"Gödel's Modernism: on Set-Theoretic Incompleteness," revisited*

Mark van Atten and Juliette Kennedy

> As to problems with the answer Yes or No, the conviction that they are always decidable remains untouched by these results.
>
> —Gödel

Contents

1 Introduction

1.1 Questions of incompleteness

On Friday, November 15, 1940, Kurt Gödel gave a talk on set theory at Brown University.[1] The topic was his recent proof of the consistency of Cantor's Continuum Hypothesis, henceforth *CH*,[2] with the axiomatic system for set theory ZFC.[3] His friend from their days in Vienna, Rudolf Carnap, was in the audience, and afterward wrote a note to himself in which he raised a number of questions on incompleteness:[4]

> (Remarks I planned to make, but *did not*)
>
> Discussion on Gödel's lecture on the Continuum Hypothesis, November 14,[5] 1940
>
> There seems to be a difference: between the *undecidable* propositions of the kind of his example [i.e., 1931] and propositions such as *the Axiom of Choice*, and *the Axiom of the Continuum* [*CH*].
>
> We used to ask: "When these two have been decided, is then everything decided?" (The Poles, Tarski I think, suspected that this would be the case.) Now we know that (on the basis of *the usual finitary rules*) there will always remain undecided propositions.

*An earlier version of this paper appeared as 'Gödel's modernism: on set-theoretic incompleteness', *Graduate Faculty Philosophy Journal*, 25(2), 2004, pp.289–349. Erratum facing page of contents in 26(1), 2005.

1. Can we nevertheless still ask an analogous question? I.e. is there an objective difference between 2 kinds of problems, or is it just a difference in degree of simplicity?
2. If so, are there grounds for a positive answer? I.e., "Now that we have accepted both axioms, all simple problems are determined?"

We recapitulate the basic facts. In 1931, Gödel proved his well-known theorem: for every ω-consistent formal system that contains arithmetic and is recursively axiomatizable, as we would say now, there exist sentences ϕ (in the language of the system) such that neither ϕ nor $\neg\phi$ is derivable in the system. Such a sentence is said to be undecidable in the system and renders it incomplete. The three conditions on a formal system mentioned in the theorem mean the following. 1. ω-consistency means that the system should not prove (for some P definable in it) $\exists x \neg P(x)$ while also proving $P(n)$ for each natural number term n. 2. Containing a sufficient amount of arithmetic means that the operations of addition, multiplication, successor, as well as the notion of an order, should be definable in the system, and that the principle of induction should be included. 3. Recursive axiomatizability means that the axioms should be either finite in number or enumerable by an effective procedure. (In 1936, J.B. Rosser showed that the requirement of ω-consistency can be weakened to consistency.)

The class of formal systems to which the incompleteness theorem applies includes all of the more ambitious formal systems that had been formulated up till 1931: Principia Mathematica, the systems devised by Hilbert and his followers, and, in particular, the system of set theory that is still the canonical system today, ZFC.

Although the theorem shows that, for each system of the type described, there are undecidable sentences, it does not show that there is a sentence that cannot be decided in any possible system of that type. However, the theorem does not exclude the existence of such a sentence either. If it exists, it could be called absolutely undecidable (we will introduce a slightly more refined terminology below).[6]

In this paper, we will be concerned with incompleteness and undecidability in ZFC and related systems for set theory. The question that will be in the foreground most, and in the background all of the time, is: do absolutely undecidable propositions exist in set theory? We will analyse specifically how Gödel's thinking about this question developed in his published and unpublished work, closing with considerations on the present situation in set theory in the light of Gödel's ideas.

1.2 Splitting of the notion of undecidability

Gödel held that the axioms of classical Zermelo-Fraenkel set theory (or some system equivalent to it) are true and evident. It must have been they that he had in mind when he said, in 1966, that the axiomatization of set theory was the greatest advance in its foundations prior to forcing.[7] But they cannot be more than an initial segment of the correct axioms for all of mathematics, as by the incompleteness theorems, there are sentences ϕ (in the language of ZFC) that are undecidable in ZFC. With Carnap (see above), one can ask whether the collection of undecidable sentences is exhausted by those constructed in the proof of the incompleteness theorem, and this question is central to the present paper.

Clearly, any ϕ undecidable in ZFC falls into at least one of the following nominally defined categories, which split the notion of undecidability:

1 sentences that are undecidable in ZFC but seen to be true (and hence decided informally) by reflecting on the proof of their undecidability in ZFC.

2 sentences that are undecidable in ZFC, and are not decided informally by reflecting on the proof.

3 sentences that are undecidable in ZFC, but are decidable in an evident extension (or series of extensions) of ZFC.

4 sentences that are undecidable in ZFC, are not decidable in any evident extension of ZFC, but can be decided by human reason.

5 sentences that are undecidable in ZFC, are not decidable in any evident extension of ZFC, and cannot be decided by human reason.

These categories are not all mutually exclusive, for example ϕ may be of both the first and third category, or of the second and the third (if one finds a new axiom by other means than reflecting on undecidability proofs), or of the second and fifth. The questions at hand are the following: of which of these five categories, if any, can we establish that they are not empty? And if a category is not empty, do its members admit of a systematic characterization? It is crucial here that "extension" is taken in a non-trivializing sense: one adds only axioms that are seen to be true or evident. Simply adding ϕ without considering its evidence would miss the point. Note that an extension of a formal system may also consist in, or also involve, adding higher types to the logic or otherwise changing the logic in some appropriate way.[8]

Category 4 seems to be necessarily empty. For, on any reasonable informal understanding of proof, a proof of a sentence (or of its negation) proceeds from evident axioms, by evident inferences, to its conclusion; it is, as Gödel put it, 'not ... a sequence of expressions satisfying certain formal conditions, but a sequence of thoughts convincing a sound mind'.[9] So conversely, for any mathematical sentence that human reason decides, it should be able to indicate the evident axioms and evident inferences on the basis of which the decision was made. But if this can be done (i.e., if oracles are not admitted), these can be formalized and used to extend ZFC. (Note that 'arguments from success' only lead to probable decisions, and such arguments are therefore not excluded by the emptiness of category 4.)

To demonstrate the non-emptiness of category 1, we can simply use an undecidable statement constructed along the lines of Gödel's proof of the incompleteness theorem. By theorems of Gödel (1938) and Cohen (1963) that will play an important role in this paper, one can take $\phi = CH$ to give an example of a statement in category 2 (and hence not in category 1). But it is at present not known whether *CH* exemplifies the non-emptiness of category 3 or 5 (excluding 4 for the reason given above). It must be in one of them, and, as 3 and 5 are disjoint, exactly one. At the end of this paper we will consider the suggestion associated with the Woodin school to the effect that *CH* is (close to being) solved now. What can be said about category 5 will depend on how strong and specific one's views are on the nature of reason as well as on the ontology of mathematics.[10]

A distinction that cuts across this classification of statements undecidable in ZFC into five categories is that between statements that do play a role in mathematical practice and those that do not. This may of course change through time and therefore unlike the five-fold classification this one is not fixed. To consider these two distinctions in tandem is motivated by the fact that the undecidable sentences constructed in the proofs of Gödel's incompleteness theorems are manifestly different from anything found in mathematical practice so far, and in that specific sense not mathematically meaningful; we will take this specific sense as our definition of mathematical meaningfulness. The greatest interest is in the question whether a statement can be found that is both mathematically meaningful and absolutely undecidable, for that would make urgent the search for a new evident axiom from a practical perspective.

To a realist, the mathematical meaningfulness of a statement simply means that it has mathematical content (or is equivalent to one that does), in the sense that the terms in the statement refer.[11] Such statements can be called "contentual" ("inhaltlich"). To a (Hilbertian) formalist a certain statement may well be relevant to mathematics without being contentual (think of any practically relevant part of classical mathematics that is not finitary). On the other hand, this kind of background commitment associated with the realist and the formalist is often lacking in the the colloquial use of the phrase "mathematically meaningful" among mathematicians. That use rather emphasizes typical aspects (often of an aesthetic nature) such as being "natural," "fundamental," "elementary," or "interesting."

The question as to the cardinality of the continuum, a decision of *CH*, exemplifies many aspects of mathematical meaningfulness. As Gödel describes it in the 1947 paper, it is one of "the most fundamental questions in the field"; a question "from the "multiplication table" of cardinal numbers."[12]

> So the analysis of the phrase "how many" unambiguously leads to a definite meaning for the question stated in the second line of this paper: The problem is to find out which one of the $\aleph$'s is the number of points of a straight line or (which is the same) of any other continuum (of any number of dimensions) in a Euclidean space.[13]

1.3 Gödel's view on undecidability in 1931

How were incompleteness phenomena understood by Gödel in 1931? Did he expect all undecidable statements to be in category 1, (coded) metamathematical statements (e.g., involving provability, its particular case consistency, or computability) or in any case equivalents of those? Or did he think there are also mathematically meaningful statements which would be in category 2?

In the paper in which Gödel published his incompleteness theorem, he does not go into questions of this type, but in a lecture text which probably is from shortly after (**1931?*), Gödel mentions a concept of "absolute undecidablity" in relation to his theorem:

> The procedure just sketched furnishes, for every system that satisfies the aforementioned assumptions, an arithmetical sentence that is undecidable in that system. That sentence is, however, not at all absolutely undecidable; rather, one can always pass to "higher" systems in which the sentence in question is decidable. (Some other sentences, of course, nevertheless remain undecidable.)[14]

This quotation motivates us to make the following terminological point. As we will see below, at different times Gödel used the term "absolutely undecidable" in different ways. Around 1940 he used it in connection with category 2, but from 1951 onward it refers strictly to category 5 (which is a sub-category of 2). The latter may be the more natural thing to do in any case, for if all we know is that a sentence is of category 2, it is not excluded that we will come to find and believe an axiom that shows the sentence is also of category 3, and the sentence will have been decided after all. No such hope can be entertained if it is somehow shown that a sentence is of category 5, and that circumstance would earn it the predicate "absolutely undecidable" with more justification. We will use "weakly absolutely undecidable" for category 2 and "strongly absolutely undecidable" for category 5.

In the quotation just given, the correct reading of "absolutely undecidable" seems to be "weakly absolutely undecidable," as the reason that Gödel goes on to present is one that distinguishes category 2 from category 1 but does not contain any element that at the same time distinguishes category 5 from its supercategory 2. The reason that Gödel gives is that the higher system in turn is incomplete, and therefore still leaves formally undecided other sentences, which then must have been undecidable in the first system as well. These are decided again in even higher systems, and the story repeats itself, ad infinitum; but it never leads out of category 1.[15] So it seems that Gödel around 1931 mentions (in effect) the notion of weak absolute undecidability only once and in passing.

2 $V = L$

2.1 1935–1940: A candidate for weak absolute undecidability

But various people had already begun to entertain the possibility that *CH* may be weakly absolutely undecidable, i.e., that it was not decidable in the known systems for set theory. As early as 1922, Skolem (in a lecture in Helsinki) had conjectured that *CH* cannot be decided from the axioms given in Zermelo 1908.[16] Hilbert's well-known attempt in 1925 (published in 1926[17]) to demonstrate *CH* was, as Gregory Moore put it, "met with widespread skepticism, in particular from Fraenkel (1928) and Luzin (1929),"[18] and in Bologna in 1928, Bernays and Tarski discussed with each other the possibility of independence of *CH* from ZFC.[19] The next year, Tarski mentioned this possibility in print; in the closing paragraph of "Geschichtliche Entwicklung und gegenwärtiger Zustand der Gleichmächtigkeitstheorie und der Kardinalzahlarithmetik"[20], he says that, although he does not have any argument to offer, he strongly suspects that *CH* will in the future be shown to be independent from ZF and ZFC. And although Gödel in the lecture **1931?* does not speculate on *CH* being formally undecidable in ZFC,[21] he too may have had it in the back of his mind then—to Wang he said in 1976 that "it must have been in the summer of 1930 when [I] began to think about the continuum problem and also heard of Hilbert's proposed solution."[22] But certainly no one at the time was in a position to turn the suspicion of independence into a convincing (partial) argument. This may explain why Gödel mentions the notion of (weakly) absolute undecidability but does not give a (possible) example. He did though have a sense where to look for a partial result, i.e. showing not (as Hilbert had attempted to do) *CH* itself, but its consistency with the axioms of ZFC. He must have arrived at a good idea quickly, for, as Kreisel reports on his conversations with Gödel, "he had the general idea for his proof of *GCH* for L as a student."[23]

What Kreisel is referring to is the hierarchy of sets L that Gödel was to define in 1935 and which enabled him to establish that, if ZFC is consistent, so is ZFC+*CH* (and, what is more, ZFC+*GCH*). The strategy is the following. One formulation of *CH* is: there are $\aleph_1$ subsets of $\aleph_0$. So one could try to find a restricted notion of set that on the one hand satisfies the axioms of ZFC but on the other is so strict that it allows one to keep count of the subsets generated from every set. This strictness can be given form in a hierarchy that starts, naturally, with the empty set at the bottom level, on top of which, in a controlled way so as not to lose count, higher and higher levels of sets are built out of the ones previously obtained. This hierarchy does perhaps not capture the full notion of set because the notion of set used may be too restricted for that; but if it is shown that within a model for ZFC one can build this hierarchy L and that in this "inner model" *CH* is true, then it has been shown that if ZFC is consistent (i.e., has a model), so is ZFC+*CH*. (Appropriately, Kreisel in his memoir of Gödel gave his section on constructible sets the subtitle "reculer pour mieux sauter".[24]) The consistency proof is relative to ZF; the consistency of ZF itself has not yet been established in the strongest sense of the word.

To obtain a precise definition of such a hierarchy, two fundamental choices have to be made: what ordinals will there be to serve as indices of the subsequent levels in the buildup of the hierarchy, and what is the method to build a higher level from the ones beneath it? For the first, Gödel introduced a notion of predicative definability in first-order logic; impredicative definitions[25] would not respect the idea of constructing the universe from the ground up and thereby make it impossible to count.[26] As for the second question, Gödel told Wang that he experimented "with more and more complex constructions [for obtaining the ordinals needed to built set-theoretical hierarchies] for some extended period between 1930 and 1935."[27] The breakthrough came in 1935 and consisted in the decision simply to take the classical ordinals as given. In particular, this means that one takes the non-definable and non-denumerable ordinals as given. This is a characteristically realist idea and was what distinguished L from Gödel's earlier efforts at constructing hierarchies of sets.[28] To take the ordinals as given does not detract from the value of the proof, as Gödel explained (in the Brown lecture):

> If you want to use [the set theory based on L] for giving an unobjectionable foundation to mathematics our procedure would of course be preposterous, but for proving the consistency of the continuum hypothesis it is perfectly all right, since what we want to prove is of course only a relative consistency of the continuum hypothesis; i.e., we want to prove its consistency under the hypothesis that set theory, including all its transfinite methods, is consistent. Therefore we are justified in using the whole set theory in the consistency proof (because if a contradiction were obtaind from the continuum hypothesis and if, on the other hand, we could prove its consistency by means of set-theoretical arguments, then these set-theoretical arguments would be contradictory).[29]

We will now be somewhat more precise. Given a set x, a first-order formula $\phi(y, a_1, a_2, \ldots a_n)$ (where all quantifiers range over x) defines a subset of x, namely $\{y \in x | \phi(y, a_1, a_2, \ldots a_n)\}$, where the a_i are specific elements of x that form a (possibly empty) list of parameters. Let $D(x)$ denote the set of all sets thus definable from the set x. Then L is defined as follows (α ranging over all the classically admissible ordinals):

$$
\begin{aligned}
L_0 &= \emptyset \\
L_\alpha &= D(L_\beta) && \text{if } \alpha = \beta + 1 \\
L_\alpha &= \bigcup_{\beta<\alpha} L_\beta && \text{if } \alpha \text{ is a limit ordinal} \\
L &= \bigcup L_\alpha
\end{aligned}
$$

The sets that occur at some L_α Gödel called "constructible." (To avoid unintended bewitchment by the terminology, one should keep in mind that this notion goes far beyond what a constructive mathematican would accept.) The idea that every set is constructible, in other words the idea that the universe of all sets V coincides with the collection L, found its formulation in the axiom $V = L$.

The use of "V" to refer to the universe of all sets has its origin, via Whitehead and Russell's Principia Mathematica, in Peano; Kreisel reports that what Gödel had meant by "L" was "lawlike."[30] But as Kreisel goes on to say that "at the time [i.e. of the consistency proof] he toyed with the idea that L contained all legitimate definitions of sets," one may also suggest that originally "L" rather stood for the German "legitim (definiert, definierbar)," legitimate in the sense that the definitions are predicative and in terms of first-order logic. That "lawlike" starts with the same letter would then be a fortunate coincidence of the linguistic kind. (In German, 'lawlike' is 'gesetzmäßig'.)

Gödel's motivation to look for hierarchies of sets, which eventually led him to L, had been to work on *CH*. But the first result he actually showed about L, after having verified that the ZF axioms hold for it, was that the axiom of choice (AC) holds in it: if ZF is consistent, so is ZF+AC. He practically kept this secret at first though he did tell von Neumann when visiting Princeton that year. Also in 1935, Gödel conjectured that $V = L \rightarrow CH$ and that therefore *CH* is consistent with ZF and with ZFC. He set out to prove this,[31] but for a long period he struggled with depression and poor health. The proof that $V = L \rightarrow (G)CH$ he essentially found during the night of 14 to 15 Juni 1937.[32] On December 15, 1937 he wrote to Karl Menger that he now was trying to prove the independence of *CH* from ZFC (for which, given his earlier result, it would suffice to show that ZFC+$\neg CH$ is also consistent), but without success so far.[33] He announced his consistency results in print in 1938.[34] He did not mention his expectation of independence, which however he did do in his lecture in Göttingen in 1939.[35] Consistency of ZFC+$\neg CH$ (and of ZF+$\neg AC$) would in fact be established by Paul Cohen in 1963 by a method called forcing.[36] A result by Shepherdson from 1953 made it clear that it is actually impossible to use the method of inner models for $\neg CH$ (or $\neg AC$).[37]

In 1938, Gödel claims that

> the consistency proof for A [$V = L$] does not break down if stronger axioms of infinity (e.g., the existence of inaccessible numbers) are adjoined to T [or to ZF]. Hence the consistency of A seems to be absolute in some sense, although it is not possible in the present state of affairs to give a precise meaning to this phrase.[38]

This turned out to be only partially correct, as many of the stronger large cardinal axioms that have later been proposed and are believed to be consistent (e.g., measurable cardinals) have been shown to imply $V \neq L$; but it is correct for inaccessible, Mahlo and the very large weakly compact cardinals. The reservation Gödel expresses refers to the fact that what was still missing is a way to make exact the notion of the whole transfinite series of possible extensions by axioms of infinity.[39] Could Gödel's reason for the suggestion he made in 1938 have something to do with a passage he in 1972 urged Wang to include in the latter's book *From Mathematics to Philosophy*?

> There used to be a confused belief that axioms of infinity cannot refute the constructibility hypothesis (and therefore even less the continuum hypothesis) since L contains by definition all ordinals. For example, if there are measurable cardinals, they must be in L. However, in L they do not satisfy the condition of being measurable. This is no defect of these cardinals, unless one were of the opinion that L is the true universe. As is well known, all kinds of strange phenomena appear in nonstandard models.[40]

Or had it simply been difficult to imagine the very possibility that large cardinals could be of such a different kind that they violate $V = L$? Indeed, when Dana Scott showed in 1961 that a measurable cardinal (introduced by Ulam in 1930) would do just that, Gödel commented that that is an axiom of infinity "of an entirely new kind," as had become clear only shortly before.[41] In a (draft) letter to Tarski of August 1961, he writes: "You probably have heard of Scott's beautiful result that $V \neq L$ follows from the existence of any such measure for any set. I have not checked this proof either but the result does *not* surprise me."[42] Presumably, this would have surprised him in 1938.

In 1939, Gödel explained his consistency proofs of AC and CH in a lecture in Göttingen;[43] on that occasion he voiced his suspicion that $V = L$ is strongly absolutely undecidable:

> The consistency of the proposition A (that every set is constructible [$V = L$]) is also of interest in its own right, especially because it is very plausible that with A one is dealing with an absolutely undecidable proposition, on which set theory bifurcates into two different systems, similar to Euclidean and non-Euclidean geometry.[44]

Gödel's remark in Göttingen about Euclidean and non-Euclidean geometry is reminiscent of his remarks in the second edition of the Cantor paper from 1964. There however he makes a comment to quite the opposite effect:

> [I]t has been suggested that, in case Cantor's continuum problem should turn out to be undecidable from the accepted axioms of set theory, the question of its truth would lose its meaning, exactly as the question of the truth of Euclid's fifth postulate by the proof of the consistency of non-Euclidean geometry became meaningless for the mathematician. I therefore would like to point out that the situation in set theory is very different from that in geometry, both from the mathematical and from the epistemological point of view.[45]

Gödel then explains this in terms of weak and strong extensions (see below). In making the comparison with geometry in the Göttingen lecture, he probably did not have the notion of inner model in mind at all, but merely the fact that there are two consistent ways of extending absolute geometry and that it does not make sense to ask which one is the correct one; similarly, he thought at the time, extending ZFC by $V = L$ or by $V \neq L$ are both consistent and it does not make sense to ask which extension is the correct one. (His conviction of the consistency of the axiom stating that nonconstructible sets exist foreshadows in a way the generic sets that Cohen would later use.)

In this lecture Gödel does not explicitly define what he means by "absolutely undecidable," but in his lecture at Brown University in 1940, when referring to the very same result, he defines the related notion of absolute consistency by saying that his consistency proof is absolute in the sense that it is "independent of the particular formal system which we choose for mathematics."[46] By the formal systems that can be chosen he evidently cannot mean just any formal system, as such a system could contain $V \neq L$ as an axiom or an axiom implying it. It is far more likely that he means first of all the systems he also had in mind in the Göttingen lecture the year before, in which he had said that "as is well known, there are different mathematical formalisms, such as the Russellian, the Hilbertian, the formalism of axiomatic set theory, and others";[47] and in addition to those, their extension as suggested by applications of the incompleteness theorem. Indeed, in the Göttingen lecture Gödel went on to mention that "today in fact we know that every mathematical formalism is necessarily incomplete and can be extended by means of new evident axioms. So, strictly speaking, there is no one mathematical formalism at all, but rather only an unsurveyable sequence of ever more comprehensive formalisms." This is a reference to his own incompleteness theorem (as it is this that justifies the adverb in "necessarily incomplete," and thereby his speaking of a "sequence"). The new evident axioms then are the undecidable sentences generated by the proof of the incompleteness theorem (which we can see to be true), in particular, consistency statements, or (more generally) corresponding axioms of infinity (adding new types or levels to the iterative hierarchy). He then says that his consistency proof of CH "is applicable to all formalisms hitherto set up, and one can show that it holds unchanged even for the aforementioned extensions by new evident axioms, so that consistency therefore holds in an absolute sense." Because of the reference to the "aforementioned extensions," "absolute consistency" here seems to mean: consistent with ZFC and any series of extensions of it that result from adding statements supplied by the incompleteness theorem. By analogy, "absolutely undecidable" then means: undecidable in ZFC and in any series of extensions of it that result from adding statements seen to be true from the proof of the incompleteness theorem.

That would put Gödel's "absolutely undecidable" statements, which he suggests here includes $V = L$, in category 2, but not in category 5. Not without further argument, that is; but that is not to be found in the papers under discussion. Why, then, did he call these statements "absolutely undecidable"? Here we stumble upon a difficulty in Gödel's writings on the theme of undecidability before 1947: besides $V = L$ or its negation, he seems to have thought up till then that all axioms to extend ZFC have to be statements seen to be true from the proof of the incompleteness theorem (generally, axioms of infinity), and to have built this into his notion of absolute undecidability. But why would whether a statement is absolutely undecidable or not depend only on ZFC and axioms of infinity? In 1947, Gödel himself suggested that axioms of another type may be needed too. We will come back to this when discussing the paper from that year below.

One is presented with probably this same difficulty by a lecture manuscript which is likely to have been written between 1938 and 1940.[48] Its year is therefore referred to in *CW III* as **193?*, and we will follow this practice. Of interest in this lecture for the present discussion is that Gödel relates his ideas on absolute undecidability explicitly to Hilbert, and that he makes conjectures about the complexity of the simplest absolutely undecidable statements. Instead of ZFC specifically, he here reasons more generally about formal systems on which the only demand is that they can express Diophantine propositions of a specific, simple type. Gödel shows, as he had first done in lectures at Princeton 1934, that the undecidable sentence exhibited in his 1931 paper can be taken to be "almost Diophantine," i.e. of "class A," which is defined as the class of sentences of the form

$$(\forall a_1, \ldots, a_m)(\exists x_1, \ldots, x_n)D$$

where D is a Diophantine equation with natural number coefficients. This is theorem 2 of the manuscript, where theorem 1 asserts the undecidability of class A, in anticipation of the solution of Hilbert's Tenth Problem in the early 1970's due to Matiyasevic, Davis, Robinson and Putnam, which obtains theorem 2 for sentences of class A but with no universal quantifiers. Gödel remarks in the manuscript that the result delineates "the smallest portion of mathematics which cannot be completely mechanized" so far known.[49] This part of the paper, almost the whole, was meant by Gödel as a collection of scattered previous results.[50] But whereas the undecidable statement of class A is of category 1, Gödel suspects that there is a statement of a very similar structure, which is related to *CH*, but behaves very differently:

> As to problems with the answer Yes or No, the conviction that they are always decidable, remains untouched by these results [i.e., the existence of undecidable statements in any system that includes class A]. However, I would not leave it unmentioned that apparently there do exist questions of a very similar structure which very likely are really undecidable in the sense which I explained first. The difference in the structure of these problems is only that also variables for real numbers appear in this polynomial. Questions connected with Cantor's continuum hypothesis lead to problems of this type. So far I have not been able to prove their undecidability, but there are considerations which make it highly plausible that they really are undecidable.[51]

What does Gödel mean here by the phrase "really undecidable in the sense which I explained first"? At the beginning of the text, Gödel recalls "Hilbert's famous words that every mathematician is convinced that for any precisely formulated mathematical question a unique answer can be found."[52] Gödel points out that, if this conviction is studied in the context of mathematical logic and proof theory, the incompleteness theorem suffices to refute it even for number theory. However, he adds:

> [I]t is clear that this negative answer may have two different meanings: (1) it may mean that the problem in its original formulation has a negative answer, or (2) it may mean that through the transition from evidence to formalism something was lost.[53] It is easily seen that actually the second is the case, since the number-theoretic questions which are undecidable in a given formalism are always decidable by evident inferences not expressible in the given formalism.[54]

The sense of undecidability that Gödel, as he says at the end of the paper, 'explained first', is the one labelled (1) in this quotation from the beginning of the paper; this means that at the end of the lecture he says that there do seem to be, contrary to Hilbert's conviction, precisely formulated mathematical questions for which no unique answer can be found.[55] Such questions would be of category 5; but to reach such a strong conclusion would seem to be beyond the means available to Gödel then (or later; but see the section on rationalistic optimism, below). Thus, Parsons comments on the closing passage of **193?* that

> It is hard to see what Gödel could have expected to "prove" concerning a statement of the form he describes other than that it is consistent with and independent of the axioms of set theory, say ZF or ZFC, and that this independence would generalize to extensions of ZFC by axioms for inaccessible cardinals in a way that Gödel asserts that his consistency result does.[56]

The puzzlement seems to be caused by Gödel's particular and limited view at the time on what absolute undecidability consists in.

We summarize the discussion so far by saying that Gödel seems to have identified for a while categories 2 and 5. In the remainder of this section, we address two questions: What could the polynomials mentioned at the end of the **193?* lecture have been? And did he ever think that $V = L$ is true?

As Gödel says that the polynomials he has in mind are connected to CH, one may at first think of equivalents of "Every real is constructible" or of $V = L$. For what other candidates for absolute undecidability could he have had in view? The passage at the end of **193?* bears a close resemblance to one in the Brown lecture of 1940:

> A [every real is constructible] is very likely a really undecidable proposition (quite different from the undecidable proposition which I constructed some years ago and which can always be decided in logics of higher types). This conjectured undecidability of A becomes particularly surprising if you investigate the structure of A in more detail. It then turns out that A is equivalent to a proposition

> of the following form: $(P)[F(x_1, \ldots, x_k, n_1, \ldots, n_l) = 0]$, where F is a polynomial with given integer coefficients and with two kinds of variables x_i, n_i, where the x_i are variables for real numbers and the n_i variables for integers, and where P is a prefix, i.e., a sequence of quantifiers composed of these variables x_i and n_i. I have not yet succeeded in proving that A, and hence this proposition about this polynomial, really is undecidable, but what I can prove owing to the results which I presented in this lecture is of course this: Either this proposition is absolutely undecidable or Cantor's continuum hypothesis is demonstrable (since A implies the continuum hypothesis). But I have not yet been able to determine which one of these two possibilities is realized.[57]

"Every set is constructible" implies "Every real number is constructible", as real numbers are conceived of as particular sets. The converse does not hold, for there exist all kinds of other sets than the reals. However, both imply *CH*, and perhaps that is why Gödel chose to "denote by A or A_n the proposition which says that every real number (and more generally) every set is constructible."[58] On the assumption that the equivalence that Gödel claims indeed exists, we have chosen to gloss A by "every real is constructible"; for by forcing arguments, for no m, n is $V = L$ equivalent to a Π^m_n statement.[59]

"Every real is constructible" does not admit of a Π^1_2-equivalent, by Shoenfield's absoluteness lemma. It is a corollary of this theorem that any Π^1_2 statement is absolute for any transitive model of ZFC that contains all countable ordinals.[60] "Every real is constructible" then cannot be Π^1_2, for there are transitive models of ZFC containing all countable ordinals and also non-constructible reals. So neither statement that Gödel denotes by "A" is equivalent to a polynomial of the form he has in mind at the end of the paper **193?*; of course it cannot be asked that Gödel had known this in the 1930's. Notice that this particular condition on the form of these polynomials is no longer made in the Brown lecture. This suggests the following possible explanation of the situation: assume that **193?* indeed was written before the Brown lecture.[61] Then it could be that while working on the former, Gödel still suspected that $V = L$ or "every real is constructible' had a Π^1_2-equivalent. In the last line of **193?*, Gödel says about the unspecified polynomials: "So far I have not been able to prove their undecidability, but there are considerations which make it highly plausible that they really are undecidable".[62] In the possible explanation that we suggest, these considerations would involve two stages: first, to establish Π^1_2-equivalents of "Every real is constructible" or of $V = L$, and second, to establish that the latter two "really are undecidable." But in the interval between the two lectures he came to realize (or strongly suspect) that the first stage cannot be completed. Moreover, or as part of this realization, in that interval he had come to see that "every real is constructible" is essentially Π^1_3. The second stage remained, and it is this one that survived in the Brown lecture (and beyond, until Cohen's work).

It has been suggested (e.g., by Martin Davis and by Gregory Moore) that upon introducing $V = L$, at first, Gödel thought that it is true.[63] To be sure, Kreisel reports that "At the time he toyed with the idea that L contained all legitimate definitions of sets";[64] the crucial step to arrive at the identification of V and L would then be to assert that, besides the classical ordinals which are taken as given, no other sets but the legitimately definable (i.e., constructible) exist. As evidence for the suggestion that Gödel indeed identified V and L, Davis and Moore point to a statement that Gödel made when announcing his consistency proof of *CH* in 1938:

> The proposition A [$V = L$] 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.[65]

Naturality is a fine thing but it does not always extend to plausibility, let alone truth; for would $V = L$ determine the vague notion in the right way? Gödel's formulation leaves this very much open. It qualifies the axiom as natural "in so far as" it sharpens the notion of an arbitrary set "in *a* definite way" (emphasis ours). Even someone who is convinced that $V = L$ is false would agree that it thus sharpens the notion of arbitrary set. Gödel's formulation does not at all exclude that there are other definite ways to determine the vague notion.

An additional suggestion offered by Davis[66] is that Gödel's use of the term "axiom" for $V = L$ in his monograph on the consistency of *CH* from 1940[67] is indicative of his holding it true; but Gödel may well have meant to use the term in a formal sense that is not related to truth, as he would do for example on p. 184 of *1947*, in particular when he writes "from an axiom in some sense directly opposite to this [axiom of constructability] the negation of Cantor's conjecture [*CH*] could perhaps be derived."[68]

In 1938, Gödel only mentions the consistency of $V = L$ and says nothing about $V \neq L$. Did the reason he gives for thinking $V \neq L$ is also absolutely consistent occur only later? In any case, during the period that Gödel considered $V = L$ as well as its negation absolutely consistent (which period includes the Göttingen lecture from 1939, arguably the lecture **193?*, and the Brown lecture from 1940), he cannot, given his views, reasonably have held $V = L$ true. For to hold $V = L$ true under those circumstances would be to claim that $V = L$ is of category 4, and as we have explained it is obvious that that category should be empty. Note that in Göttingen in 1939, after having named two "interesting consequences" of $V = L$ (one of which being *CH*), he adds merely that "besides, the consistency of A [$V = L$] has a certain interest in and of itself";[69] one would have expected a stronger formulation if he had believed that $V = L$ is moreover true.

2.2 1947: ...but not for strong absolute undecidability

As $V = L$ implies *CH*, any argument against *CH* would also be an argument against $V = L$. In 1947, in "What is Cantor's continuum problem?," Gödel adduces a number of reasons why *CH* is probably false. By implication, these are reasons why $V = L$ is probably false (and to that extent indicates that $V = L$ is not in category 5); indeed, he writes that "not one plausible proposition is known which would imply the continuum hypothesis."[70] The reasons that Gödel presents all consist in a fact and a judgement; the fact being of the form "It has been shown that *CH* has consequence P," and the judgement that P is very implausible or paradoxical.[71] Gödel mentions that these facts were "not known or not existing at Cantor's time." He then gives a list of such facts, referring to results published by Luzin in 1914, by Sierpiński between 1924 and 1935 (one of them with Braun), and by Hurewicz in 1932.[72] Given these dates, it is somewhat surprising that Gödel in his lectures in 1939–1940 instead of mentioning them suggests that $V = L$ is not only undecidable in ZFC but "absolutely undecidable." As we have seen, it is not in every case immediately obvious what Gödel meant by that term, its reference seeming to oscillate between categories 2 and 5. But in either case the facts in question might have given him pause: either because they suggest inadequacy of the label "absolutely undecidable" for category 2, or because they suggest that there are considerations leading to a decision of $V = L$ after all, on account of which it would not be in category 5. This strengthens the suspicion, noted above, that there was something missing in his notion of absolute undecidability at that time.

On the other hand, Gödel's willingness in the text **193?* to identify his (particular) notion of absolute undecidability with Hilbert's informal notion (category 5) is at odds with a conviction on which Menger reports. According to Menger's memoir, Gödel in 1939[73] had come to express "more and more emphatically" his

> early conviction that the right axioms of set theory had not yet been discovered [...] He undoubtedly meant that no one had given an adequate basic description of that world of sets in which he believed—a description that would permit us to decide the fundamental problems of cardinality such as Cantor's continuum hypothesis [...] I [Menger] myself never heard from him any indications about where he expected to find such axioms.[74]

It is difficult to see how Gödel could suggest the existence of statements that are absolutely undecidable in Hilbert's original sense if he at the same time thought that axioms were still missing. The "early conviction" Menger had described in somewhat more detail earlier on: "In 1933 he already repeatedly stressed that *the right* (*die rechten*, sometimes he said *die richtigen*) *axioms of set theory had not yet been found.*"[75]

3 1947: *CH*, conceptual incompleteness and realism

The question raised by Menger's memoir is perhaps not unanswerable, but at present we have no suggestion to make. The fact remains that, even if Menger is correct about what Gödel told him in 1939, in the lectures of 1939-1940 (the Göttingen lecture in 1939 took place on December the 15th, so after the stay at Notre Dame that Menger reports on, which lasted from January till June), Gödel certainly breathed no word about this conviction that fundamental axioms were still missing from set theory. In 1947, however, he came to communicate it publicly. From a philosophical point of view, the particular form this suggestion takes is of a much broader importance (because it pertains directly to the very foundations) than a decision of the specific problem of *CH* (by perhaps known means) would be:

> As for the continuum problem, there is little hope of solving it by means of those axioms of infinity which can be set up on the basis of principles known today (the above-mentioned proof for the undisprovability of the continuum hypothesis, e.g., goes through for all of them without any change). But probably there exist others based on hitherto unknown principles; also there may exist, besides the ordinary axioms, the axioms of infinity and the axioms mentioned in footnote 17 [axioms on higher-order properties of sets], other (hitherto unknown) axioms of set theory which a more profound understanding of the concepts underlying logic and mathematics would enable us to recognize as implied by these concepts.[76]

Most of the subsequent attention of set theorists to this passage seems to have gone into "axioms of infinity based on hitherto unknown principles." Yet the most important difference with the 1939–1940 lectures is that Gödel here has come to consider the need for new axioms whose introduction is not suggested by the incompleteness theorem but rather by conceptual analysis (this emphasizes that incompleteness cannot be considered merely an artefact of formalization). It might of course happen that justifications from the concept of set will also be found for the large cardinals based on new principles. In 1966, Gödel pointed out that so far this had not happened.[77] As Charles Parsons remarked on the lecture **193?*, "There seems to be a clear conflict with the position of 1947; it's hard to believe that at the earlier time he thought that exploration of the concept of set would yield new axioms that would decide them [i.e. the statements Gödel in **193?* suspected to be "really undecidable"]."[78] (In fact, the large cardinal program to decide *CH* has so far not provided a decisive solution.) In the 1947 paper, Gödel announced the idea of conceptual analysis a few pages before the quotation just given, as follows:

> This scarcity of results, even as to the most fundamental questions in this field, may be due to some extent to purely mathematical difficulties; it seems, however [...] that there are also deeper reasons behind it and that a complete solution of these problems can be obtained only by a more profound analysis (than mathematics is accustomed to give) of the meanings of the terms occurring in them (such as "set," "one-to-one correspondence," etc.) and of the axioms underlying their use.[79]

The suggestion, then, is that the usual systems of set theory such as ZFC, as well as being formally incomplete as shown in the incompleteness theorems, are also incomplete in another, more basic sense; they may be called "conceptually incomplete."[80] It is not at all impossible that Gödel's newly found interest in the analysis of concepts was related to his study of Leibniz, but at present we cannot be more specific. There certainly is a strong Leibnizian flavour to an item in notebook XIV which is related to the "concepts underlying logic and mathematics" that he mentioned in the quotation before the last one:[81]

> The fundamental philosophical concept is cause [...] Perhaps the other Kantian categories (that is, the logical [categories], including necessity) can be defined in terms of causality, and the logical (set-theoretical) axioms can be derived from the axioms of causality. (Property = cause of the difference of things).[82]

Perhaps it is by such metaphysical derivations that Gödel hoped to clarify a fundamental underdeterminedness of the concept of set by ZFC that he mentions in 1947: one can take as a model for ZFC either his hierarchy L, or the class of arbitrary multitudes irrespective of whether or not they are constructible or in some other sense definable. But presumably it would be an essential property of sets if they are definable. "This characteristic of sets, however, is neither formulated explicitly nor contained implicitly in the accepted axioms of set theory," Gödel comments, and to that extent ZFC is, given these two very different types of models it admits, not sharp enough an axiomatization.[83]

Gödel takes this view because he is a realist, meaning that he is "someone who believes [the axioms of set theory] describes some well-determined reality," in which, in particular, "Cantor's conjecture must be either true or false."[84] Kreisel aptly remarks that the constructible may also be taken to constitute "some well-determined reality";[85] but there is a consideration that would limit the use of that observation as an independent argument for holding that $V = L$ is true. Namely, if one holds that mathematical reality should admit of a conceptual description that is entirely self-coherent, this certainly counts against $V = L$:

> [The constructibility hypothesis] is not a conceptually pure proposition because it allows ordinal numbers definable only by impredicative definitions or not definable at all, but proceeds to reject all further uses of impredicative definitions.[86]

(Borrowing a term Gödel once used to describe Hilbert's formalism,[87] from a philosophical point of view one may describe L as "a curious hermaphroditic thing.") After his remark on mathematical reality, Gödel concludes about CH that

> its undecidability from the axioms as known today can only mean that these axioms do not contain a complete description of this reality: and such a belief is by no means chimerical, since it is possible

> to point out ways in which a decision of the question, even if it is undecidable from the axioms in their present form, might nevertheless be obtained.[88]

(It is not obvious that only realists should find this sufficient reason to look for new axioms.) He then describes two such ways, the one being that of conceptual analysis and the other that of inductive arguments. We will discuss them below, taking as our point of departure the version from 1964 (in which conceptual analysis is tied to a specific notion of intuition).

An interesting example of the possibility of such conceptual advancement Gödel gives in both versions is that of the inaccessible and the Mahlo cardinals. This example is based on the iterative conception of set:

> This concept of set [...] according to which a set is anything obtainable from the integers (or some other well-defined objects) by iterated application of the operation "set of," and not something obtained by dividing the totality of all existing things into two categories, has never led to any anitomy whatsoever; that is, the perfectly "naïve" and uncritical working with this concept of set has so far proved completely self-consistent.[89]

Indeed, to Wang he later said that the iterative concept is "simply the correct" concept of set. [90] It is this concept that he has in mind when he writes that

> the axioms of set theory by no means form a system closed in itself, but, quite on the contrary, the very concept of set on which they are based suggests their extension by new axioms.[91]

The idea is that, as soon as one has determined exact ways of forming sets, all the sets obtained by these specific means can be collected to form a set. If one thinks of the ZFC axioms as a list of such exact means and then applies this idea, one is led to inaccessible cardinals, and from there to the even larger Mahlo cardinals[92]. Both give natural extensions:

> [T]hese axioms show clearly, not only that the axiomatic system of set theory as known today is incomplete, but also that it can be supplemented without arbitrariness by new axioms which are only the natural continuation of the series of those set up so far.[93]

One might have thought that the existence of inaccessibles requires a separate assumption, involving some form of maximality, to be adjoined to the pure concept of set; but this is not the case.[94]

4 Abstract considerations about absolute undecidability

4.1 1944, 1946: absolute provability

As was only to be expected, the different developments in Gödel's thought concerning these topics did not dovetail neatly but overlapped. We take a small step back in time. Only a few years after writing the manuscript **193?*, which leaves open the possibility that there exist strongly absolutely undecidable sentences, Gödel came to think that, on the contrary, category 5 is empty. In 1946, in his remarks before the Princeton bicentennial conference on problems in mathematics,[95], Gödel commented briefly on a notion of absolute demonstrability (absolute in the sense of not depending on the formalism chosen). Such a concept of demonstrability could of course not be entirely formalizable (because of his own incompleteness theorem), but Gödel does not exclude that a concept of an appropriately different character can be found which would entail the decidability of every set-theoretic proposition:

> It is not impossible that for such a concept of demonstrability some completeness theorem would hold which would say that every proposition expressible in set theory is decidable from the present axioms plus some true assertion about the largeness of the universe of all sets.[96]

As we saw above, by 1947 Gödel thought that axioms of infinity need not be sufficient and that axioms of a different kind may also be required (as in the context of this quotation Gödel only speaks of axioms of infinity, we take it that the one time he uses "largeness" he does not also have in mind the width of the hierarchy; also note that largeness may well involve more than just cardinality). What is in any case striking is the very suggestion here that a notion of absolute provability (for set theory) is possible and moreover within reach. The philosophical attitude required to make such a remark with some confidence may well have been instilled or reinforced in Gödel by his study of Leibniz, for the remark may be considered as a further development of the claim Gödel had made at the end of the Russell paper, two years earlier:

> Leibniz did not in his writings about the *Characteristica universalis* speak of a utopian project; if we are to believe his words he had developed this calculus of reasoning to a large extent [...] He went even so far as to estimate the time which would be necessary for his calculus to be developed by a few select scientists to such an extent "that humanity would have a new kind of an instrument increasing the powers of reason far more than any optical instrument has ever aided the power of vision." [...] Furthermore, he said repeatedly that, even in the rudimentary state to which he had developed the theory himself, it was responsible for all his mathematical discoveries.[97]

What Gödel says here is amplified by a remark he is recorded to have made in 1948. Wang reports on a note by Carnap on a conversation with Gödel on March 3 of that year, according to which Gödel thought that Leibniz apparently had obtained a decision procedure for mathematics.[98] Gödel also said that, while the system cannot be completely specific (again, because of his own incompleteness theorem), it may still give sufficient indications as to what is to be done.[99]

At the same conference in Princeton in 1946, Tarski also spoke on decision problems.[100] He makes the distinction (for number theory) between undecidable statements of category 1 and 2; as for problems in set theory, he mentions Gödel's recent work on the continuum hypothesis and expresses a belief that certain problems of set theory may be independent (as we saw above, he had done the same in 1929, when the actual situation in set theory had been less clear). But unlike Gödel, he does not touch on the problem whether category 5 is empty or not.[101]

A remark Church made in the discussion at the Princeton conference should be noted as well.[102] Zermelo had in 1932 proposed a theory of infinite proofs and had hoped that all true mathematical propositions were provable in this extended sense. Church objected to proposals of this kind (as reported in the minutes) that "while such systems might have considerable interest of one kind or another, they could not properly be considered *logics*, insofar as logics explicate the notion of *proof*. For what we mean by a proof is something which carries finality of conviction to any one who admits the assumptions (axioms and rules) on which the proof is based; and this requires that there be an effective (finitary, recursive) syntactical test of the validity of proposed proofs."

4.2 1951: Strong absolute undecidability as an abstract possibility

In 1951 Gödel returns to absolute undecidability. In what has become known as the Gibbs lecture, he defines absolute undecidability to mean "undecidable, not just within some particular axiomatic system, but by *any* mathematical proof the human mind can conceive."[103] (As we already had occasion to recall, in version III of the Carnap paper Gödel characterized the notion of proof in "its original 'contensive' meaning" as "a sequence of thoughts convincing a sound mind".[104]) This time there is no ambiguity, and he clearly means strong absolute undecidability. In particular he considers the possibility that among the absolutely undecidable sentences in this sense, if there are such, will occur standard Diophantine sentences of type Π^0_2. He then goes on to establish his "disjunctive theorem":

> Either mathematics is incompletable in this sense, that its evident axioms can never be comprised in a finite rule, that is to say, the human mind (even within the realm of pure mathematics) infinitely surpasses the powers of any finite machine, or else there exist absolutely unsolvable Diophantine problems of the type specified (where the case that both terms of the disjunction are true is not excluded,

so that there are, strictly speaking, three alternatives).[105]

Truth of the first disjunct of course would not mean that category 4 in our classification of undecidable sentences is non-empty after all. It is rather based on the fact that the capacity to see the consistency of every consistent finite formal system is not a capacity that a finite machine can have; so if the human mind indeed has that capacity, it is not a finite machine. Its powers would moreover surpass that of any finite machine "infinitely," because for any finite machine there exist infinitely many others of which that machine cannot establish their consistency.

As Gödel adds, he means the disjunction to be inclusive: thereby the possibility that category 5 is non-empty is, in effect, explicitly left open. That Gödel considers it plausible that it is not empty may be inferred from his characterization, later on in the Gibbs lecture, of his platonistic view as "the view that mathematics describes a non-sensual reality, which exists independently both of the acts and [of] the dispositions of the human mind and is only perceived, and probably perceived very incompletely, by the human mind."[106] An alternative explanation would be that perception only allows us to establish basics such as the axioms, and that for more complicated cases perceptions are not available and we have to resort to logic. Below we will make some remarks about Gödel's realist views on mathematics; here we would like to emphasize a point made by Charles Parsons that the existence of (strongly) absolutely undecidable propositions would in itself not be incompatible with realism.[107]

4.3 Phenomenology and rationalistic optimism

One consequence of the disjunctive theorem is this: If the mind is a finite machine, then there are absolutely undecidable Diophantine problems. So one might try to settle the issue by attempting to establish that the mind indeed is a finite machine.

However, that was clearly not what Gödel had in mind, given the views he expressed on Leibniz in the 1940s (see above), and, consistent with Leibniz' position as the (grand)father of German Idealism, Gödel's philosophical development in the direction of idealistic philosophy; in particular, from 1959 on, to Husserl's transcendental idealism, which became the general framework for his general philosophical endeavours and for the grounding of his mathematical realism in particular.[108] One of Gödel's aims was to use phenomenology to clarify our understanding of the mind as well as of the ontology of mathematics to such an extent that it would be established that the mind is not a finite machine, and that there are no absolutely unsolvable problems.[109] In a draft letter from (June?) 1963 from Gödel to TIME Inc., regarding the upcoming publication *Mathematics* in the Life Science Library, he connects his phenomenological program to his famous "disjunctive conclusion" that either the human mind infinitely surpasses the powers of any finite machine, or there exist absolutely unsolvable Diophantine problems.[110] In that draft letter, he mentions the disjunction again, with the disjuncts in reverse order, and then comments:

> I believe, on ph[ilosophical] grounds, that the sec[ond] alternative is more probable & hope to make this evident by a syst[ematic] developm[ent] & verification of my phil[osophical] views. This dev[elopment] & ver[ification] constitutes the primary obj[ect][111] of my present work.[112]

And another version of that passage reads

> I conj[ecture] that the sec[ond] altern[ative] is true & perhaps can be verified by a phenomenol[ogical] investigat[ion] of the processes of reasoning[113].

A sign of Gödel's optimism at the time is that he saw to it that in the TIME book itself, which appeared in 1963, it was reported that

> "Either mathematics is too big for the human mind,' he says, "or the human mind is more than a machine." He hopes to prove the latter.[114]

In our discussion of the paper from 1964, we will make some comments on the importance of phenomenology for Gödel's realism. In later remarks on minds and machines, Gödel brings into play what he calls "the rationalistic attitude," in connection to which he mentions the name of Hilbert but which also takes up again the Leibnizian theme at the end of the Russell paper. In the 1970's, Gödel said to Wang:

> Our incompleteness theorem makes it likely that the mind is not mechanical, or else the mind cannot understand its own mechanism. If our result is taken together with the rationalistic attitude that Hilbert had *and which was not refuted by our results*, then (we can infer) the sharp result that the mind is not mechanical. This is so, because, if the mind were a machine, there would, contrary to this rationalistic attitude, exist number-theoretic questions undecidable for the human mind.[115]

In 1972 he went into a little more detail and gave the basic ideas of two arguments, which ideas were then published in Wang's *From Mathematics to Philosophy*:

> If it were true [that there exist number theoretical questions undecidable for the human mind] it would mean that human reason is utterly irrational by asking questions it cannot answer, while asserting emphatically that only reason can answer them. Human reason would then be very imperfect and, in some sense, even inconsistent, in glaring contradiction to the fact that those parts of mathematics which have been systematically and completely developed (such as,

> e.g. the theory of 1st and 2nd degree Diophantine equations, the latter with two unknowns) show an amazing degree of beauty and perfection. In these fields, by entirely unexpected laws and procedures (such as the quadratic law of reciprocity, the Euclidean algorithm, the development into continued fractions, etc.), means are provided not only for solving all relevant problems, but also solving them in a most beautiful and perfectly feasible manner (e.g. due to the existence of simple expressions yielding *all* solutions). These facts seem to justify what may be called "rationalistic optimism."[116]

The first argument is a deduction from the essence of reason. If one wishes to attempt such an argument, it would be natural to do so in the context of phenomenology, and this is what Gödel will have had in mind. It would go together well with his intention (see above) to apply phenomenology to establish that the mind infinitely surpasses any finite machine. Similarly, Gödel's claim that "In principle, we can know all of mathematics. It is given to us in its entirety and does not change—unlike the Milky Way."[117] is probably more easily interpreted in the context of Husserl's transcendental idealism than in others.

The second argument is a projection from very specific, highly successful theories. This is a wholly different kind of argument. It is not in obvious contradiction with phenomenological principles but it would take further work to see exactly how it fits in with them. We notice that, as he would do in the 1964 version of the Cantor paper, Gödel here gives two types of argument for a strong conviction: one based on intuition (here, of essences) and one from success (of reason in a particular area). There is a comment by Gödel that is related to this second argument and that contains a reflection on the fact that Hilbert and he shared the conviction of the decidability of all mathematics:

> We have the complete solutions of linear differential equations and second-degree Diophantine equations. We have here something extremely unusual happening to small sample; in such cases the weight of the sample is far greater than its size. The a priori probability of arriving at such complete solutions is so small that we are entitled to generalize to the large conclusion, that things are made to be completely solved. Hilbert, in his program of finitary consistency proofs of strong systems, generalized in too specialized a fashion.[118]

(We have not investigated to what extent this view on what Hilbert did is historically accurate.) We will now see how Gödel in 1964 brought the strongly rationalist position which he is likely to have held from early on but took many years to articulate to bear on *CH*.

5 1964: How to find new axioms and decide *CH*

5.1 The meaningfulness of the question

In the supplement to the 1964 edition of the Cantor paper, Gödel gives two criteria for determining whether a statement that is independent of ZFC gives rise to a decision problem that is meaningful. The first is a mathematical criterion. It is in a sense a result in meaning analysis; on the other hand, at the time Gödel could not demonstrate but only make plausible that *CH* satisfies it. The second is a philosophical criterion. If one accepts the philosophical position that motivates that criterion, then *CH* certainly satisfies it.

The mathematical criterion is based on a distinction between different kinds of extensions of axiomatic systems. Consider the parallel axiom in geometry. Both it and its negation are independent of the first four axioms (absolute geometry), which can thus be extended either way, but for both extensions one can find models in the unextended (Euclidean) system. But then the question of the truth (simpliciter) of the parallel postulate "became meaningless for the mathematician."[119] Rather, geometry bifurcates at the parallel axiom. Gödel speaks of "weak extensions." Something similar holds for questions about extensions of the real field by the addition or non-addition of complex numbers.

Gödel then considers extensions that are stronger, in the sense that they are not weak extensions and that moreover they also have consequences outside their own domain. Gödel gives the example of inaccessible cardinals. In ZFC we can define a model of ZFC + the statement "there are no inaccessible cardinals" as follows:

Case 1. Suppose there are no inaccessible cardinals. Then V can be taken to be the desired model.

Case 2. Suppose there is an inaccessible κ. Take the least such, and call it λ. Cut the universe at V_λ and take everything below for the desired model.

Note that no axiom beyond ZFC has been invoked in the construction of the model, so the statement "there are no inaccessible cardinals" is a weak extension of ZFC. It does not result in new theorems about integers. On the other hand, it is easy to see that in ZFC one cannot establish a model of ZFC + "there exists an inaccessible cardinal" this way. The latter therefore is not a weak extension, and, moreover, new theorems about integers follow from it. Hence it is an extension in a stronger sense. Gödel's mathematical criterion is then, that the question as to the truth of an independent statement is meaningful if either it or its negation (and presumably not both) would be a stronger extension of this type. Applied to *CH*, Gödel notes that models of ZF+*CH* can be obtained by an inner model construction; also *CH* is "sterile for number theory,"[120] i.e. *CH* implies no new theorems about the integers. Therefore *CH* is a weak extension of ZF. Models of ZFC+$\neg$*CH* cannot be thus obtained (Shepherdson's result, see above) and assuming that alternatives to *CH* may have consequences outside their domain, the question whether *CH* is true or false remains meaningful even though it is independent of ZFC. That would show a difference between the parallel postulate and *CH*. As it turns out, a simple forcing argument demonstrates that the negation of *CH* is also sterile for number theory, as noted by Gödel in the postscript to the paper. An asymmetry between the parallel postulate and the *CH* lying in a somewhat opposite direction has been pointed out by Kreisel: namely, in Hilbert's second order axiomatization the parallel postulate is still independent, whereas second order *CH* is decidable.[121] This demonstrates, in Kreisel's view, the significance of the first order/second order distinction. In the section on the notion of success below, we will return to this mathematical criterion.[122]

The philosophical criterion is based on Gödel's realism. In the 1947 version, he had already given the argument that *CH* has a fixed truth value because it is a proposition about a well-determined reality. In our discussion of *1947* we have alluded to this passage but did not quote it. In 1964 he repeats this passage (with some changes in the formulation, which we have argued elsewhere reflect his study of Husserl):[123]:

> For if the meaning of the primitive terms of set theory as explained on page 262 and in footnote 14 are accepted as sound, it follows that the set-theoretical concepts and theorems describe some well-determined reality, in which Cantor's conjecture must be either true or false and its undecidability from the axioms as known today can only mean that these axioms do not contain a complete description of this reality; and such a belief is by no means chimerical, since it is possible to point out ways in which a decision of the question, even if it is undecidable from the axioms in their present form, might nevertheless be obtained.[124]

Robert Tragesser, in his book *Phenomenology and Logic*, explains that this is "the extremely crucial statement" in Gödel's considerations as to how the continuum problem might be solved:

> What is so important in this statement is the tie it makes between our right to say that S [i.e., the domain of set theory] is a well-determined reality (in which, say, *CH* is decided) and the discoverability of promising ways in which open problems (e.g., *CH*) about the domain could be decided. Gödel spends the remainder of the article presenting possible paths to a decision about *CH*. As long as we can find such paths, S will seem to be the well-determined reality we initially took it to be.[125]

The tie that Gödel makes here reflects his adoption of Husserl's transcendental idealism from 1959 onward. For one aspect of Husserl's later philosophy that was of particular importance to Gödel was the way it analysed the relations (1) between the existence of (concrete as well as abstract) objects and consciousness and (2) between consciousness and reason. Briefly, the basic principle of transcendental idealism is that the objects that can be said to exist are exactly those that are in principle accessible to a consciousness that acts in accordance with the evidence it obtains for those objects. To act in that way is precisely what rationality in the most pregnant sense consists in. Thus, Gödel's realism, after having received its foundation in transcendental idealism, and his rationalism, are intimately connected. One can even say that they are two sides of the same coin.[126]

Tragesser continues:

> Gödel may be viewed as giving an analysis of the elements of the prehension of S and, on the foundations of that analysis, showing how *CH* could possibily be decided. Such analysis, because it reflects faithfully upon, and describes, the elements of an act of consciousness (a prehension, in this case), is *phenomenological analysis*. We can see here the critical importance of such analysis, viz., that it provides possible paths to reasons better than arbitrary for holding something to be true of a considered object or objective domain.[127]

Tragesser defined "prehension" as the "imperfect or incomplete "grasp" of a purportedly objective state of affairs, where it is somehow known that the state of affairs is imperfectly or incompletely given."[128] ZFC expresses a prehension of the universe, in the sense that ZFC is not a complete axiomatization in the two senses Gödel gives to the word "incomplete.' Gödel suggests that on the basis of this prehension, in other words, starting from the "incomplete description" we have gotten of the set-theoretic universe so far, it is possible to proceed in such a way as to decide *CH*. In the paper, he proposes two "truth criteria"[129] for evaluating candidate axioms extending ZFC.

5.2 Two truth criteria for new axioms

5.2.1 Intuition

The strong notion of intuition invoked in the 1964 version of the Cantor paper is one of the most conspicuous differences with the text from 1947. In the larger context of Gödel's philosophical development, it is natural that it should have appeared, given Gödel's intensive study of (and enthusiasm for) Husserl's phenomenology since 1959.[130] The key passage in the 1964 paper is this one:

> But, despite their remoteness from sense experience, we do have something like a perception also of the objects of set theory, as is seen from the fact that the axioms force themselves upon us as being true. I don't see why we should have less confidence in this kind of perception, i.e., in mathematical intuition, than in sense perception.[131]

Much has been written on the interpretation of this passage, and we refer the reader to Parsons,[132] Tieszen[133], Tragesser,[134] Van Atten and Kennedy,[135] and, for background, Husserl, whose sixth *Logical Investigation*[136] Gödel recommended to logicians in the 1960s[137].

The comparison of (abstract) intuition to (sense) perception in this passage shows that Gödel means intuition in a technical sense, and as such it is just as decisive as the intuitionists intend it to be. He is not talking about intuition as (merely) a psychological fact here. Still, Gödel allows for mistakes even in intuitions, but that is because intuition is not an all-or-nothing affair. It comes in degrees.[138] And ultimately, existence remains tied to (ideal) intuition, by the basic principle of transcendental idealism.

The work that this intuition is meant to do with respect to the continuum problem is to give a well-defined meaning to the question and indeed to decide it. In 1964, Gödel writes, "That new mathematical intuitions leading to a decision of such problems as Cantor's continuum hypothesis are perfectly possible was pointed out earlier (pages 264–265)."[139]

Gödel was aware that his talk of an objective realm of transfinite set theory and of a faculty of intuition that has access to it would probably not be well received. In fact, he had already feared that the 1947 version, which did not even contain the strong views on mathematical intuition yet, would be subjected to positivistic attacks by Benacerraf and Putnam in the introduction of their planned anthology which was to contain it.[140] Gödel not only overcame these fears but he went ahead to include views even more opposed to positivism in the revision of the 1947 paper that he went on to prepare for the occasion. But perhaps it was a residual fear that motivated him to propose also an argument from metaphysically less contentious premises that should lead to the desired conclusion that *CH* can be decided.

This argument takes the existence of intuition not as an epistemological but as just a psychological fact (where the former sense does, and the latter does not, imply access to objects in reality):

> However, the question of the objective existence of the objects of mathematical intuition [...] is not decisive for the problem under discussion here [...] The mere psychological fact of the existence of an intuition which is sufficiently clear to produce the axioms of set theory and an open series of extensions of them suffices to give meaning to the question of the truth or falsity of propositions like Cantor's continuum hypothesis.[141]

By an appeal to psychology, Gödel suggests that meaningfulness and decidability may be securable without resorting to realism. In 1975, Hao Wang characterized this statement "as asserting the possibility of recognizing meaningfulness without realism"; Gödel agreed, for he suggested that Wang would report: "He [i.e., Gödel] himself suggests an alternative to realism as ground for believing that undecided propositions in set theory are either true or false."[142] Given Gödel's avowals of realism in this paper[143] and elsewhere, it is clear that he does not actually embrace this alternative. But suggesting this alternative serves a dialectical purpose to him, that of indicating that even from alternative points of view, his idea (here, of meaningfulness of the question) is the correct one.[144]

The argument from the psychological fact however is not particularly strong, for a reason that Gödel himself had indicated before and that he, surprisingly, left out of consideration both in the 1964 paper itself and, apparently, in his discussion with Wang of the passage that we just quoted. This reason is stated in a footnote to version III of the paper on Carnap, **1953/9-III*:

> the existence, as a psychological fact, of an intuition covering the axioms of classical mathematics can hardly be doubted, not even by adherents of the Brouwerian school, except that the latter will explain this psychological fact by the circumstance that we are all subject to the same kind of errors if we are not sufficiently careful in our thinking.[145]

This seems to give a more complete view of the situation but also a somewhat more complicated one, for should the intuitionists be right, one can doubt whether an intuition that is brought about by insufficiently careful thinking can itself function as the basis for a sufficiently careful judgement (a judgement that, if it can be made, from the intuitionist's point of view will be purely hypothetical or "as if"). So for the "psychological fact" to have the force that Gödel takes it to have, the suggested intuitionistic interpretation has to be shown wrong first. It would be crucial to do this at some point, for Gödel's whole approach of an appeal to intuition to clarify and (to find axioms that) decide *CH* stands or falls with his being able to come up with a notion of intuition that is relevantly different from the intuitionist's.

Of course Gödel does not at all doubt that the intuitionist is indeed wrong, as is clear from a passage that at the same time brings out the urgency of finding an argument to that effect:

> First of all there is Brouwer's intuitionism, which is utterly destructive in its results. The whole theory of $\aleph$'s greater than $\aleph_1$ is rejected as meaningless [...] However, this negative attitude toward Cantor's set theory, and toward classical mathematics, of which it is a natural generalization, is by no means a necessary outcome of a closer examination of their foundations, but only the result of a certain philosophical conception of the nature of mathematics, which admits mathematical objects only to the extent to which they are interpretable as our own constructions or, at least, can be completely given in mathematical intuition. For someone who considers mathematical objects to exist independently of our constructions and of our having an intuition of them individually, and who requires only that the general mathematical concepts must be sufficiently clear for us to be able to recognize their soundness and the truth of the axioms concerning them, there exists, I believe, a satisfactory foundation of Cantor's set theory in its whole original extent and meaning, namely the axiomatics of set theory interpreted in the way sketched below.[146]

It is possible to have qualms with Gödel's characterization of intuitionism here. A potential unclarity in the use of "completely given" here is whether this may or may not involve certain (carefully controlled) idealizations: intuitionists do make such idealizations, and they are not finitists (as Gödel himself of course pointed out on other occasions).[147] And in any case, intuitionists accept essentially incomplete objects (choice sequences) in their ontology, which by their nature can never be completely given in either actual or (to intuitionistic standards) appropriately idealized intuition. For the contrast that Gödel wishes to draw here, these qualms do of course not make much difference. But the intuitionist can in any case point out the symmetry in the situation as sketched by Gödel in this passage, and use Gödel's own words to comment that the latter's realist views are "by no means a necessary outcome of a closer examination of [the] foundations [of Cantor's set theory and classical mathematics], but only the result of a certain philosophical conception of the nature of mathematics."

Thus, further argumentation is required. Immediately after presenting the argument from the psychological fact, Gödel introduces a second and apparently stronger argument:

> What, however, perhaps more than anything else, justifies the acceptance of this criterion of truth in set theory is the fact that continued appeals to mathematical intuition are necessary not only for obtaining unambiguous answers to the questions of transfinite set theory, but also for the solution of the problems of finitary number theory (of the type of Goldbach's conjecture), where the meaningfulness and unambiguity of the concepts entering into them can hardly be doubted. This follows from the fact that for every axiomatic system there are infinitely many undecidable propositions of this type.[148]

In particular, those undecidable propositions can take the simple form of Diophantine equations. It is through mathematical intuition that we come to see that these propositions are actually true (provided we believe in the consistency of the axioms). The intuitionist agrees, but would balk at taking the range of this intuition to be so large as to include transfinite set theory. According to the intuitionist, even if such an axiom can be shown to be consistent, there is no reason to assume that it is true.

In a third type of argument, not presented in the Cantor paper but as pertinent to the issue discussed there, Gödel aims to show not so much that intuition exists at all, but rather that, if one accepts its existence, it would be wrong to think of it as an all-or-nothing phenomenon that, for that very reason, would cancel any purported intuitive evidence for objects that is not maximal (as in most people's experience the evidence for transfinite sets is not). Rather, the evidence that intuition provides comes in degrees:

> We have no absolute knowledge of anything. There are degrees of evidence. The clearness with which we perceive something is overestimated. The simpler things are, the more they are used, the more evident they become. What is evident need not be true. If 10^{10} is already inconsistent, then there is no theoretical science.[149]

It should not be surprising that what Gödel says here—to Wang, in the 1970s—is in agreement with Husserl's views on intuition, views that Gödel began to study in 1959. In an unpublished manuscript titled "For Perspicuous Objectivity," Hao Wang reports that in November 1972, Gödel drew his attention to the following sentence in Husserl's essay "Philosophy as rigorous science": "Obviously essences can also be vaguely represented, let us say represented in symbols and falsely posited; then they are merely conjectural essences, involving contradiction, as is shown by the transition to an intuition of their inconsistency."[150] Gödel then commented that he was glad that Husserl also recognizes the possibility of error'.[151] A related way in which an essence can be represented vaguely (one which is equally important to the present discussion) is if this representation, is in part symbolic and in part intuitive. We then have evidence (in Husserl's sense) up to a certain degree; but the possibility of error remains.[152] On the other hand, as Husserl writes immediately after the sentence just quoted, '[i]t is possible, however, that their vague position will be shown to be valid by a return to the intuition of the essence in its givenness'.[153]

It is also in the recognition that intuition is not an either-or affair but comes in degrees (i.e., intuitions are, generally speaking, partial) that Gödel sees an answer to skepticism:

> We have no absolute knowledge of anything. To acknowledge what is correct in skepticism serves to take the sting out of skeptical objections. None of us is infallible. Before the paradoxes Dedekind would have said that sets are just as clear as integers.[154]

This remark about Dedekind points to a difference between the degree of evidence we have for the integers and that for sets. This can be generalized from different types of objects to different conceptions of mathematics as a whole and by doing so, one may obtain a scale such as the following one suggested by Gödel in a version of his paper on Carnap paper:

> The field of unconditional mathematical truth is delimited very differently by different mathematicians. At least eight standpoints can be distinguished. They may be characterized by the following catchwords: 1. Classical mathematics in the broad sense (i.e., set theory included), 2. Classical mathematics in the strict sense, 3. Semi-Intuitionism, 4. Intuitionism, 5. Constructivism, 6. Finitism, 7. Restricted Finitism, 8. Implicationism.[155]

That Gödel interpreted such a list as a scale of evidence is clear from the following:

> Without idealizations nothing remains: there would be no mathematics at all, except the part about small numbers. It is arbitrary to stop anywhere along the path of more and more idealizations. We move from intuitionistic to classical mathematics and then to set theory, with decreasing certainty. The increasing degree of uncertainty begins [at the region] between classical mathematics and

> set theory. Only as mathematics is developed more and more, the overall certainty goes up. The relative degrees remain the same.[156]

Moreover, Gödel claims that it would be arbitrary to draw (as the intuitionist does) a line that partitions the scale into an acceptable and an unacceptable part:

> Strictly speaking we only have clear propositions about physically given sets and then only about simple examples of them. If you give up idealization, then mathematics disappears. Consequently it is a subjective matter where you want to stop on the ladder of idealization.[157]

Behind accepting this role of subjectiveness must be the idea that the various possible idealizations are in a sense continuous with one another; and, for Gödel, even as evidence decreases with each further idealization, there is still intuition of a sufficiently high degree which remains to give a purchase on the transfinite. An intuitionist, however, will argue that it is not at all arbitrary where to stop idealizing, and that the place to stop occurs well before having reached all of classical mathematics in the strict sense (let alone transfinite set theory). An argument to this effect is found in the inaugural address of Brouwer's pupil Arend Heyting from 1949. He noticed that within the intuitionistic school, there is disagreement as to what idealizations are permitted. The disputed notions are those of negation, choice sequences, and certain proof methods (Brouwer accepted all of these). Each depends on certain idealizations; with every new idealization 'we descend a step on the stairs of clarity', as Heyting says. He then wonders if, once you are prepared to descend those stairs at all, this might not provide a justification of classical mathematics:[158]

> Those concepts and methods that are not accepted by all intuitionistically oriented mathematicians, upon introspection turn out to have different degrees of clarity, and to be accompanied by convictions of correctness of different intensities. It lies close to hand to ask whether it isn't then also the case that much can be said for accepting the independent existence of the mathematical objects, independent of our thinking, and thereby arrive at classical mathematics, even if this diminishes the clarity of the concepts a bit. For me, the answer to this question is that this step is not at all comparable to the earlier ones. So far, we remained in the realm of mental constructions, now we would all of a sudden leave that. We would be asked not merely to accept a new construction method, but a philosophical thesis, of which the sense and motivation are questionable already for the objects in daily life and which would become even less understandable when applied to mathematical objects. The stairs that slowly led down from the light of day into the darkness stops here, and the next step would be a jump into the darkness of an bottomless well.[159]

This passage illustrates Gödel's general point that "There is a choice of how much clarity and certainty you want in deciding which part of classical mathematics is regarded as satisfactory: this choice is connected to one's general philosophy.";[160] but Heyting warns that it is not just a question of "how much" and that there are also qualitative differences between the idealizations that introduce discontinuities in the scale. These discontinuities put one's notion of intuition under pressure (and hence the range of idealizations one can reasonably make on the basis of that notion).

We will not discuss possible intuitionistic criticisms of Gödel's three arguments any further, but note that the task of their assessment emphasizes the need for further study of the notion of intuition—which notion, for all their differences, is central to both Brouwer's intuitionism and Gödel's platonism—before it can be made to do the work Gödel wants it to do towards, e.g., a decision of the continuum hypothesis. Phenomenology is a systematic way of going about such a study;[161] and indeed, an unpublished draft of the 1964 supplement to the Cantor paper contains an additional paragraph at the end that starts, "Perhaps a further development of phenomenology will, some day, make it possible to decide questions regarding the soundness of primitive terms and their axioms in a completely convincing manner."[162] Of course, Gödel's intention when writing this will not primarily have been to settle the dispute with the intuitionist; but a development as Gödel hopes for here may well have that effect.

5.2.2 Success

> Secondly, even disregarding the intrinsic necessity of some new axiom, and even in case it has no intrinsic necessity at all, a probable decision about its truth is possible also in another way, namely, inductively by studying its "success." Success here means fruitfulness in consequences, in particular in "verifiable" consequences, i.e. consequences verifiable without the new axiom, whose proofs with the help of the new axiom, however, are considerably simpler and easier to discover, and make it possible to contract into one proof many different proofs. The axioms for the system of real numbers, rejected by the intuitionists, have in this sense been verified to some extent, owing to the fact that analytic number theory frequently allows one to prove number-theoretical theorems which, in a more cumbersome way, can subsequently be verified by elementary methods. A much higher degree of verification than that, however, is conceivable. There might exists axioms so abundant in their verifiable consequences, shedding so much light upon a whole field, and yielding such powerful methods for solving problems, (and even solving them constructively, as far as that is possible) that, no matter whether or not they are intrinsically necessary, they would have to be accepted at least in the same sense as any well-established physical theory.[163]

Because of its dependence on (cognitive, aesthetic) values and choices of criteria, the notion of success is much less tractable than that of intuition, which stands in a direct relation to its object. On the other hand, indications of success may be easier to recognize than sharp intuitions, which will generally take much care and effort to arrive at. Some of the precautions that should be taken when searching for intuitions of essences are described by Husserl in sections 80–98 of *Erfahrung und Urteil*.[164] Gödel observed that "appealing to intuition calls for more caution and more experience than the use of proofs—not less";[165] and, as Wang says, "this may be the reason why one could believe in this strong position [according to which we perceive mathematical objects that exist objectively] and yet not regard the criterion of pragmatic success as entirely superfluous."[166]

A clear case of a conflict of values is presented by $V = L$. If one defines success of an axiom in terms of its power to decide questions, then it seems $V = L$ should qualify for that reason. But the axiom conflicts with certain large cardinal axioms that are believed to be consistent. These may not (yet) be seen to be themselves intrinsically necessary, but they too are successful in Gödel's sense, in the sense that they decide many questions lower down (Diophantine equations).[167] Such a conflict of values can be resolved by an appeal to intuition, for example if intuition of the essence of set yields that all consistent sets should exist,[168] then, if one believes the large cardinal axioms in question to be consistent, these cardinals therefore exist. A related argument is that $V = L$ may be successful but is also conceptually impure (a judgement based on intuition), in the sense explained in section **??** above.

There clearly is an asymmetry between the two criteria of intuition and success. As a first approximation one could say that intuition is conclusive while induction is not; but this approximation needs refinement once we take the following into account. As we saw above, Husserl emphasizes that in the correct understanding of intuition, it is not an either-or phenomenon – evidence comes in degrees. On such a construal, specific intuitions will in general not be conclusive in an apodictic sense. This it has in common with inductive arguments, but there are essential differences. Intuitions are of the objects (and the states of affairs composed of them) directly. By their (horizonal) structure, intuitions lend themselves to explication; they suggest ways in which they can be unfolded.[169] Inductive arguments (in terms of fruitful consequences) on the other hand are indirect; evidence for them will not come from the objects (or states of affairs) themselves but from seeing the truth of their consequences. This induces a principled epistemological difference because inductive arguments may have heuristic value but cannot have the same regulative significance (towards the ideal of full clarity and insight, through explication) as partial intuitions. Moreover, this emphasizes the foundational priority of intuition as such over induction as such: for as Wang remarks, "The truth of these consequences, however, had also be seen by mathematical intuition, and we see certain mathematical propositions, such as numerical computations, to be true directly, without going through the axioms. Indeed, we apply our intuition at all levels of generality."[170]

An example of an axiom that extends ZFC and that can be defended on the basis of its success is the axiom stating the existence of inaccessibles:

> A closely related fact is that the assertion (but not the negation) of the axiom [of inaccessibles] implies new theorems about integers (the individual instances of which can be verified by computation). So the criterion of truth explained on page 264 is satisfied, to some extent, for the assertion, but not for the negation. Briefly speaking, only the assertion yields a "fruitful" extension, while the negation is sterile outside of its own very limited domain.[171]

This is a reference to the fact that in the theory ZFC + "there is an inaccessible cardinal" one can prove the statement "there exists a model of ZFC." But this is equivalent to Con(ZFC), by the completeness theorem, which is a Π^0_1 number-theoretic or Diophantine statement. As we saw above, Gödel also had a (much more conclusive) justification of inaccessibles from an analysis of the concept of set. Thus inaccessibles are justified in two very different ways; Gödel probably wanted to include the pragmatic argument because it may serve to convince those who are more easily impressed by formal(istic) considerations.[172]

In the early 1970's Gödel suggested to Wang that the existence of measurables may be perhaps be verified in the same way:

> The hypothesis of measurable cardinals may imply more interesting (positive in some yet to be analyzed sense) universal number-theoretical statements beyond propositions such as the ordinary consistency statements: for instance, the equality of p_n (the function whose value at n is the n-th prime number) with some easily computable function. Such consequences can be rendered probable by verifying large numbers of numerical instances.[173]

On the other hand, *CH* itself cannot have new number-theoretic consequences, as is clear from absoluteness of arithmetical statements: arithmetical statements are true in a given model M of ZFC if and only if they are true in L over M. (Kreisel[174] draws attention to the fact that Gödel did not realize this in section 3 of the 1947 paper: "Certainly [...] mere consistency leaves open the possibility that CH has new, even false arithmetic consequences; but a glance at his own definition of L [...] shows that *CH*, and even $V = L$ has none at all. Gödel's oversight is natural enough if consistency is regarded as an end in itself.")

It turned out (at about the time Gödel was writing the 1964 revision and the 1966 postscript to it) that the negation of CH does not have new number-theoretic consequences either, as a simple observation about forcing shows. Hence, different aspects of fruitfulness than arithmetic consequences (verifiable up to any given integer) will have to be taken into account if an argument from success is to lead to a probable decision of *CH*.

Gödel mentions several different aspects of fruitfulness. One is that of making proofs simpler and "easier to discover," i.e., more obvious; where without the new axiom their proofs are long and "cumbersome." It is an open question whether, to the extent that fruitfulness of this kind is indeed truth-conducive, what is operative here is a relation between truth and aesthetic factors, or between truth and exigence of resources, or both (on the possible ground that economy of resources may be, at least in part, itself an aesthetic factor). So far, no convincing arguments of this type are known that suggest a decision of *CH*. Certainly the recently proposed solution to *CH* we will discuss below is by no means "easy," "obvious," or in some other sense elementary. Should the proof eventually work, then it is of course to be expected that in due course simplifications (e.g., from conceptual insight) will be found.

In the supplement to the 1964 edition of the Cantor paper, Gödel points out that his second truth criterion has not (at the time of writing) led to concrete results:

> It was pointed out earlier [...] that, besides mathematical intuition, there exists another (though only probable) criterion of the truth of mathematical axioms, namely their fruitfulness in mathematics and, one may add, possibly also in physics. This criterion, however, though it may become decisive in the future, cannot yet be applied to the specifically set theoretical axioms (such as those referring to great cardinal numbers), because very little is known about their consequences in other fields. The simplest case of an application of the criterion under discussion arises when some set-theoretical axiom has number theoretical consequences verifiable by computation up to any given integer. On the basis of what is known today, however, it is not possible to make the truth of any set-theoretical axiom reasonably probable in this manner.[175]

It is not entirely clear what Gödel has in mind when he first says that the second criterion is "only probable" and then goes on to add that this inductive criterion "may become decisive in the future." We already quoted from an unpublished draft of this supplement, in which there is an additional final paragraph. In full, it reads:

> Perhaps a further development of phenomenology will, some day, make it possible to decide questions regarding the soundness of primitive terms and their axioms in a completely convincing manner. As of now it seems to me that the character of cogency of its axioms[176] and the success of its development are sufficient reasons for putting trust in Cantor's set theory, i.e., in mathematics in its whole extension.[177]

Gödel expects phenomenology to lead to decisions; in contrast, cogency of the axioms (clearly meant in a sense that does not amount to complete clarity obtained by phenomenological analysis) and inductive arguments from success do not go that far, but increase confidence. Of course, the phenomenological investigation into the basic concepts of mathematics will not by itself settle mathematical questions. As Gödel said to Wang in the 1970's,

> The epistemological problem is to set the primitive concepts of our thinking right. For example, even if the concept of set becomes clear, even after satisfactory axioms of infinity are found, there would remain technical (i.e. mathematical) questions of deciding the continuum hypothesis from the axioms.[178]

Some years after Gödel had written about this notion of success and the senses in which axioms can to some extent be verified, specific examples were found. We mention the following two.

A result obtained by Richard Laver in 1992 says that, if one assumes the existence of very large cardinals (larger than supercompact), one can find a decision method for the word problem of the free algebra with one generator and one left-distributive binary operator.[179] Later, Patrick Dehornoy gave a proof without the large cardinals.[180] So the large cardinals gave the correct result, which in a sense verified their use.

Another, dramatic example of this has to do with Borel Determinacy, which was proved to follow from the ZFC axioms only by D.A. Martin in 1975,[181] where an earlier proof also due to Martin[182] used a measurable cardinal.[183]

In this way Borel Determinacy is a "verifiable consequence," in Gödel's sense of the phrase here, i.e., it was proved without using measurables, and the measurables in turn were verified by their having lead to the "correct" result.[184] As Yiannis Moschavakis put it in his book *Descriptive Set Theory*: "This important result of Martin answered a long-standing question and *lent considerable respectability to the practice of adopting determinacy hypotheses*" [emphasis ours]. [185]

5.3 Gödel's interpretation of Cohen's independence result

In 1963, Paul Cohen proved the consistency of $\neg CH$ with ZFC and thereby, as Gödel had shown the consistency of *CH* with ZFC, the independence of *CH* from ZFC. Cohen sent his proof to Gödel, who (in a draft letter which may or may not have been sent) commented, "Reading your proof had a similarly pleasant effect on me as seeing a really good play."[186]

As we had occasion to mention elsewhere,[187] to Gödel this independence of *CH* from ZFC did not, by suggesting a certain relativism in set theory, pose a threat to his realism. To Church, who did interpret the independence proof along these lines, he wrote on September 29, 1966:

> You know that I disagree [i.e. with Church] about the philosophical consequences of Cohen's result. In particular I don't think realists need expect any permanent ramifications (see bottom of p. 8) as long as they are guided, in the choice of the axioms, by mathematical intuition and by other criteria of rationality.[188]

The fact that (like many others) Church was so impressed by the independence of *CH* from ZFC as to conclude from it to a kind of relativism in set theory may well have had its ground in the unprecedented scope and success of ZFC as a foundation of classical mathematics.

For Gödel, however, there is no fundamental relativity in set theory. Set theory describes a certain (part of abstract) reality, which is therefore the theory's intended model. Cognitive access to that reality is provided by mathematical intuition. Like Brouwer, Gödel holds that mathematical intuition is separate from language, and that language has a practical but not fundamental role to play in obtaining intuitions of mathematical objects:

> Language is useful and even necessary for fixing our ideas. But this is a purely practical affair. Our mind is more inclined to sensual objects, which help to fix our attention on abstract objects. This is the only importance of language.[189]

Proofs that certain statements (or classes of statements) are independent of a given formalism, or that given a given formalism admits of non-intended models, or that formalisms as such have intrinsic limitations, do not, in principle, stand in the way of our mind's capacity to obtain knowledge of the mathematical realm. This explains Gödel's remark to Church about Cohen's proof. To Gödel, the independence of *CH* from (for example) ZFC is no reason whatsoever to doubt that a decision of *CH* is in principle possible. These considerations are not meant to suggest that Gödel did not credit the practical value of formal systems and formalization. In what is now referred to as his "lecture at Zilsel's," held in 1938, he acknowledged that "If the original Hilbert program could have been carried out, that would have been without any doubt of enormous epistemological value."[190] (Note that he does not add "ontological"; it is not immediate that for a realist the epistemological success of Hilbert's program would also have entailed an ontological reduction.[191]) That program of course did not succeed in its original form, but there were other benefits to formal systems; in a letter of December 1918, 1968, to Dana Scott, Gödel remarks:

> In my opinion the formalistic spirit is *extremely important* for mathematics as a technique of solving problems. Also, I perfectly agree with you that formalization, in practice, is an *indispensable* aid to understanding.[192]

5.4 Maximality

As we have seen above, Cohen's result cleared a path for considering axioms deciding CH along grounds which were no longer constrained by knowing only that *CH* was consistent. From a practical point of view this means that axioms implying $\neg CH$ are going to be "in play" in a way they would not have been otherwise.

Though as we have seen, Gödel was interested in axioms of many different kinds, as far as his concrete attempts to solve the continuum problem however, in what can be said to be his last work of a technical nature, it was to Hausdorff's scale axioms that he turned (see the next section). These are prima facie maximality principles, in that some maximal family of functions is asserted to exist, relative to the particular scale involved. But they are at the same time of a somewhat different flavor from what were thought to be maximality principles at the time (see below).

The "set of" operation on which the iterative concept of set is based is "opened forward," so to speak. In footnote 23 of the 1964 version of the paper, while discussing the idea that *CH* might be decided on the basis of axioms about definability p. 262, Gödel remarks that this is known for the type of definability known as constructibility:

> On the other hand, from an axiom in a sense opposite to this one, the negation of Cantor's conjecture could perhaps be derived. I am thinking of an axiom which (similar to Hilbert's completeness axiom in geometry) would state some maximum property of the system of all sets, whereas axiom A [$V = L$] states a minimum property. Note that only a maximum property would seem to harmonize with the concept of set explained in footnote 14.[193]

Footnote 14, as we saw above, asserts that the existence of multitudes is independent of whether we can define them in a finite number of words or whether they are random. So sets are what they are independent of their definitions (which goes some way toward arguing against restricting the notion of set to that of constructible set). Gödel is explicit that a maximality property of this type "harmonizes" with the basic concept. It is not clear that the choice of that particular word should be taken to mean that such a property is not actually contained in the pure concept of set. Recall that although $V = L$ (which is quite the opposite of a maximality principle) is in a way a natural extension of ZFC, its particular naturalness may not persist in a further unfolding of the concept of set.

This period saw Gödel begin to concentrate on maximality principles. Already in the late 1950's Gödel wrote to Stanislaw Ulam about a maximality principle of von Neumann:

> The great interest which this axiom has lies in the fact that it is a maximality principle, somewhat similar to Hilbert's axiom of completeness in geometry. For, roughly speaking, it says that any set which does not, in a certain well defined way, imply an inconsistency exists. Its being a maximum principle also explains the fact that this axiom implies the axiom of choice. I believe that the basic problems of set theory, such as Cantor's continuum problem, will be solved satisfactorily only with the help of stronger axioms of *this* kind, which in a sense are opposite or complimentary to the constructivistic interpretation of mathematics.[194]

To Wang, Gödel reaffirmed this assent to von Neumann's principle by saying that for him, completeness means that every consistent set exists.[195]

6 1970–1975: Gödel's concrete attempts to settle *CH*

In the 1970's Gödel made several efforts at deciding *CH*. These were based on axioms (and axiom schemas) about infinite sequences of integers and scales of functions. We will not go into the technical details here, but give a brief account of this episode.[196]

¿From a letter he wrote to Cohen in January 1964,[197], it becomes clear that Gödel had the basic idea for one of these axioms already then. "I always suspected that, in contrast to the continuum hypothesis, this proposition is correct and perhaps even demonstrable from the axioms of set theory." Whether the fact that a few years later he took this to be an axiom reflects a failed attempt to prove it or simply a shortcut we do not know. Neither do we know whether the scale axioms were the outcome of a conceptual analysis of the kind described in 1947 and 1964.

Gödel wrote a manuscript based on the scale axioms titled "Some considerations leading to the probable conclusion that the true power of the continuum is $\aleph_2$." He planned to submit it to the *Proceedings of the National Academy of Sciences* and in 1970 first sent it to Tarski to solicit comments. D.A. Martin found that the argument was incorrect, as it contradicted a result of Robert Solovay. Tarski sent the paper back to Gödel, simply saying that he would soon hear more about it.

Gödel then wrote another manuscript, this time not meant for publication but for his own use ('nur für mich geschieben'), titled "A proof of Cantor's continuum hypothesis from a highly plausible axiom about orders of growth." In it, he said that the argument presented there "gives *much* more likelihood to the truth of Cantor's continuum hypothesis than any counterargument set up to now gave to its falsehood."[198] Remarkably, and for the first time, Gödel had at least briefly convinced himself that *CH* is true instead of false. The manuscript establishes, correctly, an equivalence between a particular instantiation of one of the axiom schemas and *CH*.

In an unsent letter to Tarski of 1970, Gödel said about the paper he had sent him that

> Unfortunately my paper, as it stands, is no good. I wrote it in a hurry shortly after I had been ill, had been sleeping very poorly and had been taking drugs impairing the mental functions. [...] My conviction that $2^{\aleph_0} = \aleph_2$ of course has been somewhat shaken. But it still seems plausible to me. One of my reasons is that I don't believe in any kind of irrationality such as, e.g., random sequences in an absolute sense.[199]

Interestingly, $\aleph_2$ is also the value that some set theorists today think is correct (see below).

According to the diaries of Gödel's friend Oskar Morgenstern, Gödel kept working on the problem, apparently again to show that the power of the continuum is $\aleph_2$; in what is presumably the last of Morgenstern's reports, from late 1975, Gödel had become convinced of having a correct proof, according to which the power of the continuum was "*not* $\aleph_2$, but rather 'different from $\aleph_1$.' "[200]

From Gödel's first manuscript a correct proof for a weaker proposition has been reconstructed by Paul Larson, Jörg Brendle and Stevo Todorcevic.[201] They have formulated three new axioms that are implicit in Gödel's original paper that together imply that $2^{\aleph_0} \leq \aleph_2$. The problem is that there is no reason to find these new axioms self-evident and accordingly they have not led to a new credible extension of ZFC.

7 A theme in contemporary set theory

It is a very striking fact that today one begins to encounter the view, by no means unanimously held but nevertheless expressed by a number of set theorists, that the period (in set theory) of proving the consistency of incompatible statements is coming to an end,[202] and that *CH* has been solved, or if not, is once again considered an open problem. Already in 1980, Kreisel noted that the suggestion that *CH* may fail to have a definite truth value for the intended interpretation at all—a suggestion motivated by lack of progress in spite of many attempts from different directions—simply overlooks that there are "infinitely many false starts, perhaps due to a systematic oversight, for any problem."[203] In 2000 Woodin echoed Kreisel's observation:

> There is a tendency to claim that the Continuum Hypothesis is inherently vague and that this is simply the end of the story. But any legitimate claim that *CH* is inherently vague must have a mathematical basis, at the very least a theorem or collection of theorems. My own view is that the independence of *CH* from ZFC, and from ZFC together with large cardinal axioms, does not provide this basis [...] Instead, for me, the independence results for *CH* simply show that *CH* is a difficult problem.[204]

7.1 Generic Absoluteness

The main method for studying the incompleteness phenomenon in set theory is (at present) the forcing method. Forcing does not change the first order arithmetic of integers; the arithmetic of integers is "forcing absolute," i.e., any arithmetic statement true in a model M of ZFC remains true in any generic extension $M[G]$ of M. Strictly speaking this is merely a consequence of the fact that forcing does not introduce new ordinals. As a consequence, no arithmetic statement can decide or be decided by the *CH.* For, suppose some arithmetic sentence ϕ did decide the *CH.* Then if we extend a model of *CH* to one in which *CH* is false, by forcing, then ϕ should no longer hold in the forcing extension, contradicting the forcing absoluteness of ϕ. A similar argument shows that the converse also holds, i.e. neither the *CH* nor its negation has arithmetic consequences.

This is contrary to what Gödel conjectured in the 1964 supplement to Cantor paper, namely:

> The generalized continuum hypothesis, too, can be shown to be sterile for number theory ... whereas for some *other* assumption about the power of $2^{\aleph_\alpha}$ this is perhaps not so.[205]

However Gödel notes the forcing argument in the postscript to that paper:

> Shortly after the completion of the manuscript of the second edition of this paper the question whether Cantor's continuum hypothesis is decidable from the von Neumann-Bernays axioms of set theory (the axiom of choice included) was settled in the negative by Paul J. Cohen ... It turns out that for all $\aleph_\tau$ defined by the usual devices and not excluded by König's theorem ... the equality $2^{\aleph_0} = \aleph_\tau$ is consistent in the weak sense *(i.e. it implies no new number-theoretical theorem).*[206]

The forcing absoluteness of arithmetic statements is perhaps an explanation of what may be called the "empirical completeness" of ZFC as far as arithmetic is concerned; the idea behind this being that incompleteness phenomena are reduced to "residual incompleteness" or incompleteness arising from ad hoc sentences such as the Gödel sentences.[207] That no new, non ad hoc independent arithmetical statements have emerged is evidence for empirical completeness.[208]

CH of course is a different matter. Not only it is not a question about integers, it is even not about reals,[209] but about sets of reals. As forcing determines statements like *CH* in various ways, any attempt to find an extension of ZFC which settles *CH* will have to address the "essential variability in set theory"[210] due to forcing—it must attack forcing head-on, so to speak.

In forcing, the class of formulas taken and the type of forcing are parameters and can be varied. The contemporary notion of generic absoluteness studies the preservation of different classes of formulas under different forcings.

The first formulation of a general generic absoluteness principle for other than arithmetic statements, is due to Jonathan Stavi and Jouko Väänänen and dates from the late 1970's.[211] Note that the idea that any set that can exist, does exist, may turn out to be inconsistent if the meaning of the highly ambiguous "can" is not carefully specified; for example, both the *CH* and its negation state the existence of sets which can exist in a generic extension. Stavi and Väänänen state the following principle: any formula with parameters of hereditary cardinality less than the continuum that can be made true by c.c.c. forcing and that cannot be falsified later by c.c.c. forcing, is already true. The approach was motivated by the study of generalized quantifiers and the idea that the continuum is or should be "as large as possible." One of the main results in that paper is that Martin's axiom is equivalent to a very natural weakening of this. This was discovered independently by Joan Bagaría.[212]

The maximality principle "what can be forced and remains true in further forcing, is true" was rediscovered in 2003 by Joel Hamkins[213] following an idea found in 1999 by C. Chalons, a doctoral student of Boban Velickovic in Paris, whose work has remained unpublished, apart from an electronic announcement.[214]

Finally, some fundamental results pertaining to generic absoluteness, due to Hugh Woodin, are cited below.

7.2 The Woodin Program

The approach associated with the Woodin school seeks to identify the largest possible class of forcing immune statements. In pursuit of this goal they have identified extensions of ZFC which decide a large family of statements, including *CH* itself.[215]

A key result was obtained in 1984, when Woodin proved,[216] based on work of Foreman, Magidor and Shelah,[217] that the first order theory of the structure $(H(\omega_1), \varepsilon)$ is invariant under any kind of forcing, relative to a certain large cardinal assumption, namely the existence of a proper class of so-called Woodin cardinals. Here $H(\omega_1)$ is the set of hereditarily countable sets, i.e. the set of sets which are countable, all elements are countable, all elements of elements are countable, etc. $H(\omega_1)$ can be in effect identified with the set of all real numbers. This effectively lifts the forcing absoluteness of arithmetic statements in the presence of ZFC to forcing absoluteness of the theory of the reals, more exactly of $H(\omega_1)$ in the presence of sufficiently large cardinals.

The next step was to formulate an axiom which would lift this result from $H(\omega_1)$ to $H(\omega_2)$, the set of sets of hereditary cardinality at most ω_1. Thus the new goal is to guarantee the forcing absoluteness of the first order theory of $(H(\omega_2), \varepsilon)$, in the presence of some new axioms. Note that both ω and ω_1 are elements of $H(\omega_2)$ and therefore the theory of $H(\omega_2)$ decides *CH*. Namely, *CH* states the existence of a bijection between ω_1 and all reals. Such a bijection would be an element of $H(\omega_2)$. Thus the existence can be stated as a first order property of $H(\omega_2)$. This observation leads to an important point: no large cardinal axiom can fix the theory of $H(\omega_2)$, as we can always change the value of $2^{\aleph_0}$ without affecting large cardinals. So something different was needed.

This next step was carried out by the following theorem, due to Woodin: The theory of $H(\omega_2)$ is forcing absolute relative to the theory ZFC + an axiom that Woodin calls the '$(\star)$-axiom'.[218] Moreover, the $(\star)$-axiom implies $2^{\aleph_0} = \aleph_2$. Whatever the $(\star)$-axiom is, Woodin proves that the mere existence of an axiom which fixes the theory of $H(\omega_2)$ in (Woodin's Ω-logic) violates the *CH*.

Finally, more recently Woodin has extended his approach to an argument against formalism and the view that the truth or falsity of *CH* has lost its meaning.[219]

It would be beyond the scope of this paper to indicate why Woodin's fundamental results, some of which we have cited, constitute a convincing solution to the continuum problem in the eyes of a number of set theorists. Clearly it is distinguished from other proposed extensions of ZFC, in that Woodin's extensions handle their forcing extensions already, whereas any other kind of canonical extension one may propose must still confront this kind of variability.

Interestingly, a number of different people have obtained results which point to the same conclusion about the value of the continuum. To cite just a few: Velickovic and Todorcevic have shown that the Proper Forcing Axiom PFA implies $2^{\aleph_0} = \aleph_2$.[220] PFA states that if P is a "proper forcing," where the definition of properness is somewhat technical but generalizes both CCC forcings and countably closed forcings, and D is a set of dense open subsets of P with $|D| \leq \aleph_1$, then there is a generic filter on P which meets every dense set in D. Justin Moore has recently shown that a very plausible bounded version of PFA, the so-called Bounded Proper Forcing Axiom BPFA implies the same.[221]

In the opinion of Kennedy, what is particularly interesting about BPFA in addition to its solving the continuum problem in the direction Gödel anticipated, is that it has the form of a reflection principle, principles which were very important to Gödel:

> All the principles for setting up the axioms of set theory should be reducible to Ackermann's principle: The Absolute is unknowable. The strength of this principle increases as we get stronger and stronger systems of set theory. The other principles are only heuristic principles. Hence, the central principle is the reflection principle, which presumably will be understood better as our experience increases. Meanwhile, it helps to separate out more specific principles which either give some additional information or are not yet seen clearly to be derivable from the reflection principle as we understand it now.[222]

In the opinion of Van Atten, on the other hand, reflection principles such as BPFA are not an example of what Gödel is speaking about here, for the following reasons: 1) Gödel here does not speak of reflection principles in the plural, but only about the most general principle 'The Absolute is unknowable'; 2) unlike BPFA, that principle is by its nature not completely formalizable (which makes possible the repeated application that Gödel speaks of here); 3) the general principle has, for a realist, an immediate plausibility on philosophical grounds (as was noted by Cantor, who first used it[223]) that cannot be claimed for BPFA.

It is in any case tempting to infer that the results such as the ones we have cited on PFA and BPFA are results Gödel would have ascribed clear significance to, including them under the category of inductive evidence. But a comprehensive review of recent results in set theory would be needed in order to evaluate whether these results are not anomalous.

There is a straightforward connection between a central aspect of Woodin's approach to *CH* and phenomenology. One of the ideas in Husserl's genetic analysis of judgement (as presented in *Formal and Transcendental Logic*, and in *Experience and Judgement*) is that the kinds of judgements that can legitimately be made in a domain of objects depend on the type of the objects. This leads to the idea that with different domains are associated different logics (which will be extensions of the minimal logic, defined by its applicability to all domains). (Robert Tragesser has developed this idea systematically in his book *Logic and Phenomenology*, mentioned above.) Woodin delineates the domain in which *CH* "lives," so to speak, and then searches for the most appropriate or most specific logic for that domain (more specific than first order logic which (classically speaking) is domain-independent). Thus, Woodin asks:

> Can the theory of the structure $\langle H(\omega_2), \epsilon\rangle$ be finitely axiomatized (over ZFC) in a (reasonable) logic which extends first order logic?[224]

One may even say that a positive answer to this question would be an instance of unfolding the concept of set, or the subconcept of set in the structure $H(\omega_2)$, in the sense that it determines what ways of reasoning (i.e. what logic) certain sets allow for. In other words, unfolding the notion of set need not result only in axioms (of the form "There exists a set with property P") but can also yield principles of (formal) reasoning.

The criticism of Woodin's project given so far[225] indicates that various supporting facts would need to be established in order for the solution to be universally, or perhaps, more, accepted. For example, the Ω-conjecture, a statement to the effect that Woodin's Ω-logic satisfies a natural completeness theorem, has not been proved.

Furthermore, any evaluation of the solution must take into account the privileged role that forcing plays in the construction of models of set theory. The Woodin program rests on the judgement that forcing is the only model construction technique to be considered; thus finding the canonical model for set theory means, principally, finding a reasonable way to "disable" forcing. Empirical completeness suggests that forcing plays this role already for the arithmetic statements. But this is an ambitious plan; in fact another one of Shelah's "Logical Dreams"[226], number 4.4, is: "Find additional methods for independence results (in addition to forcing and large cardinals/consistency) or prove the uniqueness of these methods." Empirical completeness suggests that this holds for the arithmetic statements. As no new independence results have come from other quarters sofar, it is reasonable to conjecture that forcing must be the only phenomenon which introduces variability.

We have suggested that Woodin's work is something of a beginning toward the project of reducing the variability in set theory to so-called residual incompleteness. But are there any convincing arguments for the notion that residual incompleteness really is "residual," that is, not a meaningful phenomenon mathematically?

We saw that Gödel gives an exact criterion for when the question of truth of an axiom A to be added to a theory T loses its meaning; just when $T + A$ and $T + \neg A$ are weak extensions of T, meaning definable in a ground model of T. It is interesting that Gödel does not consider the question whether one obtains strong or weak extensions from taking A to be $con(T)$. And perhaps if the terms weak and strong extension are interpreted loosely, $con(T)$ gives a weak extension, whereas $\neg con(T)$ does not.

That is to say, adding solutions of new Diophantine equations, i.e., elements which witness statements about inconsistency, would give strong extensions. But this means there is an asymmetry present resembling the asymmetry induced when we considered extending the ZFC axioms by inaccessibles. So the question of whether an axiom of this type A is true, where A denotes the consistency of a theory T, is meaningful by Gödel's own criterion, and therefore perhaps not "residual." But this does not seem right. Believing in those particular (weak) extensions, i.e. assuming consistency, must be warranted in any case: consistency is the minimum assumption.[227] The insight that forms the basis of Feferman's analyses of hierarchies of reflection principles is also needed here: belief in the consistency of a set of axioms is supervenient on the belief in the axioms themselves. That is to say, believing in consistency should not commit us epistemologically to any principles beyond the axioms themselves.

This very brief sketch of Woodin's approach together with some of the criticism it has provoked is necessarily only too brief. The interested reader is referred to the literature.[228]

8 Other Developments

Once the independence of all non-trivial statements about the size of $2^{\aleph_0}$ were established by Cohen, attention immediately turned to other cardinals, e.g. to $2^{\aleph_n}$ and even to $2^{\aleph_\omega}$. It turned out that the power sets of regular cardinals like $\aleph_n$ could have any non-trivial cardinality and such statements were also independent from each other[229]. The case of singular cardinals remained a puzzle. For example, if $2^{\aleph_n} = \aleph_{n+1}$ for all n, does it follow that $2^{\aleph_\omega} = \aleph_{\omega+1}$? The Singular Cardinal Hypothesis, SCH, states that if κ is a singular strong limit cardinal then $2^\kappa = \kappa^+$. Ronald Jensen showed in a penetrating study[230] that SCH cannot be decided by forcing over L, or, more exactly, without forcing over models of set theory with large cardinals. Magidor[231] then showed that the independence *can* be established by forcing over models with large cardinals in them. This perhaps vindicates Gödel's idea that large cardinals are needed and can be used to "solve" problems about cardinal arithmetic. Nowadays set-theoretical axioms are known which imply SCH. We mention as an example Chang's Conjecture.[232] $(\aleph_{\alpha+1}, \aleph_\alpha) \to (\aleph_1, \aleph_0)$ A recent result of Matteo Viale[233] shows PFA implies SCH.

9 Concluding remark: Gödel's modernism

Recently, Aki Kanamori has pointed out that after Zermelo had clearly separated set theory from logic, Gödel was the one who was prepared to take the linguistic turn and study uninterpreted formal systems from a set-theoretical point of view.[234] This is one sense in which Gödel can be called modern; here is another:

Gödel was nearing the end of his career when generic absoluteness emerged in the 1970's. On his view, reality fixes the intended model of set theory, and we have access to this mathematical reality by intuition. Results about the limits of formalization and formal systems he will therefore ultimately not interpret as revealing limits to the capability of reason to grasp mathematical reality. We saw a particular example of this viewpoint above, in his reaction to Cohen's independence result. On the other hand, formalizations of mathematics will give formal approximations to reason. In that sense, it would be natural for Gödel to see in generic absoluteness a strong justification of his realism and its correlate on the side of the subject, rationalistic optimism. This is primarily because once incompleteness can be explained away as a residual phenomenon, once the statements that we really care about are decided by a theory we in some sense "like," and once the "essential variability in set theory due to forcing" has been explained, then we are on our way to a sufficiently adequate description of the intended model. These results also say something about the robustness of ZFC: namely, it is, after all, a theory which both captures the intuitions of classical mathematicians about sets, and provides a domain for deciding high order questions about sets, in spite of what the incompleteness phenomena may have led people, in the beginning, to believe.

Once again, what is missing in the generic absoluteness approach we have considered here, is, as Shelah points out, to show the uniqueness of forcing. Still, we hoped to draw attention to this sea change in set theory; point out that things are moving along the lines that Gödel anticipated they would. His iron belief in the decidability of the Continuum Hypothesis, radical as it at times seemed, may have been vindicated—by the set theorists.

Acknowledgements. We are grateful to the staff of the Department of Rare Books and Special Collections at the Firestone Library of Princeton University, and to Marcia Tucker of the Library of the Institute for Advanced Study, for facilitating our research, and overall for ensuring such a pleasant stay in the archive; also, we are much obliged to the Institute for Advanced Study which kindly granted permission to quote from Gödel's *Nachlaß*. Many thanks to Piet Hut of the Institute for Advanced Study for inviting van Atten to participate in the Program in Interdisciplinary Studies which he directs, to Peter Goddard, the Director of the Institute, for hosting Kennedy, and to the NWO for their support during the final writing of the manuscript.

We wish to thank Joan Bagaría, Leon Horsten, Aki Kanamori, Georg Kreisel, Paolo Mancosu, Göran Sundholm, Boban Velickovic and most especially Jouko Väänänen, for helpful conversation and correspondence, as well as for drawing our attention to some of the sources.

We also thank two anonymous referees for their comments.

Kennedy presented an earlier version of this paper at the workshop "Logicism, Intuitionism and Formalism: What has become of them?", Uppsala, Sweden, August 27–29, 2004. We thank the organizers for the invitation, and the audience for their questions and comments.

Notes

[1]All of Gödel's published papers and selections from his unpublished papers and correspondence have been published in five volumes: K. Gödel, *Collected Works*, eds. S. Feferman et al. (Oxford: Oxford University Press). *I: Publications 1929–1936* (1986); *II: Publications 1938–1974* (1990); *III: Unpublished essays and lectures* (1995); *IV: Correspondence A-G* (2003); *V: Correspondence H-Z* (2003). We will refer to these volumes as *CW I*, etc. The Brown University lecture is in *CW III:175–185.*

[2]*CH* is the hypothesis that $2^{\aleph_0} = \aleph_1$. *GCH* is its generalization $2^{\aleph_\alpha} = \aleph_{\alpha+1}$. Georg Cantor had first stated a weaker form of *CH* in 1878 and then *CH* in 1883. *GCH* was formulated by Felix Hausdorff in 1908. For the history of *(G)CH*, see in particular J. Dauben, *Georg Cantor. His Mathematics and Philosophy of the Infinite* (Princeton: Princeton University Press, 1990), M. Hallett, *Cantorian Set Theory and Limitation of Size* (Oxford: Clarendon Press, 1984), and the introductions to the relevant papers in K. Gödel, *Collected Works. II: Publications 1938–1974*, eds. S. Feferman et al. (Oxford: Oxford University Press, 1990) and K. Gödel, *Collected Works. III: Unpublished essays and lectures*, eds. S. Feferman et al. (Oxford: Oxford University Press, 1995). We will henceforth refer to these volumes by their title and number only, e.g., *Collected Works II.* We will sometimes refer to papers by their year, following the system of *CW*; the year of unpublished papers is preceded by an asterisk, there is a letter suffix in case different papers appeared in the same year (e.g., **1940a*), and there is a question mark if the year is not certain. For more general histories of set theory, see A. Kanamori, "The Mathematical Development of Set Theory from Cantor to Cohen," *The Bulletin of Symbolic Logic* 2(1) (1996), pp.1–71, and J. Ferreiros Dominguez, *Labyrinth of Thought: A History of Set Theory and Its Role in Modern Mathematics* (Boston: Birkhäuser, 1999).

[3]ZFC is Zermelo-Fraenkel set theory with the Axiom of Choice.

[4]*Kurt Gödel. Wahrheit und Beweisbarkeit. Band 1: Dokumente und historische Analysen.*, eds. E. Köhler et al. (Wien: öbv & hpt, 2002), pp. 127–128; translation from the German ours.

[5]According to *CW I*, p. 41, the Brown lecture was on November 15, not 14.

[6]By their characteristic logic based on their characteristic notion of truth, according to intuitionists there are no absolutely undecidable propositions. For assume that ϕ is absolutely undecidable. Then in particular the assumption that ϕ has been proved must lead to a contradiction. But if it does, this, on the intuitionistic conception of negation, is to say that $\neg\phi$ holds, which decides ϕ and thereby contradicts the assumption. But by the same idiosyncracies of intuitionistic logic, this little argument does not show that therefore every proposition is decidable. In fact intuitionists consider that as highly unlikely,

given that on their interpretation that would mean that one knows a universal method to decide all mathematical propositions.

[7] *CW II*, pp. 269-70

[8] In an explicitly typed system, adding higher types means that, besides variables for individuals (e.g. numbers, or sets), there will also be variables for sets of individuals, sets of sets of individuals, and so on, together with appropriate axioms that govern formation of sets of these types. In the cumulative hierarchy of sets, adding higher types is just adding more levels to the hierarchy.

[9] *CW III*, p. 341n.20.

[10] See the remarkable reflections on the topic by Emil Post in the appendix to a manuscript from 1941 titled "Absolutely Unsolvable Problems and Relatively Undecidable Propositions," first published in *The Undecidable*, ed. M. Davis (Hewlett NY: Raven Press, 1965), pp. 340–433; the appendix starts on p. 41.

[11] Through the device of contextual definitions, statements may arise that are not meant to be taken at face value, yet are equivalent to ones that should. Gödel **1940a*, *CW III*, p. 176 simply defines meaningfulness as being part of mathematics proper (or translatable into mathematics proper). The relation to practice that we will be concerned with is not thematized there.

[12] *CW II*, p. 257 and p. 256.

[13] *CW II*, p. 257 and p. 256.

[14] *CW III*, p. 35.

[15] See *CW I*, p. 180, footnote 48a.

[16] T. Skolem, "Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre", in *Selected Works in Logic*, ed. J.E. Fenstad (Oslo: Universitetsforlaget, 1970), p. 149n2.

[17] D. Hilbert, "Über das Unendliche," *Mathematische Annalen* 95 (1926), pp. 161–190.

[18] *CW II*, p. 157. The references are to A. Fraenkel, *Einleitung in die Mengenlehre*, 3rd, revised edition (Berlin: Springer, 1923) and Luzin's talk at the Bologna conference, "Sur les voies de la théorie des ensembles", *Atti del Congresso Internazionale dei Matematici, Bologna 3–10 settembre 1928* (Bologna:Zanichelli), I, pp. 295–299.

[19] We thank Paolo Mancosu for this last detail.

[20] In A. Tarski, *Collected Papers. Vol.1: 1921–1934*, eds. S.R. Givant and R.N. McKenzie (Boston: Birkhäuser, 1986), pp. 233–241. We thank Göran

Sundholm for bringing this passage to our attention.

[21]Incidentally, Gödel in 1931 does refer to Hilbert's paper from 1926, but in a different context. See footnote 48a on p. 180/181 of *CW I*.

[22]H. Wang, *Popular Lectures on Mathematical Logic* (New York: Dover, 1993), p. 128. "Jetzt, Mengenlehre," Gödel is alleged to have said around that time ("And now, [on to] set theory"): J. Dawson, *Logical Dilemmas. The Life and Work of Kurt Gödel* (Wellesley: A K Peters, 1997), p. 108n18. Gödel's interest in set theory may have begun to develop as early as 1928 when he requested at the library the volume containing Skolem's talk in Helsinki (mentioned at the beginning of this section), as Dawson notes (p. 120). It is worth adding to this that very likely Gödel did not actually get the volume: the library slip in question has a large question mark at the title of the book, the stamp "Ausleihe" is missing, and the smaller part has not, as would have been usual in case of loan, been detached. This library slip therefore does not constitute evidence that Gödel had seen that book at that time (For a more comprehensive discussion of the archive material and how this corroborates Gödel's statements about his completeness theorem and Skolem, see M. van Atten, "On Gödel's awareness of Skolem's Helsinki lecture" *History and Philosophy of Logic*, 26(4) (2005) pp. 321–326.). Be that as it may, Dawson goes on to mention that in 1930 Gödel requested from libraries various works on (or related to) set theory: Hilbert's list of open problems from 1900 (published in the *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen* of the same year)—of which the first is *CH*—, Fraenkel's *Einleitung in die Mengenlehre* from 1919—in which Gödel will have noticed Fraenkel's skepticism about Hilbert's attempted proof of *CH*—, the text of Skolem's Helsinki lecture again, and von Neumann's papers "Über die Definition durch transfinite Induktion und verwandte Fragen der allgemeine Mengenlehre" and "Die Axiomatisierung der Mengenlehre", both from 1928.

[23]p. 610 of G. Kreisel, "Review of K. Gödel's Collected works, Vol. II," *Notre Dame Journal of Formal Logic* 31(4) (1990) pp. 602–641.

[24]P. 195 of G. Kreisel, "Kurt Gödel. 28 April 1906–14 January 1978," *Biographical Memoirs of Fellows of the Royal Society* (1980), pp. 149–224.

[25]For a discussion of Gödel's ideas on impredicativity (and of a number of other topics), see W. Tait, "Gödel's Unpublished Papers on Foundations of Mathematics," *Philosophia Mathematica* 9 (2001), pp. 87–126.

[26]For finite sets it makes of course no difference whether one uses this notion of definability or the classical power set as the operation to generate higher levels; one can always simply define a finite set by enumerating its elements. For the infinite case this may well be different (and is in general believed indeed to be.)

[27] Wang, *Popular Lectures*, p.129.

[28] For this paragraph, see Wang, *Popular Lectures*, p. 129, and also Gödel's remarks in H. Wang, *A Logical Journey. From Gödel to Philosophy* (Cambridge: MIT Press, 1996), p. 251: 8.1.7, 8.1.8, 8.1.9.

[29] *CW III*, p. 178.

[30] P. 158 of G. Kreisel, "Gödel's Excursions into Intuitionistic Logic," in *Gödel Remembered. Salzburg 10-12 July 1983*, eds. P. Weingartner and L. Schmetterer (Napoli: Bibliopolis, 1987), pp. 65–179

[31] K. Menger, "Memories of Kurt Gödel," in his *Reminiscences of the Vienna Circle and the Mathematical Colloquium*, eds. L. Golland, B. McGuinness and A. Sklar (Dordrecht: Kluwer, 1994), p. 214.

[32] He wrote in his notebook "Kont.Hyp. im wesentlichen gefunden in der Nacht zum 14 und 15 Juni 1937' (essentially found the [consistency proof of the generalized] Continuum Hypothesis during the night of 14 to 15 Juni 1937) *CW I*, p. 36, note s.

[33] *CW IV*, pp. 112–115.

[34] *CW II*, pp. 26–27.

[35] *CW III*, p. 155.

[36] P.J. Cohen, "The Independence of the Continuum Hypothesis. I," *Proceedings of the National Academy of Sciences, U.S.A.* 50, pp. 1143–1148.

[37] J.C. Shepherdson, "Inner Models for Set Theory III," *The Journal of Symbolic Logic* 18(2), pp. 145–167. The existence of a minimal model was rediscovered by Cohen in 1963 *CW IV*, p. 376.

[38] *CW II*, pp. 26–27.

[39] See also *1946* in *CW II*, p. 151.

[40] H. Wang, *From Mathematics to Philosophy* (London: Routledge and Kegan Paul, 1974), p.204. Also Wang, *Logical Journey*, p.263.

[41] *CW II*, p. 260n20. D.S. Scott, "Measurable Cardinals and Constructible Sets," *Bulletin de l'Académie polonaise des sciences, série des sciences mathématiques, astronomiques, et physiques* 9 (1961), pp. 521–524. S. Ulam, "Zur Masstheorie in der algemeinen Mengenlehre," *Fundamenta Mathematicae* 16 (1930), pp. 140–150.

[42] *CW V*, p. 273; original emphasis.

[43] *CW III*, **1939b*, pp. 126–155.

[44] *CW III*, p. 155.

[45] *CW II*, pp. 266–267.

[46] *CW III*, p. 184.

[47] *CW II*, p. 129.

[48] One suggestion, by John Dawson in *CW III*, p. 163, is that it was prepared with the aim of being presented to the September 1940 International Congress of Mathematicians. That meeting was cancelled due to the outbreak of World War II, and in any case, Gödel never delivered the lecture.

[49] *CW III*, p. 165.

[50] *CW III*, p. 164.

[51] *CW III*, p. 175.

[52] *CW III*, p. 164.

[53] Our footnote: In "Über die Unabhängigkeit der Kontinuumhypothese", *Dialectica* 23 (1969), pp. 66–78, Paul Finsler argues that undecidability of *CH* is a phenomenon that only presents itself in the context of a strict axiomatization, for the "formal continuum." For criticism of this proposal, see e.g. Bernays' remarks on that paper in "Zum Symposium über die Grundlagen der Mathematik", *Dialectica* 25 (1971), pp. 171–195.

[54] *CW III*, p. 164.

[55] Parsons, " Platonism and Mathematical Intuition", p.67.

[56] P. 50 of C. Parsons, "Platonism and Mathematical Intuition in Kurt Gödel's thought," *The Bulletin of Symbolic Logic* 1(1) (1995), pp. 44–74.

[57] *CW III*, p. 185.

[58] *CW III*, p. 176.

[59] $V = L$ can be violated "high up" by adding a generic, hence non-contructible, subset to a large regular cardinal, e.g., $\beth_(\omega + 1)$. The forcing notion used has closure properties which imply that no new subsets of hereditary cardinality $\beth_\omega^+$ are added. Thus all Π_n^m statements are preserved by this forcing for all m and n. This technique was used e.g. in W.B. Easton, "Powers of regular cardinals," *Annals of Mathematical Logic* 1 (1970), pp. 139–178.

[60] E.g., F. Drake, *Set Theory. An Introduction to Large Cardinals* (Amsterdam: North-Holland, 1974), p. 164.

[61] See *CW III*, p. 163.

[62] *CW III*, p. 175.

[63] *CW III*, p. 163 and *CW II*, p. 158, respectively.

[64] Kreisel, "Gödel's Excursions," p. 158.

[65] *CW II*, p. 27.

[66] *CW III*, p. 163.

[67] E.g. *CW II*, p. 81.

[68] *CW II*, p. 184n22. On Gödel's use of the term "axiom" in the Cantor papers, see also H. Wang, *Reflections on Kurt Gödel* (Cambridge: MIT Press, 1988), pp. 205,294. On p. 205, Wang offers an alternative solution to ours: '[W]hat I see as the main point in this episode is an additional flexibility (besides the allowance for new axioms to be discovered) implicit in G's concept of the axiomatic method: What is thought to be an axiom at one time may later turn out to be a false proposition and, therefore, not really an axiom'. But for the reason we have given, we do not think that Gödel held $V = L$ true in the first place.

[69] *CW III*, p. 133.

[70] *CW II*, p. 186.

[71] The judgements that Gödel goes on to make are not universally shared. See Hallett, *Cantorian Set Theory*, p. 111 for some notes of dissent and for further references.

[72] *CW II*, p. 185–186.

[73] Menger's '1938' on p. 220 of his *Reminiscences* must be a slip of the pen.

[74] Menger, *Reminiscences*, p. 222.

[75] Menger, *Reminiscences*, p. 210; original emphasis.

[76] *CW II*, p. 182.

[77] *CW II*, p. 260n20.

[78] Parsons, *Platonism and mathematical intuition*, p.50.

[79] *CW II*, p. 179.

[80] This is to be distinguished from yet another notion of incompleteness, described in R. Carnap, "Die Antinomien und die Unvollständigkeit der Mathematik," *Monatshefte für Mathematik und Physik* 41 (1934), pp. 263–284: certain metamathematical notions can be defined in a given system, but others (notably, truth) cannot. The difference with conceptual incompleteness is that such in-

definability results do not arise because of insufficient analysis of the concepts used in the sytem. Gödel's review of Carnap's paper is in *CW I*, p. 389.

[81]From the editorial note on p. 428 of *CW II*, we gather that notebook XIV was used from April 1946 till 1955.

[82]*CW III*, pp. 433–435.

[83]*CW II*, p. 183.

[84]*CW II*, p. 181.

[85]Kreisel, *Review*, p. 607.

[86]Wang, *From Mathematics to Philosophy*, p. 196.

[87]*CW III*, p. 379.

[88]*CW II*, p. 181.

[89]*CW II*, p. 180.

[90]Wang, *Logical Journey*, p. 238.

[91]*CW II*, p. 181.

[92]κ is Mahlo if $\{lambda < \kappa | \lambda$ is inaccessible$\}$ is stationary.

[93]*CW II*, p. 182.

[94]In the 1964 version, Gödel amended this sentence to read "but also that it can be supplemented without arbitrariness by new axioms which only unfold the content of the concept of set explained above," which perhaps reflects the study of phenomenology that he had begun in between the two versions of the Cantor paper; in the German phenomenological argot, 'Entfaltung', or the more or less synonymous 'Auslegung', 'Explikation', 'Auseinanderlegung', 'Explizierung' are very common terms. Note that Gödel had spoken of "explicating the content of the general concept of set" already in version III of his paper on Carnap, *CW III*, p. 353n43.

[95]*CW II*, p. 150–153.

[96]*CW II*, p. 151.

[97]*CW II*, p. 140–141.

[98]Wang, *Reflections*, p. 173.

[99]Wang, *Reflections*, p. 174.

[100]Tarski's paper (and some additional material) has been transcibed, edited

and introduced by H. Sinaceur, "Address at the Princeton University Bicentennial Conference on Problems of Mathematics (December 17–19, 1946), by Alfred Tarski," *The Bulletin of Symbolic Logic* 6(1) (2000), pp. 1–44. Without hazarding an explanation, we note the following difference between Gödel's talk and a report on a discussion session at the same meeting. In the talk, Gödel says that '[I]t has some plausibility that all things conceivable by us are denumerable, even if you disregard the question of expressibility in some language' (*CW II*, p. 152); in the report, Gödel is said to propose to allow uncountably many primitive notions in a formal system, and is then quoted as saying "I do not feel sure that the set of all things of which we can think is denumerable" (Sinaceur, "Address," p. 37). Technically, one can of course admit this extension of the notion of formal system, but the quoted remark evidently adds an epistemological concern.

[101]Related material is in Tarski's letter to Gödel of Dec.10, 1946 (one week before the conference), *CW V*, pp. 271–273, and in Feferman's remarks on p. 264.

[102]Sinaceur, "Address," pp. 33–34.

[103]*CW III*, p. 310.

[104]*CW III*, p. 341n.20.

[105]*CW III*, p. 310. Are Diophantine sentences of type Π^0_2 elementary? From the point of view of strong and interesting theories like $I\Delta_0 + \Omega_1$ the Π^0_2-theory of the natural numbers looks very powerful; highly intractable.

[106]*CW III*, p. 323.

[107]Parsons, "Platonism and Mathematical Intuition', p. 52; see also note 16 on the same page for some remarks on Gödel's realism in the light of Dummett's conception of realism.

[108]In the discussion at the Princeton conference in 1946, Tarski noticed that idealistic philosophy is congenial to some of Gödel's views on mathematics (Sinaceur, "Address," p. 34). Menger reports that Gödel in the thirties "studied a great deal of philosophy including post-Kantian German idealist metaphysics" (Menger, *Reminiscences*, p. 209). See M. van Atten and J. Kennedy, "On the Philosophical Development of Kurt Gödel," *The Bulletin of Symbolic Logic* 9(4) (2003), pp. 425–476, for further discussion of Gödel's interest in idealism.

[109]The remainder of this paragraph, until and including the two quotations from Gödel, is extracted from p. 460 of Van Atten and Kennedy, "On the Philosophical Development." For the full text of the draft letter, as well as a draft letter from Gödel to Tillich on the same subject, see M. van Atten, "Two draft letters from Gödel on self-knowledge of reason", *Philosophia Mathematica* 14(2), 2006, pp.255–261.

[110] *CW III*, p. 310.

[111] After "obj[ect]," Gödel at first wrote "matter" but then crossed it out.

[112] Gödel Nachlaß (GN), Firestone Library, Princeton, item 020514.7.

[113] Gödel first wrote "thin[king]," and then crossed it out.

[114] D. Bergamini and the editors of Life, *Mathematics* (New York: Time, 1963) p. 53.

[115] Our emphasis; Wang, *Logical Journey*, pp. 186–187.

[116] Wang, *From Mathematics to Philosophy*, pp. 324–325. See also Wang, *Logical Journey*, p. 316, item 3, and p. 317. For a connection to Kant here, see Boolos' remark in *CW III*, p. 294.

[117] Wang, *Logical Journey*, p. 151, 4.4.18.

[118] Wang, *Logical Journey*, p. 317, 9.4.21.

[119] *CW II*, p. 267.

[120] *CW II*, p. 267.

[121] See G. Kreisel. "Informal rigour and completeness proofs". In I. Lakatos, editor, *Problems in the philosophy of mathematics* (Amsterdam: North-Holland, 1967), pp. 138–186.

[122] For more about Gödel's criterion, see also Kennedy's "Gödel and Meaning," in preparation.

[123] See Van Atten and Kennedy, *On the Philosophical Development*, section 6.3.

[124] *CW II*, p. 260. The two references Gödel makes are to a description of the iterative concept of set and a remark that the existence of a set does not presuppose that it is definable in a finite number of words.

[125] R. Tragesser, *Phenomenology and Logic* (Ithaca: Cornell University Press, 1977), p. 22.

[126] For this paragraph, see Van Atten and Kennedy, *On the Philosophical Development*.

[127] Original emphasis; Tragesser, *Phenomenology and Logic*, pp. 23–24.

[128] Tragesser, *Phenomenology and Logic*, 17.

[129] *CW II*, p. 269.

[130]It is not excluded that in 1947 Gödel was playing with ideas prefiguring the notion of intuition known from the 1964 paper. Charles Parsons, ("Platonism and mathematical intuition," p. 57) has drawn attention to allusions to perception of concepts in the paper on Russell from 1944 and in the Gibbs lecture from 1951. We agree with Parsons that "it is reasonable to conjecture that although [Gödel at that time] was not yet ready to defend his notion of intuition he already had some such conception in mind" (p. 57).

[131]*CW II*, p. 268.

[132]"Platonism and Mathematical Intuition."

[133]"Gödel and the Intuition of Concepts," *Synthese* 133(3) (2002), pp. 363-391.

[134]*Phenomenology and Logic.*

[135]"On the Philosophical Development," section 6.2.

[136]E. Husserl, *Logische Untersuchungen. Zweiter Band, 2. Teil* Husserliana vol. XIX/2 (Den Haag: Martinus Nijhoff, 1984).

[137]Wang, *Logical Journey*, p. 80.

[138]See also Husserl, e.g. section 24 of Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie. Erstes Buch Husserliana vol. III/1 (Den Haag: Martinus Nijhoff, 1976).

[139]*CW II*, p. 268. Gödel's page references are to the original publication, which correspond to pp. 260–261 in *CW II.*

[140]*CW II*, p. 166

[141]*CW II*, pp. 268–9.

[142]Wang, *Logical Journey*, p. 243.

[143]Wang (*Reflections*, p. 292) writes that according to Gödel, certain sections in this paper, together with his contribution to Wang's book *From Mathematics to Philosophy*, "constitute the principal statement of his mathematical realism."

[144]For other examples of this maneuver, see *CW III*, p. 345n45, p. 356 lines 3–5, p. 361 line -14; also *Logical Journey*, p. 239, 7.4.7.

[145]*CW III*, p. 338n12. Examples of such explanations by intuitionists as Gödel is thinking of can be found in L.E.J. Brouwer, *Collected works. I: Philosophy and Foundations of Mathematics*, ed. A. Heyting (Amsterdam: North-Holland, 1975), p. 423 and p. 511.

[146]*CW III*, pp. 257–258.

[147]On idealizations in intuitionism, see M. van Atten, "Intuitionistic Remarks on Husserl's Analysis of Finite Number in the Philosophy of Arithmetic," *Graduate Faculty Philosophy Journal*, 25(2) (2004), pp. 205–225.

[148]*CW II*, pp. 268–9.

[149]Wang, *Logical Journey*, p. 302, 9.2.35.

[150]The edition that Gödel read with Wang is in E. Husserl, *Phenomenology and the Crisis of Philosophy*, trl. Q. Lauer (New York: Harper & Row, 1965), p. 112.

[151]Wang, *For Perspicuous Objectivity. Discussion with Gödel and Wittgenstein*, (unpublished manuscript), p.67.

[152]An extensive account by Husserl on how knowledge of an object is in general founded on both intuitive and non-intuitive presentations of that object is in his sixth *Logische Untersuchung* in E. Husserl, *Logische Untersuchungen. Zweiter Band, 2. Teil* Husserliana vol. XIX/2 (Den Haag: Martinus Nijhoff, 1984). See also Husserl's considerations on adequacy and apodicticity of knowledge at the beginning of his *Cartesianische Meditationen* (a book Gödel valued). E. Husserl, Cartesianische Meditationen und Pariser Vorträge Husserliana vol. I (Den Haag: Martinus Nijhoff, 1950).

[153]E. Husserl, *Phenomenology and the Crisis of Philosophy*, p. 113.

[154]Wang, *Logical Journey*, p. 302, 9.2.35.

[155]*CW III*, p. 346n32.

[156]Wang, *Logical Journey*, p. 217, 7.1.11.

[157]Wang, *Logical Journey*, p. 217, 7.1.10.

[158]Also compare Heyting's passage to Gödel's lecture from 1933, "The present situation in the foundations of mathematics," *CW III*, **1933o*, p. 51. Gödel presented it to a meeting of the Mathematical Association of America in Cambridge, MA, on December 30, 1933 (which meeting Heyting did not attend).

[159]A. Heyting, *Spanningen in de wiskunde* inaugural lecture, University of Amsterdam, 1940 (Groningen: Noordhoff, 1949), p. 13; translation ours.

[160]Wang, *Logical Journey*, p. 216, 7.1.9.

[161]*CW III*, **1961/?*, p. 383.

[162]GN, item 040311, p. 12.

[163]*CW II*, p. 261.

[164]E. Husserl, Erfahrung und Urteil (Hamburg: Meiner 1985).

[165]Wang, *Logical Journey*, p. 301.

[166]Wang, *From Mathematics to Philosophy*, p. 207.

[167]*CW II*, p. 261. Note that the large cardinal axioms do not decide "lower down" statements like the *CH*, as, roughly speaking, large cardinal properties are preserved under Cohen extensions, which of course change the value of the continuum. See D.A. Martin, "Hilbert's First Problem: The Continuum Hypothesis," *Proceedings of the Symposia in Pure Mathematics*, vol. 28 (1976).

[168]See the sections on maximality and generic absoluteness, below.

[169]E.g. Husserl, *Ideen I*, section 44; *Cartesianische Meditationen*, section 19; *Erfahrung und Urteil*.

[170]Wang, *Logical Journey*, p. 244.

[171]*CW II*, p. 267.

[172]Note that Gödel's page reference is to the original publication, whose page numbers are indicated in the margins of the Collected Works. He introduces the argument from success which we quoted at the beginning of this section only on page 265. Yet there is clearly a pragmatic aspect to this second argument for inaccessibles. In Wang's chapter "The Concept of Set," which was written in close collaboration with Gödel, this is made explicit (*From Mathematics to Philosophy*, p. 203).

[173]Wang, *Logical Journey*, p. 263, 8.3.13.

[174]Kreisel, "Kurt Gödel," p. 197.

[175]*CW II*, p. 269.

[176]footnote Gödel: "On the basis of the concept of set explained on p. ," where he will have meant to refer to his explanation of the iterative conception; in the Collected Works, this is on pp. 258–259.

[177]GN, item 040311, p. 12.

[178]GN: item 013183.5, dated March 11, 1976: a list of quotations from Gödel, corrected by himself.

[179]R. Laver, "The left distributive law and the freeness of an algebra of elementary embeddings," *Advances in Mathematics* 91(2) (1992), pp. 209–231.

[180]P. Dehornoy, "Elementary embeddings and algebra." Invited chapter in the forthcoming *Handbook of Set Theory*, eds. M. Foreman, A. Kanamori, and M. Magidor.

[181]D.A. Martin, "Borel Determinacy," *Annals of Mathematics* 102 (1975), pp. 363-371.

[182]D.A. Martin, "Measurable Cardinals and Analytic Games," *Fundamenta Mathematicae* 66 (1969/1970), pp. 287–291.

[183]We note also here Jeff Paris's result that Pi^0_4-determinacy is provable in ZFC. Paris, J. B. "ZF $\vdash \sum^0_4$ determinateness," *The Journal of Symbolic Logic* 37 (1972), pp. 661–667.

[184]The further, aesthetic issue, having to do with simplicity, we do not comment on here. The axiom asserting the existence of a measurable was simply fruitful in yielding a powerful result about Borel games.

[185]Y. Moschavakis, *Descriptive Set Theory* (Amsterdam: North-Holland 1980), p. 357.

[186]*CW IV*, p. 378.

[187]Van Atten and Kennedy, "On the Philosophical Development of Kurt Gödel," p. 470.

[188]*CW IV*, p. 372. Gödel's page reference is to the text of Church's talk "Paul J. Cohen and the Continuum Problem," published in 1968 in the Proceedings of the International Congress of Mathematicians (Moscow-1966), pp. 15–20; p. 8 of the manuscript corresponds to p. 18 of the publication.

[189]Wang, *Logical Journey*, p. 180.

[190]*CW III*, p. 113.

[191]This may explain why Gödel could, as we saw, in 1938 say that a successful completion of Hilbert's program would have been "of enormous epistemological value," and in (or around) 1947 decline an invitation to write on Hilbert's program by saying that he was not sufficiently sympathetic to it (*CW II*, p. 144).

[192]GN, item 012279; original emphasis.

[193]*CW II*, pp. 262–263.

[194]S. Ulam, "John von Neumann, 1903–1957," *Bulletin of the American Mathematical Society* 64(3) (1958), part 2, pp. 1–49, as quoted in *CW II*, p. 168; original emphasis. Note that this is different from the very similar passage in *CW V*:295.

[195]Wang, *Logical Journey*, p. 144.

[196]For the details, see *CW II*, pp. 173–175 and *CW III*, pp. 405–420.

[197]*CW IV*, pp. 383–384

[198] *CW III*, p. 423.

[199] *CW III*, p. 424.

[200] For this paragraph, see Dawson, *Logical Dilemmas*, pp. 235–236.

[201] J. Brendle, P. Larson, and S. Todorcevic, "Rectangular Axioms, Perfect Set Properties and Decomposition." Preprint at `http://www.math.toronto.edu/larson/goedel.pdf`

[202] Typical for that period is that Kunen's classic textbook *Set Theory* from 1980 has the subtitle "An introduction to independence proofs."

[203] Kreisel, "Kurt Gödel," p. 212.

[204] W.H. Woodin, "The Continuum Hypothesis" to appear in the *Proceedings of the Logic Colloquium 2000*, Paris.

[205] Italics the authors'.

[206] Italics the authors'.

[207] A different view has been expressed by Paul Cohen, who has stated the view that any interesting statement about primes will eventually be shown to be independent of ZFC.

[208] But see Harvey Friedman's work on structurally simple combinatorial statements which are equivalent to large cardinals.

[209] As a matter of interest in this connection it is now known that there are two models of set theory with the same ordinals, cardinals and reals, one satisfying *CH* and the other not. Folklore, communicated by Matt Foreman.

[210] P. Dehornoy, "Recent Progress on the Continuum Hypothesis (After Woodin)." `http://www.math.unicaen.fr/\~dehornoy/surveys.html`

[211] J. Stavi and J. Väänänen, "Reflection principles for the continuum." In *Logic and Algebra*, ed. Yi Zhang, Contemporary Mathematics, vol. 302 (AMS 2002), pp. 59-84.

[212] J. Bagaría, "A Characterization of Martin's Axiom in Terms of Absoluteness," *The Journal of Symbolic Logic* 62(2) (1997), pp. 366-372.

[213] J.D. Hamkins, "A Simple Maximality Principle," *The Journal of Symbolic Logic* 68(2) (2003), pp. 527–550.

[214] C. Chalons, "An axiom schema," circulated email announcement (1999).

[215] W.H. Woodin, "The Continuum Hypothesis (I and II)," *Notices of the American Mathematical Society* 48(6) (2001), pp. 567-576, and 48(7) (2001), pp. 681-690.

[216]See W. Hugh Woodin, "The axiom of determinacy, forcing axioms and the non-stationary ideal," Walter de Gruyter and co., Berlin, 1999.

[217]M. Foreman, M. Magidor, and S. Shelah "Martin's Maximum, Saturated Ideals, and Nonregular Ultrafilters", *Annals of Mathematics* 127(1) (1988), pp. 1-47.

[218]W. Hugh Woodin, "The Continuum Hypothesis". I: *Notices of the American Mathematical Society* 48(6) (2001), pp. 567–576. II: *Notices of the American Mathematical Society* 48(7) (2001), pp. 681–690. Correction in 49(1) (2002), p.46.

[219]See Coxeter Lectures, Fields Institute, Toronto, November 2002 at `http://www.fields.utoronto.ca`

[220]B. Velickovic, "Forcing Axioms and Stationary Sets," *Advances in Mathematics* 94(2) (1992), pp. 256–284.

[221]"Set Mapping Reflection", *Journal of Mathematical Logic* 5(1) (2005), pp. 87-98; erratum in 5(2), p. 299.

[222]Wang, *A Logical Journey*, p. 283, 8.7.9.

[223]See note 2 in G. Cantor, 'Ueber unendliche, lineare Punktmannichfaltigkeiten. Nr.5', p. 205 in the reprint in his *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts* (E. Zermelo, ed.), Berlin: Springer, 1932.

[224]Woodin, "The Continuum Hypothesis. II," p. 682.

[225]See Matthew Foreman's "Has the continuum hypothesis been settled?", in Stoltenberg-Hansen, Viggo and Väänänen, Jouko (eds), *Logic colloquium '03. Proceedings of the annual European summer meeting of the Association for Symbolic Logic (ASL), Helsinki, Finland, August 14–20, 2003.* Wellesley, MA: A K Peters; Urbana, IL: Association for Symbolic Logic (ASL). Lecture Notes in Logic 24 (2006), pp. 56–75.

[226]S. Shelah, "Logical Dreams," *Bulletin of the American Mathematical Society* 40 (2003), pp. 203-228.

[227]S. Feferman, in "Reflecting on Incompleteness," *The Journal of Symbolic Logic* 56(1) (1991), pp. 1-49, gives a precise method of measuring the "cost" of accepting consistency statements of increasing strength, in terms of reflection principles.

[228]See e.g. Patrick Dehornoy, Progres récents sur l'hypothèse du continu (d'après Woodin); Seminaire Bourbaki, exposé 915, mars 2003. An English version is available at `http://www.math.unicaen.fr/$\sim$dehornoy/surveys.html`

[229]Easton, William B. "Powers of regular cardinals." *Annals of Mathematical Logic* 1 (1970), pp. 139–178

[230]Devlin, Keith I.; Jensen, R. B. Marginalia to a theorem of Silver. ⊨ISILC Logic Conference (Proc. Internat. Summer Inst. and Logic Colloq., Kiel, 1974), pp. 115–142. Lecture Notes in Math., Vol. 499, Springer, Berlin, 1975.

[231]Magidor, Menachem. "On the singular cardinals problem. II." *Annals of Mathematics*, 2nd series, 106(3) (1977), pp. 517–547.

[232]Shelah: Cardinal Arithmetic, Oxford logic Guides, vol. 29, Oxford University Press, 1994.

[233] *The Journal of Symbolic Logic* 71(2) (2006), pp. 473–479.

[234]P. 453 of A. Kanamori, "Zermelo and set theory," *The Bulletin of Symbolic Logic* 10(4) (2004), pp. 487-553.