THE JOURNAL OF SYMBOLIC LOGIC
Volume 53, Number 2, June 1988

BELIEVING THE AXIOMS. I

PENELOPE MADDY

Nothing venture, nothing win,
Blood is thick, but water's thin.
—Gilbert & Sullivan

§0. Introduction. Ask a beginning philosophy of mathematics student why we believe the theorems of mathematics and you are likely to hear, "because we have proofs!" The more sophisticated might add that those proofs are based on true axioms, and that our rules of inference preserve truth. The next question, naturally, is why we believe the axioms, and here the response will usually be that they are "obvious", or "self-evident", that to deny them is "to contradict oneself" or "to commit a crime against the intellect". Again, the more sophisticated might prefer to say that the axioms are "laws of logic" or "implicit definitions" or "conceptual truths" or some such thing.

Unfortunately, heartwarming answers along these lines are no longer tenable (if they ever were). On the one hand, assumptions once thought to be self-evident have turned out to be debatable, like the law of the excluded middle, or outright false, like the idea that every property determines a set. Conversely, the axiomatization of set theory has led to the consideration of axiom candidates that no one finds obvious, not even their staunchest supporters. In such cases, we find the methodology has more in common with the natural scientist's hypotheses formation and testing than the caricature of the mathematician writing down a few obvious truths and proceeding to draw logical consequences.

The central problem in the philosophy of natural science is when and why the sorts of facts scientists cite as evidence really are evidence. The same is true in the case of mathematics. Historically, philosophers have given considerable attention to the question of when and why various forms of logical inference are truth-preserving. The companion question of when and why the assumption of various axioms is justified has received less attention, perhaps because versions of the "self-evidence" view live on, and perhaps because of a complacent if-thenism. For whatever reasons, there has been little attention to the understanding and

Received March 4, 1987.

0022-4812/88/5302-0014/$04.10

classification of the sorts of facts mathematical scientists cite, let alone to the philosophical question of when and why those facts constitute evidence.

The question of how the unproven can be justified is especially pressing in current set theory, where the search is on for new axioms to determine the size of the continuum. This pressing problem is also the deepest that contemporary mathematics presents to the contemporary philosopher of mathematics. Not only would progress towards understanding the process of mathematical hypothesis formation and confirmation contribute to our philosphical understanding of the nature of mathematics, it might even be of help and solace to those mathematicians actively engaged in the axiom search.

Before we can begin to investigate when and why the facts these mathematicians cite constitute good evidence, we must know what facts those are. What follows is a contribution to this preliminary empirical study (thus the reference to "believing" rather than "knowing" in my title). In particular, I will concentrate on the views of the Cabal seminar, whose work centers on determinacy and large cardinal assumptions.[1] Along the way, especially in the early sections, the views of philosophers and set theorists outside the group, and even opposed to it, will be mentioned, but my ultimate goal is a portrait of the general approach that guides the Cabal's work.[2]

Because of its length, this survey appears in two parts. The first covers the axioms of ZFC, the continuum problem, small large cardinals and measurable cardinals. The second concentrates on determinacy hypotheses and large large cardinals, and concludes with some philosophical observations.

§I. The axioms of ZFC. I will start with the well-known axioms of Zermelo-Fraenkel set theory, not so much because I or the members of the Cabal have anything particularly new to say about them, but more because I want to counteract the impression that these axioms enjoy a preferred epistemological status not shared by new axiom candidates. This erroneous view is encouraged by set theory texts that begin with "derivations" of ZFC from the iterative conception, then give more self-conscious discussions of the pros and cons of further axiom candidates as they arise. The suggestion is that the axioms of ZFC follow directly from the concept of set, that they are somehow "intrinsic" to it (obvious, self-evident), while other axiom candidates are only supported by weaker, "extrinsic" (pragmatic, heuristic) justifications, stated in terms of their consequences, or intertheoretic connections, or

[1] Naturally the various members of the Cabal do not agree on everything. When appropriate, I will take note of these disagreements.

[2] I am indebted to John Burgess, for introducing me to much of the material discussed here, to Matt Foreman, Menachem Magidor, Yiannis Moschovakis, John Steel, and Hugh Woodin, for helpful conversations, to Chris Freiling, Stewart Shapiro, John Simms, and an anonymous referee for helpful comments on earlier drafts, and, especially, to Tony Martin, without whose patience and generosity the project would have been dead in the water. Versions of this work have been delivered to helpful audiences at the University of California at Los Angeles (Department of Mathematics) and Irvine (Department of Philosophy), and at conventions of the Philosophy of Science and the American Philosophical Associations. The support of NSF Grant No. SES-8509026 and the hospitality of the UCLA Math Department are also gratefully acknowledged.

explanatory power, for example. (It is these extrinsic justifications that often mimic the techniques of natural science.) Thus some mathematicians will stand by the truth of any consequence of ZFC, but dismiss additional axioms and their consequences as metaphysical rot. Even the most cursory look at the particular axioms of ZFC will reveal that the line between intrinsic and extrinsic justification, vague as it might be, does not fall neatly between ZFC and the rest. The fact that these few axioms are commonly enshrined in the opening pages of mathematics texts should be viewed as an historical accident, not a sign of their privileged epistemological or metaphysical status.

The impulse towards axiomatization can be seen as beginning in 1883, when Cantor introduced "a law of thought",

> ... fundamental, rich in consequences, and particularly marvelous for its general validity ... It is always possible to bring any *well-defined* set into the *form* of a *well-ordered* set.

(Cantor [1883, p. 550], as translated in Moore [1982, p. 42]). Hallett [1984, pp. 156–157] traces Cantor's belief in the well-ordering principle to his underlying conviction that infinite sets are not so different from finite ones, that the most basic properties are ones they share. (This will be called "Cantorian finitism" in what follows. In this case, the basic property shared is "countability" or "enumerability".) Unfortunately, the mathematical community at large did not find it obvious that infinite sets could be well-ordered, and by 1895, Cantor himself came to the conclusion that his principle should really be a theorem.

Though Cantor made various efforts to prove this and related theorems (see, for example, his famous letter to Dedekind [1899]), the first proof was Zermelo's in [1904]. This proof, and especially the Axiom of Choice on which it was based, created a furor in the international mathematical community. Under the influence of Hilbert's axiomatics, Zermelo hoped to secure his proof by developing a precise list of the assumptions it required, and proposing them (in [1908]) as an axiomatic foundation for the theory of sets. The fascinating historical ins and outs of this development are clearly and readably described in Moore's book. The point of interest here is that the first axioms for set theory were motivated by a pragmatic desire to prove a particular theorem, not a foundational desire to avoid the paradoxes.[3]

For our purposes, it will be enough to give a brief survey of the arguments given by Zermelo and later writers in support of the various axioms of ZFC.

I.1. *Extensionality.* Extensionality appeared in Zermelo's list without comment, and before that in Dedekind's [1888, p. 45]. Of all the axioms, it seems the most "definitional" in character; it distinguishes sets from intensional entities like

[3] See Moore [1982]. Apparently Zermelo discovered the paradox some two years before Russell. On his interpretation, it shows only that no set can contain all its subsets as members (see Moore [1982, p. 89]). Recall that Cantor also took the paradoxes less seriously than the philosophers, for example, in the letter to Dedekind [1899]. Gödel also expresses the view that the paradoxes present a problem for logic, not for mathematics (Godel [1944/67, p. 474]).

properties or concepts. Most writers seem to echo the opinion of Boolos [1971, p. 501], that if any sense can be made of the distinction between analytic and synthetic, then the Axiom of Extensionality should be counted as analytic. (See also Shoenfield [1977, p. 325], and Wang [1974, p. 533].)

Fraenkel, Bar-Hillel and Levy give a bit more in their [1973, pp. 28, 87]. They argue that an extensional notion of set is preferable because it is simpler, clearer, and more convenient, because it is unique (as opposed to the many different ways intensional collections could be individuated), and because it can simulate intensional notions when the need arises (e.g. two distinguishable "copies" of an extensional set can be produced by taking its cross product with distinct singletons). Thus extrinsic reasons are offered even for this most obvious of axioms.

I.2. *Foundation.* Zermelo used a weak form of the Axiom of Foundation ($A \notin A$) to block Russell's paradox in a series of lectures in the summer of 1906, but by 1908, he apparently felt that the form of his Separation Axiom was enough by itself, and he left the earlier axiom off his published list. (See Moore [1982, p. 157]; Hallett [1984, p. 252].) Later Mirimanoff [1917] defined "ordinary sets" to be those without infinite descending epsilon chains. Using the notion of rank, he was able to formulate necessary and sufficient conditions for the existence of ordinary sets. Though he did not suggest that the ordinary sets are all the sets, he did think that restricting attention to them (in effect adopting Foundation) was a good working method.

This attitude towards Foundation is now a common one. It is described as weeding out "pathologies" or "oddities" (Boolos [1971, p. 491]) on the grounds that

> ... no field of set theory or mathematics is in any ~~general~~ need of sets which are not well-founded.

(Fraenkel, Bar-Hillel and Levy [1973, p. 88]) Von Neumann adopted it in [1925], hoping to increase the categoricity of his axioms, and Zermelo included it in [1930] because it was satisfied in all known applications of set theory and because it gives a useful understanding of the universe of sets. (Supporters of the "iterative conception" discussed below often see foundation as built into the very idea of the stages. See Boolos [1971, p. 498]; Shoenfield [1977, p. 327].)

I.3. *Pairing and Union.* Cantor first stated the Union Axiom in a letter to Dedekind in 1899 (see Moore [1982, p. 54]), and the Pairing Axiom superseded Zermelo's 1908 Axiom of Elementary Sets when he presented the modified version of his axiom system in [1930]. Both are nearly too obvious to deserve comment from most commentators. When justifications are given, they are based on one or the other of two rules of thumb. These are vague intuitions about the nature of sets, intuitions too vague to be expressed directly as axioms, but which can be used in plausibility arguments for more precise statements. We will meet with a number of these along the way, and the question of their genesis and justification is of prime importance. For now, the two in question are *limitation of size* and *the iterative conception.*

Limitation of size came first. Hallett [1984] traces it to Cantor, who held that transfinites are subject to mathematical manipulation much as finites are (as mentioned above), while the absolute infinity (all finites and transfinites) is God and incomprehensible. Later more down-to-earth versions like Fraenkel's hold that the paradoxes are generated by postulating sets that are "too large", and that set theory will be safe if it only eschews such collections. (Hallett gives a historical and philosophical treatment of the role of this rule of thumb in the development of modern axiomatic set theory.) Thus, for example, Fraenkel, Bar-Hillel and Levy [1973] argue that a pair set is of "very modest size", and that the Union Axiom will not produce any thing "too large", because

> ... the sets whose union is to be formed will not be taken arbitrarily—they must be members of a single given set. (pp. 32–34)

(Hallett, incidentally, disagrees about Union. See [1984, pp. 209–210].)

The iterative conception originated with Zermelo [1930] (prefigured perhaps in Mirimanoff [1917]). Although Cantor, Fraenkel, Russell [1906], Jordain [1904], [1905], von Neumann [1923] and others all appealed to *limitation of size*, *the iterative conception* is more prevalent today. Because of its general familiarity, I shall not pause to describe it here. (See e.g. Boolos [1971] or Shoenfield [1977].) For the record, then, given two objects a and b, let A and B be the stages at which they first appear. (On the iterative picture, everything appears at some stage.) Without loss of generality, suppose B is after A. Then the pair set of a and b appears at the stage immediately following B. Similarly, if a family of sets f appears at stage F, then all members of f, and hence all members of members of f, appear before F. Thus the union of f appears at or before F. (Arguments of this form are given in Boolos [1971, p. 496] and Shoenfield [1977, p. 325].)

I.4. *Separation.* The Axiom of Separation is in many ways the most characteristic of Zermelo's axioms. Here he sees himself as giving us as much of the naive comprehension scheme as possible without inconsistency [1908, p. 202]. We see here the emergence of another rule of thumb: *one step back from disaster*. The idea here is that our principles of set generation should be as strong as possible short of contradiction. If a natural principle leads to contradiction, this rule of thumb recommends that we weaken it just enough to block the contradiction. We shall meet this principle again in [BAII, §VI.3].

Zermelo steps back in two ways. First,

> ... sets may never be *independently defined* ... but must always be *separated* as subsets of sets already given.

[1908, p. 202]. Predictably, Fraenkel, Bar-Hillel and Levy see this as the result of applying *limitation of size* to unlimited comprehension (p. 36). Zermelo's second modification is to require that the separating property be "definite" (p. 202), which he understood as ruling out such troublesome turns of natural language as "definable in a finite number of words". The vagueness of the term "definite" brought Zermelo's Separation Axiom under considerable fire until Skolem suggested that "definite" be replaced by "formula of first-order logic". (Even then, Zermelo himself held to a second-order version. See his [1930].)

Advocates of *the iterative conception* have no trouble with Separation: all the members of a are present before a, so any subset of a appears at or before the stage at which a itself appears (Boolos [1971, p. 494]; Shoenfield [1977, p. 325]).[4]

I.5. *Infinity.* The Axiom of Infinity is a simple statement of Cantor's great breakthrough. The rather colorless idea of a collection of elements that had lurked in the background of mathematical thought since prehistory might have remained there to this day if Cantor had not had the audacity to assume that they could be infinite. This was the bold and revolutionary hypothesis that launched modern mathematics; it should be seen as nothing less.

Hallett in his historical study of Cantorian thought, enshrines Cantor's perspective into a rule of thumb called *Cantorian finitism*: infinite sets are like finite ones. (This was mentioned above in connection with Cantor's belief in the well-ordering principle.) The rule and its applications are justified in terms of their consequences. In this case:

> Dealing with natural numbers without having the set of all natural numbers does not cause more inconvenience than, say, dealing with sets without having the set of all sets. Also the arithmetic of the rational numbers can be developed in this framework. However, if one is already interested in analysis then infinite sets are indispensable since even the notion of a real number cannot be developed by means of finite sets only. Hence we have to add an existence axiom that guarantees the existence of an infinite set.

(Fraenkel, Bar-Hillel and Levy [1973, p. 45]). *Iterative conception* theorists now often take the existence of an infinite stage as part of the intuitive picture (see Boolos [1971, p. 492]; Shoenfield [1977, p. 324]), but this would hardly have come to pass if Cantor had not taken a chance and succeeded in showing that we can reason consistently about the infinite and that we have much to gain by doing so (see epigraph).

I.6. *Power set. Cantorian finiteness* yields an argument for the Power Set Axiom, as it is presumably uncontroversial that finite sets have power sets. *The iterative conception* also makes quick work. If a appears at A, then all the elements of a appear before A, so any subset of a appears at or before A. Thus the power set of a appears at the stage after A. Advocates of *limitation of size* suggest that the power set of a given set will not be large because all its members must be subsets of something small.

---

[4] Wang [1974] has a more philosophical account of the iterative picture in terms of what we can "run through in intuition". Thus his justification of Separation is:

> Since x is a set, we can run through all [redacted] members of x, and, therefore, we can do so with arbitrary omissions. In particular, we can in an idealized sense check against A and delete only those members of x which are not in A. In this way, we obtain an overview of all the objects in A and recognize A as a set. (p. 533)

Parsons [1977] points out that this puts a terrible strain on the notion of intuition, and that the problem becomes worse in the case of the Power Set Axiom. See also their exchange on Replacement.

Hallett casts some well-deserved doubt on this last form of justification for the Power Set Axiom, but he does not mean to reject the axiom entirely. Instead, he resorts to a series of extrinsic justifications, the simplest of which is reminiscent of that given above by Fraenkel, Bar-Hillel and Levy for Infinity, namely, that Power Set is indispensable for a set-theoretic account of the continuum:

> This does not prove the legitimacy of the power-set principle. For the argument is not: we have a perfectly clear intuitive picture of the continuum, and the power-set principle enables us to capture this set-theoretically. Rather, the argument is: the power-set principle ... was revealed in our attempts to make our intuitive picture of the continuum analytically clearer; in so far as these attempts are successful, then the power-set principle gains some confirmatory support. (p. 213)

Not surprisingly, a similar extrinsic support for the Power Set Axiom is to be found in Fraenkel, Bar-Hillel and Levy (pp. 34–35).

I.7. *Choice.* The Axiom of Choice has easily the most tortured history of all the set-theoretic axioms; Moore in [1982] makes it a fascinating story. In this case, intrinsic and extrinsic supports are intertwined as in no other. Zermelo, in his passionate defense, cites both. He begins:

> ... how does Peano [one of Zermelo's critics] arrive at his own fundamental principles and how does he justify their inclusion...? Evidently by analyzing the modes of inference that in the course of history have come to be recognized as valid and by pointing out that the principles are intuitively evident [intrinsic] and necessary for science [extrinsic]—considerations that can all be urged equally well in favor of [the Axiom of Choice]. [1908, p. 187]

First the intrinsic supports predominate:

> That this axiom, even though it was never formulated in textbook style, has frequently been used, and successfully at that, in the most diverse fields of mathematics, especially in set theory, by Dedekind, Cantor, F. Bernstein, Schoenflies, J. König and others is an indisputable fact... Such an extensive use of a principle can be explained only by its *self-evidence*, which, of course, must not be confused with its provability. No matter if this self-evidence is to a certain degree subjective—it is surely a necessary source of mathematical principles ...

(Zermelo [1908, p. 187]. See also Fraenkel, Bar-Hillel and Levy [1973, p. 85].) Early set theorists did indeed use Choice implicitly, and the continuing difficulty of recognizing such uses is poignantly demonstrated by Jordain's persistent and ill-starred efforts to prove the axiom (see Moore [1982, §3.8]). Ironically, Choice was even used unconsciously by several French analysts who were officially its severest critics: Baire, Borel and Lebesgue (see Moore [1982, §§1.7 and 4.1]).[5]

[5]The referee indicates that the Paris school did eventually distinguish what they considered acceptable versions of Choice from the unacceptable ones, and that they were also the first to formulate the principle of Dependent Choice, an important tool in the presence of full Determinacy (see [BAII, §V]).

The debates over the intrinsic merits of the axiom centered on the opposition between existence and construction. Modern set theory thrives on a realistic approach according to which the choice set exists, regardless of whether it can be defined, constructed, or given by a rule. Thus:

> In many cases, it appears unlikely that one can *define* a choice function for a particular collection of sets. But this is entirely unrelated to the question of whether a choice function *exists*. Once this kind of confusion is avoided, the axiom of choice appears as one of the least problematic of the set theoretic axioms.

(Martin [SAC, pp. 1–2]) *Iterative conception* theorists seem also to lean on this *realism* rather than on the iterative picture itself (see Boolos [1971, pp. 501–502]; Shoenfield [1977, pp. 335–336]). One might also revert to *Cantorian finitism* (see Hallett [1984, p. 115]). (I will discuss another rule of thumb supporting Choice in II.2 below.)

Zermelo goes on to emphasize extrinsic supports:

> But the question that can be objectively decided, whether the principle is *necessary for science*, I should now like to submit to judgment by presenting a number of elementary and fundamental theorems and problems that, in my opinion, could not be dealt with at all without the principle of choice.

[1908, pp. 187–188]. He then describes seven theorems that depend on the Axiom of Choice, including the fact that a countable union of countable sets is countable, as well as two examples from analysis. Since then it has become clear that the Axiom of Choice and its equivalents are essential not only to set theory but to analysis, topology, abstract algebra and mathematical logic as well.

To take just one example, Moore [1982, §4.5] describes the axiom's growing importance in algebra during the 20s and 30s. In 1930, van der Waerden published his *Modern Algebra*, detailing the exciting new applications of the axiom. The book was very influential, providing Zorn and Teichmüller with a proving ground for their versions of choice, but van der Waerden's Dutch colleagues persuaded him to abandon the axiom in the second edition of 1937. He did so, but the resulting limited version of abstract algebra brought such a strong protest from his fellow algebraists that he was moved to reinstate the axiom and all its consequences in the third edition of 1950. Moore summarizes, "Algebraists insisted that the Axiom had become indispensable to their discipline" (p. 235). And they were not alone.

Nowadays, intrinsic arguments for Choice in terms of intuitiveness or obviousness go hand-in-hand with extrinsic arguments in terms of its indispensability. Modern mathematics has sided firmly with Zermelo:

> . . . no one has the right to prevent the representatives of productive science from continuing to use this "hypothesis"—as one may call it for all I care—and developing its consequences to the greatest extent . . . We need merely separate the theorems that necessarily require the axiom from those that can be proved without it in order to delimit the whole of Peano's

> [choiceless] mathematics as a special branch, as an artifically mutilated science, so to speak ... principles must be judged from the point of view of science, and not science from the point of view of principles fixed once and for all. (p. 189)

I.8. *Replacement.* Early hints of the Axiom of Replacement can be found in Cantor's letter to Dedekind [1899] and in Mirimanoff [1917], but it does not appear on Zermelo's list in [1908]. This omission is due to his reductionism, that is, his belief that theorems purportedly about numbers (cardinal or ordinal) are really about sets. Since von Neumann's definition of ordinals and cardinals as sets, this position has become common doctrine, but Zermelo first proposed his axioms in the context of Cantor's belief that ordinal and cardinal numbers are separate entities produced by abstraction from sets. So, while Cantor sometimes stated the well-ordering theorem in the form "Every set is isomorphic to some ordinal number", Zermelo preferred the form "Every set can be well-ordered". As a result, he had no need for Replacement. (See Hallett[1984].)

Around 1922, both Fraenkel and Skolem noticed that Zermelo's axioms did not imply the existence of

$$\{N, \mathscr{P}(N), \mathscr{P}(\mathscr{P}(N)), \ldots\}$$

or the cardinal number $\aleph_\omega$. These were so much in the spirit of informal set theory that Skolem proposed an Axiom of Replacement to provide for them. It then took von Neumann to notice the importance of Replacement for the ordinal form of the well-ordering theorem, as well as in the justification of transfinite recursion.[6] Zermelo included it (in his second-order version) in [1930].

Replacement is made to order for the *limitation of size* theorists:

> ... our guiding principle ... is to admit only axioms which assert the existence of sets which are not too "big" compared to sets already ascertained. If we are given a set a and a collection of sets which has no more members than a it seems to be within the scope of our guiding principle to admit that collection as a new set. We still did not say exactly what we mean by saying that the collection has "no more" members than the set a. It turns out that it is most convenient to assume that the collection has "no more" members than a when there is a "function" which correlates the members of a to all the sets of the collection ...

(Fraenkel, Bar-Hillel and Levy [1973, p. 50]). *The iterative conception* does less well because the only way to guarantee stages large enough to cover the range of the given function is to assume a version of Replacement in the theory of stages (see Shoenfield [1977, p. 326]); Boolos [1971, p. 500]).

[6] Von Neumann actually used a stronger principle based on *limitation of size*, namely, "A collection is too large iff it can be put in one-to-one correspondence with the collection of all sets." This implies Separation, Replacement, Union and Choice (even Global Choice). Gödel found von Neumann's axiom attractive because it takes the form of a maximal principle (compare *maximize* in II.2 below): anything that can be a set, is. See Moore [1982, pp. 264–265].

On the extrinsic side stand the deep set theoretic theorems noted in the paragraph before last:

> ... the reason for adopting the axioms of replacement is quite simple: they have many desirable consequences and (apparently) no undesirable ones.

(Boolos [1971, p. 500]) Still, the consequences noted are all within set theory; there is nothing like the broad range of applications found in the case of Choice. Recently, however, Martin used Replacement to show that all Borel sets are determined (see Martin [1975]). Earlier work of Friedman establishes that this use of Replacement is essential (see Friedman [1971]). Thus Replacement has consequences in analysis, consequences even for the simple sets of reals favored by the French analysts. Furthermore, these consequences are welcome ones, as we shall see in [BAII, §V].

Let me end this survey here, leaving the interested reader to the more informed works of the mathematical historians. I think enough has been said to demonstrate that from the very beginning, the process of adopting set-theoretic axioms has not been a simple matter of noting down the obvious. Rather, the axioms we now hold to be self-evident were first justified by reference to vague rules of thumb and purely extrinsic consequences, in addition to intrinsic evidence. The arguments offered for the new axioms are no different. But first we should pause to look at the problem that makes new axioms vital.

§II. The continuum problem. Cantor first stated his continuum hypothesis in 1878:

> The question arises ... into how many and what classes (if we say that [sets of reals] of the same or different [cardinality] are grouped in the same or different classes respectively) do [sets of reals] fall? By a process of induction, into the further description of which we will not enter here, we are led to the theorem that the number of classes is two.

(See Cantor [1895, p. 45].) The nature of this "process of induction" is never made clear, but Hallett reconstructs it from the contents of a letter Cantor wrote to Vivanti in 1886 (see Hallett [1984, pp. 85–86]). There Cantor comments on Tannery's purported proof of the continuum hypothesis:

> He believed he had given a proof for the theorem first stated by me 9 years ago that only two [equivalence] classes [by cardinality] appear among linear pointsets, or what amounts to the same thing, that the [cardinality] of the linear continuum is just the *second*. However, he is certainly in error. The facts which he cites in support of this theorem were all known to me then, as anyone can see, and form only a part of that induction of which I say that it led me to that theorem. I was convinced at that time that this induction is *incomplete* and I still have this conviction today.
>
> ... the theorem to be proved is
>
> $$\mathfrak{c} = \aleph_1.$$

> The facts on which Herr T. believes he can base the theorem are only these:
>
> $$n + \aleph_0 = \aleph_0,\ \aleph_0 + \aleph_0 = \aleph_0,\ \aleph_0 \cdot n = \aleph_0,\ \aleph_0^2 = \aleph_0,$$
> $$\aleph_0^n = \aleph_0;\ \aleph_0^{\aleph_0} = \mathfrak{c},\ 2^{\aleph_0} = \mathfrak{c},\ 3^{\aleph_0} = \mathfrak{c}, \ldots, n^{\aleph_0} = \mathfrak{c}.$$
>
> These facts suggest the conjecture that $\mathfrak{c}$ should be the power $\aleph_1$ following next after $\aleph_0$; but they are a long way from furnishing a proof for it.

Perhaps these facts seem even less persuasive today.

II.1. *Cantor's views.* Cantor's writings suggest two other reasons he might have had for believing the continuum hypothesis.[7] In 1874, Cantor proved the first version of his famous theorem: no countable sequence of elements from a real interval can exhaust that interval. In 1883, he proved that there are more countable ordinals than finite ordinals, and that any infinite set of countable ordinals is either countable or equinumerous with the set of all countable ordinals. Three things must have struck Cantor here: first, the two proofs of nondenumerability are similar (the usual diagonal argument for the nondenumerability of the reals came only in 1891), which produces an analogy between the reals and the countable ordinals (see below); second, the property proved for infinite subsets of the countable ordinals is exactly what CH conjectures for the reals; third, that the CH could now be formulated as "the reals and the countable ordinals are equinumerous."

Cantor apparently found evidence for the CH in the structural similarities revealed by the two proofs of nondenumerability. In particular, he came to see the reals and the countable ordinals as generated by similar processes from similar raw materials; in both cases, one begins with a countable set (the rationals and the finite ordinals, respectively) and one considers countably infinite rearrangements (Cauchy sequences and well-orderings, respectively). This analogy suggests that the two sets may also share the same cardinal number. Add to this the discovery that the set of countable ordinals has exactly the property Cantor expected to hold for the reals, and the CH in its new form seems a fairly natural conjecture (see Hallett [1984, pp. 74–81]).

Of course, wherever there are analogies there are both similarities and dissimilarities. What makes the 1874 proof of nondenumerability go through is the fact that any bounded sequence of reals approaches a limit; likewise, the 1883 proof of nondenumerability depends on the fact that any countable sequence of countable ordinals has a countable ordinal as its supremum. Still, as Hallett points out (p. 81), the topologies underlying these limit properties are not really all that similar.

A second reason Cantor may have had for believing the continuum hypothesis is based on the Cantor-Bendixon theorem of 1884, that is, the result that every closed infinite set of reals is countable or has a perfect subset, and hence, that CH is true for closed sets of reals. At the end of the paper in which this result is proved, he promises a proof of the same result for nonclosed sets of reals. He may have believed at one time that the proof itself could be generalized, and in fact, it can to a certain extent. I will take up the idea that these partial results constitute evidence for CH in

[7] Apparently, the term "continuum hypothesis" was first used by Bernstein in 1901. See Moore [1982, p. 56].

II.3.1. In any case, it is clear that for a while he hoped to establish CH by finding a closed set of cardinality $\aleph_1$. Such a set would be nondenumerable, so by the Cantor-Bendixson theorem, it would have the cardinality of the continuum. But that cardinality is $2^{\aleph_0}$, so CH is true. Working along these lines, in 1884 he wrote:

> I am now in possession of an extremely simple proof for the most important theorem of set theory, that the continuum has the [cardinality] of the [set of countable ordinals] ... you see that everything reduces to defining a closed set having [cardinality $\aleph_1$]. When I have sorted it out, I will send you the details.

(See Hallett [1984, p. 92]; Moore [1982, p. 43].) Of course, the details were never sorted out.[8]

For the record, during the prehistory of CH (that is, before the consistency and independence results), opinion seems to have been divided. Hilbert and Jourdain were both in favor (Moore [1982, pp. 55, 63]), though Hilbert apparently did not expect it to be provable in ZFC alone (Wang [1981, p. 656]). König attempted to prove it false, but only because he felt the reals could not be well-ordered at all (Moore [1982, §2.1]). Finally, Gödel cites Lusin and Sierpiński as tending to disbelieve it for reasons closer to his own ([1947/64, p. 479]).

II.2. *Consistency and Independence.* A wag once suggested that if only Gödel had announced having *proved* the continuum hypothesis, instead of its mere *consistency*, there would be no [redacted] continuum problem. Strangely enough, Gödel does almost exactly that in [1938]. Of the Axiom of Constructibility, from which he did prove CH, he writes:

> The proposition ... added as a new axiom, seems to give a natural completion of the axioms of set theory, in so far as it determines the vague notion of an arbitrary infinite set in a definite way. (p. 557)

By [1944], however, he has changed his mind and come around to the view now so strongly associated with his name:

> [The] axiom [of constructibility] states a minimum property. Note that only a maximum property would seem to harmonize with the concept of set ... (p. [redacted] 479)

Perhaps Gödel's new opinion of $V = L$ was also influenced by his developing belief in the falsity of CH (see II.3.3 below).

We see here the statement of a new rule of thumb, namely *maximize*.[9] This rule is often associated with the *iterative conception* in two more specific forms:

> Intrinsic necessity depends on the concept of iterative model. In a general way, hypotheses which purport to enrich the content of power sets ... or to introduce more ordinals conform to the intuitive model. We believe that the collection of all ordinals is very 'long' and each power set (of an infinite set)

[8] Cantor's final attempt at proving the continuum hypothesis involved a new method of decomposing arbitrary sets of reals. See Moore [1982, pp. 43–44], and Hallett [1984, p. 103].

[9] Recall the earlier hint of this rule in Gödel's reaction to von Neumann's axiom.

> is very 'thick.' Hence, any axioms to such effects are in accordance with our intuitive concept.

(Wang [1974, p. 553]) For example, the Axiom of Choice is widely thought to contribute to the "thickness" of the power set (see e.g. [redacted] Drake [1974, p. 12]). I will take up the question of postulating more ordinals in the next section, but for now it is clear that the restriction to definable subsets at each stage can be seen as an unwelcome curtailment of the full power set. The view that $V = L$ contradicts *maximize* is widespread (see e.g. Drake [1974, p. 131]; Moschovakis [1980, p. 610]; Scott [1977, p. xii]).[10]

There are also extrinsic reasons for rejecting $V = L$, most prominently that it implies the existence of a Δ^1_2 well-ordering of the reals, and hence that there is a Δ^1_2 set which is not Lebesgue measurable.[11] It can be proved in ZFC that every Borel well-ordering of the reals is countable. A Δ^1_2 set can be obtained from a Borel set by one application of projection followed by one application of complement. Many find it implausible that a set as complex as a well-ordering of the real numbers could be generated by such simple operations.[12] The Axiom of Choice guarantees that there is such a well-ordering, but the proofs are highly nonconstructive, so it is considered implausible that the well-ordering should be definable at all (see Moschovakis [1980, p. 276]; Wang [1974, p. 547]; Martin [1976, p. 88] and [PSCN, p. 2]). Further extrinsic evidence against $V = L$ will be discussed in II.3.1, below, and in [BAII, §V].

After his proof of the consistency of CH, Gödel conjectured that it is independent as well. The axioms of ZFC, he argues, are true in V and in L, so one can hardly expect to decide the numerical question of the size of the continuum until one has settled "what objects are to be numered, and on the basis of which one-to-one correspondences." Even if one believes that $V = L$,

> ... he can hardly expect more than a small fraction of the problems of set theory to be solvable without making use of this, in his opinion essential, characteristic of sets. [1947/64, p. 478]

Finally, in 1963, Cohen proved him right (see his [1966]).[13]

---

[10] I should not suggest that no one supports the adoption of $V = L$ as an axiom; sentiment in favor can be found (see e.g. Fraenkel, Bar-Hillel and Levy [1973, pp. 108–109]; Devlin [1977, p. iv]). Reasons usually given are that it is simple and safe (see Moschovakis [1980, p. 609]), and that it provides answers to a great many outstanding problems. Discussion below and in [BAII, §VI] will suggest that these answers are "in the wrong direction", but that opinion is surely open to debate. Despite all this, I will stick to the anti-($V = L$) line because it is favored by the Cabal group.

[11] *Notation*: Σ^0_1 is the class of open sets of reals; Π^0_α is the class of all complements of Σ^0_α sets of reals; $\Sigma^0_{\alpha+1}$ is the class of all countable unions of Π^0_α sets; and $\Delta^0_\alpha = \Sigma^0_\alpha \cap \Pi^0_\alpha$. All these together are the Borel sets. Further, Σ^1_1 is the class of all projections of closed sets; Π^1_α is the class of all complements of Σ^1_α sets; $\Sigma^1_{\alpha+1}$ is the class of projections of Π^1_α sets; and Δ^1_α is $\Sigma^1_\alpha \cap \Pi^1_\alpha$. These are the projective sets. In 1917, Souslin proved that the Borel sets are the Δ^1_1 sets. Finally, if R is a well-ordering of the reals, then Fubini's theorem implies that R is not Lebesgue measurable.

[12] This way of putting the implausibility was suggested by Matt Foreman.

[13] See Scott [1977] and Wang [1981] for some discussion of why the independence proof was so long in coming.

While Gödel's result had a temporary discouraging effect on research in set theory (for fear that the problem in question was in fact undecidable), Cohen's invention of the forcing method led to a boom (see Martin [1976, pp. 82–83]). While the truth of CH in the constructible universe has had some influence on opinion as to its truth or falsity (see II.3.4 below), the relevance of forcing models to that question is much less clear. The plethora of different models moved Cohen himself to a version of formalism (see his [1966] and [1971]), but Scott, another innovator in the early development of forcing, writes:

> I myself cannot agree, however. I see that there are any number of contradictory set theories, all extending the Zermelo-Fraenkel axioms; but the models are all just models of the first-order axioms, and first-order logic is weak. I still feel that it ought to be possible to have strong axioms which would generate these types of models as submodels of the universe, but where the universe can be thought of as something absolute. [1977, p. xiv]

(See also Kanamori and Magidor [1978, p. 109].) Perhaps the association of CH with the restrictive $V = L$, combined with the development of this striking new technique for adding extra real numbers to models, led some to agree with Gödel that CH is false in the absolute real world.

II.3. *Informed opinion.* Despite the results of Gödel and Cohen, there remain set theorists who feel the CH is a real question, the sort of thing that is either true or false in the real world of sets. Various arguments for and against have been bandied about in their ranks. The purpose of this subsection is to summarize the most prominent of these.

II.3.1. *Partial results (in favor).* Recall that Cantor may have expected the proof of the Cantor-Bendixon theorem (that CH holds for closed sets of reals) to generalize to all sets of reals. This program was carried forward by Young in 1906 to a subset of the Π^0_2 sets, then by Hausdorff in 1914 to all Π^0_2 sets, and finally, by Hausdorff again in 1916 to all Borel sets. Still, Hausdorff himself was reluctant to count these results as evidence for the CH:

> If we knew for all sets what we know for closed [and Π^0_2]... then ... the continuum hypothesis would be decided. However, in order to see how far we still are from this goal, it is sufficient to recall that the system of sets closed or [Π^0_2] forms only a vanishingly small part of the system of all point sets.

Even after the proof had been extended to all Borel sets, he continued:

> Thus the question of power is clarified for a very inclusive category of sets. Nevertheless, one can scarcely see this as a genuine step towards the solution of the continuum problem, since the Borel sets are still very specialized, and form only a vanishingly small subsystem.

(Both translations are due to Hallett [1984, p. 107].) Of course there are $2^{2^{\aleph_0}}$ sets of reals, only $2^{\aleph_0}$ of which are Borel.

Even more damaging to the interpretation of these results as evidence in favor of the CH is something Hausdorff apparently did not realize at the time, namely that his proofs could be strengthened to the form of the original Cantor-Bendixon result, that is, that every infinite Borel set is either countable or contains a perfect subset. In 1916, Alexandroff proved the theorem in this form, and in 1917, Souslin extended it to Σ^1_1 sets. (In the presence of a measurable cardinal, this pattern can be extended to Σ^1_2. See §IV below.) The trouble arises from Bernstein's proof that there are uncountable sets of reals without perfect subsets. Thus these proofs that CH holds for restricted classes of sets all depend on establishing a stronger property, the perfect subset property, that cannot hold for all sets of reals (see Martin [1976, p. 88]; Hallet [1984, pp. 103–110]). For this reason, the technique cannot be fully generalized.

In Cantor's defense, it should be noted that he was probably unaware of the existence of uncountable sets without perfect subsets. Most of the sets of reals Cantor worked with were Σ^1_1 at worst. Furthermore, Bernstein's proof, published in 1908, made essential use of the Axiom of Choice. Though Cantor often used that axiom, he did so to form orderings, or to make simultaneous choices from many order types or cardinalities, not to form sets of reals, so he may well not have noticed this possibility.

In 1925, Lusin wondered whether every infinite Π^1_1 set is either countable or contains a perfect subset. He writes:

> My efforts towards settling this question have led to an unwelcome result:
> we do not know *and will never know*...

(Translation due to Hallet [1984, p. 108]). This may sound overly dramatic, but in a sense, Lusin was right, for the Axiom of Constructibility implies the existence of an uncountable Π^1_1 set with no perfect subset, while other hypotheses imply the opposite (see [BAII, [redacted] §V]). That such a "pathology" should occur so low in the projective hierarchy is considered another extrinsic disconfirmation of $V = L$ (e.g. Wang [1974, p. 547]).

II.3.2. *The effectiveness of* CH (*in favor*). The generalized CH is an extremely simple and powerful assumption that immediately settles all questions of cardinal arithmetic. Furthermore, it allows any power set to be well-ordered in such a way that every initial segment is no bigger than the original set. This facilitates many complex constructions, such as saturated models of every regular cardinality. Sierpiński's book *Hypothèse du Continu* deduces 82 propositions from the CH. In stark contrast, Martin and Solovay remark [1970] that not a single one of these 82 propositions is known to be decided by the negation of the continuum hypothesis.

II.3.3. *Gödel's counterintuitive consequences* (*against*). In [1947/64], Gödel argues that CH is false because it has certain "highly implausible consequences" (p. 479). Several of these assert the existence of sets of reals of cardinality $2^{\aleph_0}$ with strong "smallness" properties. For example, a subset of the unit interval is called "absolute zero" if it can be covered by any countable collection of intervals. If covering is only required when the intervals are of equal length, then the set would have Lebesgue measure zero, but would not necessarily be absolute zero. Thus Cantor's discontinuum has Lebesgue measure zero, but is not absolute zero, because it cannot be covered by countably many intervals of length $1/3^n$. In fact, no perfect set can be absolute zero, and Borel conjectured that no set of size $2^{\aleph_0}$ could be. The CH implies that there is such a set.

Commentators were quick to point out that many consequences of the set-theoretic reduction of the continuum that do not depend on CH are similarly counterintuitive, for example, Peano's space-filling curve. Gödel insists that his examples are not of this sort, because in those cases:

> ... the appearance [of counterintuitiveness] can be explained by a lack of agreement between our intuitive geometrical concepts and the set-theoretical ones occurring in the theorems. (p. 480)

While we all might be surprised at Peano curves or the uncountable Cantor set of measure zero, this surprise is presumably based on exactly the clash Gödel mentions: a disagreement between our geometric intuition and our set-theoretic geometry, or a vague feeling that sets large in cardinality should not also be small in measure. But Gödel is basing his reactions on something else.

What then? Wang suggests that

> ... it cannot be excluded that someone might have such intimate knowledge so that, for example, he can separate out the errors coming from using the pre-set-theoretical [intuition].

[1974, p. 549]. He reminds us of Gödel's view that all intuition must be cultivated. It seems to me more likely that Gödel had in mind some form of peculiarly set-theoretic intuition not connected with pre-set-theoretic geometry. In either case, we are left with Gödel's bare claims, because even our best set theorists do not share these "intuitions":

> While Gödel's intuitions should never be taken lightly, it is very hard to see that the situation *is* different from that of Peano curves, and it is even hard for some of us to see why the examples Gödel cites are implausible at all.

(Martin [1976, p. 87]; see also Martin and Solovay [1970, p. 177]).

Gödel apparently did make at least one attempt to axiomatize his views on the continuum. It appears in Ellentuck [1975] and takes its cue from a conjecture of Borel. Suppose that the functions from ω to ω are ordered so that $\mathbf{f}$ is less than $\mathbf{g}$ if and only if $\mathbf{f}(n)$ is always less than $\mathbf{g}(n)$ after a proper initial segment. Borel conjectured that there is a set S of size $\aleph_1$ which is cofinal in this ordering. The "square axiom", A, is just this conjecture; the "rectangle axioms", A_n, are generalizations of the square axiom to functions from ω_n to ω. Gödel agreed with Borel on the plausibility of A; his hope was that the A_n's could be justified by analogy with A, and that they would set bounds on the size of the continuum.

Now $\aleph_2$ is the only value for $2^{\aleph_0}$ that is known to be consistent with the nonexistence of absolute zero sets. Furthermore, A_2 implies that $2^{\aleph_0} \leq \aleph_2$. Thus it seems Gödel must have suspected that $2^{\aleph_0} = \aleph_2$. Unfortunately, the plan to derive this from a theory involving the rectangle axioms was ruined with the discovery that A_1 implies CH. Alternate versions of the square axiom turned out to be relatively consistent with a wide range of values of $2^{\aleph_0}$, and A itself implies the existence of an absolute zero set of size $\aleph_1$. While this is perhaps not so counterintuitive as an absolute zero set of size $2^{\aleph_0}$, it must have been an unwelcome result. Thus this effort of Gödel's to formalize his intuitions about the continuum was quite unsuccessful.[14]

II.3.4. CH *is restrictive* (*against*). As mentioned earlier, perhaps some of the reason CH is felt to be restrictive is because it is true in L. If this line of thought is to have any force, it must first meet a difficult challenge, namely that the Axiom of Choice, generally regarded as a maximizing principle in itself, is also true in L. If L is so impoverished, why does the additional assumption of Choice have no maximizing effect? (It doesn't, because it is not an additional assumption, after all.)

I think the answer to this question is not so difficult. Choice is true in L because there is a definable well-ordering of the constructible universe. This reverses the intuitive order of things. Why is a given set well-orderable? Because an element can be chosen for each ordinal until the set is exhausted. Why should such choices be possible? Because *realism* and *maximize* guarantee the existence of a choice set. Thus the well-ordering principle derives from Choice and not vice versa.

The maximizing force of Choice lies in its implying the existence of complex, probably undefinable sets like a well-ordering of the reals, a non-Lebesgue-measurable set, and an uncountable set with no perfect subset. That the Axiom of Constructibility forces such sets far down into the simple projective sets counts as extrinsic evidence against it. Thus the Axiom of Choice is true in L, but it does not do any maximizing work because it is true for the wrong reason. It is not true because there are complex sets; rather it is true because there is an artificially simple well-ordering.

Now what about CH? Is it truly a maximizing principle that just happens to be true in the restricted world of L because its maximizing force is masked by some unrelated pathology? For what it's worth, I see nothing in the proof of CH in L that suggests this. CH is true in L because all the constructible subsets of ω appear in L_{ω_1}, and L_{ω_1} has small cardinality. But why is that cardinality small? Because the limited procedure of subset formation in L only allows at most one new element for every formula and finite sequence of parameters. Thus CH is true in L because the formation of subsets is artificially restricted, not because some other pathological condition in L is robbing it of its maximizing force.[15]

---

[14] The history of the square and rectangle axioms is described in more detail in Moore [1982].

[15] I should note that J. Friedman [1971] argues that GCH is a maximizing rather than a restrictive principle. He does so by showing it equivalent to what he calls the "generalized maximizing principle," namely, the assumption that every "local universe" contains all its smaller-cardinality subsets. (Note the similarity to von Neumann's maximizing principle above.) The problem is that a "local universe" is defined as a collection closed under Pairing, Union and Replacement. Obviously Replacement is being maximized at the expense of Power Set. Thus Friedman is right that:

> A fundamental question is whether [the generalized maximizing principle] maximizes these operations [Pairing, Union and Replacement] at the expense of the power set operation.

But perhaps he is less than candid when he claims:

> [The Generalized Maximizing Principle] says nothing explicit about the power set operation, but as an afterthought, GCH follows from it. (p. 41)

Another way of stating the idea that CH is restrictive is to insist that the continuum is somehow too complicated to be numbered by the countable ordinals. Drake presents a version of this view:

> Of course, many mathematicians do not feel that the cumulative type structure is a well-defined, unique object, and from this point of view the independence results may have to be considered the final word on the GCH. But there are also many mathematicians who feel that the cumulative type structure is real enough, in some sense, for the GCH, or at least the CH, to be a real question. It is worth noting that amongst these mathematicians, many feel that the GCH is just too *simple* to be right. Perhaps the following illustrates this feeling:... [1974, pp. 65–66]

He goes on to point out that $\aleph_1$ is the cardinal number of the collection of all countable well-ordering types, while $2^{\aleph_0}$ is the cardinal number of the collection of all countable linear ordering types.

> To say of a linear ordering that it is a well-ordering is a very strong requirement, so that there should be many more linear orderings than well-orderings ... (p. 66)

Of course, if CH were true, it would not be the first time that a difference in complexity was not mirrored in a difference in cardinality. It is hard to see why the CH should be interpreted as saying that there are not very many subsets of ω when it could just as easily be taken to say that there are lots of countable ordinals.

The question of how complexity matches up with cardinality is further muddled by results involving Martin's Axiom (see Martin and Solovay [1970]). Recall that many of the consequences of CH are made possible by the well-ordering of $2^{\aleph_0}$ with countable initial segments. Though Martin's Axiom is relatively consistent with $\aleph_1 < 2^{\aleph_0}$, it still guarantees that the cardinals smaller than $2^{\aleph_0}$ are well enough behaved to allow complicated constructions to go through. As a result, 79 of Sierpiński's 82 consequences of CH also follow from MA + ($\aleph_1 < 2^{\aleph_0}$) (with the natural modification that the countable/uncountable distinction is replaced by the less-than-$2^{\aleph_0}/2^{\aleph_0}$ distinction). Thus advocates of the view that the continuum is complex might wonder if large cardinality alone is enough to guarantee that complexity.

II.3.5 *Power Set is stronger than Replacement* (*against*). This position is Cohen's. As a formalist, Cohen realizes he should reject the question of the truth or falsity of CH [1971, p. 12], but he feels he cannot reject the same question concerning large cardinals:

> I, for one, cannot simply dismiss these question of set theory for the simple reason of their reflections in number theory. I am aware that there would be few operational distinctions between my view and the Realist position. [1971, p. 15]

Thus he is willing to speculate on the truth value of the CH from the realist point of view:

> A point of view which the author feels may eventually come to be accepted is that CH is *obviously* false. The main reason one accepts the Axiom of

> Infinity is probably that we feel it absurd to think that the process of adding only one set at a time can exhaust the entire universe... Now $\aleph_1$ is the set of countable ordinals and this is merely a special and the simplest way of generating a higher cardinal. The set C is, in contrast, generated by a totally new and more powerful principle, namely the Power Set Axiom. It is unreasonable to expect that any description of a larger cardinal which attempts to build up that cardinal from ideas deriving from the Replacement Axiom can ever reach C. Thus C is greater than $\aleph_n$, $\aleph_\omega$, $\aleph_\alpha$ where $\alpha = \aleph_\omega$, etc. This point of view regards C as an incredibly rich set given to us by a bold new axiom, which can never be approached by any piecemeal process of construction. Perhaps later generations will see the problem more clearly and express themselves more eloquently. [1966, p. 151]

In this connection, recall that the *limitation of size* theorists had difficulties with Power Set but smooth sailing with Replacement. It should be noted that most set theorists who disbelieve CH think $2^{\aleph_0}$ is more likely to be very large, as Cohen indicates, than to be $\aleph_2$, as Gödel suggests.

II.3.6. *Finitism* (*against* GCH). Here the argument depends on analogy with the finite numbers, where $n + 1 = 2^n$ is true only for 0 and 1. This is felt by some to constitute an argument against the GCH, if not against the particular case of the CH (see Drake [1974, p. 66]).

II.3.7. *Whimsical identity* (*against* GCH). This argument depends on the same facts as the finitism argument, but it uses them in a different way. Notice that if the GCH were true, then $\aleph_0$ could be defined as that cardinal before which GCH is false and after which it is true (excepting 0 and 1, of course). But this identity would seem "accidental", like the identity between "human" and "featherless biped". While the physical universe might be too impoverished to falsify such accidental identities, the set-theoretic universe should be rich enough to rule them out. Therefore, GCH is false. (Kanamori and Magidor [1978, p. 104] and Martin [1976, p. 85] use whimsical identity arguments to support large cardinal axioms. See §III below.) Of course this line of argument faces considerable difficulties in explaining what is meant by "accidental", and how this particular identity can be seen to have that property.

II.3.8. *The delicate balance* (*against*). This argument is stated by Wang:

> Some set theorist states that if $\aleph_1 = 2^{\aleph_0}$, then there must be a surprisingly delicate balance between the reals and the countable ordinals. [1974, pp. 549–550]

As Wang goes on to point out, the balance must be delicate whatever the cardinality of the reals turns out to be. Indeed, it might seem more delicate if $2^{\aleph_0}$ were $\aleph_{17}$.

II.3.9. *Gupta's wager* (*against*). Gupta suggests, somewhat facetiously, that since $\aleph_1$ is only one among the proper-class-many values $2^{\aleph_0}$ might consistently take, it makes more sense to plump for not-CH.

II.3.10. *Freiling's darts* (*against*). Freiling [1986] suggests a thought experiment in which random darts are thrown at the real line. Suppose that a countable set $f(x)$ is associated with each real x. Now I throw two darts; the first hits a point x_0 and the second a point x_1. Given that a countable set is very sparsely distributed, the probability that my second dart will hit a member of $f(x_0)$ is vanishingly small. Thus, in all likelihood, the point x_1 is not a member of $f(x_0)$.

But the situation is symmetric: there is just as little reason to suppose that the first dart has hit a point in the set $f(x_1)$. Thus Freiling proposes that for every assignment of a countable set to each real, there are two reals neither of which is a member of the set assigned to the other. This rather innocuous-sounding statement turns out to be equivalent to not-CH.

A common objection to this line of thought is that various natural generalizations contradict the axiom of choice as well. For example, if Freiling's principle is modified to cover assignments of sets of any cardinality less than that of the reals, the result immediately implies that there is no well-ordering of the reals. Similarly if we are allowed to throw $\omega + 1$ darts. Members of the Cabal suggest that Freiling's hypotheses yields a picture more like that of the full AD-world (see [BAII, §V]) than of the choiceful universe V. Meanwhile, Freiling disputes the "naturalness" of these generalizations. He also points out that *one step back from disaster* could provide a rationale for accepting his axiom and rejecting its generalizations even if they are "natural."

II.3.11. *Not*-CH *is restrictive* (*in favor*). This argument uses the same general considerations as II.3.4 in support of the opposite conclusion. While established opinion among more mature members of the Cabal is against CH, younger members are sympathetic to this more recent argument and to the considerations raised in the next subsection. It has been suggested that the cut-off age is 40.

To see how not-CH can be considered restrictive, we imagine ourselves constructing the iterative hierarchy. By stage $\omega + 2$, we have the set of reals and we have a well-ordering of type $\aleph_1$. The question is whether or not a one-to-one correspondence between them is included at the next stage. Since it is consistent to do so, it would artifically restrict the power set operation to leave it out. The thinking behind II.3.4 sees CH as restricting the size of the power set of ω. From the present point of view, not-CH is a restriction of the power set operation at the next stage.[16]

II.3.12. *Modern forcing* (*in favor*). Practitioners of modern versions of forcing point out that it is much easier to force CH than not-CH; that is, that a wide variety of forcing conditions collapse $2^{\aleph_0}$. Since the addition of generic sets tends to make CH true, it is most likely true in the full richness of V itself.

I think this list includes most of the arguments standardly offered for and against the CH. It should be emphasized that few set theorists consider any of them conclusive, and even those with fairly strong opinions adopt a decidedly wait-and-see attitude toward CH. Let me turn now to the search for new axioms to settle the question.

---

[16]Chris Freiling points out that an argument of similar form can be presented against the Axiom of Choice. Notice that a choice function for a countable partition of the reals can be coded as a single real. At stage $\omega + 1$, we have all the reals, so we also have all codes of choice functions for countable partitions. Any countable partition can be coded as a set of reals at stage $\omega + 2$. Thus the question, at stage $\omega + 2$, is whether or not to include a countable partition without a choice function.

§III. Small large cardinals—up from below. There are those who hold that the universe of sets is not sufficiently well-defined for the continuum question to have an unambiguous answer. Little can be done to rebut this position short of coming up with an unambiguous solution, so perhaps this question should be set aside pending further developments. A less reasonable view is that the consistency and independence proofs by themselves show that the CH poses a meaningless question. It is hard to see any justification for the implicit claim that the axioms of ZFC must be taken as the final word:

> Although the ZFC axioms are insufficient to settle CH, there is nothing sacred about these axioms...
>
> ... undecidability [of CH] from the axioms being assumed today can only mean that these axioms do not contain a complete description of [set-theoretic] reality.

(Martin [1976, p. 84]; Gödel [1947/64, p. 476])

Where are we to look for the new axioms that will make our description more complete? In [1946], Gödel suggests:

> ...stronger and stronger axioms of infinity...It is not impossible...that every proposition expressible in set theory is decidable from the present axioms plus some true assertion about the largeness of the universe of sets.

(p. 85)[17] Skolem's and Fraenkel's introduction of the Axiom of Replacement can be seen in this light as they specifically wanted to generate $\aleph_\omega$. Thus the suspicion that adding larger ordinals can produce new results about sets of reals is confirmed by Martin's proof of Borel determinacy (see Hallett [1984, p. 102]; Kreisel [1980, p. 208]).

The first such new axiom of infinity is the Axiom of Inaccessibles, asserting the existence of regular, strong limit cardinals. The existence of such cardinals was first suggested by Zermelo [1930]; the axiom itself was formulated by Tarski in [1938]. Gödel presents an intrinsic defense:

> These axioms show clearly, not only that the axiomatic system of set theory as used today is incomplete, but also that it can be supplemented without arbitrariness by new axioms which only unfold the content of the concept of set explained [by the iterative conception]. [1947/64, pp. 476–477]

(See also Wang [1974, p. 554].) Of course, *maximize* presents a simple and immediate defense for the Axiom of Inaccessibles. Recall that this rule of thumb is actually a pair of admonitions: thicken the power set, and lengthen the class of ordinals. Axioms of infinity in general, and the Axiom of Inaccessibles in particular, clearly do the second of these.

The most commonly given argument is more closely tailored to the actual content of Inaccessibles (see e.g. Gödel [1947/64, p. 476]; Wang [1974, p. 554]; Drake [1974,

---

[17]This conjecture may seem less likely in light of Levy and Solovay's strong theorem [1967] on the stability of large cardinals under most forcing extensions. See §IV below.

pp. 267–268]). It depends on the widespread view that the universe of sets is too complex to be exhausted by any handful of operations, in particular by power set and replacement, the two given by the axioms of Zermelo and Fraenkel. Thus there must be an ordinal number after all the ordinals generated by replacement and power set. This is an inaccessible.[18] Similarly, the universe above a given point should not be exhausted by these two operations, so there is another inaccessible, and so on. Versions of *inexhaustibility* can also be used to defend the various hyperinaccessibles and Mahlo cardinals. All of these are generated by thinking of processes that build up larger ordinals from below.

The Axiom of Inaccessibles can also be defended by two other rules of thumb each incomparably stronger than *inexhaustibility*. The first of these is *uniformity*.[19] To understand the thrust of this rule, suppose that a certain interesting situation occurs at a low level of the iterative hierarchy. If no similar situation occurred in the remainder of the hierarchy, it would be as if the universe had lost its complexity at the higher levels, as if it had flattened out, become homogeneous and boring. *Uniformity* says that this does not happen, that situations similar to our chosen interesting one will recur at higher levels:

> We mean by [uniformity] a process of reasonable induction from familiar situations to higher orders, with the concomitant confidence in the recurring richness of the cumulative hierarchy.
>
> Uniformity of the universe of sets (analogous to the uniformity of nature): the universe of sets doesn't change its character substantially as one goes over from smaller to larger sets or cardinals, i.e., the same or analogous states of affairs reappear again and again (perhaps in more complicated versions).

(Kanamori and Magidor [1978, p. 104]; Wang [1974, p. 541]; see also Solovay, Reinhardt and Kanamori [1978], and Reinhardt [1974, p. 189]. Wang and Reinhardt attribute support for this principle to Gödel.) Thus, $\aleph_0$ is inaccessible, so there must be uncountable inaccessible cardinals. Otherwise, the universe would be sparse above $\aleph_0$, or change its character in an objectionable way.

Uniformity arguments often go hand-in-hand with *whimsical identity* arguments. In this case, for example, if there are no uncountable inaccessibles, then $\aleph_0$ can be defined as the inaccessible. But:

> It would seem rather accidental if [$\aleph_0$] can be characterized by [this] propert[y].

(Kanamori and Magidor [1978, p. 104]; see also Martin [1976, p. 85]). So there must be uncountable inaccessibles.

---

[18] Of course, Replacement must be taken in Zermelo's second-order form.

[19] Solovay, Reinhardt and Kanamori [1978] and Kanamori and Magidor [1978] call this principle *generalization*, while Wang [1974] calls it *uniformity*. I will want to retain the first of these for a slightly different rule of thumb (see [BAII, §VI]). Hallet [1984, pp. 114–115] also connects *uniformity* to the views behind *cantorian finitism*, but not in the way suggested here.

Uniformity itself is sometimes defended on the basis of *Cantorian finitism*: the sequence of natural numbers continues to produce interesting complexities arbitrarily far out, so the sequence of transfinite ordinals should do the same. Unfortunately, the premise concerning the natural numbers is debatable. While the sequence of natural numbers does continue to produce, for example, arbitrarily large prime numbers, it may or may not continue to produce twin primes, and it definitely runs out of adjacent primes after 2 and 3, and even primes after 2. This highlights the delicacy of formulating the property to be projected. Projecting properties of $\aleph_0$ is similarly chancy. I will return to this point below, in connection with weakly compact cardinals.

The second powerful rule of thumb sometimes cited in support of Inaccessibles is *reflection*: the universe of sets is so complex that it cannot be completely described; therefore, anything true of the entire universe must already be true of some initial segment of the universe. In other words, any attempt to uniquely describe V also applies to smaller R_α's that "reflect" the property ascribed to V.[20] In particular, V is closed under the operations of replacement and power set, so there is an R_κ which is also so closed. Then κ is an inaccessible. Similarly, V is closed under replacement and power set above this κ, so there is another inaccessible, and so on.

Hallett [1984, pp. 116–118], traces *reflection* to Cantor's theory that the sequence of all transfinite numbers is absolutely infinite, like God. As such, it is incomprehensible to the finite human mind, not subject to mathematical manipulation. Thus nothing we can say about it, no theory or description, could single it out; in other words, anything true of V is already true of some R_α. Reinhardt [1974, p. 191] expresses a similar sentiment, though without the reference to God. A related view is that the universe of set theory is "ever-growing", so that our attempt to speak of "all sets" actually refers to "temporary" partial universes that "approximate" the universe of all sets (Fraenkel, Bar-Hillel and Levy [1973, p. 118]; see also Parsons [1974] and Wang [1974, p. 540]). Discussions of this sort characteristically emphasize the indefiniteness or incomprehensibility or ineffability of V.

Martin strikes a somewhat different note:

> Reflection principles are based on the idea that the class ON of [redacted] ordinal numbers is so large that, for any reasonable property P of the universe [V], ON is not the first stage α such that R_α has P. [1976, pp. 85[redacted]]

Here the emphasis is on the largeness and complexity of the class of ordinals rather than some mysterious indefinability V; it is not that V is so inscrutable that nothing can describe it, but that ON is so vast that whatever happens at the top must already have happened before.

In any case, *reflection* is probably the most universally accepted rule of thumb in higher set theory (in addition to references already cited, see Solovay, Reinhardt and Kanamori [1978, p. 15[redacted]], and many others). It is partially confirmed by weak, single formula versions that are provable in ZFC (see Levy [1960]). More powerful applications attempt to use stronger properties involving infinite sets of formulas,

[20] Notice that *reflection* implies *inexhaustibility*.

and/or higher order properties, while avoiding "nonstructural" properties, like "$x = V$", which lead to contradiction.

It should be mentioned that the Axiom of Inaccessibles also has a few extrinsic merits. It implies that ZFC has a standard model in the iterative hierarchy, and thus, that ZFC is consistent. This last is an arithmetic fact, and the Axiom of Inaccessibles, like other axioms of infinity, also implies the solvability of new Diophantine equations. (These facts are often cited. See e.g. Gödel [1947/64, p. 477], the quotation from Cohen [1971, p. 15], cited in II.3.5 above, and Kanamori and Magidor [1978, p. 103], to name a few.) In addition, there are the impressive relative consistency results of Solovay [1970]. Assuming a model of "ZFC + The Axiom of Inaccessibles", Solovay uses forcing to collapse the inaccessible and obtains models of ZFC in which all or many sets of reals are Lebesgue measureable ("many" means those constructible from the reals). Thus, these conditions are refutable only in the unlikely event that inaccessibles are refutable. Solovay's theorem:

> ... even today rivals any other as the most mathematically significant result obtained by forcing since Cohen's initial work.

(Kanamori and Magidor [1978, pp. 204–205]). And it presupposes the consistency, if not the existence, of an uncountable inaccessible.

There are larger small large cardinals, but nothing new appears in the usual defenses.[21] An exception is weakly compact cardinals, from discussions of which two morals can be derived. These cardinals can be defined in terms of a generalization of Ramsey's theorem; that is, κ is weakly compact iff every partition of the two-element subsets of κ into two groups has a homogeneous set of size κ. Because of Ramsey's theorem on $\aleph_0$, the existence of an uncountable weakly compact cardinals can be defended by *uniformity* or by *whimsical identity*. The first point of interest is that the proof of Ramsey's theorem also gives a homogeneous set for partitions of n-element set into m groups, but this property cannot be consistently generalized to an uncountable cardinal (see Drake [1974, p. 315]). This dramatically spotlights the difficulty of knowing when *uniformity* and *whimsical identity* can be applied without ill effect.

Second, the property of weakly compactness is equivalent to the compactness of the language $L_{\kappa\kappa}$, and to a certain tree property, and to an indescribability property, and to several other natural properties (see Drake [1974, §10.2]). This convergence has led some writers to *diversity*, another rule of thumb:

> It turned out that weak compactness has many diverse characterizations, which is good evidence for the naturalness and efficacy of the concept.

(Kanamori and Magidor [1978, p. 113]). Recall that similar arguments were once given for the naturalness of the notion of general recursiveness.

[21]Though extrinsic defenses are nothing new, Harvey Friedman has extended the range of such defenses for small large cardinals. His [1981] contains nonmetamathematical statements, statements not involving such "abstract" notions as uncountable ordinals or arbitrary sets of reals, which are provable with and not without the assumption of Mahlo cardinals. See Drake [1974] for an account of Mahlo cardinals, and Stanley [1985] for a description of recent extensions of Friedman [1981].

§IV. Measurable cardinals. Measurable cardinals were introduced by Ulam in [1930], where he proved that they are inaccessible. They are now known to be much larger than that, larger than all the hyperinaccessibles, Mahlos and weakly compacts. Indeed, because of their power, they are probably the best known large cardinals of all. The voice of caution reminds us that they were invented by the same fellow who invented the hydrogen bomb.[22]

Unlike the small large cardinals suggested by *inexhaustibility*, measurable cardinals are not usually held to follow naturally from the concept of set or the nature of the iterative hierarchy:

> ...that these axioms are implied by the general concept of set in the same sense as Mahlo's has not been made clear yet.

(Gödel [1947/64, pp. 476–477])[23] Some wish for an *inexhaustibility* defense:

> What we would really like to do (but are presently unable to do) is to reformulate the definition of a measurable cardinal to look like this: κ is measurable iff R_κ is closed under certain operations.

(Shoenfield [1977, p. 343]) Others are more harsh:

> Also there are axioms such as that of the Measurable Cardinal which are more powerful than the most general Axiom of Infinity yet considered, but for which there seems absolutely no intuitively convincing evidence for either rejection or acceptance.

(Cohen [1971, pp. 11–12]) Against this we should point out that the very fact that the Axiom of Measurable Cardinals implies the existence of so many small large cardinals provides evidence based on *maximize.*

The rule of thumb most commonly cited in discussions of measurable cardinals is *uniformity* (see Wang [1974, p. 555]; Drake [1974, p. 177]; Kanamori and Magidor [1978, pp. 108–109]; Martin [PSCN, p. 8]). A measure on a cardinal κ is a division of its subsets into large and small in such a way that κ is large, $\varnothing$ and singletons are small, complements of large sets are small and small sets large, and intersections of fewer than κ large sets remain large. A measure on $\aleph_0$ is formed by extending the cofinite filter to an ultrafilter. Thus $\aleph_0$ is measurable, so *uniformity* implies that there are uncountable measurable cardinals. To apply *whimsical identity* instead, notice that if there were no uncountable measurable cardinals, then $\aleph_0$ could be defined as the infinite measurable cardinal. (2 is also measurable.)

Unfortunately, as pointed out in connection with weakly compact cardinals, *uniformity* can lead to inconsistencies. Thus in cases where this is the main rule of thumb used, extrinsic evidence and evidence for relative consistency are both extremely important. Before turning to these, I should mention that Reinhardt

---

[22] This particular voice of caution belonged to my thesis advisor, John Burgess.

[23] Moore [198?] points out that Gödel's attitude towards measurable cardinals had softened by 1966 when he thought their existence "followed from the existence of generalizations of Stone's representation theorem to Boolean algebras with operations on infinitely many elements".

[1974] has proved the existence of a measurable cardinal from a system which embodies what he claims to be a version of *reflection*. Martin, however, calls these "pseudo-reflection principles" [PSCN, p. 8] and Wang remarks that:

> ... reflection principles of diverse forms which are strong enough to justify measurable cardinals (by way of 1-extendible [cardinals]) no longer appear to be clearly implied by the iterative concept of set. [1974, p. 555]

I will take this up Reinhardt's ideas later, in [BAII, §VI], in connection with his closely-related defense of supercompact cardinals.

Given how seriously the Axiom of Measurable Cardinals has been pursued, it may seem surprising that the intrinsic and rule of thumb evidence is so scarce, but in this case the extrinsic evidence is extremely persuasive. The two most impressive consequences of the existence of measurable cardinals are that $V \neq L$ and that Σ^1_1 sets of reals are determined (Martin [1970]). I will discuss the second of these in [BAII, §V].

The first indication that the presence of measurable cardinals rules out $V = L$ came in Scott [1961], where he shows (using an ultrapower construction) that the measure on a measurable cardinal cannot be constructible. This is a welcome result ("so much the worse for the 'unnatural' constructible sets!" says Scott [1977, p. xii]), but perhaps not completely surprising given how complicated a measure must be. The nonconstructibility was brought closer to home by Rowbottom [1964] when he showed that the presence of a measurable cardinal guarantees a nonconstructible set of integers. Even further improvement came in Silver [1966] and Solovay [1967], where the nonconstructible set of integers is shown to be as simple as Δ^1_3. Notice that these results (as well as Martin's on the determinacy of analytic sets) confirm Gödel's prediction that the postulation of large cardinal numbers might yield new facts about sets of integers and reals.

Silver's model theoretic results show that $V \neq L$ can actually be derived from the existence of one particular Δ^1_3 set of integers, $0^\#$. $0^\#$ codes a set of formulas which in turn show how the constructible universe is generated by a proper class of order indiscernibles that contains all uncountable cardinals and more. Thus, not only does Silver's theory show *that* L goes wrong, it shows *how* L goes wrong: the process of taking only definable subsets at each stage yields a model statisfying ZFC at some countable stage, and all the further stages make no difference (this countable structure is an elementary substructure of L). The range of L's quantifiers is so deficient that L cannot tell one uncountable cardinal from another, or even from some countable ones. In purpler terms:

> L takes on the character of a very thin inner model indeed, bare ruined choirs appended to the slender life-giving spine which is the class of ordinals.

(Kanamori and Magidor [1978, p. 131]) The point is that $0^\#$ yields a rich explanatory theory of exactly where and why L goes wrong. Before Silver, many mathematicians believed that $V \neq L$, but after Silver they knew why.[24]

[24] Ironically, Silver's subsequent efforts have been to prove that measurable cardinals are inconsistent.

Thus the assumption that $0^{\#}$ exists presents a very attractive form of $V \neq L$. (Actually, the most attractive assumption is that $x^{\#}$ exists for every real number x, where $x^{\#}$ codes the indiscernible construction that shows $V \neq L[x]$.) This assumption turns out to be equivalent to a determinacy assumption and to an elementary embedding assumption (see [BAII, §§V and VI], respectively). This prompts Kanamori and Magidor [1978] to another application of diversity:

> [This] is a list of equivalences, much deeper than the confluence seen at weak compactness. (p. 140)

Thus the implication of the existence of the sharps provides a very appealing extrinsic support for the Axiom of Measurable Cardinals.

Another sort of extrinsic support comes from the fact that measurable cardinals allow patterns of results provable in ZFC alone to be extended. For example, as mentioned in II.3.1, Souslin's theorem that Σ^1_1 sets have the perfect subset property can be extended, in the presence of a measurable cardinal, to the Σ^1_2 sets (Solovay [1969]). The same goes for Lebesgue measurability and the Baire Property. (Borel determinacy and Σ^1_1 determinacy provide another example. See [BAII, §V].) Kanamori and Magidor emphasize yet another form of extrinsic support when they stress:

> ...the fruitfulness of the methods introduced in the context of large cardinal theory in leading to new,'standard' theorems of ZFC. (p. 105)

Many of these new methods (e.g. Silver forcing (p. 152), Ulam matrices (p. 162)), arose in studies of measurable cardinals. They also mention connections with other branches of mathematics (p. 109).

Finally, as promised, I should mention some of the evidence presented for the relative consistency of measurable cardinals. One idea is to show that various strong consequences of the Axiom of Measurable Cardinals are relatively consistent themselves; then we know at least that any inconsistency that follows from the existence of a measurable cardinal is not to be reached by those particular routes. So, for example, we know that $V \neq L$ is relatively consistent (Cohen [1966]), and furthermore, that the existence of a Δ^1_3 nonconstructible set of integers with properties much like those of $0^{\#}$ is relatively consistent (Jensen [1970] and Jensen and Solovay [1970]). Another line of argument runs that:

> ...some comfort can be gained from the fact that any number of attempts at showing that measurable cardinals do not exist have failed even though much cleverness was expended.

(Scott [1977, p. xii]). (See footnote 24.)

This is already a quite impressive list of extrinsic supports, but at least two more can be added. First, there is the inner model $L[U]$, where U is a normal measure on some uncountable cardinal κ. This model is the smallest in which κ is measurable, and it does not depend on the particular choice of U. Surprisingly, $L[U]$ shares many of the simplifying structural features of L: GCH is true, and there is a Δ^1_3 well-ordering of the reals (Silver [1971]). But this is only the beginning. Kunen's analysis of $L[U]$ via iterated ultrapowers [1970], and the work of Solovay and Dodd and

Jensen [1977] on the fine structure of $L[U]$ have revealed the "uniform generation and combinatorial clarity" of this inner model in considerable detail (Kanamori and Magidor [1978, p. 147]). The familiarity and depth of understanding provided by this inner model theory leads modern researchers to the view that:

> One of the main plausibility arguments for measurable cardinals is that they have natural inner models.

(Kanamori and Magidor [1978, p. 147]) The canonical inner model makes measurable cardinals much less mysterious.

Second, concentration on the nontrivial elementary embedding of V into a transitive model M that is provided by Scott's ultrapower construction revealed that the existence of such an embedding is in fact equivalent to the existence of a measurable cardinal. (The first ordinal moved by such an embedding must be measurable.)

> Thus, the really structural characterization of measurable cardinals in set theory emerged.

(Kanamori and Magidor [1978, p. 110]) Though this is not quite the *inexhaustibility* characterization that was hoped for, it is simple and basic, and it does lead to many fruitful generalizations (see [BAII, §VI]). If the definition in terms of measures or ultrafilters had once seemed unmotivated, the connection with elementary embedding via ultrapowers revealed its true nature. Furthermore, elementary embedding cardinals are more amenable to the rule of thumb justifications that elude measurable cardinals in their original guise (see [BAII, §VI]).

Given this wide range of support for the Axiom of Measurable Cardinals, it is perhaps not surprising that proof from that axiom, at least among members of the Cabal group, has come to be treated as tantamount to proof outright. Here we have, as Gödel predicted, an axiom so rich in extrinsic supports that:

> ... whether or not [it is] intrinsically necessary, [it is] accepted at least in the same sense as any well-established physical theory. [1947/64, p. 477]

Unfortunately, for all that, it cannot answer the question we had hoped it would. Levy and Solovay [1967] have shown that measurable cardinals, and indeed all large cardinals of the sort developed so far, are preserved under most forcing extensions, and thus, that they can be shown to be relatively consistent with both the continuum hypothesis and its negation.

In [BAII], I will consider axiom candidates of a completely different sort—hypotheses on determinacy—along with the larger large cardinals.

REFERENCES

Y. Bar-Hillel, editor

[1970] ***Mathematical logic and foundations of set theory***, North-Holland, Amsterdam, 1970.

J. Barwise, editor

[1977] ***Handbook of mathematical logic***, North-Holland, Amsterdam, 1977.

P. Benacerraf and H. Putnam, editors

[1983] ***Philosophy of mathematics***, 2nd ed., Cambridge University Press, Cambridge, 1983.

G. Boolos

[1971] *The iterative conception of set*, in Benacerraf and Putnam [1983], pp. 486–502.

L. E. J. BROUWER

[1912] *Intuitionism and formalism*, in Benacerraf and Putnam [1983], pp. 77–89

G. CANTOR

[1883] *Über unendliche, lineare Punktmannigfaltigkeiten*. V, ***Mathematische Annalen***, vol. 21 (1883), pp. 545–591.

[1895] ***Contributions to the founding of the theory of transfinite numbers*** (translated and with an introduction by P.E.B. Jourdain), Open Court Press, Chicago, Illinois, 1915; reprint, Dover, New York, 1952.

[1899] *Letter to Dedekind*, in van Heijenoort [1967], pp. 113–117.

P. COHEN

[1966] ***Set theory and the continuum hypothesis***, Benjamin, ~~Reading, Massachusetts~~ New York, 1966.

[1971] *Comments on the foundations of set theory*, in Scott [1971], pp. 9–15.

R. DEDEKIND

[1888] ***Essays on the theory of numbers*** (translated by W. W. Beman), Open Court Press, Chicago, Illinois, 1901; reprint, Dover, New York, 1963.

K. J. DEVLIN

[1977] ***The axiom of constructibility***, Springer-Verlag, Berlin, 1977.

A. DODD and R. B. JENSEN

[1977] *The core model*, ***Annals of Mathematical Logic***, vol. 20 (1981), pp. 43–75.

F. DRAKE

[1974] ***Set theory***, North-Holland, Amsterdam, 1974.

E. ELLENTUCK

[1975] *Gödel's square axioms for the continuum*, ***Mathematische Annalen***, vol. 216 (1975), pp. 29–33.

A. A. FRAENKEL, Y. BAR-HILLEL, and A. LEVY

[1973] ***Foundations of set theory***, 2nd ed., North-Holland, Amsterdam, 1973.

C. FREILING

[1986] *Axioms of symmetry: throwing darts at the real number line*, this JOURNAL, vol. 51 (1986), pp. 190–200.

H. M. FRIEDMAN

[1971] *Higher set theory and mathematical practice*, ***Annals of Mathematical Logic***, vol. 2 (1971), pp. 325–357.

[1981] 'On the necessary use of abstract set theory', Advances in Mathematics 41 (1981), pp. 209–280.

J. I. FRIEDMAN

[1971] *The generalized continuum hypothesis is equivalent to the generalized maximization principle*, this JOURNAL, vol. 36 (1971), pp. 39–54.

K. GÖDEL

[1938] *The consistency of the axiom of choice and of the generalized continuum-hypothesis*, ***Proceedings of the National Academy of Sciences of the United States of America***, vol. 24 (1938), pp. 556–557.

[1944] *Russell's mathematical logic*, in Benacerraf and Putnam [1983], pp. 447–469.

[1946] *Remarks before the Princeton bicentennial conference on problems in mathematics*, ***The undecidable*** (M. Davis, editor), Raven Press, Hewlett, New York, 1965, pp. 84–88.

[1947/64] *What is Cantor's continuum problem?*, in Benacerraf and Putnam [1983], pp. 470–485.

H. HALLETT

[1984] ***Cantorian set theory and limitation of size***, Oxford University Press, Oxford, 1984.

T. J. JECH, editor

[1974] ***Axiomatic set theory***, Proceedings of Symposia in Pure Mathematics, vol. 13, part II, American Mathematical Society, Providence, Rhode Island, 1974.

R. B. JENSEN

[1970] *Definable sets of minimal degree*, in Bar-Hillel [1970], pp. 122–128.

R. B. JENSEN and R. M. SOLOVAY

[1970] *Some applications of almost disjoint sets*, in BarHillel [1970], pp. 84–104.

P. JOURDAIN

[1904] *On the transfinite cardinal numbers of well-ordered aggregates*, ***Philosophical Magazine***, vol. 7 (1904), pp. 61–75.

[1905] *On transfinite numbers of the exponential form*, ***Philosophical Magazine***, vol. 9 (1905), pp. 42–56.

K. KANAMORI and M. MAGIDOR
[1978] *The evolution of large cardinal axioms in set theory*, ***Higher set theory*** (G. H. Müller and D. S. Scott, editors), Lecture Notes in Mathematics, vol. 669, Springer-Verlag, Berlin, 1978, pp. 99–275.

G. KREISEL
[1980] *Kurt Gödel*, ***Biographical memoirs of fellows of the Royal Society***, vol. 26 (1980), pp. [redacted] 149–224.

K. KUNEN
[1970] *Some applications of iterated ultrapowers in set theory*, ***Annals of Mathematical Logic***, vol. 1 (1970), pp. 179–227.

A. LEVY
[1960] *Axiom schemata of strong infinity in axiomatic set theory*, ***Pacific Journal of Mathematics***, vol. 10 (1960), pp. 223–238.

A. LEVY and R. M. SOLOVAY
[1967] *Measurable cardinals and the continuum hypothesis*, ***Israel Journal of Mathematics***, vol. 5 (1967), pp. 234–248.

P. MADDY
[BAII] *Believing the axioms*. II, this JOURNAL (to appear).

D. A. MARTIN
[PSCN] *Projective sets and cardinal numbers*, circulated photocopy.
[SAC] *Sets versus classes*, circulated photocopy.
[1970] *Measurable cardinals and analytic games*, ***Fundamenta Mathematicae***, vol. 66 (1970), pp. 287–291.
[1975] *Borel determinacy*, ***Annals of Mathematics***, [redacted] vol. 102 (1975), pp. 363–371.
[1976] *Hilbert's first problem: the continuum hypothesis*, *Mathematical developments arising from Hilbert problems* (F. E. Browder, editor), Proceedings of Symposia in Pure Mathematics, vol. 28, American Mathematical Society, Providence, Rhode Island, 1976, pp. 81–92.

D. A. MARTIN and R. M. SOLOVAY
[1970] *Internal Cohen extensions*, ***Annals of Mathematical Logic***, vol. 2 (1970), pp. 143–178.

D. MIRIMANOFF
[1917] *Les antinomies de Russell et de Burali-Forti et le problème fondamental de la théorie des ensembles*, ***L'Enseignement Mathématique***, vol. 19 (1917), pp. 37–52.
Remarques sur la théorie des ensembles et les antinomies Cantoriennes. I, ***L'Enseignement Mathématique***, vol. 19 (1917), pp. 208–217.

G. H. MOORE
[1982] ***Zermelo's axioms of choice***, Springer-Verlag, Berlin, 1982.
[198?] *Introductory note to 1947 and 1964*, ***The collected works of Kurt Gödel***. Vol. II, Oxford University Press, Oxford (to appear).

Y. N. MOSCHOVAKIS
[1980] ***Descriptive set theory***, North-Holland, Amsterdam, 1980.

C. PARSONS
[1974] *Sets and classes*, ***Noûs***, vol. 8 (1974), pp. 1–12.
[1977] *What is the iterative conception of set?*, in Benacerraf and Putnam [1983], pp. 503–529.

W. N. REINHARDT
[1974] *Remarks on reflection principles, large cardinals and elementary embeddings*, in Jech [1974], pp. 189–205.
Set existence principles of Shoenfield, Ackermann, and Powell, ***Fundamenta Mathematicae***, vol. 84 (1974), pp. 5–34.

R. ROWBOTTOM
[1964] ***Some strong axioms of infinity incompatible with the axiom of constructibility***, Ph.D. dissertation, University of Wisconsin, Madison, Wisconsin, 1964; reprinted in ***Annals of Mathematical Logic***, vol. 3 (1971), pp. 1–44.

B. RUSSELL
[1906] *On some difficulties in the theory of transfinite numbers and order types*, ***Proceedings of the London Mathematical Society***, ser. 2, vol. 4 (1906), pp. 29–53.

D. S. SCOTT
[1961] *Measurable cardinals and constructible sets*, ***Bulletin de l'Académie Polonaise des Sciences. Série des Sciences Mathématiques, Astronomiques et Physiques***, vol. 7 (1961), pp. [redacted] 521–524.

[1977] Foreword to J. L. Bell, ***Boolean-valued Models and independence proofs in set theory***, Clarendon Press, Oxford, 1977.

D. S. SCOTT, editor

[1971] ***Axiomatic set theory***, Proceedings of Symposia in Pure Mathematics, vol. 13, part I, American Mathematical Society, Providence, Rhode Island, 1971.

J. R. SHOENFIELD

[1977] *Axioms of set theory*, in Barwise [1977], pp. 321–344.

J. SILVER

[1966] ***Some applications of model theory in set theory***, Ph.D. dissertation, University of California, Berkeley, California, 1966; reprinted in ***Annals of Mathematical Logic***, vol. 3 (1971), pp. 45–110.

[1971] *The consistency of the GCH with the existence of a measurable cardinal*, in Scott [1971], pp. 391–395.

R. M. SOLOVAY

[1967] *A nonconstructible Δ^1_3 set of integers*, ***Transactions of the American Mathematical Society***, vol. 127 (1967), pp. 50–75.

[1969] *The cardinality of Σ^1_2 sets*, ***Foundations of Mathematics*** (J. J. Bulloff et al., editors), Symposium papers commemorating the sixtieth birthday of Kurt Gödel, Springer-Verlag, Berlin, 1969, pp. 58–73.

[1970] *A model of set-theory in which every set of reals is Lebesgue measurable*, ***Annals of Mathematics***, ser. 2, vol. 92 (1970), pp. 1–56.

R. M. SOLOVAY, W. N. REINHARDT, and A. KANAMORI

[1978] *Strong axioms of infinity and elementary embeddings*, ***Annals of Mathematical Logic***, vol. 13 (1978), pp. 73–116.

L. J. STANLEY

[1985] *Borel diagonalization and abstract set theory: recent results of Harvey Friedman*, ***Harvey Friedman's research on the foundations of mathematics*** (L. A. Harrington et al., editors), North-Holland, Amsterdam, 1985, pp. 11–86.

A. TARSKI

[1938] *Über unerreichbare Kardinalzahlen*, ***Fundamenta Mathematicae***, vol. 30 (1938), pp. 68–89.

S. ULAM

[1930] *Zur Masstheorie in der allgemeinen Mengenlehre*, ***Fundamenta Mathematicae***, vol. 16 (1930), pp. 140–150.

J. VAN HEIJENOORT, editor

[1967] ***From Frege to Gödel. A source-book in mathematical logic, 1879–1931***, Harvard University Press, Cambridge, Massachusetts, 1967.

J. VON NEUMANN

[1923] *On the introduction of transfinite numbers*, in van Heijenoort [1967], pp. 346–354.

[1925] *An axiomatization of set theory*, in von Heijenoort [1967], pp. 393–413.

H. WANG

[1974] *The concept of set*, in Benacerraf and Putnam [1983], pp. 530–570.

[1981] *Some facts about Kurt Gödel*, this JOURNAL, vol. 46 (1981), pp. 653–659.

E. ZERMELO

[1904] *Proof that every set can be well-ordered*, in van Heijenoort [1967], pp. 139–141.

[1908] *A new proof of the possibility of a well-ordering*, in van Heijenoort [1967], pp. 183–198.

Investigations in the foundations of set theory. I, in van Heijenoort [1967], pp. 199–215.

[1930] *Über Grenzzahlen und Mengenbereiche*, ***Fundamenta Mathematicae***, vol. 16 (1930), pp. 29–47.

DEPARTMENT OF PHILOSOPHY
UNIVERSITY OF CALIFORNIA AT IRVINE
IRVINE, CALIFORNIA 92717

THE JOURNAL OF SYMBOLIC LOGIC
Volume 53, Number 3, Sept. 1988

BELIEVING THE AXIOMS. II

PENELOPE MADDY

This is a continuation of *Believing the axioms.* I,[1] in which nondemonstrative arguments for and against the axioms of ZFC, the continuum hypothesis, small large cardinals and measurable cardinals were discussed. I turn now to determinacy hypotheses and large large cardinals, and conclude with some philosophical remarks.

§V. **Determinacy.** Determinacy is a property of sets of reals.[2] If A is such a set, we imagine an infinite game $G(A)$ between two players I and II. The players take turns choosing natural numbers. In the end, they have generated a real number $\mathbf{r}$ (actually a member of the Baire space ${}^{\omega}\omega$). If $\mathbf{r}$ is in A, I wins; otherwise, II wins. The set A is said to be determined if one player or the other has a winning strategy (that is, a function from finite sequences of natural numbers to natural numbers that guarantees the player a win if he uses it to decide his moves).

Determinacy is a "regularity" property (see Martin [1977, p. 807]), a property of well-behaved sets, that implies the more familiar regularity properties like Lebesgue measurability, the Baire property (see Mycielski [1964] and [1966], and Mycielski and Swierczkowski [1964]), and the perfect subset property (Davis [1964]). Infinitary games were first considered by the Polish descriptive set theorists Mazur and Banach in the mid-30s; Gale and Stewart [1953] introduced them into the literature, proving that open sets are determined and that the axiom of choice can be used to construct an undetermined set.

Gale and Stewart also raised the question of whether or not all Borel sets are determined, but the answer was long in coming. Wolfe [1955] quickly established the determinacy of Σ^0_2 games, but it was not until [1964] that Davis showed the same for Σ^0_3 games. It was [1972] before Paris was able to extend the result to Σ^0_4, and by that point the proof had become fiendishly complex. Martin then capped the whole enterprise with his surprising proof of Borel Determinacy in [1975]. (This result was

Received March 4, 1987.

[1] Maddy [BAI].

[2] My [1984] contains an earlier version of some of the material in this section. (The support of the NEH during that work is again gratefully acknowledged.) The standard (and excellent) reference for both historical and mathematical information on determinacy is Moschovakis [1980]. Notice that by "reals" modern descriptive set theorists mean members of the Baire space ${}^{\omega}\omega$, which is homeomorphic to the irrationals.

0022-4812/88/5303-0005/$03.90

mentioned in [BAI, §§I.8 and III] because of its essential use of Replacement.) Given that $V = L$ implies the existence of a nonmeasurable Δ^1_2 set, this result is the best possible; any further determinacy goes beyond ZFC.[3]

The first determinacy hypothesis, suggested by Steinhaus in Mycielski and Steinhaus [1962], was the full Axiom of Determinacy (AD), that is, the assumption that every set of reals is determined. Given that it contradicts the Axiom of Choice, the authors did not propose AD as a truth about V, but rather as applying to some substructure thereof:

> It is not the purpose of this paper to depreciate the classical mathematics with its fundamental 'absolute' intuitions on the universum of sets (to which belongs the axiom of choice), but only to propose another theory which seems very interesting ... Our axiom can be considered as a restriction of the classical notion of a set leading to a smaller universum, say of determined sets, which reflect some physical intuitions which are not fulfilled by the classical sets ([the various pathologies implied by Choice] are eliminated). Our axiom could be considered as an axiom added to the classical set theory claiming the existence of a class of sets satisfying [AD] and the classical axioms (without the axiom of choice). (p. 2)

Though the Axiom of Choice implies the existence of various extremely complex sets (for example, non-Lebesgue measurable sets, uncountable sets without perfect subsets, well-orderings of the reals, etc.), the Axiom of Determinacy might still hold in some inner model of ZF (ZFC, without Choice). This inner model would then consist only of regular sets; the irregular sets would appear in the more remote parts of V:

> We can only hope that some submodels of the natural models of [ZF] are models of [ZF + AD] ... It would be still more pleasant if such a submodel contains all the real numbers...In that case [AD] may be considered as a limitation of the notion of a set excluding some 'pathological' [ZF] sets.

(Mycielski [1964, p. 205]) The smallest such model is $L[R]$, and the Axiom of Quasi-Projective Determinacy (QPD)[4] is the assumption that all sets of reals in this submodel are determined. This is the live axiom candidate (see e.g. Moschovakis [1970, p. 31] and [1980, pp. 422 and 605]; Martin [PSCN, p. 8]).[5]

---

[3] I mean, full Σ^1_1 or Π^1_1 determinacy go beyond ZFC. Modest gains beyond Borel determinacy are possible without additional assumptions. See, for example, Wolf [1985].

[4] This assumption is usually written symbolically as $AD^{L[R]}$ and otherwise unnamed. In [1969], Solovay uses the term "quasiprojective" for the sets of reals in $L[R]$, so I have adopted his terminology.

[5] A weaker assumption, the Axiom of Projective Determinacy (PD), is also discussed in the literature. (PD is naturally the assumption that all projective sets of reals are determined; it is weaker than QPD because all projective sets appear in $L[R]$.) QPD is the better axiom candidate because the projective hierarchy is only the second of a series of hierarchies, while $L[R]$ is a transitive model of ZFC generated in a natural way.

It is worth noting that QPD has a form reminiscent of a number of mathematical implications (see Fenstad [1971, p. 42]; Addison and Moschovakis [1968, p. 710]). The quantifier switch

$$\exists x \forall y Rxy \supset \forall y \exists x Rxy$$

is a theorem of logic, but the other direction

$$\forall x \exists y Rxy \supset \exists y \forall x Rxy$$

is nontrivial. Determinacy assumptions have the second form: if for every strategy for I, there is a way for II to play that results in a win for II, then there is a strategy for II that results in a win for II no matter what I plays. (In other words, if I has no winning strategy, then II does.) This same nontrivial quantifier switch is seen in various mathematical contexts, for example, in the implication from continuity to uniform continuity. An implication of this sort usually requires a simplifying assumption—in the continuity example, that the space in question is compact. In the case of QPD, the simplifying assumption is that the set in question is constructible from the reals. So QPD at least has a general form that is familiar from other parts of mathematics.

Still, as far as intrinsic evidence for QPD is concerned, even its staunchest supporters are emphatic in their denials:

> No one claims direct intuitions...either for or against determinacy hypotheses.
>
> There is no *a priori* evidence for [Q]PD.
>
> Is [Q]PD true? It is certainly not self-evident.

(Moschovakis [1980, p. 610]; Martin [1976, p. 90]; Martin [1977, p. 813]; see also Wang [1974, p. 554]). What sets QPD apart (or what did set it apart before the recent discoveries discussed in the next section) is that its defense has been purely extrinsic. Even the most skeptical among the supporters of large cardinals admit that extending the sequences of ordinals is intrinsic to the iterative conception of set (for example, *maximize*). Nothing of this type whatsoever was offered for QPD from its origin until the mid-80s.[6] Yet it has been taken very seriously as an axiom candidate.

---

[6]To be accurate, I should admit one exception:

> The reader who knows the Zermelo-von Neumann theorem on the strict determinateness of finite positional games could accept perhaps the following 'intuitive justification' of [AD]. Suppose that both players I and II are infinitely clever and that they know perfectly well what [the set of reals] P is, then owing to the complete information during every play, the result of the play cannot depend on chance. [AD] expresses exactly this.

(Mycielski and Steinhaus [1962, p. 1]) I ignore this argument for two reasons. First, it supports, if anything, the full, false, AD. Second, it has not been adopted by subsequent researchers.

On the other hand, the appeal to extrinsic supports has been quite explicit:

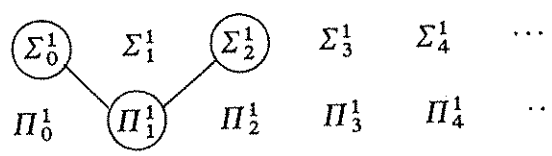
> The author regards [Q]PD as an hypothesis with a status similar to that of a theoretical hypothesis in physics ... quasi-empirical evidence for [Q]PD [has] been produced.
>
> ... those who have come to favor these hypotheses as plausible, argue from their consequences ... the richness and internal harmony of these consequences.
>
> In the case of [Q]PD the evidence is mostly *a posteriori*: its consequences look right.
>
> There is a good deal of *a posteriori* evidence for it.

(Martin [1977, p. 814]; Moschovakis [1980, p. 610]; Martin [PSCN, p. 8]; Martin [1976, p. 90]). I will sketch the three main types of extrinsic evidence—from consequences, from intertheoretic connections, and from the "naturalness" of game theoretic proofs—in the next three subsections. In the final subsection, I will consider the relevance of determinacy assumptions to the continuum problem.

V.1. *Welcome consequences.* Recall that the existence of a Δ_2^1 well-ordering of the reals, a Δ_2^1 non-Lebesgue measurable set, and an uncountable Π_1^1 set with no perfect subset were all counted as extrinsic disconfirmation for $V = L$ (see BAI, §§ II.2 and II.3.1]). It was felt that these Choice-generated oddities should not appear among the simpler sets, that they should probably not be definable at all. This might be counted as a rule of thumb in favor of the *banishment* of such sets to remote regions of V far beyond the simple sets. Of course, QPD does exactly this; it contradicts $V = L$ by forcing these "irregular" sets of reals out of the projective hierarchy, and indeed, out of $L[R]$. Thus these consequences of $V = L$ are called "a defect in that theory" (Fenstad [1971, p. 59]) or "unpleasant ~~consequences~~ aspects of that theory" (Martin [1977, p. 806]), while the corresponding consequences of QPD are "pleasing consequences ~~for~~ about the behavior of projective sets" (Martin [1976, p. 90]; see also Martin [1977, p. 811]).

Another set of welcome consequences concerns the "structural" properties of the projective sets: separation, reduction and uniformization. In the 30s, Kuratowski, Lusin and Novikov established reduction and uniformization for Σ_0^1, Π_1^1, and Σ_2^1, and separation for their opposites, Π_0^1, Σ_1^1, and Π_2^1.[7] Thus the reduction and uniformization principles hold for the circled classes, while separation holds on the opposite side:

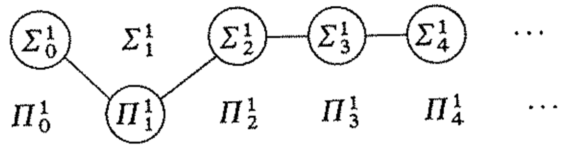
ZFC



[7] For definitions, proofs and references, see Moschovakis [1980, pp. 33–35 and 4B.10, 4B.11, and 4D.4]. Here Σ_0^1 is taken to be the open sets of reals and Π_0^1 the closed. More on this choice of notation below.

Nothing more was known until [1959], when Addison used the Δ^1_2 well-ordering of the reals in L to show that reduction and uniformization continue on the Σ side, and separation on the Π side, for the remainder of the projective hierarchy:

$V = L$



There the matter stood for nearly ten years.

The question of what pattern of structural properties to expect at the Σ^1_3/Π^1_3 level and beyond was somewhat overshadowed in the early 60s by the new forcing industry. Among those interested in the truth of this matter, rather than in relative consistency results, opinion varied. Some felt that the use of a measurable cardinal at the next level should push through the same pattern as $V = L$. Speaking in 1967, Addison considers this possibility:

> This [Σ-side] pattern at level 3 follows from the existence of a Δ^1_2 well-ordering of [the reals], which follows in turn from the axiom of constructibility. One possibility is that higher "axioms of infinity" such as the axiom of measurable cardinals might imply this pattern at the third level. From results of Silver it is known that this pattern at the third level is at least consistent with the axiom of measurable cardinals. On the other hand the axiom of measurable cardinals wipes out some nice well-orderings of [the reals] and it is thought by some that still higher axioms of infinity may be found which wipe out all projective well-orderings of [the reals]. Although nice well-orderings can be viewed as pushing in the direction of [the Σ-side] pattern, weaker principles not ruled out by higher axioms of infinity might still be enough to force it.

(Addison [1974, p. 9]) Others expected the pattern to continue alternating. Addison again:

> On the other hand if there is indeed some pressure, not yet understood, pushing for the separation principle to hold on one side or the other then it might be sufficient... to push through a[n alternating] pattern at level 3.... This might look surprising, but at least one respected logician has suggested it. It has the advantage of prolonging the alternation... (p. 10)

It should be noted that the second group outnumbered the first, and that it included Gödel (Addison [1974, p. 10]).

What reasons could be given for or against the alternating pattern? The structural properties at level three and above were strongly suspected of being independent, although this was not proved until much later (see Moschovakis [1980, p. 284]). Those who expected the continuation of the Σ-side pattern of $V = L$ had a powerful new hypothesis to work with (MC), one that had only recently begun producing results about projective sets (Solovay [1969]). Silver's work on $L[U]$ showed that their conjecture was relatively consistent, and the similarity of that model to L made them expect the same pattern. Meanwhile, those favoring the alternating picture were without a new assumption, but they were supported by the brute fact that almost any human being will judge WWW to be a "more natural" continuation of ∨ than √‾‾‾‾. (This fact is slightly compromised by its dependence on the identification of Σ^1_0 with Σ^0_1. Without this, the first "zig" of the "zigzag" is lost.[8] But see the next quotation from Addison below.)

Moschovakis had a deeper reason for expecting the alternation to continue. In the mid-60s, he showed how the prewellordering property could be used to lift the structural theory of Π^1_1 sets to Σ^1_2. A prewellordering misses being a full wellordering by lacking antisymmetry; equivalently, it is a mapping onto an ordinal. A class of sets of reals has the prewellordering property (PWO) if every set in it admits a prewellordering that meets a delicate definability condition (see Moschovakis [1980, 4B] for details). PWO(Π^1_1) is essentially a classical theorem proved by Lusin and Sierpiński in 1923 using something called the Lusin-Sierpiński ordering.[9]

Since the prewellordering property is the key to the structural properties of the projective classes, Moschovakis's idea was to prove:

$$\mathrm{PWO}(\Pi^1_1) \Rightarrow \mathrm{PWO}(\Sigma^1_2)$$

thus effectively lifting the theory of Π^1_1 to Σ^1_2. The proof takes a simple form. Suppose A is Σ^1_2. Then there is a Π^1_1 B such that

$$A = \{x \mid \exists z((x,z) \in B\}.$$

If f maps B onto an ordinal as PWO(Π^1_1) requires, then a suitably definable prewellordering of A is achieved by taking infimums: for x, y in A,

$$x \lesssim y \quad \text{iff} \quad \inf\{f(z) \mid (x,z) \in B\} \leq \inf\{f(z) \mid (y,z) \in B\}.$$

In fact, this proof is perfectly general; whenever PWO holds at Π^1_n it can be lifted to Σ^1_{n+1}.

If a proof using infimums moves the prewellordering property from Π^1_n to Σ^1_{n+1}, shouldn't a proof using supremums move it from Σ^1_n to Π^1_{n+1}? Of course this form of proof cannot work without a new hypothesis, because Addison's results from $V = L$ show that PWO(Σ^1_3) and not-PWO(Π^1_3) are relatively consistent. The trouble is that the prewellordering defined using supremums is often trivial, so the definability condition does not hold unless the set in question is more special than Π^1_3. Still, Moschovakis felt the failed argument was a "false, but natural" proof, too reasonable to be completely off-base, too natural to be a totally wrong idea. The flaw seemed akin to dividing by zero in a proof that is otherwise in order; some minor

[8] Opinion on this matter varies. Some would find it more natural to start the projective hierarchy from the Borel sets, which would destory the first leg of the alternation. It is worth noting that the Σ^0_1 beginning does not work for the actual reals (see Martin [1977, p. 790]). (Recall that modern descriptive set theory is done on the Baire space ${}^{\omega}\omega$ instead.)

[9] The "lightface" or effective version of the same theorem was proved in 1955 by Kleene. There the ordering is called the Brouwer-Kleene ordering.

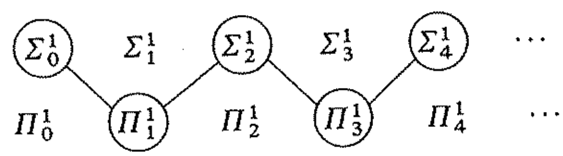
adjustment—multiplying through by a factor before dividing, or some such thing—should be enough to make it go through. This line of thought led Moschovakis to the conjecture that $\mathrm{PWO}(\Pi^1_3)$, and to favor the alternating pattern.[10]

Then came Blackwell [1967], in which an elegant new proof of Lusin's classical theorem on the separability of Σ^1_1 sets is derived from the determinacy of open sets. Addison and Martin quickly adopted the method to show that the reduction property for Π^1_3 sets could be derived directly from the determinacy of Δ^1_2 sets. When Moschovakis heard of this work, he had his additional hypothesis. The "false but natural" proof could be revived by requiring one supremum to be "effectively" smaller than the other, where "effectively" is parsed out in terms of a determined game. In general, then, we get the periodicity theorem:

$$\mathrm{Det}(\Delta^1_n) \Rightarrow (\mathrm{PWO}(\Sigma^1_n) \Rightarrow \mathrm{PWO}(\Pi^1_{n+1})).$$

(See Addison and Moschovakis [1968]. Martin proved the same thing independently, using degree theoretic methods, in [1968].) Thus under the assumption of QPD, the alternating pattern of structural properties continues for the remainder of the projective hierarchy:

QPD



This result—the extension of the alternating pattern—is now considered strong extrinsic evidence in favor of QPD. In a footnote added to the pro-alternation paragraph in the printed version of his talk, Addison remarks:

> This paragraph turned out to be prophetic. Only a month or so after the talk was given it was shown that [QPD] does indeed give the alternation of ... patterns discussed here ... Moreover the reasons mentioned above for the plausibility of this hypothesis actually lie behind the proof of the alternating pattern from ... determinateness. Furthermore the idea of considering Σ^0_1 and Π^0_1 as the first level of the projective hierarchy is not only completely vindicated by the outcome but indeed actually suggested the structure of the proof of the alternating pattern. [1974, p. 10]

(This last because Blackwell's proof depended on the determinacy of Δ^0_1 sets.) Nowadays we read:

> Which is the correct picture is perhaps not absolutely clear yet, but it is fair to say that many people working in this area and prone to speak of truth in set theory (ourselves included) tend to favor the alternating picture.
>
> ... there is something odd about [redacted] the sequence Σ^1_0, Π^1_1, Σ^1_2, Σ^1_3, $\Sigma^1_4, \ldots$, the sequence Σ^1_0, Π^1_1, Σ^1_2, Π^1_3, [redacted] ... seems more plausible.

[10]Notice that it was the structure of a proof (or an attempted proof) that produced this conjecture. Philosophers are often tempted to think that conjectures are formed by a process quite independent of proof, but this case suggests otherwise.

(Moschovakis [1980, pp. 33–34]; Martin [1977, pp. 806 and 811]; see also Martin [PSCN, p. 8]; Fenstad [1971, p. 59]; Wang [1974, pp. 547 and 553–554]).

To summarize these extrinsic supports:

> [Q]PD has pleasing consequences about the behavior of projective sets, such as: Every projective set is Lebesgue measurable; Every uncountable projective set has a perfect subset. More impressive is the fact that [Q]PD allows one to extend the classical structural theory of projective sets, which dealt only with the first two levels of the projective hierarchy, to a very elegant and essentially complete theory of the projective sets. [Q]PD cannot be proved in ZFC... but it is not unreasonable to suspect that it may be true.

(Martin [1976, p. 90]) The full extent of this beautiful and remarkably detailed theory of the projective sets is clearly laid out in Moschovakis's [1980].[11]

V.2. *Intertheoretic connections.* Despite the early (erroneous) suggestion that the Axiom of Measurable Cardinals (MC) might force the "wrong" resolution for the structural properties at the third level, measurable cardinals and determinacy hypotheses were soon found to point in the same direction. Often QPD will extend a result provable from MC which in turn extends a result provable in ZFC. To use an example that has already been discussed, ZFC implies that every uncountable Σ^1_1 set has a perfect subset, MC implies every uncountable Σ^1_2 set has a perfect subset, and QPD implies every uncountable quasiprojective set has a perfect subset. Another case in point:

$$\mathrm{ZF} \Rightarrow \text{Every } \Sigma^1_2 \text{ set is the union of } \aleph_1 \text{ Borel sets,}$$
$$\mathrm{MC} \Rightarrow \text{Every } \Sigma^1_3 \text{ set is the union of } \aleph_2 \text{ Borel sets,}$$
$$\mathrm{QPD} \Rightarrow \text{Every } \Sigma^1_4 \text{ set is the union of } \aleph_3 \text{ Borel sets.}$$

(Sierpinski [1925]; Martin [PSCN]) In other cases, the same result can be proved using either MC or a determinacy hypothesis:

> To take one example, the fact that Π^1_2 sets can be uniformized by Π^1_3 sets follows both from MC and from Det(Δ^1_2), but by proofs which (at least on the surface) are totally unrelated; one tends to believe the result then and consequently to take both proofs seriously and to feel a little more sympathetic towards their respective hypotheses.

(Moschovakis [1980, p. 610]) But these observations mark only the beginning of the deep connection between measurable cardinals and lower forms of determinacy.

The first two results suggesting this connection were Solovay's of 1967:

$$\mathrm{AD} \Rightarrow \aleph_1 \text{ is a measurable cardinal}$$

[11] It has been conjectured that these implications might be reversed, that is, that a strong determinacy hypothesis might be derivable from the assumption that this rich theory holds of the projective sets. This would obviously provide considerable additional support for determinacy.

and Martin's in [1970]:

$$\text{MC} \Rightarrow \text{Det}(\Sigma^1_1).$$

In fact, Martin's theorem only depends on the existence of the sharps, and by [1978], Harrington had proved the converse:

$$\forall x(x^{\#} \text{ exists}) \equiv \text{Det}(\Sigma^1_1).$$

Meanwhile, Solovay (building on work of Martin and Friedman) improved his result to:

$$\text{Det}(\Delta^1_2) \Rightarrow \text{There are inner models with many MCs.}$$

Martin saw the development of his and Harrington's results as following a pattern: a large cardinal axiom (MC) implies some determinacy assumption ($\text{Det}(\Sigma^1_1)$); careful analysis reveals that the hypothesis can be weakened to the existence of an inner model with a slightly smaller large cardinal and indiscernibles (here ZFC itself is viewed as "large cardinal assumption"); finally, the implication is improved to an equivalence. In view of Solovay's result, this pattern might be extended within the Δ^1_2 sets.

To get a finer breakdown of the Δ^1_2 sets, Martin turned to the "difference hierarchy" of Π^1_1 sets: A is α-Π^1_1 iff there is a sequence $(A_\beta: \beta < \alpha)$ such that each A_β is Π^1_1 and $x \in A$ iff the least β such that ($\beta = \alpha$ or $x \notin A_\beta$) is odd. (Limit ordinals are even.) Thus A is 1-Π^1_1 iff A is Π^1_1; A is 2-Π^1_1 iff A is a difference of Π^1_1 sets; A is 3-Π^1_1 iff A is the union of a difference of Π^1_1 sets and a Π^1_1 set; A is 4-Π^1_1 iff A is the union of two differences of Π^1_1 sets; and so on. The finite levels of this hierarchy generate all the Boolean combinations of Π^1_1 sets.

The theorem, then, is:

$$\text{Det}((\omega^2 \cdot \alpha + 1)\text{-}\Pi^1_1) \equiv \forall x(\text{there is an inner model of ZFC containing } x \text{ with indiscernibles and } \alpha \text{ MCs}).$$

For $\alpha = 0$, this is exactly the Martin/Harrington equivalence. For $\alpha = 1$, it is:

$$\text{Det}((\omega^2 + 1)\text{-}\Pi^1_1) \equiv \forall x(\text{there is an inner model of ZFC containing } x \text{ with indiscernibles and one MC}).$$

The canonical model $L[U]$ has one measurable cardinal and indiscernibles, and the set of formulas that codes its construction (just as $0^{\#}$ codes the construction of L) is called $0^{\dagger}$. If $x^{\dagger}$ is defined analogously with $x^{\#}$, the theorem for $\alpha = 1$ can be written

$$\text{Det}((\omega^2 + 1)\text{-}\Pi^1_1) \equiv \forall x(x^{\dagger} \text{ exists}).$$

Thus the pattern continues: the existence of two measurable cardinals implies $\text{Det}((\omega^2 + 1)\text{-}\Pi^1_1)$. Careful analysis reveals that the hypothesis can be reduced to the existence of an inner model with one measurable cardinal and indiscernibles. Finally, the implication can be reversed. The general form of Martin's theorem shows that this pattern continues through the entire difference hierarchy of Π^1_1 sets.

Thus simple game-theoretic hypotheses are equivalent to the inner model versions of measurable cardinal hypotheses for many natural classes of sets of reals within Δ^1_2.[12] This wonderful and surprising correspondence between powerful and well-supported hypotheses of such different character counts as extrinsic evidence for both.

V.3. *The naturalness of game-theoretic proofs.* Finally, there is what those involved call the "naturalness" of the proofs from QPD:

> In fact, the most persuasive argument for accepting [quasi]-projective determinacy (aside from Martin's proof of Det(Σ^1_1)) is the naturalness of the known proofs of [the periodicity theorem], both Martin's and ours.

(Moschovakis [1970, p. 34]) Given our look at the development of that theorem in V.1, it is easy to see what Moschovakis has in mind here. Not only does QPD imply PWO(Π^1_3) and the rest, it does so by means of an argument that was previously thought to be of the correct sort. The proof is "natural".

Another aspect of "naturalness" is revealed when the new game-theoretic proofs yield new, simpler proofs of old theorems, and recast them as special cases of new more powerful theorems:

> One [reason for believing QPD] is the *naturalness* of the proofs from determinacy—in each instance where we prove a property of Π^1_3 (say from Det(Δ^1_2)), the same argument gives a new proof of the same (known) property for Π^1_1, using only the determinacy of clopen sets (which is a theorem of ZF). Thus the new results appear to be natural generalizations of known results and their proofs shed new light on classical descriptive set theory. (This is not the case with the proofs from $V = L$ which all depend on the [Δ^1_2] well-ordering of [the reals] and shed no light on Π^1_1.)

(Moschovakis [1980, p. 610]) The periodicity theorem itself gives an example of this phenomenon. Recall that the classical proof of PWO(Π^1_1) involved the special properties of Π^1_1, in particular, the Lusin-Sierpiński ordering. Now for $n = 0$, the periodicity theorem is

$$\mathrm{Det}(\Delta^1_0) \Rightarrow (\mathrm{PWO}(\Sigma^1_0) \Rightarrow \mathrm{PWO}(\Pi^1_1)).$$

But determinacy of Δ^1_0 sets is just the Gale-Stewart theorem, and it is simple to show PWO(Σ^1_0). Thus the periodicity theorem provides a new proof of PWO(Π^1_1) that avoids such complexities as the Lusin-Sierpiński ordering. (See also Moschovakis [1980, p. 309].)

Another example is provided by Solovay's proofs [1969] that the regularity properties of Σ^1_1 sets could be lifted to Σ^1_2 assuming the Axiom of Measurable

---

[12] This sequence of results can be extended further. For example, Simms has shown that the existence of a measurable cardinal which is the limit of measurably many measurable cardinals implies the determinacy of countable unions of Boolean combinations of Π^1_1 sets. The hypothesis can be improved to the existence of an inner model with indiscernibles and a proper class of measurable cardinals. Then the implication can be reversed.

Of course it would be nicer if the determinacy assumptions could prove the large cardinal hypotheses outright, but this is impossible. If, for example, κ is the first measurable cardinal, then R_κ is a model of Det(Σ^1_1) but not of MC. Thus the inner model equivalences are the best possible.

Cardinals. When Martin showed that Det(Σ_1^1) could be derived from MC, he opened the way for game-theoretic proofs of these results. These new proofs avoid the complex forcing constructions of Solovay's original versions (see Moschovakis [1980, pp. 375–378, 544–546, 611]).

V.4. *Relevance to the continuum problem.* QPD gives us lots of information about the projective sets; what can it tell us about the size of the continuum? The quick answer is that it cannot settle the continuum hypothesis. (It will be easy to see why from the result of the next section.) Still, it might give us evidence for or against, or, even better, it might lead us in the direction of a larger theory that does decide the question.

Under the assumption of QPD, the perfect subset property is extended to cover the entire quasiprojective hierarchy, so the CH holds for all quasiprojective sets. As mentioned earlier [BAI, §II.3.1], the fact that CH holds for many simple sets might have been considered as evidence in its favor, except that the perfect subset property is known not to hold for all sets of reals. Thus this consequence of QPD does not really provide evidence in favor of CH.

What is at issue here is the length of the shortest well-ordering of the reals. Since a definable well-ordering yields a definable non-Lebesgue measurable set, and QPD implies that all quasiprojective sets are Lebesgue measurable, it also implies that there is no quasiprojective well-ordering of the reals. This is as it should be (see V.1). In fact, the perfect subset property implies that every quasiprojective well-ordering of a set of reals is countable. This means that no projective well-ordering can provide a counterexample to CH; we cannot test the CH by looking at the projective well-orderings.

What we can do is look into the lengths of projective prewellorderings:

> Now every Σ_1^1 prewellordering has countable length, but there is a Π_1^1 prewellordering of [the reals] of length [$\aleph_1$]... This already shows that our simple sets are more typical with respect to prewellorderings than with respect to well-orderings.

(Martin [1976, p. 89]) In particular, consider: $\delta_n^1 = \sup\{\text{length of } R \mid R \text{ is a } \Delta_n^1$ prewellordering of the reals$\}$. Information about these "projective ordinals" is information about the length of the continuum. It is a classical theorem that $\delta_1^1 = \aleph_1$; if any δ_n^1 is greater than $\aleph_1$, then the continuum hypothesis is false.

The best way to approach the question of the size of the projective ordinals under QPD is to investigate them first under the full, false, AD, then to transfer the results to $L[R]$. In the strange world of full AD, it is known that the projective ordinals form a strictly increasing sequence of regular cardinals, in particular:

$$\begin{aligned} AD \Rightarrow \delta_1^1 &= \aleph_1, \\ \delta_2^1 &= \aleph_2, \\ \delta_3^1 &= \aleph_{\omega+1}, \\ \delta_4^1 &= \aleph_{\omega+2}, \\ \delta_5^1 &= \aleph_{\omega(\omega^\omega)+1}, \\ \delta_6^1 &= \aleph_{\omega(\omega^\omega)+2}. \end{aligned}$$

(These results are due to many researchers, among them Martin, Solovay, Kunen, Mansfield, Shoenfield and Jackson. See Moschovakis [1980, 7D.11]. Incidentally, the cardinals between $\aleph_2$ and $\aleph_{\omega+1}$ are all singular assuming AD; the projective ordinals are not only regular, but measurable.) This means that in the strange world of full determinacy, the continuum hypothesis is false in the sense that the reals can be mapped onto very large ordinals. In the real world, the Axiom of Choice would then yield very large subsets of the reals, but Choice does not hold in the AD world. There, remember, all uncountable sets have perfect subsets. Thus the CH is true in the sense that there are no sets of reals of intermediate cardinality, but false in the sense that the reals can be mapped onto large ordinals. As far as the actual cardinality of the reals is concerned, in the world of full AD it is not an aleph at all, because the reals cannot be well-ordered.

What does this mean for the real world, on the assumption that both QPD and Choice hold there? Since QPD is the hypothesis that AD holds in $L[R]$, the results above hold unchanged in that inner model. From this it follows that:[13]

$$\begin{aligned} \text{QPD} \Rightarrow \delta^1_1 &= \aleph_1, \\ \delta^1_2 &= (\aleph_2)^{L[R]} \le \aleph_2, \\ \delta^1_3 &= (\aleph_{\omega+1})^{L[R]} \\ &= (\text{the first regular cardinal after } \aleph_2)^{L[R]} \le \aleph_3, \\ \delta^1_4 &\le (\aleph_3^+)^{L[R]} \le \aleph_4. \end{aligned}$$

Recently, Jackson has shown that AD implies that there are exactly three regular cardinals between δ^1_3 and δ^1_5. By reasoning similar to what gave us the above, this means that

$$\text{QPD} \Rightarrow \delta^1_5 \le \aleph_7.$$

Of course it is relatively consistent that all these inequalities are strict, and that all the projective ordinals are in fact $\aleph_1$. On the other hand, $\delta^1_2 = \aleph_2$ is also relatively consistent, and for someone looking for a theory to imply the falsity of CH, QPD would seem to make a good beginning:

> ... while our simple sets have not provably given us a counterexample to CH, the possibility that they *are* counterexamples definitely arises.

(Martin [1976, p. 89]) Working with Jackson's three intermediate cardinals in the context of QPD, Martin came to conjecture that the true picture might be

[13] Here we have another example of QPD extending a pattern that begins in ZFC and continues under MC:

$$\begin{aligned} \text{ZFC} \Rightarrow \delta^1_1 &= \aleph_1, \\ \delta^1_2 &\le \aleph_2, \\ \text{MC} \Rightarrow \delta^1_3 &\le \aleph_3, \\ \text{QPD} \Rightarrow \delta^1_4 &\le \aleph_4. \end{aligned}$$

something like this:

Regular Cardinals in $L[R]$		Regular Cardinals in V
$\aleph_1 =$	δ^1_1	$= \aleph_1$
$\aleph_2 =$	δ^1_2	$= \aleph_2$
$\aleph_{\omega+1} =$	δ^1_3	$= \aleph_3$
$\aleph_{\omega+2} =$	δ^1_4	
	α	
	β	$= \aleph_4$
$\aleph_{\omega(\omega^\omega)+1} =$	δ^1_5	$= \aleph_5$

where α and β are the two other regular cardinals between δ^1_3 and δ^1_5 in the AD world of $L[R]$.

This complex conjecture can be partly confirmed by an assumption on saturated ideals developed independently by Foreman and others (for related principles, see Foreman, Magidor and Shelah [MM]). The saturated ideal hypothesis, along with QPD, implies that $\delta^1_5 \leq \aleph_5$. It remains possible that $\delta^1_n = \aleph_n$ for odd n, but a new hypothesis would be needed, presumably one that would help us understand why $L[R]$ produces so many false cardinals, both regular and singular, between the first few regular cardinals of V. Thus the best that can be said is that the rich theory of the projective ordinals provided by determinacy hypotheses might one day contribute to a theory that could falsify CH. Of course, QPD might eventually play a role in a theory that verifies CH instead, and some members of the Cabal lean toward this possibility.

§VI. Large large cardinals—down from above. By the early 70s, then, the most productive and appealing new axiom candidate, QPD, was supported exclusively by extrinsic evidence. Still, there was hope that an intrinsic connection could be found:

> Some set theorists consider large cardinal axioms self-evident, or at least as following from *a priori* principles [rules of thumb?] implied by the concept of set. [Det(Σ^1_1)] follows from large cardinal axioms. It is possible that [Q]PD itself follows from large cardinal axioms, but this remains unproved.
>
> One way to increase the evidence for [Q]PD would be to prove it from large cardinal axioms ...

(Martin [1977, p. 813]; Martin [PSCN, p. 8]; see also Martin [1976, p. 90]). The inner model $L[U]$ discussed in [BAI, §IV], contains a measurable cardinal and a Δ^1_3 well-ordering of the reals, so its Δ^1_3 sets are not all determined. Thus it was clear that a more powerful large cardinal would be needed.

Meanwhile, inspired by the success of measurable cardinals, and the isolation of the simpler, structural characterization in terms of elementary embeddings, Solovay and Reinhardt produced stronger large cardinal axioms. I will discuss the first of these in the next subsection, and two rule of thumb arguments for its existence in the subsection following. The very largest of the large cardinals will then be introduced, and the final subsection traces the recently-revealed connections with determinacy assumptions.

VI.1. *Supercompactness.* Recall that the ultrafilter on a measurable cardinal κ generates a nontrivial elementary embedding of V into a transitive M, and conversely, that the first ordinal moved by such an elementary embedding, the "critical point", must be measurable. Many of the strong properties of measurable cardinals spring from M's closure under arbitrary sequences of length κ, but M is not closed under longer sequences. Thus, the search for a strengthening of the Axiom of Measurable Cardinals naturally led Solovay and Reinhardt to try imposing stronger closure conditions on the range of the elementary embedding. This idea led to the notion of supercompactness:

> Then all the desired fruit, suddenly ripened, were easily plucked, and appropriately enough, the new concept was dubbed *supercompactness.*

(Kanamori and Magidor [1978, p. 183]) Specifically, a cardinal κ is λ-supercompact (for $\lambda \geq \kappa$) iff there is a nontrivial elementary embedding of V into a transitive M with κ critical and M closed under arbitrary sequences of length λ; a cardinal κ is supercompact iff κ is λ-supercompact for all $\lambda \geq \kappa$.

The connections between measurability and supercompactness are quite simple: κ is measurable iff it is κ-supercompact, and below a supercompact κ there are κ measurable cardinals. Furthermore, like measurability, supercompactness also has an ultrafilter characterization. Thus supercompact cardinals are thought of as "the proper generalization of measurability" (Solovay, Reinhardt and Kanamori [1978, p. 83]). The rule of thumb involved here, *generalization*, seems to be a presumption in favor of a natural strengthening of a well-supported axiom. Of course any large cardinal axiom also acquires intrinsic support from *maximize.* (Other rules of thumb favoring the Axiom of Supercompact Cardinals are discussed in VI.2 below.)

Until recently (see VI.4 below), the only significant consequences of the Axiom of Supercompact Cardinals were various relative consistency results. When a statement is too strong to be proved consistent relative to ZFC alone, its consistency can sometimes be derived from the assumption that ZFC plus some further axiom is consistent. (For example, recall Solovay's results from the consistency of "ZFC + The Axiom of Inaccessibles Cardinals" mentioned in [BAI, §III].) Several strong results of this sort follow from the consistency of "ZFC + The Axiom of Supercompact Cardinals" (see e.g. Foreman, Magidor and Shelah [MM]).

Notice that relative consistency results of this sort involving large cardinals are among the most useful applications of these axioms:

> ... large cardinals via the method of forcing turn out to be the natural measures of the consistency strength of ZFC + φ for various statements φ in the language of set theory.

(Kanamori and Magidor [1978, p. 105]) Large cardinals provide such a yardstick because they fit into an ordering:

> As our edifice grew, we saw how one by one the large cardinals fell into place in a *linear* hierarchy. This is especially remarkable in view of the ostensibly

> disparate ideas that motivate their formulation. As remarked by H. Friedman, this hierarchical aspect of the theory of large cardinals is somewhat of a mystery.

(Kanamori and Magidor [1978, p. 264]; see also Parsons [1983, p. 297] and Wang [1974, p. 555]). This unexpected pattern suggests that large cardinal axioms are straightforward ways of saying that the iterative hierarchy contains more and more levels, that is, that they are implementations of *maximize*:

> ... the neat hierarchical structure of the large cardinals and the extensive equi-consistency results that have already been demonstrated to date are strong plausibility arguments for the inevitability of the theory of large cardinals as *the* natural [extension of] ZFC.

(Kanamori and Magidor [1978, p. 264]) Thus the relative consistency results and the linear ordering of the large cardinal axioms provide extrinsic evidence for the large cardinal program in general.

VI.2. *Arguments for supercompact cardinals.* Two further rule of thumb based arguments have been offered in favor of Supercompact Cardinals. The first is a fairly simple set-theoretic argument based on the model theoretic Vopěnka's principle. The second, via extendibles,[14] is a more elaborate argument due to Reinhardt. As mentioned in §IV above, it depends on somewhat dubious pseudo-*reflection* principles. Vopěnka first.

The most general version of Vopěnka's principle states that any proper class of structures for the same language will contain two members, one of which can be elementarily embedded in the other. The rule of thumb usually cited as lying behind this principle is the idea that the proper class of ordinals is extremely rich (Kanamori and Magidor [1978, p. 196]). Suppose, for example, that a process is repeated once for each ordinal—Ord-many times, we might say—and every step produces a structure. Then *richness* implies that no matter how closely we keep track of the structures generated, there are so many ordinals that some will be indistinguishable.[15] A similar idea can be developed from *reflection*: Anything true of V is already true of some R_α, that is, there is an R_α that resembles V. This property of V should also be reflected, that is, there is an R_α with a smaller R_β that resembles it.

[14]For details of extendibles, see Solovay, Reinhardt and Kanamori [1978, §5], or Kanamori and Magidor [1978, §16]. On the relationship between supercompacts and extendibles, Kanamori and Magidor remark:

> All in all, supercompactness and extendibility have similar features... Supercompactness has the flavor of a generalization from measurability, but extendibility reflects more ethereal ambitions. (pp. 196, 192)

The nature of these "ethereal ambitions" will emerge from Reinhardt's argument, below. As I find this argument flawed, I will keep the emphasis here on supercompacts, rather than extendibles. It is also worth noting that supercompacts seem to occur more naturally in the hypotheses of theorems.

[15]Notice that the thinking behind *richness* is very close to that behind Martin's version of *reflection* in [BAI, §III].

Either way, we get a new rule of thumb, *resemblance*:

... there are R_α's that resemble each other.

... there should be stages R_α and R_β which look very much alike.

(Solovay, Reinhardt and Kanamori [1978, p. 75]; Martin [1976, p. 86]; see also Kanamori and Magidor [1978, p. 104]). The trick, of course, comes in spelling out "resembles".

To do this, let us go back to *richness* and imagine ourselves in an Ord-long process, generating an R_α at each stage, one for each ordinal.[16] Suppose we step several ranks at a time, so that by step α, we are already to R_{γ_α}, for some $\gamma_\alpha > \alpha$. We keep careful track of the structures at each stage by making copious notations on a clipboard, one scoresheet for every stage; we note down every detail of the structure we have just generated, along with every detail of the process that got us there. *Richness* then implies that with so many stages, our scoresheets cannot all be different. At step one, we record the complete diagram of

$$(R_{\gamma_0}, \in, \langle R_{\gamma_\beta} : \beta < 0 \rangle).$$

At step two, we look to see if that scoresheet is satisfied by

$$(R_{\gamma_1}, \in, \langle R_{\gamma_\beta} : \beta < 1 \rangle).$$

Of course it is not, so we write down the complete diagram of this new structure. And so on. At each step, we generate a new structure, then check to see if any of our old scoresheets will do; if not, we prepare a new one.

Richness then guarantees that we will eventually reach a step α' where one of our old scoresheets will match up. That is, we will reach a step α' where

$$(R_{\gamma_{\alpha'}}, \in, \langle R_{\gamma_\beta} : \beta < \alpha' \rangle)$$

is a model of the complete diagram of

$$(R_{\gamma_\alpha}, \in, \langle R_{\gamma_\beta} : \beta < \alpha \rangle)$$

for some $\alpha < \alpha'$. This means that the smaller structure can be elementarily embedded in the larger; that is:

$$\exists j : (R_{\gamma_\alpha}, \in, \langle R_{\gamma_\beta} : \beta < \alpha \rangle) \xrightarrow{\text{e.e.}} (R_{\gamma_{\alpha'}}, \in, \langle R_{\gamma_\beta} : \beta < \alpha' \rangle).$$

This embedding must be nontrivial, because:

$$\begin{aligned} \alpha' &= \text{length}(\langle R_{\gamma_\beta} : \beta < \alpha' \rangle) \\ &= \text{length}(j(\langle R_{\gamma_\beta} : \beta < \alpha \rangle) \\ &= j(\text{length}(\langle R_{\gamma_\beta} : \beta < \alpha \rangle) \\ &= j(\alpha). \end{aligned}$$

[16] This treatment of the argument from Vopěnka's principle was suggested by Magidor and clarified in discussion with Martin.

We thus have a nontrivial elementary embedding of R_α into $R_{\alpha'}$. Our conclusion is a special case of Vopenka's principle, namely, that in any proper class of R_α's, there is a nontrivial elementary embedding of one into another. We get a supercompact by applying this to the class of all R_α's for limit α that reflect supercompactness. Some large cardinal theorists see this blend of *richness*, *reflection* and *resemblance* as providing strong intrinsic evidence for the larger large cardinal axioms.

Reinhardt's argument also begins with the consideration of the proper class ORD of all ordinals. The universe V is then R_{ORD}. Reinhardt now asks that we look at V and ORD as it were "from the outside", in which case we see that there would be further ordinals and ranks of the form ORD + 1, ORD + ORD, $R_{\mathrm{ORD}+1}$, $R_{\mathrm{ORD}+\mathrm{ORD}}$, and so on. At this point, we seem to have introduced things other than sets, which threatens the universality of set theory, but Reinhardt proposes that we

> ... mitigate this sorrow by seeing the universality [of set theory] not in the extension of the concept set, but in the applicability of the theory of sets. [1974, p. 198]

In other words, we assume that our theory of sets is *the* universal theory of collections, and hence that it applies to these new objects. This gesture produces lots and lots of these class-like entities, lots of ordinal-like objects greater than ORD, lots of stages of construction after V; and they all obey the axioms of set theory.

This treatment is neat, so neat that we begin to wonder if these new layers really consist of entities of a new and different type; perhaps we just forgot to finish the iterative hierarchy in the first place. To this Reinhardt replies by drawing a distinction between sets and classes that depends on their behavior in counterfactual situations. For example, the set consisting of the current members of congress would be the same in any case, but the class of current members of congress would have been different if the voters had favored the Republicans instead of the Democrats. Reinhardt also imagines that there might be more ordinals in some counterfactual situation, and hence, that there might have been more stages and more sets than there are. Granted this assumption, a set is completely determined by its members—it has the same members in every possible world—but a class might have more members in another possible world—as, for example, the class ORD has more members in a counterfactual situation with more ordinals.

Now let us imagine one of these counterfactual situations, a projected universe with more ordinals. Reinhardt calls these extra ordinals, and the sets in the stages they number, imaginary ordinals and imaginary sets. Our ORD becomes, in the projected world, a new class $j(\mathrm{ORD})$ that consists of real and imaginary ordinals, while the old ORD is just an imaginary set, that is, $j(\mathrm{ORD}) > \mathrm{ORD}$. Sets, on the other hand, do not change their membership in counterfactual situations, so $j(x) = x$, for all $x \in V$.

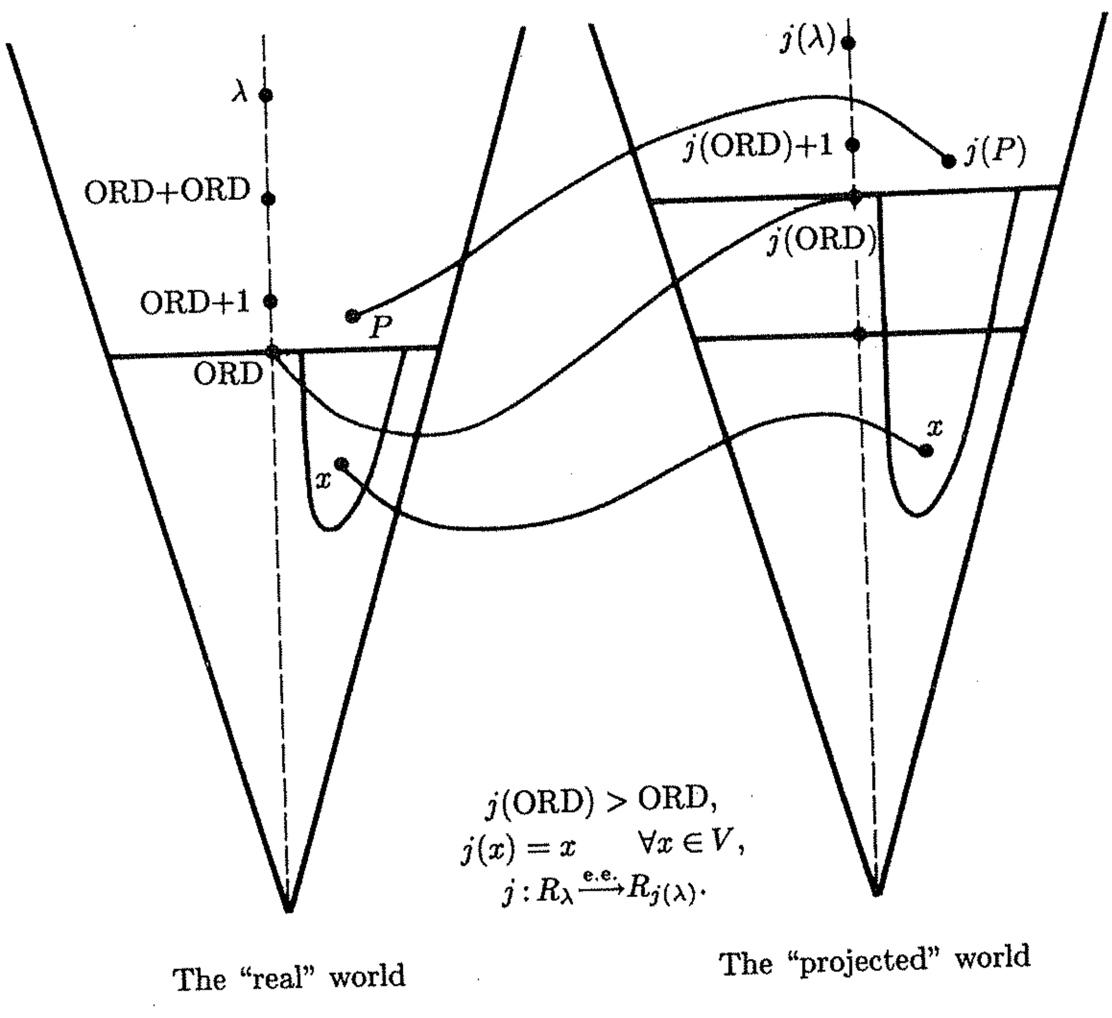
What about truth? Well, consider a proper class P. P consists of sets; it is a subset of R_{ORD}, a member of $R_{\mathrm{ORD}+1}$. Thus $j(P)$, in the projected world, consists of sets and imaginary sets; it is a subset of $R_{j(\mathrm{ORD})}$, and a member of $R_{j(\mathrm{ORD})+1}$. Now set theory is the universal theory of collections, so what is true of P in $R_{\mathrm{ORD}+1}$ should be true of $j(P)$ in $R_{j(\mathrm{ORD})+1}$, that is,

$$j\colon R_{\mathrm{ORD}+1} \xrightarrow{\text{e.e.}} R_{j(\mathrm{ORD})+1}.$$

And the same should be true of the extra layers of proper classes. That is, if λ is some ordinal-like object greater than ORD, then

$$j: R_\lambda \xrightarrow{\text{e.e.}} R_{j(\lambda)}.$$

Thus for any ordinal-like λ greater than ORD, we have argued that there is an elementary embedding of R_λ into $R_{j(\lambda)}$ with ORD as critical point, as shown in the figure:



This is just to say that ORD is extendible. All that is needed now is an application of *reflection*: if ORD is extendible, then there should be an extendible cardinal κ. From this we get our supercompact.

This argument has a number of severe shortcomings. The first arises even if we accept Reinhardt's premises; it is internal to the argument itself. Consider once again the purported universality of set theory. This is first applied to guarantee that the theory of the "real" world, with its extra stages, is identical to the theory of V. So far, so good. The second application comes when the "real" world is compared with the projected world. Here it would seem fair to conclude that the theory of the projected world is the same as that of the "real" world, that is, that they are elementarily equivalent. This is enough to assure us an embedding that preserves truth for the definable proper classes, but Reinhardt needs the full force of the elementary embedding j. It is hard to see how the universality of set theory will do this job.

Several other objections arise once we allow ourselves to question Reinhardt's premises. First, there are the alarming entities ORD + 1 and $R_{\text{ORD+ORD}}$. The Vopěnka argument involves thinking of proper classes of sets, but nothing so extravagant and potentially treacherous as these. Second, there is the use of counterfactual situations to distinguish these new entities from sets. I think even those with strong modal intuitions will have trouble imagining how there might be more pure sets and ordinals than there are. After all, V is supposed to contain all the sets and ordinals there could possibly be.

Finally, there is a pernicious ambiguity in Reinhardt's notion of a proper class. Everyone grants that collections can be thought of in two quite different ways: as extensions of concepts on the Fregean model, or as combinatorially generated in stages on the iterative model. These are sometimes called the "logical" and the "mathematical" notion of collection, respectively. When Reinhardt argues that classes differ from sets in their behavior in counterfactual situations, he is playing on the logical notion of extension; the extension of the concept "ordinal" is different in the projected world. On the other hand, when he argues that set theory should apply to all collections, classes included, he is thinking of classes on the mathematical or combinatorial model. On the logical notion of class, there is little reason to think that set theory should apply to entities so different from sets, and abundant reason to think that it should not. For example, it seems that logical classes, like the class of all infinite classes, can be self-membered.[17]

But even if Reinhardt's argument is flawed, we retain the Vopěnka argument based on *richness*, *reflection* and *resemblance*, as well as the earlier defenses in terms of *maximize* and *generalize*.

VI.3. *Huge cardinals and beyond.* If strengthening the closure condition on the range of the elementary embedding gives a natural *generalization* of measurability, *generalization* itself suggests the closure conditions might be strengthened even further. Indeed, why should the range of the elementary embedding not be completely closed, that is, why should it not be V itself?:

> In the first flush of experience with these ideas, Reinhardt speculated on the possibility of an ultimate extension: *Could there be an elementary embedding $j: V \to V$?*

(Kanamori and Magidor [1978, p. 200]) It was not long before Kunen destroyed this hope. In [1971], he showed that if j is an elementary embedding of V into a transitive M, if κ is the critical point of j, and if $j(\kappa) = \kappa_1$, $j(\kappa_1) = \kappa_2, \ldots$ and $\kappa_\omega = \lim_{n\to\infty} \kappa_n$, then there is a subset of κ_ω that is not in M. Thus M is not V.

[17] Some aspects of the set/class distinction are discussed in my [1983].

Kunen's theorem shows when a large large cardinal axiom is too large, so large that it contradicts ZFC (specifically, the Axiom of Choice):

> Kunen's result will limit our efforts in that we cannot embed the universe into too "fat" an inner model.

(Solovay, Reinhardt and Kanamori [1978, p. 83]) Modern set theorists have reacted much as Zermelo did to the inconsistencies of his day, that is, by applying the rule of thumb *one step back from disaster* (see [BAI, §1.4]). Thus they consider n-huge cardinals:

> ... they assert stronger and stronger closure properties, until their natural ω-ary extension turns out to be inconsistent.

(Kanamori and Magidor [1978, p. 202]) A cardinal κ is *n-huge* iff there is an elementary embedding j of V into a transitive M such that M is closed under arbitrary sequences of length κ_n. Kunen's theorem says that there is no such thing as an ω-huge cardinal.

Notice that 0-hugeness is just measurability. In addition:

> The n-huge cardinals certainly have an analogous flavor to λ-supercompact cardinals, but there is an important difference: While λ-supercompactness is hypothesized with an *a priori* λ in mind as a proposed degree of closure for M, n-hugeness has closure properties only *a posteriori*: M here is to be closed under κ_n-sequences, however large the κ_n turn out to be. This is a strengthening of an essential kind.

(Kanamori and Magidor [1978, p. 198]) Thus the comparison with supercompacts tends to tarnish the image of the n-huge cardinals:

> Indeed, it is not clear how to motivate n-hugeness ... at all.

(Kanamori and Magidor [1978, p. 198]) But if intrinsic support is lacking, at least n-huge cardinals do have a familiar sort of ultrafilter characterization, and they have played a role in some relative consistency results. Both these are cited as weak extrinsic evidence (Kanamori and Magidor [1978, pp. 198, 200]).

Another method of applying *one step back from disaster* is suggested by the form of Kunen's proof. The argument depends on the occurrence of a certain function in the domain of the elementary embedding. The domain of the function is the set of ω-sequences from κ_ω, so the function itself first occurs at the level $R_{\kappa_\omega+2}$. Thus Kunen actually shows that there is no nontrivial elementary $j: R_{\kappa_\omega+2} \to R_{\kappa_\omega+2}$. This leaves two possible forms of "there is a nontrivial elementary embedding of some R_λ into itself":

EE(I) $$\exists j: R_{\kappa_\omega+1} \xrightarrow{\text{e.e.}} R_{\kappa_\omega+1};$$

EE(II) $$\exists j: R_{\kappa_\omega} \xrightarrow{\text{e.e.}} R_{\kappa_\omega}.$$

It is known that EE(I) implies EE(II) and that EE(II) implies the existence of n-huge cardinals for every n. Indeed, the large cardinal property which EE(I) asserts of the critical point of its embedding is so strong that the existence of such a cardinal implies the existence of an even larger cardinal with the same property (Kanamori and Magidor [1978, p. 203]).

Even the defenders of large large cardinals express discomfort over axioms this strong:

> It seems likely that [EE(I) and EE(II) are] inconsistent since they appear to differ from the proposition proved inconsistent by Kunen, in an inessential technical way. The axioms asserting the existence of n-huge cardinals, for $n > 1$, seem (to our unpracticed eyes) essentially equivalent in plausibility: far more plausible than [EE(II)], but far less plausible than say extendibility.

(Solovay, Reinhardt and Kanamori [1978, §7]; see also Kanamori and Magidor [1978, p. 202]). Notice also that, given the reformulation of Kunen's result, EE(I), if consistent, would seem to be the largest possible large cardinal axiom. Some set theorists feel that for every large cardinal axiom there should be a larger, and this sentiment counts for them against EE(I).

VI.4. *Connections with Determinacy.* Recall that the existence of a measurable cardinal implies the determinacy of Σ^1_1 sets of reals (see V.2). This sort of result was extended one level further, to the determinacy of Σ^1_2 sets, by Martin in [1978], but the large cardinal axiom this proof requires is EE(I). This result caused some soul-searching among those who had hoped to increase the intrinsic support of determinacy hypotheses by deriving them from large cardinal axioms, but who also felt uncomfortable with EE(I). Furthermore, if the "last" large cardinal axiom was indeed necessary to prove Det(Σ^1_2), then the program of proving all of QPD from such axioms seemed hopeless. Still, the fact that EE(I) implied more determinacy, and the naturalness of the proof, led to something of a softening in the attitude towards this axiom.

There the situation remained until 1984, when consideration of the sets constructible from $R_{\kappa_\omega+1}$ led Woodin to an elementary embedding condition between EE(I) and Kunen's inconsistency:

EE(0) $$\exists j: L[R_{\kappa_\omega+1}] \xrightarrow{\text{e.e.}} L[R_{\kappa_\omega+1}].$$

Then came the result that everyone had been hoping for; Woodin went on to derive the full QPD from EE(0). With the discovery of EE(0), EE(I) no longer seemed the "last" large cardinal axiom, and EE(0) produced a natural and detailed theory of $L[R_{\kappa_\omega+1}]$ that resembled the theory of $L[R]$ on the assumption QPD. All this was counted as extrinsic evidence in their favor.

Recall that in the wake of Martin's earlier theorem deriving Det(Σ^1_1) from the existence of a measurable cardinal, various determinacy assumptions were proved equivalent to the inner model versions of the corresponding large cardinal axiom (see V.2). In addition to indiscernibles and a slightly smaller large cardinal, these inner models have well-orderings of the reals that are as simple as their level of determinacy allows. For example, the existence of the sharps (the inner model version of the Axiom of One Measurable Cardinal) guarantees the existence of an inner model of ZFC with indiscernibles; that model has Δ^1_1 determinacy, so it cannot have a Δ^1_1 well-ordering of the reals, but it does have a Δ^1_2 well-ordering. Similarly, the inner model of ZFC + 2MC (the inner model version of the Axiom of Three Measurable Cardinals) has $(\omega^2 + 1) - \Pi^1_1$ determinacy, so it cannot have a Δ^1_2 well-ordering of a certain special sort, but it does have a Δ^1_3 well-ordering. If this pattern were to continue, as most set theorists concerned with the problem expected that it would, then there should be inner models of all large cardinal axioms up to EE(I) with various degrees of Δ^1_2 determinacy and Δ^1_3 well-orderings of the reals. Alas, the inner model theorists, Mitchell, Dodd, Steel and others, were unable to find such a model; their efforts failed before they reached a supercompact cardinal.

The reasons for this failure were soon clarified from another quarter. Working on the development of further relative consistency results, Foreman, Magidor and Shelah were able to improve an older result of Kunen's by reducing the hypothesis from the consistency of a huge cardinal to the consistency of a supercompact cardinal (see [MM]). Shelah and Woodin then managed to reduce the hypothesis even more, to something between measurable and supercompact, and along the way, Woodin realized their methods led to another surprising result: if there is a supercompact cardinal, then every quasiprojective set of reals is Lebesgue measurable, has the Baire and perfect subset properties, and so on. Thus, the model the inner model theorists were seeking—an inner model with a supercompact cardinal and a Δ^1_3 well-ordering of the reals—does not exist. Indeed there is no inner model with a supercompact cardinal and *any* quasiprojective well-ordering of the reals. The neat inner model theory that did so much to familiarize measurable cardinals cannot be duplicated for supercompacts.[18]

But what about determinacy? There were two possibilities. Up to this point, the old-fashioned regularity properties like Lebesgue measurability had gone hand-in-hand with determinacy. Now the determinacy of quasiprojective sets seemed to require the somewhat staggering assumption of EE(0), while their other regularity properties required only a supercompact cardinal. The first possibility was that determinacy and Lebesgue measurability do in fact diverge here, and the inner model equivalences possible within Δ^1_2 cannot be extended. The second possibility was that QPD could actually be proved from the far weaker assumption of a supercompact cardinal.

Two who believed in the second possibility were Martin and Steel. Woodin had shown that his theorem on the Lebesgue measurability of quasiprojective sets could actually be derived from a complex hypothesis slightly weaker than the existence of a full supercompact cardinal, so Martin and Steel felt they had an exact formulation of the hypothesis that should yield QPD. Further, Steel had extensive experience with the sort of phenomena that had blocked the development of the inner model theory before it was known to be impossible. He and Martin theorized that whatever blocked the construction of a nice inner model might be closely connected with determinacy. Reasoning in this way, they were able to prove PD, and (using another result of Woodin) QPD, from Woodin's hypothesis, and hence, from the existence of a supercompact cardinal.

---

[18] As the existence of simple inner models made some set theorists more comfortable with measurable cardinals, the nonexistence of such inner models makes supercompacts appear more mysterious, perhaps even dangerous.

This sudden and unexpected reduction in the ante required for QPD naturally contributes strongly to the attractiveness of the theory. All the determinacy needed for descriptive set theory can be viewed as a theorem of "ZFC + The Axiom of Supercompact Cardinals". Indeed, the theory of $L[R]$ under these axioms is in some sense "complete": it is invariant under most forcing extensions (see [MM]). Thus supercompact cardinals gain a tremendous amount of extrinsic evidence, and QPD inherits various intrinsic and rule of thumb support (*maximize*, *generalize*, *richness* and *reflection*) from the Axiom of Supercompact Cardinals. And both are extrinsically supported by the impressive strength of their intertheoretic connections. Thus it is not surprising that some Cabal members view the Martin/Steel theorem as proving the detailed descriptive set theory described in Moschovakis's book [1980].

Of course, now that QPD is seen to follow from the existence of a supercompact cardinal, the Levy/Solovay theorem of [1967] immediately implies that QPD cannot decide the size of the continuum. The next step would be to investigate hypotheses on the structure of $L[\mathscr{P}(R)]$.

§VII. Concluding philosophical remarks. As this is not a history paper, and even more obviously not a logic paper, I feel I owe at least a few philosophical reflections. Of course the motivating force behind the presentation of all this material has been a philosophical one: I hope to display the role of nondemonstrative arguments in set theory, especially in the search for new axioms, and to pose the philosophical task, for epistemologists and philosophers of mathematics, of describing and accounting for this role. In this final section, I will summarize and lightly categorize the data, then address a few random remarks to the serious philosophical questions raised.

The defenses given here for set-theoretic axiom candidates have been roughly divided into three categories: intrinsic, extrinsic and rules of thumb. So far, I have not tried to classify particular rules of thumb as intrinsic, extrinsic or other, but it should be clear that there is considerably variation within that group. Let me begin with a rather stylized discussion of intrinsic justification.

I have argued elsewhere [1980] that we acquire our most primitive physical and set-theoretic beliefs when we learn to perceive individual objects and sets of these. We come to believe, for example, that objects do not disappear when we are not looking at them, and that the number of objects in a set does not change when we move the objects around. These intuitive beliefs are not incorrigible—consider, for example, our erstwhile convictions that objects are solid, or that every property determines a set—but they do provide a starting point for our physical and mathematical sciences. The simplest axioms of set theory, like Pairing, have their source in this sort of intuition. If they are not strictly part of the concept (whatever that comes to), they are acquired along with the concept. Given its origin in prelinguistic experience, the best indication of intuitiveness is when a claim strikes us as obvious, or in Gödel's words, when the axioms "force themselves upon us as being true" [1947/64, p. 484].

The extrinsic evidence cited in previous sections came in a bewildering variety of forms, among them: (1) confirmation by instances (the implication of known lower-level results, as, for example, *reflection* implies weaker reflection principles known to be provable in ZFC); (2) prediction (the implication of previously unknown lower level results, as, for example, the Axiom of Measurable Cardinals implies the determinacy of Borel sets which is later proved from ZFC alone); (3) providing new proofs of old theorems (as, for example, game-theoretic methods give new proofs of Solovay's older set-theoretic results); (4) unifying new results with old, so that the old results become special cases of the new (as, for example, the proof of PWO(Π^1_1) becomes a special case of the periodicity theorem); (5) extending patterns begun in weaker theories (as, for example, the Axiom of Measurable Cardinals allows Souslin's theorem on the perfect subset property to be extended from Σ^1_1 to Σ^1_2); (6) providing powerful new ways of solving old problems (as, for example, QPD settles questions left open by Lusin and Souslin); (7) providing proofs of statements previously conjectured (as, for example, QPD implies there are no definable well-orderings of the reals); (8) filling a gap in a previously conjectured "false, but natural proof" (as, for example, Det(Δ^1_2) filled the gap in Moschovakis's erroneous "sup" proof of PWO(Π^1_3)); (9) explanatory power (as, for example, Silver's account of the indiscernibles in L provides an explanation of how and why $V \neq L$); (10) intertheoretic connections (as, for example, the connections between determinacy hypotheses and large cardinal assumptions count as evidence for each).

All of these more or less correspond to forms of confirmation recognized in the physical sciences. I would like very much to give an account of their rationality, but even our best philosophers of science, from Hempel [1945] to Glymour [1980], have so far been satisfied with predominantly descriptive accounts. A careful analysis of the structure of such arguments must precede what we hope will be an explanation of why they lead us toward truth (cf. Glymour [1980, p. [redacted] 377]).

Finally, rules of thumb. When uncritical, intuitive work with sets was interrupted by the appearance of the paradoxes, examination of previously unexamined practice revealed that full Comprehension was not in fact used. Rather, sets were thought of as being formed from objects already available. This lead to the separation of sets from classes, and eventually, to the development of the rule *iterative conception*. The source of this rule of thumb in pretheoretical practice, and the overwhelming impression of its naturalness once it was specified, suggest that its origin is at least partly intuitive (see, e.g. Shoenfield [1967, p. 238]). *Realism*, *maximize*, and its companion, *richness*, are all closely tied to *iterative conception*. Finally, *reflection* is often claimed to be intuitive, perhaps with grounds in *maximize* as well. *Inexhaustibility* is just a special case of *reflection*, and *resemblance* is a consequence.

In contrast, the evidence for the boldest of our rules of thumb—*Cantorian finitism*—is predominantly extrinsic, lying in the depth, breadth and effectiveness of the subject it launched. Other rules have the flavor of general methodological maxims, principles that express our higher-order preferences for theories of one sort or another. An example from physical science is Maxwell's principle, which states that a law of nature should be valid at all points in space and time (see Wilson [1979] for discussion). *Diversity* and *generalization* are rules of thumb at a similar level of abstraction. *One step back from disaster*, and its special case *limitation of size*, might also be viewed as methodological, though they share something of the spirit of *maximize*. *Banishment*, on the other hand, seems neither intrinsic, nor extrinsic, nor methodological, but rather based in seat-of-the-pants experience with the theory in question, like most conjectures.

Finally, *uniformity* and its capriciousness companion *whimsical identity* have been defended both as methodological principles akin to Maxwell's—a good theory does not single out particular locations—and as intuitions about the nature of the iterative hierarchy connected with *richness* and *resemblance*. Either way, we have seen the dangers inherent in applications of these two related rules. Perhaps what is needed is a theory of exactly what sorts of properties are allowable in *uniformity* and *whimsical identity* arguments, much as only so-called "structural" properties are allowed in *reflection* arguments.[19] Another possibility would be to grant evidential status to *uniformity* and *whimsical identity* arguments only in the presence of good evidence for consistency, or perhaps to relegate them to the status of heuristic devices for generating hypotheses that must then be justified by other, probably extrinsic, means.[20]

If, as we have seen, the practice of mathematics can be understood as analogous to that of the physical sciences in a great many respects, it must also be admitted that there is a striking difference: mathematicians rarely rely on observations in their nondemonstrative testing. This can be understood if we revert to our perceptual story. When we learn to see sets of things, we learn to see number properties, and from this we develop the humblest of our mathematical sciences: arithmetic. If our rudimentary physical science is the study of things qua stuff, arithmetic is the study of things qua individuals, the study of sets of things, and as such it is independent of the make-up of a set's elements qua stuff. As far as arithmetic is concerned, the particular things in its sets are irrelevant, as is their stuff; a set of a given cardinality is interchangeable, for arithmetical purposes, with a wide range of others, sets with different particular elements but the same cardinality, even sets of symbols.[21] Thus, once our perceptual relation to the physical world has produced our ability to see sets and our basic intuitions about them, the further observation of particulars is unimportant.

This is not to say that the physical world remains entirely irrelevant after this initial stage, but before I mention its further incursions, I should say a word about the ramifications of a naturalistic, empirical, perception-based account of mathematical knowledge.[22] Such views face an unavoidable challenge from the venerable philosophical observation that while our various perceptual, neurological and

---

[19] We might try something along the following lines: in both forms of argument, the crucial property, the one that is to recur or the one that appears in the whimsical identity, must be "natural". Obviously a natural property cannot involve notions like "first" or "smallest", and it cannot involve proper names. After this, it is hard to know what to say, except that the failures of *uniformity* and *whimsical identity* could be explained away as involving instances of unnatural properties. That is, for example, "2 is *the* even prime" shows that "prime" would be more naturally defined as "odd number not divisible by anything but itself and 1", and "$\aleph_0$ is *the* cardinal κ such that $\forall n_1 m < \kappa\ \kappa \to \kappa^n_m$" shows that arrow properties should not be formulated to allow infinite exponents.

[20] Representatives of each these various opinions on *uniformity* and *whimsical identity* arguments can be found within the Cabal.

[21] Even sets of appearances. This is why the threat of sensory illusion is less pressing for mathematics than for physical science: even if there were an evil demon systematically deceiving us as to the structure of the external world, arithmetic would still apply to the world of appearances.

[22] The idea that perception is involved in the genesis of mathematical knowledge is fairly popular these days. See, for example, Resnik [1982], Kitcher [1983], Parsons [19[illegible]80].

evolutionary interactions with the world might well tell us what is true, they cannot tell us what must be true. This, coupled with the equally venerable assumption that mathematical truths are necessary, creates a mystery. We begin to ask ourselves odd questions: if our world (or the evil demon's illusion) were different, would we have a different arithmetic? Of course, it is much easier to imagine a world with a different physical make-up than ours, or even different physical laws, than to imagine one to which our arithmetic does not apply. But then again, if objects systematically appeared and disappeared during counting, perhaps we would calculate differently; at least it seems likely that the ancient Babylonians (or whoever) would have lost interest in the subject. Still, it seems that once a world has two objects, it has a potential infinity of which arithmetic is true: the apple, the orange, the set of these two, the set of the preceding three, etc.[23] Perhaps only a world with absolutely no differentiation, a world completely homogeneous, the eternal oneness of the mystics, would be without number properties. But even if we leave aside the irritating inconclusiveness of musings along these lines, I think we must question their moral, their importance, their significance. We lack so much as a clear understanding of what it means to say that something is necessary: true in all possible worlds? true due to some irreducibly modal property of this world? At this point, it seems to me that the most reasonable answer to the old question—how do we know that mathematical truths are necessary?—must be that we do not know.[24]

It is worth noting that the same goes for certainty. This obvious point should not need belaboring, except when a mathematical epistemologist attempts to find arguments strong enough to "convince the skeptic". Philosophers gave up the search for such arguments in natural science long ago; its retention in the philosophy of mathematics can only be traced to an outmoded vision of the nature of mathematical knowledge. No one would expect even the best scientific arguments to be absolutely justifying. Our epistemological inquiries in mathematics will be hampered if we set an unreasonably high standard.

What, then, is the post-perceptual evidential connection of set theory with the physical world? I would suggest that it is the profound applicability of set theory's twin pillars: number theory and geometry/analysis. While number theory has its origin in counting, geometry arises from the study of the shapes of things (things as individuated objects, that is, not as amorphous arrays of physical stuff) and analysis from the study of their motions. Set theory systematizes and explains these two extravagantly useful branches of mathematics, and in so doing, gains much of its own justification (recall the extrinsic argument for the Power Set Axiom in [BAI, §I.6]). Notice that the continuum problem, whose independence prompted the search for new axioms, and whose solution would provide the most impressive extrinsic evidence, is itself a question about the real numbers of physical science. This is a central reason why many set theorists are confident of its meaningfulness, and thus of the propriety of the search for new axioms herein described.

---

[23] Perhaps this is the truth behind Brouwer's obscure "two-oneness". See his [1912]. In any case, one object would do, as long as it was differentiated from its background: the it and the not-it.

[24] I do not mean by this that we know mathematics to be contingent, either, but that we have no dependable information whatever on the question (assuming it is well-formed).

The success of set theory—its objectivity and its applicability—confirm the enterprise and its justificatory practices as a whole, but within that whole, the particular methods can be analyzed, supported or criticized individually. Not only would a clear account of the structure and rationality of nondemonstrative set theoretic arguments provide solace for the practitioners and philosophers of the subject, but it might even help with the very real problem of locating new rules of thumb and new axiom candidates for the solution of the continuum problem. I should emphasize that this is not a project of importance only to those with a Platonistic bent. It is central to any philosophical position for which the size of the continuum is a real issue: all realistic philosophies of set theory, even those that eschew mathematical objects (like Kitcher's [1983], Resnik's [1981], [1982], or Shapiro's [1983]); modalist accounts that depend on full second-order models (like Putnam's [1967] and Hellman's [1986]); and even some versions of Field's nominalism (the second-order option of [1985] where only one of "ZFC + $(V = L)$" and "ZFC + QPD" can be conservative (see his footnote 16)). This strongly suggests that in this area at least, we would do well to drop the ingrained philosophical tendency to concentrate of the differences (however minute) between positions, and to engage in a cooperative effort.

REFERENCES

J. W. ADDISON

[1959] *Some consequences of the axiom of constructibility*, ***Fundamenta Mathematicae***, vol. 46 (1959), pp. 337–357.

[1974] *Current problems in descriptive set theory*, in Jech [1974], pp. 1–10.

J. W. ADDISON and Y. N. MOSCHOVAKIS

[1968] *Some consequences of the axiom of definable determinateness*, ***Proceedings of the National Academy of Sciences of the United States of America***, vol. 59 (1968), pp. 708–712.

Y. BAR-HILLEL, editor

[1970] ***Mathematical logic and foundations of set theory***, North Holland, Amsterdam, 1970.

J. BARWISE, editor

[1977] ***Handbook of mathematical logic***, North-Holland, Amsterdam, 1977.

P. BENACERRAF and H. PUTNAM, editors

[1983] ***Philosophy of mathematics***, 2nd ed., Cambridge University Press, Cambridge. 1983.

D. BLACKWELL

[1967] *Infinite games and analytic sets*, ***Proceedings of the National Academy of Sciences of the United States of America***, vol. 58 (1967), pp. 1836–1837.

L. E. J. BROUWER

[1912] *Intuitionism and formalism*, in Benacerraf and Putnam [1983], pp. 77–89

M. DAVIS

[1964] *Infinite games of perfect information*, ***Advances in game theory*** (M. Dresher et al., editors), Annals of Mathematics Studies, vol. 52, Princeton University Press, Princeton, New Jersey, 1964, pp. 85–101.

J. E. FENSTAD

[1971] *The axiom of determinateness*, ***Proceedings of the second Scandinavian logic symposium*** (J. F. Fenstad, editor), North-Holland, Amsterdam, 1971, pp. 41–61.

H. FIELD

[1985] *On conservativeness and incompleteness*, ***Journal of Philosophy***, vol. 82 (1985), pp. 239–260.

M. FOREMAN, M. MAGIDOR, and S. SHELAH

[MM] *Martin's Maximum, saturated ideals, and nonregular ultrafilters* (to appear).

D. GALE AND F. M. STEWART

[1953] *Infinite games with perfect information*, **Contributions to the theory of games**, vol. 2, Annals of Mathematics Studies, vol. 28, Princeton University Press, Princeton, New Jersey, 1953, pp. 245–266.

C. GLYMOUR

[1980] **Theory and evidence**, Princeton University Press, Princeton, New Jersey, 1980.

K. GÖDEL

[1947/64] *What is Cantor's continuum problem?*, in Benacerraf and Putnam [1983], pp. 470–485.

L. HARRINGTON

[1978] *Analytic determinacy and* $0^{\#}$, this JOURNAL, vol. 43 (1978), pp. 685–693.

G. HELLMAN

[1986] *A modal interpretation of mathematics*, circulated photocopy.

C. HEMPEL

[1945] *Studies in the logic of confirmation*, **Mind**, vol. 54 (1945), pp. 1–26, 97–121; reprinted in C. W. Hempel, **Aspects of scientific explanation and other essays in the philosophy of science**, MacMillan, 1965, pp. 3–46.

T. J. JECH, editor

[1974] **Axiomatic set theory**, Proceedings of Symposia in Pure Mathematics, vol. 13, part II, American Mathematical Society, Providence, Rhode Island, 1974.

K. KANAMORI and M. MAGIDOR

[1978] *The evolution of large cardinal axioms in set theory*, Higher set theory (G. H. Müller and D. S. Scott, editors), Lecture Notes in Mathematics, vol. 669, Springer-Verlag, Berlin, 1978, pp. 99–275.

P. KITCHER

[1983] **The nature of mathematical knowledge**, Oxford University Press, Oxford, 1983.

K. KUNEN

[1971] *Elementary embeddings and infinitary combinatorics*, this JOURNAL, vol. 36 (1971), pp. 407–413.

A. LEVY and R. M. SOLOVAY

[1967] *Measurable cardinals and the continuum hypothesis*, **Israel Journal of Mathematics**, vol. 5 (1967), pp. 234–248.

P. MADDY

[1980] *Perception and mathematical intuition*, **Philosophical Review** vol. 89 (1980), pp. 163–196.

[1983] *Proper classes*, this JOURNAL, vol. 48 (1983), pp. 113–139.

[1984] *New directions in the philosophy of mathematics*, **PSA 1984** (P. Kitcher, editor), Philosophy of Science Association, 1985, pp. 425–447.

[BAI] *Believing the axioms.* I, this JOURNAL, vol. 53 (1988), pp. 481–511.

D. A. MARTIN

[PSCN] *Projective sets and cardinal numbers*, circulated photocopy.

[1968] *The axiom of determinateness and reduction principles in the analytic hierarchy*, **Bulletin of the American Mathematical Society**, vol. 74 (1968), pp. 687–689.

[1970] *Measurable cardinals and analytic games*, **Fundamenta Mathematicae**, vol. 66 (1970), pp. 287–291.

[1975] *Borel determinacy*, **Annals of Mathematics**, ser. 2, vol. 102 (1975), pp. 363–371.

[1976] *Hilbert's first problem: the continuum hypothesis*, **Mathematical developments arising from Hilbert problems** (F. E. Browder, editor), Proceedings of Symposia in Pure Mathematics, vol. 28, American Mathematical Society, Providence, Rhode Island, 1976, pp. 81–92.

[1977] *Descriptive set theory*, in Barwise [1977], pp. 783–815.

[1978] *Infinite games*, **Proceedings of the International Congress of Mathematicians (Helsinki, 1978)**, vol. 1, Academia Scientiarum Fennica, Helsinki, 1980, pp. 269–273.

Y. N. MOSCHOVAKIS

[1970] *Determinacy and prewellorderings of the continuum*, in Bar-Hillel [1970], pp. 24–62.

[1980] **Descriptive set theory**, North-Holland, Amsterdam, 1980.

J. MYCIELSKI

[1964] *On the axiom of determinateness*, **Fundamenta Mathematicae**, vol. 53 (1964), pp. 205–224.

[1966] *On the axiom of determinateness.* II, ***Fundamenta Mathematicae***, vol. 59 (1966), pp. 203–212.

J. MYCIELSKI and H. STEINHAUS

[1962] *A mathematical axiom contradicting the axiom of choice*, ***Bulletin de l'Académie Polonaise des Sciences, Série des Sciences Mathématiques, Astronomiques et Physiques***, vol. 10 (1962), pp. 1–3.

J. MYCIELSKI and S. SWIERCZKOWSKI

[1964] *On the Lebesgue measurability and the axiom of determinateness*, ***Fundamenta Mathematicae***, vol. 54 (1964), pp. 67–71.

J. B. PARIS

[1972] *$ZF \vdash \Sigma^0_4$ determinateness*, this JOURNAL, vol. 37 (1972), pp. 661–667.

C. PARSONS

[19[redacted]80] *Mathematical intuition*, ***Proceedings of the Aristotelian Society***, vol. 80 ([redacted] 1980), pp. 142–168.

[1983] ***Mathematics in philosophy***, Cornell University Press, Ithaca, New York, 1983.

H. PUTNAM

[1967] *Mathematics without foundation*, in Benacerraf and Putnam [1983], pp. 295–311.

W. N. REINHARDT

[1974] *Remarks on reflection principles, large cardinals and elementary embeddings*, in Jech [1974], pp. 189–205.

Set existence principles of Shoenfield, Ackermann, and Powell, ***Fundamenta Mathematicae***, vol. 84 (1974), pp. 5–34.

M. RESNIK

[1981] *Mathematics as a science of patterns: ontology and reference*, ***Noûs***, vol. 15 (1981), pp. 529–550.

[1982] *Mathematics as a science of patterns: epistemology*, ***Noûs***, vol. 16 (1982), pp. 95–105.

S. SHAPIRO

[1983] *Mathematics and reality*, ***Philosophy of Science***, vol. 50 (1983), pp. 523–548.

J. R. SHOENFIELD

[1967] ***Mathematical logic***, Addison-Wesley, Reading, Massachusetts, 1967.

W. SIERPINSKI

[1925] *Sur une classe d'ensembles*, ***Fundamenta Mathematicae***, vol. 7 (1925), pp. 237–243.

R. M. SOLOVAY

[1969] On *the cardinality of Σ^1_2 sets* of reals, ***Foundations of Mathematics*** (J. J. Bulloff et al., editors), Symposium papers commemorating the sixtieth birthday of Kurt Gödel, Springer-Verlag, Berlin, 1969, pp. 58–73.

R. M. SOLOVAY, W. N. REINHARDT, and A. KANAMORI

[1978] *Strong axioms of infinity and elementary embeddings*, ***Annals of Mathematical Logic***, vol. 13 (1978), pp. 73–116.

H. WANG

[1974] *The concept of set*, in Benacerraf and Putnam [1983], pp. 530–570.

M. WILSON

[1979] *Maxwell's condition—Goodman's problem*, ***British Journal for the Philosophy of Science***, vol. 30 (1979), pp. 107–123.

R. WOLF

[1985] *Determinateness of certain almost Borel games*, this JOURNAL, vol. 50 (1985), pp. 569–579.

P. WOLFE

[1955] [redacted] *The strict determinateness of certain infinite games*, ***Pacific Journal of Mathematics***, vol. 5 (1955), pp. 841–847.

DEPARTMENT OF PHILOSOPHY
UNIVERSITY OF CALIFORNIA AT IRVINE
IRVINE, CALIFORNIA 92717