The Bulletin of Symbolic Logic
Volume 1, Number 4, Dec. 1995

INNER MODELS AND LARGE CARDINALS

RONALD JENSEN

In this paper, we sketch the development of two important themes of modern set theory,[1] both of which can be regarded as growing out of work of Kurt Gödel. We begin with a review of some basic concepts and conventions of set theory.

§0. The ordinal numbers were Georg Cantor's deepest contribution to mathematics. After the natural numbers 0, 1, ..., n, ... comes the first infinite ordinal number ω, followed by $\omega+1, \omega+2, \dots, \omega+\omega, \dots$ and so forth. ω is the first *limit ordinal* as it is neither 0 nor a successor ordinal. We follow the von Neumann convention, according to which each ordinal number α is identified with the set $\{\nu \mid \nu < \alpha\}$ of its predecessors. The $\in$ relation on ordinals thus coincides with $<$. We have $0 = \emptyset$ and $\alpha + 1 = \alpha \cup \{\alpha\}$. According to the usual set-theoretic conventions, ω is identified with the first infinite cardinal $\aleph_0$, similarly for the first uncountable ordinal number ω_1 and the first uncountable cardinal number $\aleph_1$, etc.[2] We thus arrive at the following picture:

$$\begin{array}{ccccccc} 0,1,\dots, & \omega, & \omega+1,\dots,\omega+\omega,\dots, & \omega_1, & \dots, & \omega_2, & \dots \\ & \| & & \| & & \| & \\ & \aleph_0 & & \aleph_1 & & \aleph_2 & \end{array}$$

The *von Neumann hierarchy*[3] divides the class V of all sets into a hierarchy of sets V_α indexed by the ordinal numbers. The recursive definition reads: $V_0 = \emptyset$; $V_{\alpha+1} = \mathscr{P}(V_\alpha)$ (where $\mathscr{P}(x) = \{z \mid z \subseteq x\}$ is the *power set* of x);

This article is based on the Gödel Lecture given at the meeting of the Association for Symbolic Logic at Berkeley in January, 1990. A German version of this paper, "Innere Modelle und Große Kardinalzahlen," was published in Jahresbericht der Deutschen Mathematiker-Vereinigung, Jubiläumstagung 100 Jahre DMV (B. G. Teubner, Stuttgart, 1992), pp. 265–280. This English translation was prepared by Andreas Blass. It is published by permission of B. G. Teubner.

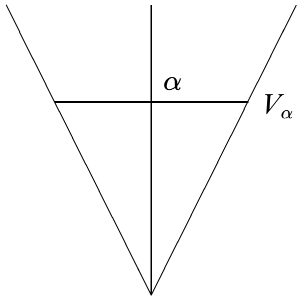
[1] This paper and the lecture on which it is based are not intended as historical surveys. In particular, no attempt has been made to give proper credit for many of the early contributions to the subject.

[2] ω_ν is the νth infinite ordinal number of cardinality greater than all its predecessors and is identified with the corresponding cardinal number $\aleph_\nu$.

[3] This hierarchy was introduced independently by Zermelo.

1079-8986/95/0104-0001/$02.50

$V_\lambda = \bigcup_{\nu<\lambda} V_\nu$ for limit ordinals λ.[4] We can represent this hierarchy by the following picture.



The V_α increase in width as α increases. The ordinal numbers, here drawn as a straight line, grow linearly, since in the transition from α to $\alpha + 1 = \alpha \cup \{\alpha\}$ exactly one new element is added. The *axiom of foundation* in set theory asserts that $V = \bigcup_\alpha V_\alpha$.

This structure has two complementary motivations. The iteration can be understood as a generalization of analysis, since the real numbers can be represented in $\mathscr{P}(\omega)$, subsets and functions on $\mathbb{R}$ in $\mathscr{P}\mathscr{P}(\omega)$, etc. On the other hand, set theory can be regarded as a relativization of finite combinatorics, where the concepts "finite" and "infinite" are replaced by "set" and "proper class." From this point of view it is particularly easy to motivate the usual axioms of set theory,[5] for example the axiom of separation which says that a subclass of a set is a set or the axiom of replacement which says that the image of a set under a function is a set. It is perhaps not a coincidence that the young Cantor was influenced by Kronecker as well as by Weierstraß.

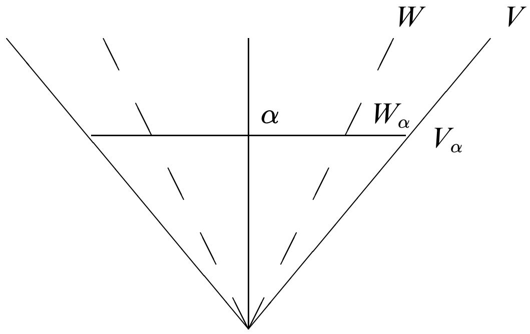
§1. The most famous question in set theory is the continuum problem. The *continuum hypothesis* (CH) asserts that the cardinal number Card($\mathbb{R}$) of $\mathbb{R}$ has the smallest possible value ω_1. An equivalent formulation is $\mathrm{Card}(\mathscr{P}(\omega)) = \omega_1$. The *generalized continuum hypothesis* (GCH) asserts that $\mathrm{Card}(\mathscr{P}(\omega_\nu)) = \omega_{\nu+1}$ for all ν. GCH implies the axiom of choice (AC). Despite great efforts, Cantor did not succeed in deciding whether CH is true, and neither did his successors. In this century, the suspicion grew that CH might be independent of the axioms of set theory, i.e., that these axioms do not suffice to prove either the continuum hypothesis or its negation. This suspicion was based on the apparent insufficiency of the axioms to adequately characterize the power set operation. It proved to be correct, and the two

[4] *Urelements* or non-sets are not needed in mathematics. If we want to have them, we must modify the definition so that they occur in V_0.

[5] By this I mean the axioms of Zermelo-Fraenkel (ZF), possibly extended by the axiom of choice (ZFC).

research directions described in this article can be regarded as attempts to understand power sets better.

The first progress came in 1939, when Gödel showed that GCH is consistent with the axioms. Then the unprovability of CH was shown by Cohen in 1963. Thus the independence was proved. Gödel's proof was the first application of the concept of *inner models*. This concept arises from an attempt to circumscribe the amorphous power set operation: one replaces the V_α-hierarchy with a new hierarchy W_α, defined just like V_α except that, in the transition from W_α to $W_{\alpha+1}$, not all subsets of W_α are adjoined but only those that are obtained by means of certain "well-understood" and clearly specified operations. $W = \bigcup_\alpha W_\alpha$ is then a "slice" of V, as indicated in the following picture.



We can hope that, for a suitable choice of these operations, W will *formally* satisfy all the axioms of set theory. In this case we call W an inner model.[6] Gödel defined the inner model $L = \bigcup_\alpha L_\alpha$ of *constructible* sets. $L_{\alpha+1}$ is the collection of subsets of L_α that can be obtained from the members of $L_\alpha \cup \{L_\alpha\}$ by *constructive* set operations. A pleasant property of this concept is its invariance with respect to different notions of constructivity. (For constructivists disagree as to what "constructive" means.) The ordinal numbers form a non-constructive ingredient in the definition of constructibility, and so we arrive at the same class L by using any "reasonable" notion of constructivity. It follows that many versions of the hierarchy L_α are possible. One frequently used definition sets $L_{\alpha+1} = \mathrm{Def}(L_\alpha)$ = the set of subsets of L_α definable in first-order logic over the structure $\langle L_\alpha, \in\rangle$ with the use of arbitrary parameters from L_α.

Gödel showed that L is an inner model. He also showed that the definition of L is *absolute* in every inner model W in the following sense. If we interpret the definition of L in W rather than in V, then we obtain a class $L_W \subseteq W$. Absoluteness means that $L_W = L$. In particular, it follows that $L_L = L$.

[6]The language of set theory is a first-order language with the predicates $=$ and $\in$. We interpret this language in W by restricting all quantifiers to range over W while the atomic formulas retain the same meaning in W that they have in V.

But then the statement $V = L$ holds in L, for this means just that $V_L = L_L$, where $L_L = L$ and $V_L = \{x \mid x = x\}_L = L$. So the axioms of set theory remain consistent if we adjoin the additional axiom $V = L$, called the *axiom of constructibility*. Gödel deduced GCH from $V = L$. So GCH is consistent relative to the other axioms. From GCH follows AC, which asserts that every set can be well-ordered. Gödel went further and showed, as a consequence of $V = L$, that the real continuum $\mathbb{R}$ has a *definable* well-ordering, and in fact one whose definition can be formulated in second-order number theory.[7] (The precise complexity of this definition is specified by its level, Σ^1_2, in the projective hierarchy.)

Gödel's discovery came shortly before the start of World War II and constituted a coda for the development of naive set theory in the early decades of this century. In the next twenty years, set theory was quiescent. In the sixties, after Cohen's discovery had revived interest in set theory, we discovered a series of strong and applicable combinatorial principles that follow from the axiom $V = L$. These principles, which can be regarded as natural strengthenings of GCH, are used in the infinitary parts of algebra, topology, and combinatorics. (The best known of these principles, called $\diamondsuit$, permits constructions in which the $\geq \omega_2$ subsets of ω_1 are diagonalized in ω_1 steps.)

§2. According to Gödel's theorem, the negation of CH is not provable. It remained to show that CH is not provable. But this could not be shown with the method of inner models, since precisely the success of this method precluded such an application: By absoluteness, we have $L = L_W \subseteq W$ for every inner model W. So $V = L$ says that V is the only inner model. As $V = L$ is consistent and implies CH, we cannot prove the existence of an inner model in which CH is false. To negate CH, we must apparently find an "outer model" that extends V. But how is this possible if V already contains all sets? The solution to this puzzle was given by Cohen's *forcing method*. Although forcing is somewhat outside the scope of this paper, I shall try to give a sketch of it (which the reader may skip without loss of continuity).[8] Instead of working with sets, one can consider their characteristic functions. This suggests an alternative to the V_α hierarchy: We define a hierarchy V^2_α of partial characteristic functions. As before, we set $V^2_0 = \emptyset$ and $V^2_\lambda = \bigcup_{\nu<\lambda} V^2_\nu$ for limit ordinals λ. At successor ordinals, we set $V^2_{\alpha+1} =$ the set of all $f : V^2_\alpha \to 2$. Finally, we define $V^2 = \bigcup_\alpha V^2_\alpha$. Then V^2 is "essentially" isomorphic to V, which can be made precise as follows. Each $f \in V^2$ is defined on a subset V^2_α of V^2. Extend f to a total function f^* on V^2 by

[7] Second-order number theory is the theory of the structure $\langle \mathscr{P}(\omega), \omega, +, \cdot, \in \rangle$. (Real numbers can be identified with elements of $\mathscr{P}(\omega)$.)

[8] Cohen's original version looks somewhat different from the version presented here, which is based on a later analysis by Solovay.

setting $f^*(x) = 0$ wherever $f(x)$ is undefined. Define an *equality relation* I on V^2 by letting fIg iff $f^* = g^*$, and define a pseudo-$\in$ relation E by fEg iff $g^*(f) = 1$. Then I is a congruence relation on $\langle V^2, E\rangle$, and the quotient structure $\langle V^2, E\rangle/I$ is isomorphic to V.

This analysis suggests a strategy for extending V^2. $2 = \{0, 1\}$ is the smallest Boolean algebra. If we carry out the same construction with another complete Boolean algebra $\mathbb{B}$ in place of 2, we obtain a hierarchy $V^{\mathbb{B}}_\alpha$. ($V^{\mathbb{B}}_{\alpha+1}$ is then defined as the set of all $f : V^{\mathbb{B}}_\alpha \to \mathbb{B}$.) We set $V^{\mathbb{B}} = \bigcup_\alpha V^{\mathbb{B}}_\alpha$. Then we have $V^2 \subseteq V^{\mathbb{B}}$ because $2 \subseteq \mathbb{B}$. Somewhat more complicated definitions than before give an "equality relation" I and a "pseudo-$\in$ relation" E. These are, however, not relations in the usual sense but "$\mathbb{B}$-valued relations," i.e., binary functions whose values lie in $\mathbb{B}$. In other words, we now assign to each statement of the form $x = y$ or $x \in y$ (with $x, y \in V^{\mathbb{B}}$) a *truth value* in $\mathbb{B}$. Then we can recursively define, for each statement φ of the language of set theory, a truth value $\|\varphi\| \in \mathbb{B}$. [9] ($\|\varphi\|$ may lie strictly between 0 and 1.) It turns out that all the axioms of set theory have truth value 1 (including AC if we assume AC in V) and the property of having truth value 1 is preserved by logical deduction. However, a suitable choice of $\mathbb{B}$ makes $\|\mathrm{CH}\| = 0$. This shows that CH does not follow from the axioms. (The consistency of CH with the axioms could also have been established by the forcing method, for one can choose $\mathbb{B}$ so that $\|\mathrm{CH}\| = 1$.)

Forcing also led to the discovery of useful combinatorial principles. The best known of these is Martin's Axiom (MA). The strength of MA depends on the size of the continuum. MA is a trivial consequence of CH. If CH is false, then MA has far-reaching consequences (but remains consistent relative to the axioms). MA is often a useful complement to the combinatorial principles that hold in L. For example, the independence of various statements can be proved by showing that one direction follows from $\diamondsuit$ and the other from MA and $\neg$CH.

§3. The forcing method proved to be an extraordinarily powerful tool for obtaining independence results in set theory and led to a renaissance of this field. It was an exciting time for the young mathematicians who flocked into this field. Major classical problems were solved with gratifying regularity. In the long run, though, it is somewhat unsatisfying to prove only that statements are unprovable. Increasingly, people asked whether the weakness of the set-theoretic axioms could be alleviated by accepting new axioms. An obvious candidate would be Gödel's axiom of constructibility, $V = L$. This

---

[9] For example, we set $\|\neg\varphi\| = \neg\|\varphi\|$ =the complement of $\|\varphi\|$ in $\mathbb{B}$; $\|\varphi \wedge \psi\| = \|\varphi\| \cap \|\psi\|$; $\|\forall v\, \varphi(v)\| = \bigcap_{x \in V^{\mathbb{B}}} \|\varphi(x)\|$. Once this definition is given, the $\mathbb{B}$-valued relations are uniquely characterized by the requirements that the axioms of equality logic and the axiom of extensionality have truth value 1 and that $\|x \in y\| \subseteq \bigcup_{z \in \mathrm{dom}(y)} \|x = z\|$ for $x, y \in V^{\mathbb{B}}$.

axiom makes a clear statement about the nature of the mathematical universe. It is mathematically fruitful in that it solves many problems and leads to interesting new concepts and theories. It is philosophically attractive for adherents of "Ockham's razor," which says that one should avoid superfluous existence assumptions. I personally find it a very attractive axiom. Nevertheless, it has been rejected by the majority of set theorists, beginning with Gödel himself. One could perhaps summarize Gödel's objection as follows. The universe of sets is too large and varied to be fully grasped by the human spirit. Over time, however, mathematical intuition improves and with it our ability to describe this universe. The steps of this approximation process are marked by the axioms that we recognize as true. The axiom of constructibility presents, modulo the ordinal numbers, a complete description of sets and is thus a limiting principle that would halt the approximation process. Instead, one should look for additional existence axioms, called strong axioms of infinity. A typical axiom of this sort asserts the existence of a large cardinal number. As an example, Gödel suggests the inaccessible cardinals. As the universe satisfies all axioms of set theory, it should be so big that even some initial segment V_κ also satisfies the axioms. If V_κ satisfies the axioms (in the sense of second order logic[10]), then we call κ *inaccessible*. But if an inaccessible cardinal exists, then this statement also should be true in an initial segment V_τ, i.e., there should be two inaccessible cardinals κ and τ. By continuing this argument, we see that the inaccessible cardinals are unbounded in the ordinal numbers. But then also this assertion should hold in an initial segment, i.e., there is an inaccessible limit of inaccessible cardinals, etc.

Inaccessible cardinals are, however, not an alternative to the axiom of constructibility, for they are consistent with this axiom.[11] For this we need a principle saying that the universe is not only "long" but also "wide." The first large cardinal axiom proved to imply $V \neq L$ asserts the existence of a *measurable* cardinal. As the name implies, these cardinals originated from questions in measure theory. κ is measurable if and only if there is a κ-complete, non-trivial, 0,1-valued measure defined on all subsets of κ. (Or, equivalently, there is a κ-complete non-principal ultrafilter on κ.) The set-theoretic content of this concept becomes clearer in the following definition.

Definition. κ is *measurable* if and only if there exist M and π such that
(1) M is an inner model.
(2) $\pi : V \to M$ is an elementary embedding.[12]
(3) κ is the critical point of π (i.e., $\pi \restriction \kappa = \mathrm{id}$, $\pi(\kappa) > \kappa$).

[10] This means that V_κ is used as the domain of sets and $V_{\kappa+1}$ as the domain of classes.
[11] If κ is inaccessible in V, then also in L.
[12] That is, $\varphi(x_1, \dots, x_n)$ is true in V if and only if $\varphi(\pi(x_1), \dots, \pi(x_n))$ is true in M.

In the late fifties, Dana Scott showed that the existence of a measurable cardinal implies $V \neq L$. Later, it was shown that these cardinals have interesting consequences for the real continuum, e.g., that every Σ^1_2 set is Lebesgue measurable and has the Baire property. (Consequently, there can be no Σ^1_2 well-ordering of $\mathbb{R}$.)

We can generate a hierarchy of even stronger large cardinal principles by imposing additional conditions on the model M in the definition above. For example, we call κ *α-strong* iff $V_\alpha \subseteq M$. We call κ *strong* if it is α-strong for all α. κ is *superstrong* iff $V_{\pi(\kappa)} \subseteq M$. (This axiom is stronger than the previous one in the sense of *relative consistency* , i.e., the existence of superstrong cardinals implies the consistency of strong cardinals, but not conversely.)

These cardinals are the first stages of a hierarchy that has been intensively investigated in the last 20 years. Among the later stages are the compact and huge cardinals. The strongest principle of this sort was proposed by Reinhardt. We call κ a *Reinhardt cardinal* if $M = V$ in the definition above. At first, Reinhardt cardinals were greeted by many set theorists, including the author of this article, as the embodiment of an intuitively especially satisfactory existence principle. But after a few years Kunen proved that these cardinals do not exist. This set an upper limit for the hierarchy.

§4. Strong axioms of infinity say that the universe of sets is very large, and, as we have seen, they have important consequences for the real continuum. But one can also formulate axioms saying very directly that the continuum is large. The most successful principles of this sort are the so-called *axioms of determinacy*.

Let ω^ω be the set of all infinite sequences of natural numbers. To every set $A \subseteq \omega^\omega$ we associate a game as follows: Player I chooses a number $a_0 \in \omega$; Player II then chooses $a_1 \in \omega$; Player I chooses a_2; etc. After ω moves, the players have defined a sequence $a = \langle a_i | i < \omega\rangle$. Player I wins if $a \in A$; otherwise II wins. We say that A is *determined* if one of the two players has a *winning strategy*. A winning strategy for I, for example, is a function f, from finite sequences of natural numbers to natural numbers, such that I always wins if he plays $a_{2i} = f(a_1, a_3, \dots, a_{2i-1})$ for all i. Strategies can be coded by real numbers, so the existence of a strategy for A is an existence postulate about $\mathbb{R}$. The simplest form of the axiom of determinacy (AD) asserts that every $A \subseteq \omega^\omega$ is determined. But AD implies that $\mathbb{R}$ cannot be well-ordered, so it cannot be accepted as an axiom. There are, however, various axioms of *definable* determinacy, saying that a certain class of "nicely definable" $A \subset \omega^\omega$ are determined. These seem to be consistent with the axiom of choice (AC) and have been intensively studied in the last twenty years. The axiom of *projective determinacy* (PD) says that the projective, i.e., $\mathbf{\Sigma}^1_\omega$, sets are determined. Naturally, $\mathbf{\Sigma}^1_n$-AD means that the $\mathbf{\Sigma}^1_n$ sets are determined.[13] An especially attractive axiom is AD in $L[\mathbb{R}]$, which says that AD holds in the inner model $L[\mathbb{R}]$ that is defined exactly like L except that $L_0[\mathbb{R}] = \mathbb{R}$. This axiom is equivalent to the statement that every $A \in \mathscr{P}(\omega^\omega) \cap L[\mathbb{R}]$ is determined. It turns out that under this assumption the model $L[\mathbb{R}]$ becomes a paradise for analysts: Every set of real numbers is Lebesgue measurable and has the Baire property. AC is false there, but if we assume AC in V then $L[\mathbb{R}]$ still satisfies the weaker forms of AC—such as the axiom of dependent choice—that are needed by analysts.

It was soon discovered that there is a connection between determinacy and large cardinals. For example, AD implies the existence of an inner model with a measurable cardinal. (In fact, this cardinal is ω_1 in V.) In the other direction, $\mathbf{\Sigma}^1_1$-AD follows from the existence of a measurable cardinal. The great hope of researchers in both fields was to deduce definable determinacy from the existence of sufficiently large cardinals. Until recently, however, this hope remained unfulfilled, and the three main areas of set theory—inner models, large cardinals, and determinacy—went their separate ways. Large cardinals probably attracted the most attention and were certainly the subject of very deep investigations. Nevertheless, this area remained in a certain sense the least satisfactory: It lacked—until recently—both the deep structure theory and the occasional dramatic shifts of perspective of the other two areas.

§5. The axiom of constructibility on the one hand and large cardinals and determinacy on the other embody two radically different conceptions of the universe of sets. How can these conceptions be justified? Most proponents of $V = L$ and similar axioms support their belief with a mild version of Ockham's razor. L is adequate for all of mathematics; it gives clear answers to deep questions; it leads to interesting mathematics. Why should one assume more? The proponents of strong axioms of infinity usually call themselves Platonists and use this to support their belief. I do not understand, however, why a belief in the objective existence of sets obligates one to seek ever stronger existence postulates. Why isn't Platonism compatible with the mild form of Ockham's razor cited above? Something else must be going on. I would like to propose a—necessarily somewhat speculative—hypothesis.[14] Could it be that the duality in modern set theory is nothing but a new manifestation of an ancient conflict between two points of view—I almost want to say two emotional states—which have always existed in mathematics? I call them the arithmetical and geometrical points of view. I also call the first

[13] $\mathbf{\Sigma}^1_0$ consists of the Borel sets. $\mathbf{\Sigma}^1_{n+1}$ consists of projections of complements of $\mathbf{\Sigma}^1_n$ sets. $\mathbf{\Sigma}^1_\omega = \bigcup_n \mathbf{\Sigma}^1_n$.

[14] The reader can skip the rest of §5 without loss of continuity.

one the Pythagorean point of view, for Pythagoras expressed it in its purest form: Everything consists of numbers. In other words, every mathematical structure can be interpreted in the natural structure of the positive integers. This idea is naturally very attractive; it gives to all of mathematics the intuitive clarity of the natural numbers. I conjecture that, if it could really be carried out, it would still be the dominant point of view today. In reality, however, the geometric point of view has been dominant since the rise of analysis. Thus I also call it the Newtonian point of view.[15] The Newtonian directs his gaze to the real rather than the natural numbers. He is less impressed by their clarity than by their boundless multiplicity. The real numbers constitute a gigantic, unfathomable sea. For every principle that generates real numbers, there must be a number not attainable by that principle. This excludes the possibility of an interpretation of the real numbers within the natural numbers.

If one accepts set theory, then there is no doubt that Cantor has refuted Pythagoreanism in the strongest terms by showing that there are more real than natural numbers. But Cantor also introduced the ordinal numbers, which are in every sense the transfinite continuation of the natural numbers. They share much of the intuitive clarity of the natural numbers. Thus Cantor, who refuted the old Pythagoreanism, made possible a new Pythagoreanism in which the ordinal numbers take over the role of the natural numbers. In this sense, Gödel's axiom of constructibility seems to me to embody an entirely coherent Pythagorean picture of the world. And this picture cannot be refuted, for Gödel showed that $V = L$ is consistent if the other axioms are. But this axiom provides—modulo the ordinal numbers—a complete description of all sets and is therefore unacceptable for a Newtonian. For him, there must be a real number not generated by constructible processes. Since one cannot prove in set theory that such a number exists, one must seek new axioms. Thus, the ancient conflict is fought in a new arena.

Whether or not one accepts this analysis, there is no doubt that the quest for new axioms of infinity has led to very interesting mathematics. In addition, the proponents of these axioms have, through hard technical work, assembled a series of impressive plausibility arguments.[16] Nevertheless, I doubt that one could, with the sort of evidence that we have, convert the mathematical world to one or the other point of view.[17] Deeply rooted differences in mathematical taste are too strong and would persist.[18]

[15]Of course one could also call it Leibnizian. But perhaps Newton thought more geometrically.

[16]See §7, §8.

[17]The author confesses to being emotionally a Pythagorean.

[18]In the course of time I have had numerous opportunities to discuss the alternatives in set theory with mathematicians from other disciplines. I am always surprised how rapidly and

§6. Strong axioms of infinity should prevent the universe from being "constructed from below" like L. But do they really prevent this? It turns out that we can have a measurable cardinal in a very L-like model. The measurability of κ means that there is a normal ultrafilter U in $\mathscr{P}(\kappa)$.[19] Let the hierarchy L^U_α be defined like L_α except that $L^U_{\alpha+1} = \mathrm{Def}(\langle L^U_\alpha, \varepsilon, U \cap L^U_\alpha \rangle)$, and set $L^U = \bigcup_\alpha L^U_\alpha$. Then L^U is an inner model. In addition, κ is measurable in L^U, for $U \cap L^U \in L^U$ appears in L^U to be a normal ultrafilter. Silver and others have shown that L^U is thoroughly similar to L. It satisfies GCH. It has the combinatorial properties of L. (In fact, until now, no purely combinatorial difference between L and L^U has been discovered.) It has a well-ordering of the reals definable in second-order arithmetic (in this case Σ^1_3 instead of Σ^1_2). Even the absoluteness of L has an analog: L^U has a definition (using κ as a parameter) that is absolute in every inner model M with $L^U \subseteq M$.

But our definition of L^U does not by any means show that this model can be "built from below by well-understood operations," for the object U is explicitly used in the construction. That a construction from below is nevertheless possible to a certain extent was proved only later in the so-called *core model theory*. This theory is somewhat complicated, but I shall try to give a rough sketch of it. Let us assume not only that $V \neq L$ but also that there is a non-trivial elementary embedding $\pi : L \to L$.[20] It turns out that this assumption is equivalent to the existence of a certain set, denoted $0^\sharp$. In a very concrete way—which unfortunately cannot be described here—$0^\sharp$ encodes a complete description of the structure of L together with a canonical elementary embedding $\pi : L \to L$. The definition of $0^\sharp$ is absolute in every inner model that contains $0^\sharp$. Although it unfortunately cannot be justified here, there is no doubt that $0^\sharp$ is a "well-understood" mathematical object. If $0^\sharp$ does not exist, then the global structure of V does not deviate too much from that of L. This is the main content of the so-called "covering theorem": If $0^\sharp$ does not exist then every uncountable set X of ordinal numbers has a "covering set" $Y \supseteq X$ with $Y \in L$ and $\mathrm{Card}(Y) = \mathrm{Card}(X)$. (From this follows, for example, that every singular cardinal number has the same cardinal successor in L as in V.) But if $0^\sharp$ exists, then the structure of L-cardinals is "wiped out" in V. (For example, every uncountable cardinal looks inaccessible in L.) So $0^\sharp$ may be regarded as the "next larger construction step" after L. Then we form the inner model $L[0^\sharp]$, the constructible closure of

---

with what certainty they express an opinion. Their opinions are divided and reflect, in my view, a pre-existing orientation.

[19] U is *normal* if for every regressive function $f : X \to \kappa$ (where "regressive" means $f(\xi) < \xi$ for all $\xi \in X$) with $X \in U$ there is an $\eta < \kappa$ with $\{\xi \mid f(\xi) = \eta\} \in U$. Normality implies κ-completeness.

[20] Non-trivial means $\pi \neq \mathrm{id}$.

$0^\sharp$.[21] Now let us assume that $L[0^\sharp]$ is non-trivially elementarily embeddable in itself by $\pi : L[0^\sharp] \to L[0^\sharp]$. This leads to a set $0^{\sharp\sharp}$ encoding $L[0^\sharp]$ and a canonical embedding π. Then we form $L[0^{\sharp\sharp}]$ and so on. After ω steps, we have a sequence $0^\sharp, 0^{\sharp\sharp}, \ldots, 0^{(n)}, \ldots$. We let $0^{(\omega)}$ encode this sequence and form $L[0^{(\omega)}]$. An embedding $\pi : L[0^{(\omega)}] \to L[0^{(\omega)}]$ yields $0^{(\omega+1)}$, and so on. When we have $0^{(\alpha)}$ for all ordinal numbers α, we form the constructible closure $L^\sharp$ of all the $0^{(\alpha)}$.[22] A non-trivial elementary embedding then yields a set $0^{(\infty)}$, and so forth.[23] We arrive at a sequence: $0^\sharp, 0^{\sharp\sharp}, \ldots, 0^{(\alpha)}, \ldots, 0^{(\infty)}, 0^{(\infty+1)}, \ldots$. The terms in this sequence are called *mice*. With some effort, one can define the class of all mice along with its natural ordering. (This ordering can be much longer than ∞.) But then we can build the constructible closure of all mice, which we denote by K.[24] K is called the *core model*. K is an inner model. The definition is absolute in the sense that $K_M = K$ for every inner model $M \supseteq K$. K is L-like in the same sense as L^U. What happens if there is a non-trivial elementary embedding $\pi : K \to K$? A new mouse cannot result, as all mice are already present in K. What we get is an inner model L^U, where $U \in L^U$ witnesses the measurability in L^U of some ordinal κ. We have $\mathscr{P}(\kappa) \cap K = \mathscr{P}(\kappa) \cap L^U$. U can be obtained in a fairly direct way from the embedding π. So we have, after all, obtained L^U "from below" via the long march through the mice.

The model L^U is uniquely determined by the number κ. Let us now choose κ as small as possible. What happens if there is a non-trivial elementary embedding $\pi : L^U \to L^U$? We obtain a set $0^\dagger$ coding L^U and a canonical embedding π. So, after all, $0^\dagger$ is "the next mouse." It differs from the previous mice in that, to obtain it, we needed an intermediate step marked not with a mouse but with an inner model L^U. This suggests the possibility of a continuing process, producing inner models for stronger and stronger axioms of infinity. How far can we go? Could it be that the directions mentioned in §5, Newtonianism and Pythagoreanism, are compatible after all, with Newtonianism indicating possibilities that then turn out to be realizable in a Pythagorean universe? Progress so far has been relatively modest. The next major step after a measurable cardinal is a strong cardinal. A few years ago, Dodd and the author set themselves the goal of defining a core model for a strong cardinal. Mitchell worked in the same direction but with a more modest goal. The technical difficulties turned out to be very great. Dodd and I barely got past the fundamental fine structure theory. Mitchell got closer

[21] $0^\sharp$ can be regarded as a set of ordinal numbers (even as a subset of ω). Then $L[0^\sharp]$ is defined from $0^\sharp$ as L^U was defined from U.

[22] $L^\sharp = \bigcup_\alpha L[0^{(\alpha)}]$. (Remark: $0^{(\nu)} \in L[0^{(\alpha)}]$ for $\nu \leq \alpha$.)

[23] ∞ is the class of all ordinal numbers and is also regarded as the "largest ordinal number."

[24] K is the union of all $L[m]$ such that m is a mouse.

to his goal. Recently, Mitchell—building on a suggestion of Baldwin—proposed a new definition of mice, which made immense simplifications possible.

The theory is still very long, and not all the details have been written down yet. Nevertheless, I believe I can claim that we now have an extended core model K in which a strong cardinal can exist. How big K is depends, naturally, on what there is. If there is no inner model with a measurable cardinal, then the new K coincides with the old one. If there is such an inner model but $0^\dagger$ does not exist, then $K = L^U$, etc. We again have $K_M = K$ if M is an inner model with $K \subseteq M$. K is again thoroughly L-like. The existence of a strong cardinal in V ensures the existence of one in K. If there is no elementary embedding $\pi : K \to K$, then again much of the global structure of K is preserved in V.[25] But if $\pi : K \to K$ exists, then we obtain the next mouse. So we can go further. But it turned out that one cannot go much further without seriously modifying the program.

§7. A few years ago, Hugh Woodin, in the wake of a deep result of Foreman, Magidor, and Shelah, proved the following theorem: If there is a superstrong cardinal, then there is no well-ordering of $\mathbb{R}$ in $L[\mathbb{R}]$. In particular, $\mathbb{R}$ has no well-ordering definable in second-order arithmetic. This result was surprising, because the concept of a superstrong cardinal had previously been regarded as a rather modest extension of the concept of measurability. The theorem set off an avalanche of results that have fundamentally altered the landscape of set theory. It also set a sharp limit for the conventional core model theory, because the methods that we had been using for the construction of these models always yielded a Σ^1_3 well-ordering of $\mathbb{R}$. Even more serious for this theory was another consequence of Woodin's proof: In a possible modified core model theory, one would have to give up the absoluteness theorem; if K contains a superstrong cardinal, then one cannot exclude the possibility that $K_M \neq K$ for some inner model $M \supseteq K$.[26]

A rapid sequence of discoveries followed. Martin and Steel produced the long-desired proof that determinacy axioms follow from strong axioms of infinity. A typical result—due to Martin, Steel, and Woodin—is that the existence of a superstrong cardinal implies AD in $L[\mathbb{R}]$. An even weaker cardinal—the so-called Woodin cardinal—played a key role in the new developments. Call a cardinal τ *strong with respect to a class A* if for each

[25] But this can no longer be expressed by so strong and simple a statement as the covering theorem.

[26] Using forcing, Woodin showed that, given an inner model M with a superstrong cardinal, one can consistently assume the existence of a strictly larger model M' and an elementary embedding $\pi : M \to M'$. Now let $M = K$. Then $V = K$ holds in M and hence also in M'. Absoluteness would imply $M' = K = M$, a contradiction.

β there exist an inner model M, a class B, and an elementary embedding $\pi : \langle V, A\rangle \to \langle M, B\rangle$ such that τ is the critical point, $V_\beta \subseteq M$, and $B \cap V_\beta = A \cap V_\beta$. If κ is inaccessible, $A \subseteq V_\kappa$, and $\tau < \kappa$, then we can relativize this concept to V_κ. We call κ a *Woodin cardinal* if for each $A \subseteq V_\kappa$ there exists a $\tau < \kappa$ that is strong with respect to A in V_κ. Martin and Steel showed that $\mathbf{\Sigma}^1_{n+1}$-determinacy follows from the existence of n Woodin cardinals and a measurable cardinal above them. Then Woodin showed that the existence of ω Woodin cardinals and a measurable cardinal above them implies AD in $L[\mathbb{R}]$. Woodin also has interesting equiconsistency results: Σ^1_2-determinacy is equiconsistent with the existence of a Woodin cardinal. AD in $L[\mathbb{R}]$ is equiconsistent with the existence of ω Woodin cardinals.

§**8.** The two areas, large cardinals and determinacy, had come together. What about inner models? At first glance, the core model program seemed to have reached a dead end. For two essential attributes of this theory would have to be sacrificed in any extension: a nicely definable well-ordering of $\mathbb{R}$ and the absoluteness theorem. The second sacrifice seemed especially serious to me. K is defined as the union of its construction steps, where a construction step (e.g., a mouse) should be a "well understood" object. For me, this requirement always meant that one can recognize in an arbitrary inner model M whether a given element of M is a construction step. That implies $K_M = K$ if $K \subseteq M$. But this conclusion is false if a superstrong (or just a Woodin) cardinal occurs in K.

Fortunately, Martin, Mitchell, and Steel did not let these considerations trouble them but set out to find a concept of mouse suitable for the continuation of the theory. The first step in this extended core model program was to realize a Woodin cardinal in a core model. Martin and Steel first defined a "model of type L^U." Assuming a Woodin cardinal κ, they defined an inner model W. The construction of W is not really "from below" but uses a set whose existence follows from the assumed property of κ. They showed that, in W, κ remains a Woodin cardinal, CH holds, and $\mathbb{R}$ has a Σ^1_3 well-ordering. (So $\mathbf{\Sigma}^1_2$-determinacy is false in W.) But the internal structure of this model remained quite mysterious; for example we do not know whether GCH holds in it. Mitchell and Steel then created a revised W using the fine structure techniques of core model theory. Their W satisfies GCH and is quite generally L-like. Finally, Steel managed to construct a genuine core model K. But to carry out his construction, he needs third order set theory and a large cardinal property (namely measurability) for ∞ (the class of all ordinals).[27] When one has this, one finds a first order definition of K that

[27]Steel thinks that second order set theory will suffice and that the measurability assumption can be weakened.

is absolute in K. At the moment, only a rather weak analog of the covering theorem has been proved.[28]

The mice studied in this extended theory are still "well understood" objects, but the definition of mice is of such high logical complexity that the absoluteness requirement is no longer satisfied.[29] The possibility that there are models M with $K \subseteq M$ and $K_M \neq K$ turned out to be less troublesome than one had expected. The increasing complexity of the concept of mouse also explains why we lose the nicely definable well-orderings of $\mathbb{R}$. The existence of such a well-ordering depends on the fact that every mouse that constructs a new real number is itself a countable object and so can be coded by a real number. The earlier mice were so simple that one could define the set of all such codes—together with their natural well-ordering—in second order arithmetic. With increasing complexity, this is no longer possible.

Will it be possible to realize all reasonable large cardinals in "well understood" inner models? This goal is important not only for the reasons indicated in §6 but also because of a consideration connected with Gödel's justification for strong axioms of infinity. A common plausibility argument for these axioms is that they are—as far as we can see—linearly ordered by the relation of relative consistency.[30] Historically, the axioms of infinity have many different mathematical sources, and there is no *prima facie* reason to suppose that they are linearly ordered. The existence of such an order—insofar as we can confirm it—is surely a strong argument for regarding these axioms as successive approximations to a "final" universe. The axioms in the sequence mentioned earlier—measurable, strong, superstrong, etc.—arise from weaker and stronger variations of a single theme. So it is not surprising that they are linearly ordered. We can regard this sequence as a yardstick with which to compare other strong axioms of infinity. (It turns out that such comparison is sensible even for some assertions that seem to have nothing to do with large cardinals, for example the existence of a saturated ideal on ω_1.) Very many open questions remain. Perhaps the best known is whether compact and supercompact cardinals are equiconsistent. If we go "down" a bit and consider principles of infinity that are weaker than measurability yet strong enough to imply $V \neq L$, then the so-called Erdős cardinals form an analogous yardstick. Here the core model is available, and it frequently allows us to compare other principles with this yardstick.

[28]By this I mean theorems saying that the global structure of K is preserved in V if K is rigid with respect to elementary embeddings.

[29]That the absoluteness condition held previously is due to the fact that the concept of mouse was Δ_1 in the parameter ω_1.

[30]An axiom A is *consistent relative to B* in the axiom system of set theory (ZFC) if it is provable in first order arithmetic that the consistency of ZFC+A follows from the consistency of ZFC+B. Modulo equiconsistency, this relation is a partial order.

(Typical examples are the Chang conjecture[31] and the existence of Jónsson cardinals.[32]) The only strategy we have for attaining comparable results for larger cardinals is a corresponding generalization of the inner model theory.

I would like to close with a general comment on the direction of research in set theory. In recent years I have occasionally heard other mathematicians complain that set theory has turned inward, i.e., that set theorists are occupied too much with their own structural questions and not enough with applications. This reproach does not by any means apply to all set theorists, but this paper does little to refute it. I hope, however, to have made it plausible that a subject should from time to time turn inward. The deep inner problems have become steadily clearer in recent years, and the tools needed to overcome them have been developed. There has been enormous progress recently. If set theory now finds itself in a rather introspective phase, I am convinced that it will emerge from this phase stronger than ever and of greater use for other areas of mathematics.

REFERENCES

[1] K. DEVLIN, ***Constructability***, Springer-Verlag, 1984.

[2] F. R. DRAKE, ***Set theory: An introduction to large cardinals***, North-Holland, 1974.

[3] R. JENSEN, *Measures of order 0*, unpublished manuscript.

[4] K. KUNEN, ***Set theory: An introduction to independence proofs***, North-Holland, 1980.

[5] D. A. MARTIN and J. R. STEEL, *A proof of projective determinacy*, ***Journal of the American Mathematical Society***, vol. 2 (1989), pp. 71–125.

[6] ———, *Iteration trees*, ***Journal of the American Mathematical Society***, vol. 7 (1994), pp. 1–73.

[7] W. MITCHELL and J. R. STEEL, *Fine structure and iteration trees*, ***Lecture notes in logic 3***, Springer-Verlag, 1994.

[8] J. R. STEEL, *The core model iterability problem*, unpublished manuscript.

[9] J. R. STEEL, *Inner models with many Woodin cardinals*, ***Annals of Pure and Applied Logic***, vol. 65 (1993), pp. 185–209.

ALL SOULS COLLEGE
OXFORD OX1 4AL
GREAT BRITAIN

[31] The Chang conjecture is the following model-theoretic assertion: If $\langle A, B, \dots \rangle$ is a model for a countable language with $\mathrm{Card}(A) = \omega_2$ and $\mathrm{Card}(B) = \omega_1$, then there is an elementary submodel $\langle \bar{A}, \bar{B}, \dots \rangle$ with $\mathrm{Card}(\bar{A}) = \omega_1$ and $\mathrm{Card}(\bar{B}) = \omega$.

[32] κ is a Jónsson cardinal if every algebra $A = \langle |A|, f_1, f_2, \dots, f_n \rangle$ of cardinality κ has a proper subalgebra of the same cardinality.