UNDERSTANDING THE INFINITE

Shaughan Lavine

Understanding the Infinite

Understanding the Infinite

Shaughan Lavine

Harvard University Press
Cambridge, Massachusetts
London, England

Printed in the United States of America
Second printing, 1998

First Harvard University Press paperback edition, 1998

Library of Congress Cataloging-in-Publication Data

Lavine, Shaughan.
Understanding the infinite / Shaughan Lavine.
p. cm.
Includes bibliographical references and index.
ISBN 0-674-92096-1 (cloth)
ISBN 0-674-92117-8 (pbk.)
1. Mathematics—Philosophy. I. Title.
QA8.4.L38 1994
511.3'22—dc20
93-49697

Preface

In writing this book I have tried to keep mathematical prerequisites to a minimum. The reader who is essentially innocent of mathematical knowledge beyond that taught in high school should be able to read at least halfway through Chapter VIII plus parts of the rest of the book, though such a reader will need to skip the occasional formula. That is enough of the book for all of the major ideas to be presented. The Introduction may seem daunting since it refers to ideas that are not explained until later—trust me, they *are* explained. A reader who learned freshman calculus once, but perhaps does not remember it very well, and who has had a logic course that included a proof of the completeness theorem will be in fine shape throughout the book, except for various "technical remarks," an appendix to Chapter VI, and a few parts of Chapter IX. Those few technical discussions require varying degrees of mathematical sophistication and knowledge of general mathematical logic plus occasional knowledge of elementary recursion theory, model theory, or modal logic.

Thanks are due Bonnie Kent, Vann McGee, Sidney Morgenbesser, and Sarah Stebbins for their infinite patience in listening to my many half-baked ideas and for their substantial help in culling and completing them while I was writing this book. As they learned, I cannot think without the give and take of conversation. Thanks also to Ti-Grace Atkinson, Jeff Barrett, William Boos, Hartry Field, Alan Gabbey, Haim Gaifman, Alexander George, Allen Hazen, Gregory Landini, Penelope Maddy, Robert Miller, Edward Nelson, Ahmet Omurtag, David Owen, Charles Parsons, Thomas Pogge, Vincent Renzi, Scott Shapiro, Mark Steiner, and Robert Vaught for their thoughtful comments on an early version of the book. Those comments have led to significant improvements. And thanks to Thomas Pogge for his substantial help in correcting my translations from German. Any remaining mistakes are, of course, my own.

Thanks are also due my parents, Dorothy and Leroy Lavine, not only for their moral support, which I very much appreciated, but also for their generous financial support, without which the preparation of this book would not have been possible. My wife, Caroline, and daughter, Caila, deserve the most special thanks of all, for tolerating with such understanding my absences and the stresses on our family life that the writing of a book inevitably required. This book is dedicated to them.

Contents

I

Introduction

In the latter half of the nineteenth century Georg Cantor introduced the infinite into mathematics. The Cantorian infinite has been one of the main nutrients for the spectacular flowering of mathematics in the twentieth century, and yet it remains mysterious and ill understood.

At some point during the 1870s Cantor realized that sets—that is, collections in a familiar sense that had always been a part of mathematics—were worthy of study in their own right. He developed a theory of the sizes of infinite collections and an infinite arithmetic to serve as a generalization of ordinary arithmetic. He generalized his theory of sets so that it could encompass all of mathematics. The theory has become crucial for both mathematics and the philosophy of mathematics as a result. Unfortunately, Cantor had been naive, as Cantor himself and Cesare Burali-Forti realized late in the nineteenth century and as Bertrand Russell realized early in the twentieth. His simple and elegant set theory was inconsistent—it was subject to paradoxes.

The history of set theory ever since the discovery of the paradoxes has been one of attempting to salvage as much as possible of Cantor's naive theory. Formal axiom systems have been developed in order to codify a somewhat arbitrarily restricted part of Cantor's simple theory, formal systems that have two virtues: they permit a reconstruction of much of Cantor's positive work, and they are, we hope, consistent. At least the axiomatic theories have been formulated to avoid all of the known pitfalls. Nonetheless, they involve certain undesirable features: First, the Axiom of Choice is a part of the theories not so much because it seems true—it is at best controversial—but because it seems to be required to get the desired results. Second, since present-day set theory is *ad hoc*, the result of retreat from disaster, we cannot expect it to correspond in any very simple way to our uneducated intuitions about collections. Those are what got Cantor into trouble in the first place.

We can never rely on our intuitions again. The fundamental axioms of mathematics—those of the set theory that is its modern basis—are to a large extent arbitrary and historically determined. They are the remote and imperfectly inferred remnants of Cantor's beautiful but tragically flawed paradise.

The story I have just told is a common one, widely believed. Not one word of it is true. That is important, not just for the history of mathematics but for the philosophy of mathematics and many other parts of philosophy as well. The story has influenced our ideas about the mathematical infinite, and hence our ideas about mathematics and about abstract knowledge in general, in many deep ways.

Both elementary number theory and the geometry of the Greeks, for all that they are abstract, have clear ties to experience. They are, in fact, often thought to result from idealizing that experience. Modern mathematics, including much of the mathematics of physics, is frequently thought to be abstract in a much more thoroughgoing sense. As I shall put it, modern mathematics is not only abstract but also remote, because it is set-theoretic:[1] The story tells us that modern axiomatic set theory is the product not of idealization but of the failure of an attempted idealization.

Since science and often mathematics are thought of as quintessential examples of human knowledge, modern epistemology tries to come to grips with scientific and mathematical knowledge, to see it as knowledge of a typical or core kind. That poses a serious problem for epistemology, since mathematical knowledge and the scientific knowledge that incorporates it is thought to be so remote.

The whole picture of mathematical knowledge that drives the epistemology is wrong. As this book will demonstrate, set theory, as Cantor and Ernst Zermelo developed it, is connected to a kind of idealization from human experience much like that connected to the numbers or to Euclidean geometry.

Cantor studied the theory of trigonometric series during the 1870s. He became interested in arbitrary sets of real numbers in the process of making

1. When I say that modern mathematics is set-theoretic, I am not referring to the so-called set-theoretic foundations of mathematics, which play little role in this book. What I have in mind is the ubiquitous use of set-theoretic concepts in mathematics, concepts like open set, closed set, countable set, abstract structure, and so on and on. The concepts mentioned were, as we shall see in Chapter III, introduced by Cantor in the course of the same investigations in which he introduced his theory of infinite numbers and their arithmetic.

that theory apply to more general classes of functions. His work was part of a long historical development that had in his day culminated in the idea that a function from the real numbers to the real numbers is just any association—however arbitrary—from each real number to a single other real number, the value of the function. The term *arbitrary* is to make it clear that no rule or method of computation need be involved. That notion of a function is the one we use today.

Cantor's study of the theory of trigonometric series led him to this progression of transfinite "indexes":

$$0, 1, \ldots, \infty, \infty + 1, \infty + 2, \ldots, \infty \cdot 2, \ldots, \infty \cdot 3, \ldots,$$
$$\infty^2, \ldots, \infty^3, \ldots, \infty^\infty, \ldots, \infty^{\infty^\infty}, \ldots.$$

Cantor's set theory began as, and always remained, an attempt to work out the consequences of the progression, especially the consequences for sets of real numbers. Despite the usual story, Cantor's set theory was a theory not of collections in some familiar sense but of collections that can be counted using the indexes—the finite and transfinite ordinal numbers, as he came to call them. Though Cantor came to realize the general utility of his theory for codifying a large part of mathematics, that was never his main goal.

Cantor's original set theory was neither naive nor subject to paradoxes. It grew seamlessly out of a single coherent idea: sets are collections *that can be counted*. He treated infinite collections as if they were finite to such an extent that the most sensitive historian of Cantor's work, Michael Hallett, wrote of Cantor's "finitism." Cantor's theory is a part of the one we use today.

Russell was the inventor of the naive set theory so often attributed to Cantor. Russell was building on work of Giuseppe Peano. Russell was also the one to discover paradoxes in the naive set theory he had invented. Cantor, when he learned of the paradoxes, simply observed that they did not apply to his own theory. He never worried about them, since they had nothing to do with him. Burali-Forti didn't discover any paradoxes either, though his work suggested a paradox to Russell.

Cantor's theory had other problems. It did not, in its original form, include the real numbers as a set. Cantor had, for good reason, believed until the 1890s—very late in his career—that it would include them. (Most everything else I am saying here is known to one or another historian or mathematician, but the claim that Cantor had a smooth theory that broke down in the 1890s is new here. It is argued in detail in §IV.2.)[2] Cantor grafted a new assumption on to his theory as soon as he realized he needed it, an assumption that allowed him to incorporate the real numbers, but the assumption caused big trouble.

The new assumption was his version of what is today the Power Set Axiom. The trouble it caused was that his theory was supposed to be a theory of collections that can be counted, but he did not know how to count the new collections to which the Power Set Axiom gave rise. The whole theory was therefore thrown into doubt, but not, let me emphasize, into contradiction and paradox. It seemed that counting could no longer serve as the key idea. Cantor did not know how to replace it.

Zermelo came to the rescue of Cantor's theory of sets in 1904. He isolated a principle inherent in the notion of an arbitrary function, a principle that had been used without special note by many mathematicians, including Cantor, in the study of functions and that had also been used by Cantor in his study of the ordinal numbers. Zermelo named that principle the Axiom of Choice. Though the principle had been used before Zermelo without special notice, no oversight had been involved: the principle really is inherent in the notion of an arbitrary function. What Zermelo noted was that the principle could be used to "count," in the Cantorian sense, those collections that had given Cantor so much trouble, which restored a certain unity to set theory.

The Axiom of Choice was never, despite the usual story, a source of controversy. Everyone agreed that it is a part of the notion of an arbitrary function. The brouhaha that attended Zermelo's introduction of Choice was a dispute about whether the notion of an arbitrary function was the appropriate one to use in mathematics (and indeed about whether it was a coherent notion). The rival idea was that functions should be taken to be given only by rules, an idea that would put Choice in doubt. The controversy was between advocates of taking mathematics to be about arbitrary functions and advocates of taking mathematics to be about functions given by rules—not about Choice *per se*, but about the correct notion of function. Arbitrary functions have won, and Choice comes with them. There is, therefore, no longer any reason to think of the Axiom of Choice as in any way questionable.

Zermelo's work was widely criticized. One important criticism was that he had used principles that, like Russell's, led to known contradictions. He hadn't. In order to defend his theorem that the real numbers can be "counted,"

---

2. The reference is to Chapter IV, Section 2. A reference to §2 would be a reference to Section 2 of the present chapter.

Zermelo gave an axiomatic presentation of set theory and a new proof of the theorem on the basis of his axioms. The axioms were to help make it clear that he had been working on the basis of a straightforwardly consistent picture all along. That is a far cry from the common view that he axiomatized set theory to provide a consistent theory in the absence of any apparent way out of the paradoxes.

There *was* a theory developed as a retreat from the disastrous Russellian theory and its precursor in Gottlob Frege, namely, the theory of types. But it never had much to do with Cantorian set theory. I discuss it only in so far as that is necessary to distinguish it from Cantorian set theory. In the process of discussing it, I introduce a distinctive use Russell suggested for something like schemas,[3] a use that shows that schemas have useful properties deserving of more serious study. Such a study is a running subtheme of this book.

It did not take long for Thoralf Skolem and Abraham Fraenkel to note that Zermelo's axioms, while they served Zermelo's purpose of defending his theorem, were missing an important principle of Cantorian set theory—what is now the Replacement Axiom. The universal agreement about the truth of the Replacement Axiom that followed is remarkable, since the axiom wasn't good for anything. That is, at a time when Replacement was not known to have any consequences about anything except the properties of the higher reaches of the Cantorian infinite, it was nonetheless immediately and universally accepted as a correct principle about Cantorian sets.

Chapters II–V establish in considerable detail that it is the historical sketch just given that is correct, not the usual one I parodied above, and they include other details of the development of set theory. Just one more sample—the iterative conception of set, which is today often taken to be the conception that motivated the development of set theory and to be the one that justifies the axioms, was not so much as suggested, let alone advocated by anyone, until 1947.

There are three main philosophical purposes for telling the story just sketched. The first is to counteract the baneful influence of the standard account, which seems to have convinced many philosophers of mathematics that our intuitions are seriously defective and not to be relied on and that the axioms of mathematics are therefore to a large extent arbitrary, historically

3. A *schema* is a statement form used to suggest a list of statements. For example, $X = X$, where the substitution class for X is numerals, is a schema that has as instances, among others, $0 = 0$, $1 = 1$, and $2 = 2$.

determined, conventional, and so forth. The details vary, but the pejoratives multiply.

On the contrary, set theory is not riddled with paradoxes. It was never in such dire straits. It developed in a fairly direct way as the unfolding of a more or less coherent conception. (Actually, I think there have been two main strands in the development of the theory, symbolized above by the notion of counting and by Power Set. As I discuss in §V.5, it could be clearer how they fit together. One symptom of our lack of clarity on the issue is the independence of the Continuum Hypothesis. But that is a far cry from the usual tale of woe.)

The second purpose is to show what as a matter of historical fact we know about the Cantorian infinite on the basis of clear and universal intuitions that distinctively concern the infinite. The two most striking cases of things we know about the Cantorian infinite on the basis of intuition are codified as Choice and Replacement. How we could know such things? It seems completely mysterious. The verdict has often been that we do not—our use of Choice and Replacement is to a large extent arbitrary, historically determined, conventional, and so forth. But that is not true to the historical facts of mathematical practice, facts that any adequate philosophy of mathematics must confront. (Allow me to take the liberty of ignoring constructivist skepticism about such matters in the Introduction. I shall confront it in the text.)

The third purpose is to make clearer the nature of intuition—the basis on which we know what we do. I have been using the term *intuition* because it is so familiar, but I do not mean the sort of armchair contemplation of a Platonic heaven or the occult form of perception that the term conjures up for many. Whatever intuition is, it is very important to mathematics:

> In mathematics, as in any scientific research, we find two tendencies present. On the one hand, the tendency toward *abstraction* . . . On the other hand, the tendency toward *intuitive understanding* fosters a more immediate grasp of the objects one studies, a live *rapport* with them, so to speak, which stresses the concrete meaning of their relations.
>
> . . . It is still as true today as it ever was that *intuitive* understanding plays a major role in geometry. And such concrete intuition is of great value not only for the research worker, but also for anyone who wishes to study and appreciate the results of research in geometry. (Page iii of David Hilbert's preface to [HCV52].)

The quotation is from a book about geometry, but the point is far more general.

Just as one scientific theory can displace another because of its superior ability to systematize, one mathematical theory can displace another. Unexpected developments can spawn new theories, which can in turn lead to fruitful developments in old theories and become so intertwined with them that the new and the old become indistinguishable. We shall see examples of those things: The modern notion of a function evolved gradually out of the desire to see what curves can be represented as trigonometric series. The study of arbitrary functions, in the modern sense, led Cantor to the ordinal numbers, which led to set theory. And set theory became so intertwined with the theories of functions and of the real numbers as to transform them completely. That is all a part of the story told in Chapters II and III. Mathematics does not have the same ties to experiment as science, but the way mathematics evolves is nonetheless very similar to the way that science evolves.

The view of mathematics just outlined is usually thought to be antithetical to the possibility of any distinctive sort of mathematical intuition. New mathematics has been thought to evolve out of old without any further constraint than what can be proved. But that cannot have been right for most of the history of modern mathematics: from, say, the first half of the seventeenth century until the second half of the nineteenth there was no coherent systematization or axiomatization for much of mathematics and certainly no adequate notion of proof.

Mathematicians necessarily saw themselves as working on the basis of an intuitive conception, relying to some extent on what was obvious, to some extent on connections with physics, and to some extent—but only to some extent, since proof was not a completely reliable procedure—on proof. (See Chapter II.) I believe that most mathematicians today still see themselves as working in much the same conceptually based and quasi-intuitive way, though that is much harder to show, since rigorous standards of proof and precise axiomatizations are now available. The intuitive conceptions that underlie mathematical theories evolve, as do the theories, but the intuitions both constrain the theories and suggest new developments in them in unexpected ways.

The development of set theory is an excellent example of the positive and necessary role intuition plays in mathematics. Because set theory is in so many respects unlike the mathematics that had gone before, it is clear that prior training was far from an adequate guide for Cantor. Besides, the progression that he found does, in some sense, have clear intuitive content. There is a great and mysterious puzzle in the suggestiveness of Cantor's progression that can hardly be overstated. The progression is infinite, and we have absolutely no experience of any kind of the infinite. So what method are we using—what method did Cantor use—to make sense of the progression? The question is another version of the one raised above about Choice and Replacement.

It is difficult to understand how we can know any mathematical truths at all, since the subject matter of mathematics is so abstract. But the problem is particularly acute for truths about the infinite. There is no doubt that we know that $2 + 2 = 4$ in some sense or other, and that that knowledge is somehow connected to our experience that disjoint pairs combine to form a quadruple. The facts are indisputable and have multifarious connections to human experience. But there *is* genuine doubt about the truth of, say, $\aleph_2 + \aleph_2 = \aleph_2$, because, for example, there is doubt about whether there could be $\aleph_2$ things.[4] Everyone agrees we must in some sense accept that $2 + 2 = 4$, but it is reasonable to be altogether skeptical about the infinite. Worse still, it is not clear what connections to human experience truths about the infinite might have. A modern philosopher of mathematics put it this way:

> The human mind is finite and the set theoretic hierarchy is infinite. Presumably any contact between my mind and the iterative hierarchy can involve at most finitely much of the latter structure. But in that case, I might just as well be related to any one of a host of other structures that agree with the standard hierarchy only on the minuscule finite portion I've managed to grasp. [Mad90, p. 79]

There is a general philosophical problem about knowledge of abstract objects, mathematical objects in particular. But the special case of knowledge of infinite mathematical objects is a distinctive problem for which distinctive solutions have been suggested. Chapters VI and VII are concerned with that problem of the infinite. In Chapter VI, I survey various accounts of mathematical knowledge of the infinite that attempt to show how it can come out of experience. They begin with a theory of knowledge and try to fit mathematics to it. Intuitionism, various forms of formalism, and one version of David Hilbert's program are discussed. I use a Russellian picture of schemas to clarify how Hilbert's finitary mathematics could avoid any commitment to the infinite. It is a consequence of each of the philosophies surveyed that we could not know what we in fact do.

4. The symbol is a capital Hebrew aleph. $\aleph_2$ (pronounced "aleph two") stands for one of Cantor's infinite numbers.