5. A *formula* Φ is built up from *atomic* formulas $x \in y$ and $x = y$ using the five operations in the table.

Some minimal familiarity with logical notions and notations must be assumed here (for a quick review, see Boolos, Burgess, and Jeffrey, 2002, chapters 9 and 10), including an ability to recognize simple logical laws. In particular, familiarity is assumed with the distinction between "free" and "bound" occurrences of variables in a formula, those that are not and those that are caught by a quantifier. For example, in the formula asserting the nonemptiness of x, namely $\exists y(y \in x)$, the x is free but the y is bound. The latter could be changed to z without changing the meaning. Other logical and set-theoretic notions may be defined in terms of what we have so far, as in Tables 6 and 7, but officially these are mere abbreviations.

© in this web service Cambridge University Press www.cambridge.org

Table 6 Defined logical notions

Abbreviation	Definition	Operation	Reading
$\Phi \supset \Psi$	$\neg\Phi V \Psi$	Conditional	"if Φ then Ψ"
$\Phi \equiv \Psi$	$(\Phi \supset \Psi) \wedge (\Psi \supset \Phi)$	Biconditional	"Φ if and only if Ψ" or " Φ iff Ψ"
$x \neq y$	$\neg x = y$	Nonidentity	"*x* is distinct from *y*"
$\exists! x\Phi(x)$	$\exists x \forall y(\Phi(y) \equiv x = y)$	Unique existence	"there exists a unique"

Table 7 Defined set-theoretic notions

Abbreviation	Definition	Reading
$x \notin y$	$\neg\, x \in y$	"*x* is not an element of [or not in] *y*"
$x \subseteq y$	$\forall z\, (z \in x \supset z \in y)$	"*x* is a subset of [or included in] *y*"
$\forall x \in y\, \Phi(x)$	$\forall x(x \in y \supset \Phi(x))$	"for all *x* in *y* . . ."
$\exists x \in y\, \Phi(x)$	$\exists \mathrm{x}(x \in y \Lambda \Phi(x))$	"for some *x* in *y* . . ."

3 The Zermelo–Fraenkel Axioms

The axioms of the system ZFC will be presented next, in both words and symbols, to be assumed without proof, but not without something in the way of informal, intuitive justification.

3.1 Statement

The first axiom says sets with the same elements are the same. It has two equivalent formulations:

Extensionality $(\mathbf{1})\ \forall z(z \in x \equiv z \in y) \supset x = y, \quad (\mathbf{2})\ x \subseteq y \wedge y \subseteq x \supset x = y.$

By convention, in displaying formulas initial universal quantifiers are omitted, so what is meant is really $\forall x \forall y$(__) where what is explicitly written is __. As (2) suggests, proofs of identities most often come in two parts, proving inclusion in two directions. Extensionality implies that if there is a set *y* whose elements are all and only the sets *x* satisfying a condition Φ, it is unique. That unique set, if it exists, is denoted $\{x \mid \Phi(x)\}$, and we have $z \in \{x \mid \Phi(x)\}$ if Φ(*z*). Frege's inconsistent assumption would be an axiom of *comprehension*, according to which

© in this web service Cambridge University Press www.cambridge.org

$\{x|\,\Phi(x)\}$ *always* exists for *any* condition Φ. Applied to the condition $x\notin x$ this would give the Russell paradox, and it is not assumed in ZFC.

The second axiom says that if we *already have* some set u, we can at least separate out from u those of its elements that satisfy a condition Φ to form $\{x\in u|\Phi(x)\}$:

Separation $\exists y\forall x(x\in y\equiv(x\in u\wedge\Phi(x)))$.

This is not a single formula, but rather a rule to the effect that anything of a certain *form* counts as an axiom. The cases for different Φ are called *instances* of the *scheme* of separation. (Zermelo's original formulation was vaguer.) Note that separation implies there is no *universal set* of all sets $\mathrm{V}=\{x|\,x=x\}$. If there were, we could, by separation, obtain comprehension.

Further axioms state the existence of certain specific sets:

Pairing $\exists y\,(u\in y\wedge v\in y)$.
Union $\exists y\,\forall z\in X\,\forall x\in z(x\in y)$.

With what we have so far, some basic existence results then become deducible, those in Table 8. (The expression "family" used in the table may be used for any set of sets.)

Separation gives us the empty set, since given any set u at all – and even pure logic assumes there is at least one item in the domain our quantifiers range over, which in the present case consists of sets – separation gives $\{x\in u|x\notin u\}$, which is empty. It also gives twofold intersections, and by the alternative definition, family intersections, if the family X has at least one member u; also differences. Now given y containing u and v, we can separate out the elements of y identical to one of those two, so pairing with separation gives the unordered pair. Union with separation gives us family union. The unordered triple and twofold union we then get using the alternative definitions. The difference $u-v$ is also called the *relative* complement of v in u. An *absolute* complement $-v=\{x|\,x\notin v\}$ cannot exist, because $v\cup-v$ would be the nonexistent V.

The next two axioms are these:

Power $\exists y\forall x(x\subseteq u\supset x\in y)$.
Infinitity $\exists y(\varnothing\in y\wedge\forall x\in y(\{x\}\in y))$.

Power with separation gives the *power set* $\mathcal{P}(x)=\{y|y\subseteq x\}$ and also

$$\{y\subseteq x|\;\Phi(y)\}=\{y\in\mathcal{P}(x)|\Phi(y)\}.$$