About Logic and Logicians

A palimpsest of essays by Georg Kreisel
Selected and arranged by Piergiorgio Odifreddi

Volume 2: Mathematics

Lógica no Avião

ABOUT LOGIC AND LOGICIANS

A palimpsest of essays by

Georg KREISEL

Selected and arranged by

Piergiorgio ODIFREDDI

Volume II. Mathematics

Editorial Board

Piergiorgio Odifreddi, About Logic and Logicians, Volume 2. Brasília: Lógica no Avião, 2019.

Série A, Volume 2

I.S.B.N. 978-65-900390-1-9

Prefixo Editorial 900390

Obra publicada com o apoio do PPGFIL/UnB.

UnB

Editor's Preface

These books are a first version of Odifreddi's collection of Kreisel's expository papers, which together constitute an extensive, scholarly account of the philosophical and mathematical development of many of the most important figures of modern logic; some of those papers are published here for the first time.

Odifreddi and Kreisel worked together on these books for several years, and they are the product of long discussions. They finally decided that they would collect those essays of a more expository nature, such as the biographical memoirs of the fellows of the Royal Society (of which Kreisel himself was a member) and other related works. Also included are lecture notes that Kreisel distributed in his classes, such as the first essay printed here, which is on the philosophy of mathematics and geometry.

Kreisel himself wrote all the texts, but Odifreddi has made some substantial editorial interventions, rearranging some of the material, breaking the text into sections and paragraphs, inserting titles, moving or removing some notes, and eliminating some digressions. These interventions were made in order to give the essays some of their original freshness and linearity, qualities that were lost in later versions.

Some other minor modifications were made here and there, consisting basically in the correction of a small number of erroneous references in the original manuscripts. Kreisel's expository works are invaluable to logicians, and we hope that the reader may find the present edition to his advantage, even if there is still some editorial work to do.

Rodrigo Freire.
Brasília, April 2019.

Contents

Part I

ON BROUWER

Chapter 1

Brouwer's Foundations

To understand Brouwer's particular brand of so-called *constructive foundations* one must compare its merits and defects with those of other better known versions (including, incidentally, the bulk of Brouwer's early writings on constructive foundations).

To put first things first: Brouwer's final version is incomparably *more imaginative*. The commonplace versions are preoccupied with the business of 'pure' existence theorems $\exists x A(x)$ and the search for 'explicit' realizations t such that $A(t)$. For the silent majority of mathematicians this business is hardly dramatic: there is nothing to stop one from presenting such t even if one does not reject pure existence theorems. What is more, mathematics has developed a whole arsenal of notions for stating significant differences between such t (much more pertinent than the crude idea of an 'explicit' t,[1] or the crude distinction between 'constructive' and 'nonconstructive' definitions of t): for example, if $\exists x A(x)$ expresses the existence of a zero of a polynomial of odd degree, the continuity of t in the coefficients. And once the attention of mathematicians is drawn to such notions, their relevance is plain without any foundational preoccupation.

Secondly, commonplace constructive foundations constitute a *restriction*, and thus form a proper part of ordinary mathematics (usually accompanied by grand,

[0]Originally published in *Bulletin of the Americal Mathematical Society*, 83 (1977) 86–93, as 'Brouwer's Collected Works, Volume I'.

[1]A happy coincidence shows the *appreciation* by the Mathematical Establishment of significant 'explicit' realizations. Without much exaggeration: a Fields Medal was awarded in 1958 (to Roth) for the 'pure' existence theorem

$$\forall n \exists q_0 \forall p \forall q (q > q_0 \to |\sqrt[3]{2} - \frac{p}{q}| > q^{-2-1/n}),$$

and another one in 1970 (to Baker) for the 'worse' result

$$\exists q_0 \forall p \forall q (q > q_0 \to |\sqrt[3]{2} - \frac{p}{q}| > q^{-3+0.05})$$

where, however, a (manageable) value for q_0 was supplied. So much for blind prejudice against an appropriate search for explicit realizations.

but dubious foundations (cl)aims, to which we return later on). Brouwer's version of constructive foundations is instead *incomparable* with ordinary mathematics. On the one hand it does not contain higher set theory with the (transfinite) iteration of the power set operation applied to infinite sets. On the other it includes as principal objects of mathematical study:

1. *choice sequences* of various kinds, for example, (the idealization of) the random sequences of throws of a die

2. *proofs* which enter into a new meaning of the familiar logical operations, the new meaning being used to state laws or 'axioms' concerning 1.

In contrast, ordinary mathematics which (of course) *uses* proofs as tools, does not make them explicit objects of study, and *paraphrases* properties of random sequences (for example, in measure-theoretic or set-theoretic terms).

Intuitionistic notions

Brouwer himself did not give full fledged axiomatic theories of choice sequences, let alone of proofs. But it may fairly be said that modern axiomatic theories, especially of the former, stand in much the same relation to Brouwer's writings, as modern axiomatic theories of sets stand to Cantor's.[2]

While it would be quite inappropriate to go here into the details of such axiomatic theories,[3] it seems worthwhile (and easy!) to present the general idea.

Quite naively: *freely chosen* sequences s (say, of natural numbers) are thought of as (necessarily) 'incomplete', only finite initial segments being 'given'; so *all*

[2] The next remarks show that the parallel between (theories of) sets and choice sequences goes further [∞] (the mark "[∞]" indicates that more precise references will be presented in future editions of this work).

The (currently) most successful theories do not treat the *most general notions* involved, but rather the cumulative hierarchy (of those sets which are generated from $\emptyset$ by iterating the power set operation) on the one hand, and lawless sequences on the other. And there is no evidence that, even if there are such things as 'the most general' notions, (of set or choice sequence), they would lend themselves to a rewarding theory.

In fact, our knowledge of the cumulative hierarchy and of lawless sequences is (at the present stage) most effective when applied to *other notions* defined in terms of those things: we know more about the so-called constructible sets than about the full cumulative hierarchy (used to define them), and have more applications of so-called projections of lawless sequences than of the latter.

Incidentally, there is a little-known overlap in the interests of Cantor and Brouwer, the founders of the theories of sets and choice sequences. Ever since 1877, Dedekind and Cantor speculated that regions of different dimension are not in one-one *bicontinuous* correspondence (only, Cantor's attempted proof [1879] was defective). Brouwer thought of euclidean spaces as including all points given by freely chosen sequences (considered immediately below), and proved that regions of different dimensions are in one-one correspondence, bicontinuity now being a consequence of the (new) conception of euclidean space.

[3] Readable accounts are in Troelstra [1977] and Troelstra and Van Dalen [1988].

operations on such s must be continuous for the product topology. More generally, if P is an arbitrary predicate of s we expect

$$P(s) \to \exists n \forall s' [(\forall m \leq n)(s(m) = s'(m)) \to P(s')], \tag{1.1}$$

inasmuch as $P(s)$ can only 'depend' on a finite segment of s.

Evidently, in 1.1 the logical particles cannot have the usual truth functional meaning, since the following instance of the law of the excluded middle

$$\exists n[s(n) = 0] \vee \neg \exists n[s(n) = 0]$$

fails for any s such that, for all 'given' initial segment, $s(n) \neq 0$.[4] There is nothing dramatic about all this: in ordinary reasoning we relatively rarely use the truth functional meaning (such as 'q is true or p is false' for 'p implies q'). The latter meaning happens to have a particularly simple theory. Brouwer indicated (and Heyting developed) another meaning of the logical particles, well adapted for an analysis of 1.1 but perhaps even further removed from most ordinary reasoning than the truth functional meaning (and satisfying different formal laws, as shown above).

The new meaning is well illustrated by the case of implication. First of all, the *data determining a proposition* p are not simply truth values (this would obviously be inappropriate for $P(s)$ when s is 'incomplete'): instead, we have a condition C_p determining what are *proofs* of p (not: whether or not there is a proof of p). Then $C_{p \to q}$ is built up from C_p and C_q as follows: by definition, we require an operation I, and an argument π_0 establishing

$$\text{for any } \pi,\ C_p(\pi) \Rightarrow C_q(I(\pi)) \tag{1.2}$$

where π consists, hereditarily, of operations and arguments.[5] The reader can guess the corresponding explanations for other particles.

Clearly, 1.2 must be expected to be quite sensitive to the domain Π of the π: by restricting Π, one restricts the domain on which I must satisfy 1.2, and thus increases the possibilities of proving $p \to q$; but one also restricts the permitted range of I, and thus decreases those possibilities. Because of that sensitivity, no one set of formal logical laws can be expected to be 'fundamental' (in contrast to

[4] $\exists n[s(n) = 0]$ fails by choice of s. And $\neg \exists n[s(n) \neq 0]$ cannot hold otherwise, by 1.1, it would also hold for *all* sequences s' agreeing with a certain finite initial segment of s (while there are such sequences s' such that, for some n, $s'(n) = 0$).

[5] There is an obvious (though not necessarily vicious!) circularity here, unless $\Rightarrow$ in 1.2 is different in 'kind' from $\to$. The usual idea is that the conditions C_p are decidable, and so $\Rightarrow$ has simply its truth functional meaning. For coherence, this then requires that it be decidable, for any pair (I, π_0), whether or not π_0 establishes 1.2 for variable π. All this would not only be pretentious, but genuinely dubious if we were realistically thinking of arbitrary proofs and operations. But it makes good sense when applied to a wide range of proofs and operations, hereditarily formalizable in various 'logic-free' systems.

ordinary first order logic).[6] As a so-to-speak positive counterpart, the particular formal laws first stated by Heyting for the intuitionistic meaning apply also to situations only quite vaguely related to constructions in the literal sense (meant by Brouwer); for example, they apply to particular 'explicit' definitions in axiomatic set theory involved in (weak) forcing, or to certain 'uniform' definitions in category theory applied to sheaves or Cartesian closed categories. It is fair to say that at least the elementary exposition of such 'constructions' benefitted from experience with formal intuitionistic logic.

So much then for the extension of ordinary mathematics by detailed systematic developments of specifically intuitionistic notions. In my opinion they have substance and some mathematical wit. But it can hardly be claimed that they (ought to) have a central place in the 'mainstream' of mathematics. As hinted at the beginning, the vague general ideas which preceded those developments have been 'absorbed' in ordinary mathematics (without any logic-chopping). Those ideas certainly fired the imagination of mathematicians like Poincaré and the young Brouwer.[7] These two constructivists are associated with the switch to algebraic topology operating on finite 'pieces' from set theoretic topology in the style of Schoenflies.[8] But also more modern developments in ordinary mathematics are related to vague, general preoccupations of constructivists, for example, how objects are 'given' to us: elementary category theory points out the consequences of 'giving' a function by its graph together with a bound on its range (even though the exact range is determined by the graph, the passage involved may require an operation not in the category considered). Bishop's book [1967] illustrates this state of affairs very well (in effect if not by intention): leaving aside the introduction, the style is perfectly familiar to the modern mathematician.

Brouwer's foundational critique

As everybody knows, Brouwer's own case for his logical work had little to do with *extending* our ordinary (view on our knowledge of) mathematics, but with *correcting* it. There is widespread misunderstanding concerning his specific critique:

[6]The sensitivity of 1.2 to the choice of domain Π should be compared to the sensitivity of (ordinary) *second order logic* to the class C of sets involved in the (set theoretically explained) meaning of logical formulas. It was a discovery that the validity of first order formulas (which is of course also defined set theoretically) is remarkably insensitive to C: once the set of natural numbers and so-called Δ^0_2 subsets are included in C, the validity of first order formulas is stable.

Without any evidence to the contrary, the patent sensitivity of validity in the case of intuitionistic logic suggests that the latter cannot be expected to be often useful (in intuitionistic mathematics). Without going into the particular consideration above, Brouwer was certainly skeptical of the role of logic (tacitly, in his kind of mathematics).

[7]Contrary to an almost universal misunderstanding, Brouwer's work in topology was *preceded* by his interests in constructivity (for example, in his dissertation [1907]), and *followed* by his work in choice sequences.

[8]For an interesting account, see Newman [1969].

1. He by no means confined himself to *finite objects*: in fact, he introduced (in connection with his proposed notion of arbitrary operation on choice sequences) the idea of 'fully analyzed', possibly *infinite* proofs.

2. He was not tempted by hackneyed generalized doubts about the legitimacy of *abstract notions*: he studied proofs (that is, thoughts), which he distinguished emphatically from the linguistic objects used to represent them (for example, but not necessarily, formal derivations of some specific formal system).

3. Though he saw defects in the theory of sets (tacitly: as presented in the first decade of this century), the *antinomies* were not particularly prominent in his critique.[9]

What Brouwer did do in his foundational critique was really absolutely orthodox (at least since Kant, and particularly in the first quarter of this century). According to Brouwer, our ordinary view neglects the *role of the subject* (called 'observer' in physical contexts[10]). A natural and, in the short run, effective reaction to the neglect of anything is to make it the sole object of study. Brouwer's version of constructive foundations is an instance of this: as described at the end

[9] It is perhaps natural that the antinomies are often used (in effect if not by intention) to introduce a bit of drama into foundations, a subject by and large devoted to the undramatic business of 'analyzing' what (we believe) we know anyway. But it is simply historically false to think that the antinomies provide evidence for any failure of the 'logical intuitions' of Cantor, let alone of his contemporaries (who were oversuspicious of his notions). Here are the facts.

Back in [1885], in a review (of Frege's *Grundlagen*) easily accessible in Cantor's *Collected works*, he objected to Frege's:

$$\exists x \forall y [y \in x \leftrightarrow P(y)] \tag{1.3}$$

for the precise reason that precautions are needed to ensure that the predicate P has an extension which can be comprehended (as a 'unity'). Frege himself, in the introduction to the *Gundgesetze*, discussed the possibility that 1.3 might be (not only false for the intended meaning, but) formally contradictory. His own malaise is apparent from the thoughtless 'evidence' proposed there for 1.3, namely its wonderful consequences (which, at best, provide a reason for our interest in 1.3, certainly not for its validity or consistency).

As I read Brouwer's diatribes against set theory, for example, in his dissertation, the main source of his 'gut reaction' seems to have been less the topic of (infinite) sets than the extraordinarily pretentious claim for set-theoretic foundations in Russell's *Principles of mathematics*, as providing the true analysis of all mathematical concepts (and that the formal deductions in axiomatic set theory analyze all mathematical reasoning). Brouwer surely had a point. Though even today, set theory is better known as a *general framework* for mathematics than as a *branch* of mathematics, the value of this or any other 'general framework' is dubious: axioms are given in the first chapter of a text, but hardly ever has one occasion to refer to them later (in any detail).

[10] It goes without saying that this stress on the subject acquiring knowledge got a boost from Einstein's singularly successful use of the observer in his special theory of relativity, shortly before Brouwer's dissertation.

of the previous subsection, he made *proofs* (that is, the activity of the subject, also called *creative subject* by Brouwer and *ideal mathematician* by others, as in 'ideal fluid'), part of the meaning of mathematical assertions.

Although similar 'subjectivist' analyses of scientific knowledge have been proposed in the philosophical literature for all sorts of other sciences, nothing in that literature seems as imaginative as the notions Brouwer introduced in his attempt at a purely subjectivist analysis of mathematics. If the latter has not gone very far, this is surely partly due to inherent weaknesses in the scheme itself; but probably even more because he was preoccupied with hackneyed traditional questions such as:

> Is mathematics about an external reality or about our own 'free' constructions?[11]

More specifically, Brouwer remained hung up on the validity of (principles of) proofs, neglecting more 'structural' relations between proofs, and between proofs and other things.

If the reference above to the business about external reality and our own constructions (discovery and invention) appears irreverent, the reader should stop to give second thoughts to the favorite implications attributed to this matter, for example, concerning *certainty* (of mathematical knowledge): we are supposed to be peculiarly certain of our own (mental, presumably not necessarily also of our physical) productions. Is the idea of all possible proofs of any one theorem (or all possible definitions of any one object, say the empty set) as clear, let alone clearer than the idea of the collection of all subsets of say ω?

A second favorite is the would-be dramatic *conflict* between external reality and our free constructions. Getting knowledge of any (external) reality requires activity or constructions on the part of the subject; the bit about their being 'free' is particularly unconvincing since they are certainly not made by consciously arbitrary choices, no more so than constructions of material tools (which are limited by the properties of the material available, quite apart from the intended purpose). Besides, why expect a conflict between our own possibilities and the external reality in which we have evolved?[12]

[11] Those questions are so banal that we ask and understand them when (ontogenetically or phylogenetically speaking) we know next to nothing (about mathematics); reflecting only, as somebody said, *wie sich der kleine Moritz die Dinge und das Denken vorstellt* (on Simple Simon's ideas about things and thoughts). Of course, for this very reason these questions have a perennial pedagogic interest (at least, for the untamed spirits among us).

[12] To avoid misunderstanding: all this pretentiousness does not discredit, by itself, all foundational questions *like* those that excited Brouwer. Early speculations (which used to be considered philosophical) on the question:

What is matter (made of)?

were pretentious too, and spiced with such 'conflicts' as: 'Matter is atomic' versus 'All is flux'. Even using hindsight we would be hard put to find the 'conflict' in what Born called: the

The famous dispute between Brouwer and Hilbert

To put first things first, Brouwer and Hilbert were in the same camp, accepting only constructive principles as *prima facie* legitimate. The difference was elsewhere. Brouwer's principal concern was to *develop constructive mathematics*, without the distraction of studying metamathematically (by constructive means) the principles of ordinary mathematics. Hilbert wanted to *justify ordinary mathematics* by:

1. using formalizations $\mathcal{F}$ of valid principles for then-current mathematical concepts

2. proving the consistency of $\mathcal{F}$ (tacitly, constructively).

He was convinced that 2 did not need any development of constructive mathematics because he thought that so-called *finitist* methods (of which proofs in 'elementary number theory' are typical) would be enough.[13]

Brouwer was in any case dubious about the *adequacy of any formalizations* (without, however, getting anywhere near the precise incompleteness results established by Gödel). Why then all the fuss about a consistency proof for (a

restless universe of atoms.

Incidentally, it hardly seems an accident that the great interest in (set theoretic or constructive) foundations in the first quarter of this century coincided with the huge success of the atomic theory: if the physical world can be built up from a few basic elements, why not mathematical concepts (Whitehead and Russell) or proofs from a few basic intuitions (Brouwer)? But mathematical foundational schemes lack some of the most obviously essential features of modern atomic theory (not: of early generalities about atoms).

1. The basic foundational elements, such as sets, are really quite close to objects of ordinary mathematical experience: do they even *look* fundamental enough for analyzing in any depth the great diversity of mathematics?

2. Where are the analogues to geometric relations and binding forces between atoms so essential for refining (crude chemical) atomic theory?

3. On the 'phenomenological' level, foundations lack the analogues to such prerequisites of the atomic theory as the isolation of chemically pure substances, let alone the periodic table.

One wonders whether our experience of mathematics is at a comparable stage to that of physics and chemistry which was, patently, needed for progress on the structure of matter.

[13]It is to be stressed that the *general* conclusions are independent of any precise analysis of the notion of finitist proof. Besides, there is no evidence for any particular *reliability* of finitist methods, Hilbert's principal claim for them (tacitly, at the present time; of course, 100 years ago mathematicians had to treat nonfinitist, logically compound expressions quite gingerly, such as the negation of uniform convergence!) On the contrary, inasmuch as nonfinitist proofs are often simpler than finitist ones (of the same theorem), and the nonfinitist *principles* equally reliable, the actual probability of error in a finitist proof is likely to be higher. This is borne out by the literature on finitist consistency proofs, which contains remarkably many oversights. No other compelling virtue of finitist methods has turned up either (except the fact that they were one of the first that occurred to us).

necessarily) incomplete $\mathcal{F}$, if tomorrow we can think of stronger principles $\mathcal{F}^+$ which are also valid? More importantly, Brouwer objected strongly to *consistency as a sufficient condition* on $\mathcal{F}$, quite apart from the secondary matter of the methods used in a consistency proof. And he surely had a point.

What *does* consistency of $\mathcal{F}$ insure (for $\mathcal{F}$ of the kind considered in Gödel's incompleteness theorem[14])? Only the truth of (formally derived) *purely universal* arithmetical theorems. Suppose A is a purely universal proposition: typically (by Matyasevic [1970]) $(\forall\vec{x})(p(\vec{x}) \neq 0)$, where p is a polynomial in the variables $\vec{x}$ with integral coefficients. If (the translation in $\mathcal{F}$ of) A is derivable in (or even only formally independent of!) $\mathcal{F}$ then the diophantine equation $p(\vec{x}) = 0$ has no solutions: if it had, a counterexample $\vec{n}$ to A could be computed, and so $p(\vec{n}) = 0$ and hence $\neg A$ would be formally derivable in $\mathcal{F}$. But *this is all*, in the following precise sense. If such an A is not derivable in (a necessarily consistent) $\mathcal{F}$[15] then $\mathcal{F} \cup \{\neg A\}$ is also consistent, though, as we have just seen, $\neg A$ is false. Thus consistency by itself does not even insure the truth of (formally derived) *purely existential* arithmetic theorems.

Hilbert's pious rhetoric, as saviour of classical analysis against Brouwer's Bolshevic revolution, has a hollow ring. All that is 'saved' is (as Hilbert put it) a formal game $\mathcal{F}$ (where $\mathcal{F}$ is one of the formalizations of analysis developed in the first quarter of the century). Except for purely universal propositions, by no means the whole content of (the ordinary interpretation of) ordinary analysis, the latter is not 'saved' by the consistency of $\mathcal{F}$.

Actually Brouwer's attack, by way of 'contradictions' with ordinary mathematics, was hardly disturbed since he got them by changing the meaning of the logical operations and the domain of variables (for example, by replacing sequences in the sense of ordinary mathematics by choice sequences). Naturally, all supplemented by grand foundational doubts about our ordinary notions (doubts which, by the end of the previous subsection, are generally more dubious than the notions themselves).

Ironically, if (after recognizing the inadequacy of the consistency *criterion*) one actually looks at consistency *proofs* one finds that, properly formulated, they do 'save' a remarkable amount of ordinary mathematics (for use in constructive mathematics). In fact, progress over the last 40 years allows a precise formulation of the issue whether the methods developed in work on Hilbert's consistency program or those developed from Brouwer's ideas on choice sequences are more effective for this purpose.

[14]Conditions on such systems require, in particular, that numerical computations can be mimicked in $\mathcal{F}$. So if a diophantine equation $p(\vec{x}) = 0$ has a solution $\vec{n}$, $p(\vec{n}) = 0$ can be verified by computation.

[15]That this can happen is the content of Gödel's incompleteness theorem

Mysticism

Brouwer's first publication [1905] was titled *Life, art and mysticism.*[16] Curiously, Brouwer never points out the relevance of some of the mystical business to the 'main stream' of foundations, which *assumes* that we must be capable of making the grounds for our knowledge conscious to ourselves, and that this would be rewarding to boot. In sober terms, the 'mystical' alternative would be that we have a lot of knowledge, also in mathematics, which is simply more convincing than any proposed analysis (of our actual grounds for this knowledge, let alone of possible grounds).

[16]Excerpts from it have been translated in Brouwer [1975]. The editor of the latter volume has given reasons for omitting others, for example (on p. 565): 'In many places Brouwer runs on inconsiderately, for instance, on the position of women in society.' Some of the omissions are translated in van Stigt's dissertation: for example, one expressing Brouwer's view that every woman is more like a lioness than a twin is like his brother.

Chapter 2

Luitzen Brouwer

This chapter is divided into two parts. Section 1 describes Brouwer's life and that part of his work which develops a general philosophy of mathematics; since the latter is regarded as obscure, but also as important, its broad outlines are presented quite fully and without technicalities. Section 2 contains more precise accounts of Brouwer's contributions to logic.

2.1 General Development

Life, family, general interests

Brouwer was born on 27 February 1881 at Overschie, near Rotterdam. His father Egbert was village schoolmaster and lived to be 90. His mother's maiden name was Hendrika Poutsma. Two of his brothers made reputable careers in education, one as a teacher of French in a secondary school, the other as a university professor of geology; both at Amsterdam. Brouwer himself wrote poems all his life; incidentally, his first publication ([1905]) was partly an anthology of (other people's) poetry. On 31 August 1904 he married Elisabeth de Holl; the marriage brought him a stepdaughter, Miss Pijper, who remained the only child, and the income of some chemist shops. He died on 2 December 1966 in a car accident at Blaricum near the house where he had lived for many years. His wife had died about five years before him.

Even into his seventies Brouwer travelled a great deal, visiting universities where he lectured frequently and sometimes several hours at a time, but not persuasively. On his travels he often stayed with various mathematicians and logicians. By general agreement he was an accomplished conversationalist in a small circle, when he was sure of a sympathetic audience. He had several of the (rewarding) qualities of a professional entertainer in society: a large fund of

[0]Originally published in the *Biographical Memoirs of Fellows of the Royal Society*, 15 (1969) 39–68, as part I and III of 'Luitzen Egbertus Jan Brouwer'.

anecdotes and miscellaneous information, a desire to perform well and, of course, practice.

Education and academic career

Brouwer attended primary schools at Medemblik and Hoorn, and a secondary school at Haarlem. He studied mathematics and science at the University of Amsterdam, and was in close contact with the philosopher Mannoury, for whom he retained a lifelong affection (see Brouwer [1947]). Brouwer's dissertation was accepted by the applied mathematician Korteweg, who worked on surface waves and, in collaboration with van der Waals, on molecular forces. The dissertation did not contain any results, but rather Brouwer's views on what was important in mathematics, and polemics against then current set theoretic foundations. In short, the work was 'immature' for someone of his age, but full of drive. Hindsight shows that, in Brouwer's case, the style of the work could also be interpreted as the outward sign of an exceptionally prolonged and vigorous intellectual development, which was soon to bear fruit.

From 1909 to 1912 Brouwer was *privaat-docent*. In 1912 he was elected to the Chair for set theory, function theory and axiomatics at the University of Amsterdam, which he held till 1951. In view of Brouwer's critical attitude to set theory and axiomatics, the choice of name seems odd; but perhaps in those days one thought of topology as an axiomatic theory and of point set topology as part of set theory. Also in 1912 Brouwer was elected Member of the Dutch Royal Academy of Science. He was knighted in 1932 (*Ridder in de Nederlandse Leeuw*), and elected a Foreign Member of the Royal Society in 1948.

Constructivity

This was a lifelong interest dating from his thesis. A thorough understanding of the matter requires the sort of experience in mathematics which will be assumed in Section 2, but it is possible to get a satisfactory general idea without knowing more than Aristotle: the issue goes back to the Greeks.

> Is mathematics our own construction and about such constructions, or does mathematics discover truths about an abstract reality external to ourselves?

Much is made of certain obvious ambiguities in the issue. But the real difficulty, as would be expected of old rival alternatives, is that each of them is consistent with familiar facts (at least as long as one does not look at them too closely). Brouwer found a place where the issue becomes manageable, namely in the *analysis of the usual logical operations*. The considerations are so simple that they are not affected by the ambiguities mentioned. What, from a constructive point of view, do we have to know to assert a proposition P? We must give a proof. And to assert the negation of P, we require a refutation of P: it is not enough that, so to speak as a matter of historical accident, we have not so far found a proof of P. Brouwer considered the *law of the excluded middle*, which states that either P or non P, for all propositions P. Now for Brouwer's interpretation of the logical operations we should have reason to assert this law in full generality only if, here and now, we either know how to prove any given P or how to refute P. Brouwer had no difficulty in showing that there was no convincing reason for this assertion. Though for a long time he had no refutation either, he spoke of the 'unreliability' of this law already in [1908]. However, much later he obtained a refutation (for *his* interpretation of the logical operations) after the theory of constructivity was developed to include propositions about sufficiently 'sophisticated' concepts.

Brouwer himself did not stress the fact that a new interpretation of the logical operations was involved. He rejected the more usual 'truth functional' interpretation because he rejected the conception of mathematics which suggested the latter. This is the critical, so to speak, 'negative' and best-known side of his work. In this chapter the positive side will be stressed: the extent to which a coherent and sophisticated, if problematic, area of mathematics can be developed from the notions of *proof* and *rule* (construction). Brouwer's own doctrinaire presentation was not only philosophically dubious but practically unsuccessful because it did not really convey his views. It is, however, highly likely that, at an early stage, his own work benefited greatly from two very usual consequences of any doctrinaire position: he was able to develop his ideas vigorously, first because he had put out of his mind all but the matter in hand; and, second, because weaknesses of a position are less 'disturbing' if (one thinks) there is no alternative. As we shall see, things were different twenty years later.

Topology

Soon after his thesis, Brouwer published an outstanding series of papers in topology, a subject which deals with the most basic properties of geometric figures or, as the later Brouwer would have had to say, of visual perception. This work is not described here,[1] but it is pertinent to note a link with constructivity. Brouwer, like most of his contemporaries, used as a tool approximations by geometric figures which are determined by a finite number of points: polygons in the plane, polyhedra in space. But, unlike most of them (except for example Poincaré, another constructivist), his use of those figures was very elementary ('algebraic', as we should say now). In his style of work one never thinks of a line or surface as made up of an infinite number of points, in contrast to point set topology.

Inevitably one wonders about the heuristic value of Brouwer's (or Poincaré's) general logical ideas for his topology and, on a higher level, about the value of his

[1] See Newman [1969] (originally published as part II of the present chapter) for an account.

general philosophical ideas for his logic.[2] Now his logical ideas (which he published several years before his topological work) were not only novel, but almost detailed enough to deduce rigorously some of his topological innovations from them, though he himself had not done so explicitly. In contrast his philosophical views, though quite consistent with his logic,[3] were undeveloped and, literally, commonplace: the observer, language and linguistic conventions had a big role in the epistemology of the time. Of course it can happen that general philosophical views need only be formulated to be useful. But this is doubtful in the case of the epistemological ideas above, when we compare Einstein's fruitful use of them in relativity theory with Poincaré's sterile conventionalism in geometry.

Intuitionism and formalism

Brouwer's *intuitionistic* mathematics (as he called it) treats propositions, rules and proofs as objects of thought, more or less as they present themselves to us naïvely. These objects are represented (for instance, for the purpose of communication) by finite concrete signs such as words or symbols.

The *formalist* conception of mathematics holds that all mathematically significant relations involving proofs and propositions can be adequately analyzed in terms of their representations; of course, the properties of signs to be used here must be external or formal, they must refer only to mechanical manipulations of signs, not to their meaning. In short, formalism is a particular mechanistic theory of reasoning. Formalist reasoning about manipulations with finite strings of symbols has the same elementary combinatorial character as that used in topological manipulations of finite configurations of points.

Quite generally, most familiar mathematics which is constructive at all, turns out to be combinatorial: it does not use constructions on abstract, typically intuitionistic objects such as proofs or rules. For this reason it is practically easy enough to recognize elementary arguments as such when we see them. But people differ on what is essential to this elementary character: the words *combinatorial* and *finitist* correspond to the two most important views on the matter.[4]

Development of constructive mathematics

Even during the period of his topological discoveries, Brouwer thought about a constructive reworking of mathematics and reported some results in [1913]. In accordance with the last subsection, Brouwer's specific logical ideas do not

[2] In a few philosophical papers on knowledge and its origins Brouwer elaborated on the primacy of the will over the intellect.

[3] His stress on the will corresponds to his interest in constructions, in what we do ourselves.

[4] Of course, reasoning *about* manipulations with strings of symbols must be distinguished from the manipulations themselves (which do not involve mathematical reasoning, properly speaking, at all); at best such formal manipulations are concrete representations of reasoning.

yet come into play and his results can be quite adequately stated in perfectly straightforward terms: he decides alternatives which had been previously left undecided, or gives additional conditions which allow a decision.

The publications of 1918–1928 changed all this. For currently popular ideas, Brouwer's most striking innovation was his conception of proofs as *infinite* objects. The idea is not far fetched, particularly if the statements proved are about infinitely many instances, for instance in justifying the principle of induction in arithmetic. On the contrary, the opposite (that is, the formalist conception mentioned above) is 'daring', in that it assumes that the basic properties of our (mathematical) thoughts can be intelligibly analyzed in terms of the signs we use to represent these thoughts. But Brouwer went beyond mere plausibility and discovered an area where his ideas of infinite proofs led (him) to striking formal laws: the analysis of constructive operations on *arbitrary* sequences of natural numbers. Of course, there are such operations, for instance the rule which associates with any sequence $a_1, a_2, a_3 \ldots$ its first element a_1. Brouwer's aim was to analyze *all* possibilities of proving constructively that an operation is well-defined on arbitrary sequences. He came up with his *bar theorem*, and a corollary which he liked to formulate as a startling 'theorem': all functions of real numbers in the unit interval are uniformly continuous. By use of propositions about arbitrary sequences he was also able to refute the law of the excluded middle mentioned earlier. More details, in particular a sober reformulation, will be found in Section 2.

Brouwer's controversy with Hilbert

About the turn of the century Hilbert, Brouwer's senior by twenty years, began (in [1904]) his pursuit of formalist foundations. He formulated his famous programme stating adequacy conditions on such a foundation (see Section 2). The well-known controversy between Hilbert and Brouwer was, contrary to a widespread misunderstanding, an internal schism among constructivists, since Hilbert too adopted a constructive view of mathematical reasoning; in fact his was stricter than Brouwer's, since he wanted to reduce mathematics to principles that were not only constructive but *finitist*.

One difference between Hilbert and Brouwer was that Brouwer certainly regarded the finitist restriction as unnecessary for constructive foundations and, probably, as difficult to formulate in a precise and convincing fashion. But the principal difference was in their attitude towards ordinary, nonconstructive mathematical practice. Brouwer proposed to ignore this practice except at best as an heuristic guide. Hilbert, in his programme, proposed to establish its coherence by means of his finitist methods and, generally, use it as a tool for getting finitist results.

Hilbert's ideas had great appeal for mathematicians and attracted the active collaboration of several gifted people, especially in the twenties. One reason was the truly childlike simplicity of Hilbert's programme, another was his straightforward and direct presentation. Brouwer, as we saw above, certainly succeeded in startling the mathematical world, by alienation in the jargon of the theatre. In this connexion he quoted often, and with evident approval, George Bernard Shaw's advice to the effect that one has to exaggerate to make an impression. (A pedant might have pointed out that Shaw never promised a favorable impression).

But, objectively, it seems fair to say that Brouwer was right on almost every major issue on which he disagreed with Hilbert.

The half-century mark

Just as Brouwer was turning fifty, Heyting [1930] and Gödel [1931] appeared which, though they did not, objectively, create new problems for Brouwer's work in logic, must have been equally disturbing; one in an obvious, the other in a subtler way.

Heyting [1930] is best known for an elegant set of *formal laws* which are valid for Brouwer's constructive interpretation of the logical operations. Outsiders were excited by these simple laws which suggested algebraic and other mathematical manipulations, without too much thought whether such work produced misunderstanding or understanding of Brouwer's intentions.[5] Heyting himself warned against the possibilities of such misunderstandings in the clearest possible terms. But he did much more: he formulated the intended interpretation in the natural way, in terms of the concepts of *proof* and *rule*. Now, the formal laws are not *immediately* evident if one restricts oneself to some simple list of proofs and rules such as those represented in familiar formal systems. What is needed are the general abstract concepts of (constructive) proof and of rule, including, for instance, rules which associate proofs to proofs or to rules. Brouwer himself had never set out anything like Heyting's analysis. Though one may have been uncomfortable about some of Brouwer's informal explanations, this malaise was, psychologically, far less alarming than seeing in full detail what difficult abstract notions are involved in the very meaning of the logical operations. Brouwer's doctrinaire polemics against alternative foundational schemes had a hollow ring.

Gödel's famous work [1931] on the *incompleteness* of formal systems established, roughly speaking, the inadequacy of Hilbert's formalist foundations; or, more precisely, it refuted some of Hilbert's specific assumptions behind the programme. Since Brouwer had denied these assumptions, one might simplemindedly expect him to have been gratified by Gödel's results. But both the nature of Gödel's argument and Brouwer's conduct during the decade following Gödel's paper give one reason to doubt this simpleminded interpretation.

[5] Heyting's laws concerned only elementary logic, a small part of intuitionistic mathematics; like any other severely limited set of facts these laws admit a great number of quite distinct interpretations, some of which are totally removed from Brouwer's ideas.

Gödel's immensely natural proofs did not need anything like the heavy distinctions which Brouwer thought essential to his development of constructive mathematics (nor, incidentally, the ingenious constructions that Hilbert had introduced into proof theory). What they did involve was a clear perception of aims and principles, a philosophical, nondoctrinaire analysis, alien to the logical activists of the time. Seeing the incomparable superiority of this kind of analysis Brouwer had to face the question, consciously or unconsciously, to what extent he had even begun to master his own logical ideas. Actually, the same question may well have paralysed the formalist camp too, who did not draw the natural consequence from Gödel's work either: all they had to do was to reformulate Hilbert' programme more carefully in the light of Gödel's result to open up new lines of research.

Be that as it may, it is a fact that Brouwer himself devoted the ten years after publication of the papers by Heyting and Gödel to nonscientific activities, although, objectively, these papers concerned essential points of his life's work, and although he was only fifty and in good health. It does not seem unreasonable to suppose that he had received an intellectual shock.

Perhaps a biographical study of this remarkable man will bring light on this matter, when the details of his life in the thirties have been sorted out. The anecdotes about this period are fragmentary and even contradictory; but Brouwer's views on Life expressed in [1905] may help to interpret this scattered information. Since, by general agreement, Brouwer was self-willed and impulsive, the views of his youth are probably a valid guide to his later personality; just as the views in his dissertation [1907], another 'immature' work, are certainly a good guide to his later scientific interests.

Solipsism

In his sixties Brouwer took up a side of the constructive philosophy of mathematics which had not previously played any explicit role in his work: the *thinking subject* as Brouwer put it (or the *ideal mathematician*, as he will be called in Section 2). To emphasize this aspect, Brouwer began to call himself an *introspective psychologist* because, he thought, this term described the mathematician's proper task. Moreover, he dismissed the role of communication between mathematicians, comparing other people to vacuum cleaners (a phrase which suggests that Brouwer found contact with other people exhausting).

Applied to mathematical reasoning, the solipsist view is not farfetched. After all, once our attention has been drawn to certain ideas or methods, understanding them is very much a private matter; we follow proofs, and do not merely repeat the words of other people. But does the view affect (constructive) mathematical practice? Brouwer found a positive answer by considering a mathematician who gives himself a rule referring to the stages of his own mathematical activity or, as one says, who employs a *private language*. From a solipsist view of mathematics such a rule belongs to mathematics because it is as well determined (for him) as any other rule. But from an intersubjective view it does not belong to mathematics because it cannot be communicated, and the possibility of communication is required here (in contrast to the solipsist position). What is particularly interesting is that Brouwer's considerations lead to consequences that can be formulated in the ordinary language of mathematics, which does not explicitly refer to the mathematician's activity. He found a property R and a 'private' rule which evidently satisfies R, but one does not know any intersubjectively well-determined rule which does (see p. 27 for details). So, once again, Brouwer had succeeded in finding a way of making an old distinction (between a *solipsist* and an *intersubjective* view of mathematical knowledge) manageable or, at least, held out a promise of doing so. We cannot yet judge the intrinsic interest of his idea; but if it is positive at all, the value of Brouwer's contribution is surely very high when measured by the ratio of its interest to the probability of its discovery.

In his last years Brouwer felt that his contributions were not adequately appreciated. He may have exaggerated a bit; but, by and large, he was probably right. Without, of course, explaining cause and effect, Brouwer's solipsism fits in quite well with his failure to convey his ideas: while, as was said above, solipsism seems an excellent first approximation for an analysis of mathematical reasoning, it would not be expected to be equally sound in public relations.

2.2 Foundations of Mathematics

Because of his technical knowledge, the working mathematician generally has a more concrete idea of foundational problems than the 'educated outsider'. Just because of this he almost inevitably asks too soon what foundations can do for him, a typical reaction of the practical man *vis-à-vis* any fundamental theory in his own field of study. Popular accounts, including those of Brouwer on his own contributions, tend to overdramatize the role of fundamental theory, and only add to the difficulty. Above all, the mathematician should remember that his familiar experience is generally more reliable than fundamental (that is, in the present case, logical) theory; except very occasionally, the latter does not correct dramatic errors of ordinary ideas (in fact, being more precise than these ideas it is more liable to be erroneous). More generally, a new fundamental theory rarely changes the way we actually see the world; in particular, logic does not change the way we think of the integers, ordered pairs and other familiar concepts. In all sciences the bulk of day-to-day work requires the introduction of ideas which, though generally consistent with fundamental theory, are independent of it; in mathematics, these ideas are defined in terms of already current concepts and therefore logically dependent on the latter. What the working mathematician can expect to find is that some previously inaccessible area of experience becomes manageable by means of the new theory; at least in an 'abstract' subject like mathematics the area will almost inevitably be marginal, because research tends to neglect unmanageable areas. Above all, the new theory can be expected to give a sensible analysis and solution to broad general questions, such as fundamental physical theory gives to the question: What is matter? While such achievements dominate the outsider's impression of the subject, they are not uppermost in the minds of working scientists. Once a theoretical idea has been introduced the scientist uses it without going back to the considerations which led to its discovery, particularly if these considerations had a recondite philosophical character. It will be best if the reader follows the account of Brouwer's work in a relaxed, detached spirit.

Constructivity on an elementary level

It is easy to recognize the difference between the kind of mathematics of which old-fashioned school mathematics is typical and, for instance, the ϵ-δ methods in analysis. (A precise formulation of the difference is an object of research here, not its starting point.) A striking feature of the former is the very limited use of *logic*: one has variables, in the first place for integers or rational numbers, symbols for computation rules (or 'functions' in the old-fashioned sense), and considers equations built up from these symbols. The proofs are limited to two kinds:

1. recognizing that the computation rules are 'well-defined' for the arguments considered (for instance, addition or multiplication in the case of the integers);

2. the deduction process itself, which consists in mere substitution and, in the case of arithmetic, in induction.

Thus (modulo 1) every proof is, in an obvious sense, a *schema* for computations. A general proposition such as an identity containing variables is unambiguously *about* computations of particular cases (obtained by substituting numerals for the variables in the identity). The restriction to numerical variables is by no means essential, as shown by the following example (also considered by Brouwer and de Loot [1924]):

1. *zeros of polynomials with complex coefficients*
   In modern terms: the zeros are continuous functions of the coefficients for the usual topology of the complex plane.

   Given $z^n + a_1 z^{n-1} + \cdots + a_n$, approximations $\xi_{i,p}$ $(1 \leq i \leq n)$ to the set of zeros $(\xi_1, \ldots, \xi_n)$, with $|\xi_i - \xi_{i,p}| < 2^{-p}$, can be computed from suitable approximations to the coefficients. The coefficients need not be given by a rule, and certainly need not be rational: the computation applies also if, for instance,

$$a_i = q_i + \sum_{r=1}^{\infty} a_{ir} 2^{-r}$$

where q_i is an integer and the a_{ir} $(r = 1, 2, \ldots)$ are given by a *random sequence* of 0 and 1.[6] Given q_i, Brouwer determines effectively $r(p)$ such that $\xi_{1,p}$ can be computed from $\{a_{ir}\}_{1 \le i \le n, 1 \le r \le r(p)}$.

2. *real zeros of polynomials with real coefficients*

   In modern terms: the existence of a real zero does not depend continuously on the coefficients.

   We consider $z^2 + a$ in the neighborhood of $a = 0$. If $a = \sum a_r 2^{-r}$ $(a_r = 0$ or $a_r = 1)$ there is a real zero just in case $\forall r(a_r = 0)$. Quite naively, given a rule for computing the sequence of a_r's there is no reason why one should be able to decide $\forall r(a_r = 0)$. In the case of random sequences the situation is a little more delicate: by the nature of the case only a finite number of elements $a_1, \ldots, a_n$ of the sequence are available for any computation, so both $\forall r(a_r = 0)$ and the existence of a real zero of $z^2 + a$ for the a above could *never* be proved (naturally, one could not be sure that $a_r \neq 0$ for some r either).

   To avoid this fine point (which will, however, be taken up immediately) consider $a = \sum b_r 2^{-r}$ where $b_r = 1, 0,$ or -1. It cannot be excluded that the random sequence $b_1, b_2, \ldots$ consists only of 0; if all the available values $b_1, \ldots, b_n$ are 0, a can be > 0 (for instance, if $b_{n+1} = 1$ and $b_{n+2} \ge 0$) or ≤ 0 (for instance, if $b_{n+1} = -1$): in the former case $z^2 + a$ has no real zero, in the latter it does. This applies to all n ('however large', as one says). Thus, even on a quite elementary understanding of the issue it is clear that there is no general method for deciding whether a polynomial with coefficients of the kind considered has real zeros.

Brouwer used a related, but slightly simpler, example to refute the *law of the excluded middle* for his interpretation of the logical operations. Let s be a variable for a random sequence $(s_1, s_2, \ldots)$ of 0's and 1's, write A for $\exists n(s_n = 0)$ and consider A or not A, also written $A \vee \neg A$, hence

$$\exists n(s_n = 0) \vee \forall m(s_m = 1)$$

and so

$$\exists n \forall m(s_n = 0 \vee s_m = 1).$$

If this held for all s, we should have a general method, say f, to compute n from s. But the value $f(s)$ of f applied to s can depend only on a finite initial

[6] Brouwer called such sequences *freely chosen*; reasons for his terminology are considered at the end of this chapter.

segment $s_1, \dots, s_p$ of the sequences, since this is all that is ever available for a computation. But there can be no such f; for if $f(\mathbf{1}) = n_0$ (where $\mathbf{1}$ is the sequence consisting only of 1) and $f(\mathbf{1})$ depends on the first p_0 elements of $\mathbf{1}$, we get a contradiction by taking the sequence s where $s_n = 1$ for $n \leq \max(n_0, p_0)$, and $s_n = 0$ otherwise. This is a refutation of the *general* law of the excluded middle, since it is meaningful to ask for constructive operations on random sequences, as shown by the example on complex zeros.

These matters are so simple that it is a little hard to remember why they created excitement; even if one allows for exaggerations to make an impression (see p. 16) and for people's habit to attach great significance to their own oversights. But, quite objectively, the example corrects a widely current *misstatement*. When the ϵ-δ machinery was introduced into analysis, it was said to have *reduced* the subject to operations on finite or rational approximations to real numbers. The example on real zeros shows that, in the strict natural sense of the word, knowledge of such approximations is not sufficient to follow up the operations actually occurring in analysis. Put more technically, the reduction provided by the ϵ-δ formalism *does not eliminate the abstract functions involved in the familiar logical operations* (but absent from old-fashioned school mathematics). Whatever one may think of these functions, their essential role in this reduction was not emphasized.

Constructive logic

What, then, is one to think of these functions (logical operations)? Brouwer himself rejected them and, in particular, the natural justification of the law of the excluded middle. To justify the application of this law considered above, we should simply add:

> for any sequence s, $\exists n(s_n = 0)$ is well defined (true or false); this has nothing to do with whether *we* know how to make the decision.

Brouwer criticized this justification, claiming that a (false) analogy with the finite case is used; for instance, he would accept the argument if $(\exists n \leq n_0)(s_n = 0)$ were substituted for $\exists n(s_n = 0)$. Actually his criticism involves a *petitio principii* because the essential point here is whether, when such an n_0 is given, we apply the law of the excluded middle *because* we imagine a process of listing the s_n for $n \leq n_0$ and deciding $s_n = 0$ or $s_n = 1$. Granted that such a process is possible in principle, if we do not actually do what is possible, the actual evidence for our conclusion is elsewhere, and the same evidence may justify the general law too.

If the justification above is accepted, one also accepts the fact that some mathematics is about situations independent of ourselves. (Or, a little more cautiously, some mathematics concerns concepts about situations which we conceive as being independent of ourselves). But there still remains the legitimate problem of developing that part of mathematics which is, so to speak hereditarily, about our own constructions and effective decisions; not as a series of *ad hoc* remarks (pointing out that this or that proof is effective), but as a systematic theory (explaining, for instance, why certain prima facie nonconstructive proofs 'happen' to be effective).

The straightforward way would be to go back to the elementary kind of constructivity described above, and to *avoid* logical operations altogether which, as we have just seen, are naturally interpreted in a nonconstructive manner. Brouwer had the much more original idea of giving a constructive *reinterpretation* of the logical operations. Actually, he himself never formulated the interpretation explicitly; this was done by Heyting (p. 16). But it may fairly be said that the meaning given to the logical operations is uniquely determined by Brouwer's indications, if one wants to give such a meaning at all.

The cases of *implication* and *negation* are typical. Naturally, since our knowledge is involved, we think of *proofs*. To prove $A \to B$ one needs two things: a mapping π from proofs to proofs, and a proof (say p_0) establishing that *if* any p proves A *then* $\pi(p)$ proves B; to prove $\neg A$, we need a proof (say p_1) of the assertion: for any p, p does not prove A (the variable p ranges over the objects of constructive mathematics). Clearly, this explanation is noncircular only if '*if ... then ...*' and '*not*' applied to statements of the special form 'p proves A' are essentially simpler than when applied to arbitrary propositions.[7] Quite early in his work ([1924]), Brouwer noted the general validity of:

$$A \to \neg\neg A, \qquad \neg A \leftrightarrow \neg\neg\neg A, \qquad \neg\neg(A \vee \neg A).$$

Hence (taking $A \vee \neg A$ for B) $\neg\neg B \to B$ cannot be generally valid.

Thus, for Brouwer's interpretation we have to understand proofs and constructions involving proofs, while in 'elementary' constructivity we only have to understand proofs of the special kind of proposition asserting that a given rule terminates. Perhaps, on closer analysis, we shouldn't be doing mathematics at all if we really didn't understand Brouwer's interpretation. But in the present state of knowledge it is easier to give an explicit theory for elementary constructivity.

Mathematical practice is full of examples of elementary constructivity, even if nowadays they are not singled out systematically. In contrast, at least at first sight, there seems to be nothing recondite to say about constructions on proofs in general! Of course, Heyting's formal laws (mentioned on p. 16) *implicitly* state properties of proofs, since they are valid for Brouwer's interpretation (which is stated in terms of proofs). But so poor was people's general judgement of the situation that it came as a great surprise when Gödel [1933] showed that a rich part of mathematics could be developed from these laws (incidentally, by essential use of implication or negation).

[7]One condition is that, for any p and A, we can decide whether or not p proves A, while in general we cannot decide A; this condition is reasonable because if *we* do not recognize p as a proof, it isn't one *for us*.

Perhaps because of all this experience or for intrinsic reasons, nobody seems ever to have been much tempted to put down false principles in elementary constructivity. In contrast, if one actually wants to formulate explicit properties of proofs, one has to keep one's wits about one to avoid errors which are (formally) similar to Russell's paradox in set theory. This is not surprising, inasmuch as Russell's paradox involves some kind of self application and, as seen from the example of implication, proofs *obviously* are about themselves (specifically, the proof p_0 is involved in some values of the variable p).[8]

The last two paragraphs give an idea of the problems mentioned in Part 1 (p. 16). They are not solved. But what we know already in this area is a substantial contribution to foundations. Also from a purely formal point of view, Heyting's systems (which were derived from Brouwer's interpretation) are remarkably interesting objects.

Infinite proofs

Naturally, further progress depends to a large extent on clearer knowledge of possible proofs of a proposition A: the more we know about such proofs, the better a chance we have to establish $A \to B$. Brouwer himself peppered his publications with words that suggest interesting directions of research, such as *fully analyzed* or *canonical* proofs and, at least implicitly, *irredundant* proofs. Though working in the context of formal systems, Gentzen (who was familiar with Brouwer's ideas) may well have had these ideas in mind when he developed his important analysis [1935] of *derivations without detour or cuts*, which dominate current proof theory. But perhaps the most striking idea was Brouwer's insistence that proofs, that is the mental acts involved in establishing a general mathematical proposition, should be analyzed as a *transfinite* sequence of steps. This is natural enough, if we remember how each of us first convinced himself at school of the validity of the principle of induction:

$$\text{infer } \forall n A(n) \text{ if } A(0) \text{ and } \forall n[A(n) \to A(n+1)] \text{ are proved}$$

or, put differently (if we are accustomed to think of formal systems), if we remember how we convinced ourselves that the corresponding formal rule expresses a valid principle of inference.

Brouwer himself published a proposed analysis of possible proofs only for one kind of proposition; somewhat paradoxically, for a kind that also occurs in elementary constructivity, namely for propositions asserting that a rule ρ is well defined for random sequences. By the fundamental continuity requirement, ρ must operate on finite initial segments $\overline{s}_n = (s_1, \ldots, s_n)$ of s. Thus ρ can be

[8] Incidentally, it is one of the peculiarities of constructive logic that, for some A, a *natural* formal proof of A goes *via* proofs of $A \to B$ and of $(A \to B) \to A$: such a proof of A actually contains a proof of $A \to B$.

described by an operation r on finite sequences c as follows:

$$r(\overline{s}_n) = \begin{cases} 0 & \text{if the values } \overline{s}_n \text{ are not sufficient to decide } \rho(s) \\ \rho(s)+1 & \text{otherwise} \end{cases}$$

(+1 for the case in which $\rho(s) = 0$). Such a rule r satisfies the following conditions:

1. $\forall s \exists n[r(\overline{s}_n) \neq 0]$
2. if $c \prec c'$ and $r(c) \neq 0$ then $r(c') = r(c)$

(where $c \prec c'$ means that c is a proper initial segment of c'). Conversely, any r satisfying 1 and 2 determines an operation ρ.

The crucial condition is 1.[9] Brouwer [1927] concluded that *canonical proofs* of 1 are built up by means of a transfinite succession of particularly simple steps. Two corollaries follow unquestionably from his conclusion:

3. *uniform continuity*
   For sequences s taking a bounded number of values, ρ is *uniformly* continuous (pointwise continuity was assumed from the start).

4. *bar theorem*
   For elementary orderings R, *two familiar definitions of well-ordering are equivalent*. Specifically:

   - *well-foundedness*
     For any countable domain D, if $R \subseteq D^2$ then any descending random sequence s of elements in D is finite, i.e.

     $$\forall s \exists n[\neg R(s_{n+1}, s_n) \wedge (\forall m < n) R(s_{m+1}, s_m)]$$

   - *transfinite induction*
     For any property $P \subseteq D$, transfinite induction holds w.r.t. the order R, i.e.

     $$\forall x[(\forall y)(R(y,x) \rightarrow P(y)) \rightarrow P(x)] \rightarrow \forall x P(x),$$

     where x and y range over D.

Both conclusions are (formally) familiar enough from ordinary mathematics, the first being an immediate consequence of compactness. Why did Brouwer have to introduce sophisticated ideas? A glance at the *usual* proofs of these conclusions shows that the methods are patently nonconstructive. It is a commonplace that

[9] By replacing r by r^*, where

$$r^*(c) = \begin{cases} r(c) & \text{if } (\forall c' \prec c)(r(c') = 0) \\ r(c'') & \text{if } c'' \prec c \wedge (\forall c' \prec c'')(r(c') = 0) \wedge r(c'') \neq 0, \end{cases}$$

2 is automatically satisfied.

a generalization of a theorem (that is the same formal statement interpreted for a wider class of structures) may require a much more sophisticated proof; the analogue applies if the theorem is to be constructively valid.

There may be doubt about Brouwer's doctrine on the form of canonical proofs of 1; at least a closer analysis should be possible. But even at the present stage the doctrine has served a very real purpose in explaining why the operations that happen to turn up in elementary constructivity are uniformly continuous (without uniformity being explicitly required): for all constructive methods of proof so far formulated, proofs of 1 can (demonstrably) be brought to Brouwer's normal form.

Hilbert's formalist foundations

This subject is not only interesting for its important role in Brouwer's scientific life, but also for the objective issues involved. The two principal elements of Hilbert's scheme are the discovery of *formalization*, which had been established by massive case studies, and Hilbert's own formulation (in the language of combinatorial mathematics, which is a quite narrow part of elementary constructive mathematics) of the *autonomy of combinatorial mathematics*. What he aimed to show was that the abstract elements (in particular the logical operations of nonconstructive mathematics, and the operations on proofs in Brouwer's interpretation) are not needed for establishing combinatorial theorems. Hilbert's idea was very much the same as the widely current idea that an arithmetic theorem must have an arithmetic proof; or even that a 'simple' statement, if it can be proved at all, must have a simple proof.[10]

A little more explicitly, Hilbert's scheme may be described as follows. Examination of mathematical practice, in particular the reduction to set theory by Whitehead and Russell, provided a formal system $\mathcal{F}$ in which ordinary mathematical reasoning can be described compactly.[11] To explain how usual elementary combinatorial mathematics can be 'developed' in $\mathcal{F}$, let $P_{\mathcal{F}}(c, n)$ mean that the sequence of formulae c is a formal derivation in $\mathcal{F}$ of n; the property $P_{\mathcal{F}}$, by the very nature of formal systems, is quite elementary. Now to each elementary assertion (identity) N, possibly containing variables, is associated a formula n of $\mathcal{F}$ which, for the intended meaning of $\mathcal{F}$, expresses that N is true. Then there is a function α such that, for the N considered,

$$N \to P_{\mathcal{F}}(\alpha(n), n) \tag{2.1}$$

holds (when N contains no variables, and a slightly more complicated form if it

[10]It should, however, be noted here that Hilbert, in contrast to Brouwer, believed that nonconstructive notions were needed to make mathematics intelligible.

[11]The formal rules of $\mathcal{F}$ are not, contrary to *some* passages in Hilbert's writings, the essentials of mathematical reasoning; it is not enough to manipulate the formal rules of $\mathcal{F}$, but one must recognize that they are valid for their intended meaning.

does). This last statement no longer refers to the interpretation of $\mathcal{F}$. Finally, to show that the development is sound we must show that, for any c,

$$P_{\mathcal{F}}(c, n) \to N \tag{2.2}$$

is valid (that is, even if n is formally derived using abstract arguments reflected in c, N must be true). This last formula, too, is expressed in wholly elementary terms and so there is at least the possibility of an elementary proof. (If one had such a proof, and c is a derivation of n in $\mathcal{F}$, one would also have an elementary proof of N). Being sure that he would find one (for reasonable $\mathcal{F}$'s) Hilbert, of course, did not expect to have to analyze the idea of elementary proof: he would recognize one when he saw it.

Originally for purely technical reasons, Hilbert observed that 2.2 holds if we can establish the formal *consistency* of $\mathcal{F}$; that is if we can show that, for any c, not $P_{\mathcal{F}}(c, n_0)$, where n_0 is the formula of $\mathcal{F}$ expressing $0 = 1$. (Hilbert's observation uses in an essential way that combinatorial arithmetic can be developed in $\mathcal{F}$ in the sense explained above).

The Hilbert-Brouwer controversy can nowadays be put very simply. As time went on Hilbert tried to give a much more central significance to consistency; as if it were the only thing that mattered, and not merely that it was sufficient for what was called above the *autonomy of combinatorial mathematics*. But suppose $\mathcal{F}$ is such that, for some true identity N, the corresponding n is not derivable in $\mathcal{F}$; then the system obtained from $\mathcal{F}$ by adding the formal negation of n is consistent. Thus a patently false statement is derivable in the extended system, namely the negation of N. Brouwer's main criticism of Hilbert concerned this kind of inadequacy of the consistency criterion; he never attacked directly the principal claim of the autonomy of combinatorial mathematics. But, in contrast to Hilbert, his judgement on the matter was sound: he didn't believe the autonomy.[12]

There is a natural modification of Hilbert's project: to prove 2.2 not necessarily by *combinatorial* methods, but by other suitable *constructive* methods. The principles discovered and developed by Brouwer dominate in this work.

Ideal mathematician

In connection with a specific problem on random sequences, namely (in the notation of p. 20) the proof of

$$\neg\forall s[\neg\neg\exists n(s_n = 0) \to \exists n(s_n = 0)],$$

[12]Though Hilbert never explicitly said so, the only really convincing proof of his programme would have been to show, independently of any particular system $\mathcal{F}$, that any combinatorial statement can either be proved by combinatorial methods or refuted. Even without closer analysis of the notion of combinatorial proof, this last conjecture is very implausible in view of Gödel's incompleteness theorems.

Brouwer [1948] considered the case of a *thinking subject* or *ideal mathematician* Σ proving theorems in stages $1, 2, \ldots$ in order ω. Much of this is dubious, as will be seen below; besides we have nowadays better ways of dealing with his specific problem. But a few, interesting points are quite clear. Let $\vdash_n A$ mean that Σ has proved A by stage n. Then:

1. $(\vdash_n A) \vee \neg(\vdash_n A)$ (for each n, Σ knows if he has proved A or if he has not), and hence Σ has a well-defined rule r with values 0 or 1 such that
$$\forall n[r(n) = 0 \leftrightarrow \neg(\vdash_n A)].$$
2. $(\exists n)(\vdash_n A) \to A$ (whether 'A' means that Σ has already a proof of A, or will have a proof, or that somebody has a proof!) and hence, by contrapositive,
$$\neg A \to \neg(\exists n)(\vdash_n A).$$
3. $A \to \neg\neg\exists n(\vdash_n A)$ (if somebody has a proof there is no mathematical reason why Σ should not have one; for the other meanings even $A \to \exists n(\vdash_n A)$ can be asserted) and hence, by contrapositive and the triple negation law,
$$\neg\exists n(\vdash_n A) \to \neg A.$$

These conditions are clearly sufficient to conclude

$$\exists r(\forall n[r(n) = 0] \leftrightarrow \neg A).$$

This last statement no longer contains the relation $\vdash_n$, and

$$\forall n[r(n) = 0] \leftrightarrow \neg A$$

is the property R (of r) mentioned on p. 18. The proposition A is quite arbitrary: if propositions containing quantifiers over rules are included, the axiom above implies quite a strong form of *comprehension principle*. Once again, as in the case of the *bar theorem*, we find a familiar formal statement with a quite different justification.

The whole idea has a certain air of unreality, of playing with words. This impression is justified in so far as it is not clear what is meant by a stage, or whether the mathematician is supposed to prove one or infinitely many theorems at stage n; also, there is no clear reason to restrict oneself to ω stages when the canonical proofs on p. 24 consist of a transfinite sequence. But probably the actual reason behind the impression is that the argument is too 'easy'; that to get 'real' knowledge of mathematical reasoning one should study (it would be said) the physics and chemistry of the nervous system, instead of the *ideal mathematician*. This is no more convincing than having the founders of hydrodynamics study the atomic structure of matter before setting down equations of continuity and other properties of *ideal fluids*; equations which give detailed quantitative results derived from quite general qualitative assumptions about, that is from impressions of, the world. Indeed, even after one knows atomic theory one rarely uses it to understand the motion of fluids around us. Similarly, there is at present no reason to reject the possibility of a useful mathematical theory of the ideal mathematician.

Looking back

Certainly, if one wants to know at all about the constructive aspects of mathematics, Brouwer's work has been invaluable. Instead of bits and pieces of isolated 'constructivizations' one has a more or less coherent theory, made systematic by a suitably abstract conception of constructive operations. Particularly in connection with the principles of proof and definition by transfinite induction, Brouwer's constructive logic explains why formally minor variants completely alter the constructive aspects. Not surprisingly, Heyting's formal rules have interesting applications beyond the intended one. Roughly speaking, whenever mathematicians loosely speak of *constructions* (for instance, if they mean no more than some kind of explicit definition), there is a reason for their informal choice of words; the 'logic' of their concepts, as one says, is the logic of constructions in the proper sense of the word; they satisfy Heyting's rules. This is consistent with the remark on p. 16, that these rules use only very simple properties of the notion of construction in its strict sense.

Is mathematics about our own constructions or is it about an external reality?

It would have been idle to go into the ambiguities referred to on p. 12, before analyzing and developing the part of mathematics which *is* about our own constructions. But since, largely through the work of Brouwer, substantial progress has been made in this subject, it is now worth while to go into some distinctions.

First of all, some mathematics treats concepts which quite plainly are intended to be about an external reality, about sets and random sequences given, for example, by throwing dice. (Only if one thinks of them as sequences freely *chosen* by a 'thinking' *subject* are they possibly our own constructions.) Inasmuch as we understand these concepts and can reason about them, not all mathematics is about concepts which refer, hereditarily, to our own constructions. It is not relevant here whether an actual reality is involved or possibilities.

Second, since mathematics is constantly used in the description of nature, it would be splitting hairs to deny that a good deal of mathematics is also about actual external reality; not only finite mathematics used in stating results of measurements, but the highly abstract concepts used in formulating theoretical laws which connect such measurements.

But a real issue remains:

> Can the mathematics that presents itself to us as being about an external reality be construed as being about our own constructions?
>
> Or, as one says, can it be *accounted for* in terms of our own constructions?

Here the answer is negative (at the present stage of the development of constructive mathematics) if one means all principles of nonconstructive mathematics so far *formulated*. For example, we do not have a constructive consistency proof for the formal principles of analysis, and this would be a *minimum* requirement for a positive answer, since those principles are evidently valid for the (nonconstructive) concepts of the *set of natural numbers* and of its *power set*.

On the other hand (at the present stage of the development of ordinary mathematics) the bulk of mathematical *practice* can indeed be construed constructively, if Brouwer's *bar theorem* is included in the latter. Actual practice uses nothing like all the principles of nonconstructive mathematics so far formulated. Naturally the interpretation is not on a purely formal level, but the difference in meaning between the constructive and nonconstructive logical operations must be taken into account.

Chapter 3

Brouwer's 'Cambridge Lectures on Intuitionism'

This sad little book recalls vividly the pathetic spectacle of the original lectures, a couple of which I attended some 45 years ago. Even for a beginner in logic the lack of movement in Brouwer's thought, on the mathematics of the continuum in terms of choice sequences, was depressing. The most obvious novelty, compared to his publications in the mid-twenties, was a torrent of even more bizarre and stilted terminology than in the original German papers. Most significantly, the lectures do not benefit from the obviously relevant logical work in the thirties; neither for improving Brouwer's own ideas, nor by using them to enrich logic.[1] For a less parochial, not purely logical view of Brouwer's topic, progress in algebraic topology in the thirties would also be relevant. Brouwer's neglect of the literature was certainly in harmony with his idea that he was 'revolutionizing' the subject, not simply commenting on a particular aspect.

Two of the would-be revolutionary tactics which are more prominent in the lectures than in the twenties concern formal counter examples:

1. a more frequent use of socalled *fleeing properties*
2. references to the *creative subject*, so to speak, 'the' idealized mathematician.

The counter examples (to classical analysis) are purely formal, since the meaning of the logical operations and the ranges of the quantifiers used by Brouwer are different.

[0]Originally published in *Canadian Philosophical Reviews*, 2 (1982) 249–251.

[1]Even before the lectures Kleene had stressed relations to recursion theory in print. More pertinent still is the notion of cut-free proof discovered by Gentzen in the early thirties: applied to Σ^0_1-theorems with parameters for choice sequences, it gives a very satisfactory formal version of Brouwer's idea of a 'fully analyzed proof' of well-foundedness, taking the mystery out of the business about the socalled bar theorem.

As to 1, the idea involved is familiar enough from (unquestionably relevant) continuity requirements which ensure that small uncertainties in the data, defined by fleeing properties, do not spoil the outcome of deductions or operations. (There is a positive side to this matter which will be taken up below).

As to 2, it would indeed be a bad *philosophical* error to reject out of hand the very mention of the creative subject (as a topic for mathematical theory, without recourse to quantitative 'empirical' study). The error would be the failure to see the place of idealizations in the arsenal of scientific methods; plausible, but not very successful ones like ideal fluids, and spectacular ones like Einstein's observer viewing the world while traveling with the speed of light. In fact, Gödel's second incompleteness theorem tells us something not altogether trivial about formalized in some system or are discovered to do so, perhaps unconsciously. (There is still plenty of room for using creativity or ingenuity when *selecting* sensible theorems and efficient proofs *among* all legitimate ones). But equally it would be a *scientific*, if not philosophical, error to be paralyzed by awe before such limited successes; not to ask oneself just how much we can expect from (the mere concept of) the creative subject, without a realistic look at the actual possibilities of the mathematical imagination.[2]

Experience has shown that just those people who have the kind of logical sensibility needed to see and develop fruitful points of the lectures under consideration, are simply put off by Brouwer's own assessment of their place in the scheme of things; for example, the pretentious (and hence simpleminded, not merely simple!) inventory of inner experience on p. 90, beginning with 'twoity, giving rise [first] to invariable unity, and [then] to threeity.' This sort of jargon recalls the less convincing products of the Wisdom of the East. Brouwer tells us here about as much about the processes of mathematical reasoning as does (say) the *Kamasutra*, with its elephant women and rabbit men, about the processes of evolution by sexual selection.

The pity of it all is that the theory of choice sequences answers a perfectly good question, implicit already in Aristotle's *Metaphysica* Γ 7, 1012a, 21–4:

> Is there anything like ordinary logic which applies to propositions about incompletely defined terms?

The theory gives an elegant positive answer. Of course, elegance is no guarantee for relevance. It is a separate question whether, scientifically, we do better by *paraphrasing* the incompletely defined terms (here: by means of continuity requirements). But within the logical tradition Aristotle's question has, rather obviously, a permanent place.[3]

---

[2] Ignoring this principal limitation of the concept the editor complains on p. ix that Brouwer uses it only for weak, not strong counter examples!

[3] *Remark.* In the broad sense of the word, it would be a philosophical error to overlook the parallel between Aristotle's question and such algebraic questions as:

For the record, readers are warned that the editor's historical interpretations (on pp. vii-ix), with their air of judicious scholarship, are no more perceptive than most of currently popular history of mathematics; in particular, not even perceptive enough to see the exceptional degree of sensibility needed to comment at all significantly on such exceptional talents as Brouwer. What can sometimes be done seems to me very well illustrated by Newman ([1969], p. 53): he, who incidentally knew this subject well, answered (without asking) the question why Brouwer did not pursue n-dimensional topology after he had opened up that area. Newman noted that the kind of tools objectively needed for further progress in higher dimensions were consistently neglected in Brouwer's own work even where they were available. Knowing as little as we do about mathematical temperaments, one can hardly be sure that Brouwer's neglect was a basic element of his temperament. It is just a simple historical fact which had struck Newman. And he was able to use it to help us remember a significant mathematical fact about the character of n-dimensional topology. The beauty of this literary device is, to me, that allows us to think imaginatively about matters of temperament without tempting us into 'theories' out of all proportion to the data. Of course, we want to think about those matters. After all, they are of wider interest, than topology or choice sequences.

For which dimensions do we have anything like ordinary algebra?

(with the famous answer involving the magic numbers 1,2,4,8). Here again, a paraphrase (of vectors in terms of coordinates) is possible. The philosophical discovery is that, contrary to first impressions, both questions admit natural and precise formulations at all (and convincing answers).

Chapter 4

Some Corresponding High Spots of Early Classical and Intuitionistic Logic

The periods meant are roughly 1880-1930 for classical logic, and 1930-1960 for intuitionistic logic. The 'correspondence' concerns:

- completeness (tacitly, for intended meanings)
- concocted (or, more reasonably, discovered) meanings
- relevance of intended and concocted meanings.

Readers are advised to regard the material as an exposition of the logical topics above, leavened a little with names and dates for so-called human interest; and *not* as history of logic. At the end there is a brief section on uses and abuses of such material in the notorious history of ideas, in line with the reservations of experienced historians about this subject.

Completeness with respect to intended meanings

The question of *completeness* is as old as the hills or, more precisely, as modern logic. Those coming from the philosophical tradition immediately saw all sorts of difficulties, and engaged in longwinded 'analyses' of the question (in contrast to the scientific tradition, which is cavalier towards its question, but tests its answers and their corollaries with a vengeance):

- How can we ask questions about the language in which we have our intellectual being as it were?

[0]Originally published in *Gödel remembered*, Weingartner and Schmettered eds., Bibliopolis, 1987, pp. 132–142, as Appendix II to 'Gödel's excursions into intuitionistic logic'.

- Do we have our intellectual being in elementary logical language?

Real completeness of a calculus would require that it 'embraces' (that is, formalizes) all valid methods of proof (cf. p. 91 of Chapter 6, with the reminder about the mind-boggling totality of all proofs of $0 = 0$). The conclusion is not to bother since, if that's 'real' completeness, it just does not lend itself to rewarding study.

Granted that the word 'completeness' is catchy, just where is the property relevant? After all, if it is considered at all, one must understand *both* the meaning intended *and* the calculus considered. In which situations is it appropriate to *combine* the two?

For example, I happen to be familiar both with the model-theoretic (intended) meaning of classical predicate logic and quite a number of its formalizations. Around 1960 I asked myself: What are we left with it we forget about the latter (for the time being)? My answer was the text Kreisel and Krivine [1966] (cf. the long introduction to the second English or to the German translation).

Other people have since tried out expositions of intuitionistic logic using primarily Kripke models, one of the concocted meanings of intuitionistic logic (incidentally, not considered below). But since formal rules are mentioned (for example, in Gabbay [1981]), it is not altogether clear what we are left with if we forget them altogether in intuitionistic logic.

This matter is of course not settled by Kreisel and Krivine, since the *relative* role of an intended meaning and one or other of its formalizations will not be the same in all branches of logic at all stages of their development:

1. As long as the relevance of any (in particular, the intended) meaning is not tested, it may be scientifically frivolous to consider the completeness of a calculus for it.

   But for the philosophical tradition it is perfectly proper to investigate the 'nature' of the matter; specifically, how much (or, rather, how little) need be known of that intended meaning to settle completeness; cf. Chapter 11.

2. For classical logic, and also for intuitionistic logic (but particularly in the case of propositions about lawless sequences), I have the impression that the intended meanings have so far been almost indispensable sources of conjectures and cross checks, even when the calculus was the primary object of study.

   But it should also be remembered that interpolation and other results were first *established* proof-theoretically, and conjectured (by Craig) in connection with a so-called empiricist philosophy of science (which eliminates from the theoretical premise abstract notions not occurring in the empirical consequences).

   Finally, it should also be remembered that in ordinary algebra formal considerations preceded interpretations in such cases as the $\sqrt{-1}$ or quaternions, but presumably not with matrix multiplication (corresponding to the product of coordinate transformations).

As to a 'correspondence' between completeness for classical and intuitionistic logic, the proofs for the intended meanings came at the end of the periods considered (with *in*completeness for predicate calculus of intuitionistic logic, modulo Church's Thesis, in the 60's).

The most obvious difference was that the expositions for intuitionistic logic benefitted generally from experience with the classical case, and from such notions as 'basis' for a more concise formulations.

Another difference is that the properties of the propositional and first order parts of intuitionistic logic are objectively more complex than in the classical case.

Some early concocted meanings

By tradition, the requirement on the concocted meanings is that they should satisfy the particular laws that happen to have been formulated for the (originally) intended meanings.

This tradition is dominated by such ideals as *rational reconstruction*; for example, on the ground that the original meaning is not intelligible. Those of us not limited by this particular intellectual handicap find the ideal 'unintelligible'! After all, to reword 2 above, if the intended meaning is in doubt so are the laws in question, unless they have been tested in some other way. The standard claim that they represent common usage is doubly suspect, since the latter:

- differs from all formal systems (for example, with respect to natural sense; cf. p. 232 of Chapter 10)
- is rarely obviously optimal for reasoning well.

Be that as it may, here are some items produced by that tradition.

We consider first concocted meanings of classical logic in constructivist terms:

- *Skolem's second proof of Löwenheim's theorem*
  Skolem's work [1922] comes under the heading 'constructivity' in the sense that attention is given to *definability* (in particular, arithmetic definability). So the 'concocted' meaning is that of *validity for arithmetically defined structures* (over ω).

  Viewed this way, there is a point to the otherwise quite silly stress by Skolem on 'avoiding the axioms of choice'; by showing that a *superstition* about the latter is involved. Specifically, the 'essential' aspect of the axiom of choice was widely seen in the fact that the other (existential) axioms of set theory implied *uniqueness* (and hence explicit definability) of the sets asserted to exist. This overlooks the fact that not even pure (classical) logic preserves this property.[1]

- *Herbrand's théorème fondamental*
  Herbrand's *champs finis* (which are suitable sequences of expanding finite structures) are, almost as they stand, a paraphrase of logical validity; for example, in terms of an infinite Herbrand disjunction.[2]

  Herbrand mentioned in his thesis that he knew how to deduce the ordinary completeness theorem from his analysis, and there is no reason to doubt this. He also added that the ordinary notion of validity was not precise enough. Here he made a philosophical mistake.

  The question is not 'why' Herbrand failed to prove the completeness theorem, but which kind of mistake(s) he made. The feeble terminology *théorème fondamental* fits the feeble uses he made of the theorem; by proving some eminently forgettable prefix classes of predicate logic to be decidable. Almost 50 years later Dreben and Goldfarb [1979], quite touchingly, continued this line; though some really rewarding (and, of course, quite different) areas had been found, at least 30 years earlier, to which Herbrand's *théorème* is relevant.

  If anything, the point above is enhanced by the occasional success of turning the line through 180°; for example, in *un*decidability results on certain prefix classes (without any use of Herbrand's Theorem); cf. Goldfarb [1984].

In the case of intuitionistic logic, concocted meanings (in model-theoretic terms) appeared fairly soon after Heyting's formalization, while in the classical case they had come more slowly (as noted above, completeness for the intended meaning came instead at the end of the period considered, as in the classical case).

The pattern is general: different authors were struck by (one of the relatively many) different elements of the (wild) rhetoric on intuitionistic logic (cf. p. 215 of Chapter 10). Here are some samples:

- There was the alleged issue between the *truth and provability*. This is evidently reflected in Gödel's *modal meaning* (cf. p. 223 of Chapter 10). Much the same applies to Tarski's *calculus of systems*, where provability (or consequence) from single axioms in classical predicate logic is meant; cf. also Gabbay's update ([1976]) in terms of Post systems.

---

[1] For example, the application of the law of excluded middle to $\exists x(x \in L)$: $\exists x \forall y(x \in L \vee y \notin L)$. Cf. also the failure of $\exists$-theorems in some 'intuitionistic' set theories.

[2] Cf. the notion of interpretation in Kreisel [1951] and [1952a] and, in particular, of the no-counterexample-interpretation for a more compact formulation by use of function variables (avoided by Herbrand for the sake of - his particular version of - finitist ideology).

Incidentally, the meaning assigned is not suitable for formulas that are not valid, in the sense that such formulas do not generally imply their own Herbrand disjunction; [∞].

- Another catchy element of the rhetoric involved *choice sequences*, and such hot news as: all functions on the reals, tacitly thought of as given by choice sequences, are continuous. (Pity that Brouwer never thought of the variant: all mappings between spaces are topological!)

  Tarski's *topological interpretation* certainly incorporates some features of that element. More precisely, as we should say now, not so much of the choice sequences Brouwer had in mind, but of lawless sequences (which lend themselves to a smoother theory).

- Perhaps the best known concocted meaning is *recursive realizability* and its variants; in the first place with stress on the constructivity of operations, not of proofs.

  Kleene was (consciously) most taken by a felicitous formulation of Hilbert about implication involving partial information. As far as content is concerned, this is barely distinguishable from formulations by Brouwer and Heyting (and not even appropriate for the finitist case, Hilbert's avowed concern, where the formulas A, B are Π^0_1 and $A \to B$ is realized by a total recursive functions). But Hilbert's wording fits the need to use partial recursive functions for realizing the logical laws.

  Stretching matters just a shade perhaps, one might look at the enrichment of realizations by formal derivability in Kleene's *q- realizability*, and especially by formal derivations in Beeson's *fp- realizability* some 30 years later. Both reflect the concern of (some of) the rhetoric with proofs over and above definitions; specifically, proofs of the fact that the definitions do what they are supposed to do.

In summary, here we have examples of exploiting an intended meaning without unduly meticulous attention to it. This style is clearly appropriate, if one is convinced that the details of the intended meaning do not serve the intended purposes, or that the latter are misguided.

Philosophical assessment of intended and concocted meanings

First of all, *generalization* (the most commonplace activity in mathematics) is a discovery of a new meaning (among other things). For example,

$$a^2 - b^2 = (a + b)(a - b)$$

may originally have been intended for a and b ranging over $\mathcal{Z}$, but is discovered to be valid for all commutative rings.

Secondly, the philosophical tradition of working up dramatic *conflicts*, or at least *puzzles* (here, concerning the meanings mentioned), is quite misguided;

- There is no general conflict, but there is a very real problem:

  Where is which of these meanings, if any, relevant?

- There is no puzzle; for example, about different meanings being equivalent (in the here usual sense of having the same set of valid sentences in the languages considered), since those languages obviously do not exhaust the possibilities of talking about those meanings.

The tradition is also misguided (unless it simply wants to preserve the *status quo*) in a more practical sense, since those antics about conflicts and puzzles attract philosophical cripples; for example, among (both gifted and other) logicians.

It is equally misguided to dub the choice between those equivalent meanings a mere *matter of convenience*, compared to the allegedly primary matter of adequacy-in-principle. This is generally simply not good enough for *effective knowledge*; cf. 1–3 below, or consider the choice between walking and taking a car to, say, the nearest source of food; if the latter is 1 km away the choice may be a matter of convenience, if 100 it is a matter of survival. So much for adequacy-in-principle.

The next items, though still general, concern the intended meanings of classical and intuitionistic logic. Quite simply, the usual logical languages express very little about those meanings.

Thus nothing about the *power set* construction even enters into the explanation of the (classical) meaning of first order formulas, although that construction with its iteration may fairly be said to be the heart of the subject of sets, in terms of which the *set-theoretic* (alias model-theoretic) meaning is defined. At best, knowledge of that construction may be used in *logically impure proofs* of logical validity.

In the case of intuitionistic logic, the intended meaning involves *constructions*, with their (decidable) properties and proofs of assertions that any (that is, an arbitrary) construction has some given property. But, for logic, the principal topic is the relation between a proof and the assertion proved. Realistically speaking, this relation is just too meager to be rewarding (or even to save one from oversights like paradoxes, because it just does not provide enough cross checks.) Taken literally, it leaves no room at all for more delicate ideas about proofs such as (those on p. 92 of Chapter 6, nor even for) Brouwer's about 'fully analysed' proofs (and their extension in Gentzen's proofs 'without detour'). If these ideas are to be pursued in a purely logical context at all, some concocted meaning is *necessary* (for example, so-called operational semantics). And then at least the scientifically experienced will ask *whether* the most fruitful aspects of those ideas are relevant to logical contexts at all. It would be pure ideology to assume this just because the word 'logic' is glamorous.

To conclude, here are some specific examples (chosen at random) that are related to the generalities above:

1. Elements from the constructivist tradition (especially, the part concerned with definability) are practically unavoidable even in classical logic, as soon as the languages considered are restricted (for example, to finite formulas).

   Herbrand's pre-occupations described in 2 on p. 36 have turned out to be useful even when applied to Σ_1 formulas; admittedly, once stated, the form of a Herbrand disjunction for $\exists x A$:

   $$A[x/t_1] \vee \cdots \vee A[x/t_n]$$

   is memorable without (the no-counterexample) interpretation (let alone, the more tortuous *champs finis*).

   Recently [∞] applications of those ideas to Σ_2 formulas have been found, where already the wording of the Herbrand disjunctions benefits from functional interpretations.[3]

2. It is a common place that many formal results about intuitionistic logic have been obtained by use of concocted meanings (cf. the previous subsection). But it is worth stressing that also the (first) completeness proofs for the *intended* meaning used the topological interpretation concocted by Tarski (for an exposition using a meaning concocted by Kripke, cf. Burgess [1981]). The proof for the intended meaning required two further steps.

   First, the observation that, instead of 'associating' an open *subset* A_p of a certain topological space S with the propositional symbol p, the latter can be literally interpreted (that is, replaced) by a *proposition* A_p with parameter α over S.

   Secondly, if S is chosen as a (topological) space of lawless sequences α, then Tarski's 'association' is proved to respect the laws of lawless sequences for propositional operators $\circ$:

   $$\alpha \in A_{p \circ q} \;\Leftrightarrow\; (\alpha \in A_p) \circ (\alpha \in A_q).$$

   In short, far from having a 'conflict' between the concocted and intended meanings, one has a *proof of equivalence* in the relevant context.

3. This is a speculation on a possible use of the many concocted meanings for formal intuitionistic logic (in accordance with the rhetoric about formal laws not determining meaning, but at odds with the wailing about this truism).

[3] Once the wording is available, model theory is quite adequate to infer the *validity* of some Herbrand disjunction from the validity of the formula in question. Herbrand himself aimed at, but did not achieve himself, *quantitative estimates* for a suitable disjunction from richer data (namely, the validity of the formula enriched by a suitable proof).

It is addressed to those who wonder whether $P = NP$, and believe that the question is 'logical'.[4]

Now, with each concocted meaning of intuitionistic logic comes a whole body of knowledge: of the concepts used to define that meaning. So there is a chance of finding a problem about intuitionistic logic that can be proved *NP*-complete from knowledge of one concocted meaning, and decided to be polynomial (or not) from knowledge of another.

Not the *similarities* between different meanings, so dear to logicians' hearts, are useful here, but *differences*.

Warnings against uses for history or methodology

It would be unrealistic to rely on the disclaimer at the beginning of this chapter as a safeguard against misinterpretations of the material above.

- On the one hand there are *stories about individuals*, often chronologically ordered; the sort of thing associated with *history*.
- On the other hand there are *subdivisions*, and subdivisions within them, associated with so-called *scholarly history*; as opposed to, say, historical novels or to pep talks using historical illustrations (practiced by politicians in all spheres of life).

In the long run, any attempt to use the snippets in this Chapter as a conventional kind of history would draw attention away from the more modest, but surely worthwhile, uses actually made of them here. The reason is by no means marginal. It will now be considered at leisure.

Without exaggeration, our ordinary view of men and their doings is extremely primitive. Sure, all the world complains about lack of 'progress' in our understanding of human nature and human society, compared to progress in the natural sciences.

- Some intrinsic obstacles are similar to those found in the uphill fight of natural history (cf. p. 233 of Chapter 10).
- But the overwhelming obstacle is probably the obstinacy with which one clings to those aspects that strike our untutored attention. In fact, literally to the *Wisdom of the Ancients*; here, to the Bible with its list of sins (another word for: principal human motives).

  If somebody comes along and selects one sin as the dominant force, this is regarded as an intellectual revolution (one used to think of Marx and

[4]For example, as opposed to: belonging to transcendence theory; here one would look for some familiar real number with, say, a decimal expansion shown to be both *NP*-complete and in P or not (as the case may be).

> Freud, but one might as well throw in the analogue to purity of method).[5] Just imagine one had made a virtue of clinging to ideas of the same vintage about the world around us! Surely, there is room for a dialogue in which honest *Simplicio* tells us how he looks at human phenomena, even if he only says what everybody (still) says.

Now, of course we live with this sort of thing, having intellectual and other reflexes that operate independently of traditional commentaries. What the so-called history of ideas adds to this mess is the pretense that *painstaking scholarly documentation* (comparable to the most simpleminded painstaking classifications in natural history, cf. p. 233 of Chapter 10) improves matters. On the contrary, one adds to the simple-minded view a crass imbalance between the degree of accuracy implicit in the claims and that of the data.

Superficially, all this applies to all history. Not so; gifted historians have discovered (either, again, by exercise of sound intellectual reflexes or, occasionally, by following a 'philosophy' of history) aspects that do lend themselves to rewarding study. Most spectacularly perhaps in archaeology, obviously related to history (even if separated, partly for ideological reasons); in particular, there are uses of archaeology in connection with Babylonian mathematics

In accordance with common experience (so to speak, as a lesson of history), a subject like the history of ideas will be populated by people who are prepared to make little progress provided only few others make more; in particular, those one-eyed who have such a good time in the kingdom of the blind. Besides, it is a *pleasant* subject. Even if no new knowledge results, at least one is paid for reading material by gifted people; comparable to natural historians who spend their time looking at pretty butterflies or flowers, instead of bacteria looked at in scientific genetics.

Far be it from me to claim that molecular biology is the sole key to all biological (including human) phenomena. But I at least am grateful to have lived through a period when the conventional threadbare literature on human nature in so-called social sciences was replaced by a fresh kind of speculation, even if the latter is often only a *jeu d'esprit*.[6]

To conclude, it is salutary to remember one of the great successes of the natural sciences in the area of historical research. Here I shall take the *history of the planets*, since most readers will already have thought of the (origin and) history of species on earth. Naturally, *Simplicio*'s future dialogue (mentioned above) would tell us that planets are not conscious beings who intend to leave a record of their doings (and he will feel that he has made a profound and, above

[5] Of course, genuine revolutions came about when this kind of rehash was presented with uncommon conviction in uncommon circumstances and, perhaps above all, with uncommon political (including literary) skill.

[6] Cf. p. 128 of Chapter 6 for a new twist on *Genesis* by Crick and Orgel. It concerns history, but not the history of ideas.

all, decisive point). It never occurs to him even to try and *test* how much help this is.

There are two aspects to the history of the planets:

1. The history of their *outward* behaviour (in other words, their motion). This admits, as one says, astronomical precision.[7]

2. The history of their *inner* life (in particular, their chemical evolution). This not only can be, but has been asked and speculated about for a very long time. But it is fair to say that without a rather advanced knowledge of subatomic processes one had not even reached a *threshold* for informed discussion.

The other great success, the *history of species on earth*, seems (as far as I am competent to judge) more delicate.

- No doubt some evolutionary phenomena (the spread of epidemics, or development of resistance to drugs in bacteria) lend themselves to a theoretical study comparable to 1; that is, using little more than Mendel's laws and more or less sophisticated mathematics (as in particle or continuum mechanics).

- But knowledge comparable to 2 (on the modecular, if not subatomic, level) is liable to be relevant to the evolution of outward forms, even if only quite rough approximations are wanted.

  For example, just how many one-step mutations are needed for some morphological change? The answer would seem essential for giving an even approximately sensible meaning to 'missing link'.

Readers are asked to forgive the traces of 'metaphysical anger' that they may find disturbing in this subsection. But what more do those glib would-be historians (here, of logic) think they know of the relevant intellectual mechanisms than about mutations?

[7]Actually, not their apparent motion (so to speak, what they say literally), but the motion corrected for parallax (by Tycho Brahe), since only the latter lends itself to theoretical analysis.

Part II

ON GÖDEL

Chapter 5

Gödel's Election to the Royal Society

The following statement in support of Kurt Gödel for Foreign Membership of the Royal Society was submitted November 22, 1966, and Gödel's was elected April 25, 1968. His letter (originally in German) is of May 1, 1968.

Statement for the Royal Society

Far more successfully than anyone else, past or present, Gödel has been able to draw sharp consequences from philosophical conceptions of mathematics, and then to decide between rival views. Since these decisions involved detailed mathematical analysis, his contribution is just the sort of things that the R.S. is enjoined by its statute to promote.

The methods introduced by him for the purposes described above have turned out to be of intrinsic mathematical interest, and have developed into flourishing new branches of mathematical logic.

Formalism (in the sense of Hilbert), which holds that all mathematical reasoning can be 'reduced' to purely mathematical operations on symbols, was supported by the famous 'formalization of mathematics' in *Principia Mathematica*: this empirical discovery provided a formal language in which known mathematics could be expressed, and formal rules by which all theorems could be formally derived. Gödel was the first to analyze this precisely. He established it theoretically for a particular domain of mathematics, namely traditional 'logical' reasoning (concerning truth functions and quantifiers), in his *completeness theorem* (1930). In contrast, by his *incompleteness theorem* (1931), the view is false for reasoning concerning specific mathematical notions such as the natural numbers; in particular, he constructed a true statement, built up logically from equations between polynomials with variables for natural numbers, that cannot be derived by the formal rules of *Principia*. He indicated that the same method applied generally to formal systems (and this was proved as soon as the notion of formal system was analyzed precisely by Turing). This last result refutes the formalist conception as originally understood, e.g., by Hilbert. The methods used in the two theorems are essential for, respectively, model theory (i.e. the theory of general axiomatic systems) and recursion theory (i.e. the theory of mechanical operations).

Predicative foundations of set theory (associated with Poincaré and Russell): roughly speaking, on this conception one restricts oneself to sets that can be defined from 'previously introduced' sets. Gödel made this somewhat elusive idea precise by means of his notion of *constructible set*, and showed that the usual formal axioms of set theory are satisfied by the constructible sets (although these axioms were originally derived by Zermelo for another, namely Cantor's 'realist', conception of set). Gödel also verified that the constructible sets satisfy the axiom of choice, the generalized continuum hypothesis, and many old conjectures of descriptive set theory. As a corollary he obtained the (relative) consistency of the latter statements with the usual axioms of set theory.

The recent work of Cohen for which he received one of the Fields Medals in 1966 is, as he himself emphasizes, a delicate mathematical refinement of Gödel's theory of constructible sets.

When looking for a crucial test between predicative and realist foundations of set theory Gödel made two highly original suggestions, first in 1948.

1. The restriction to constructible sets may be inconsistent (with the usual axioms), when one adds socalled strong axioms of infinity which are valid for the more general notion of set (these are axioms asserting the existence of very large cardinals).

2. These axioms of infinity may be needed to decide open questions in ordinary analysis, and perhaps even in ordinary arithmetic. (The proof of the incompleteness theorem above shows that such axioms can be needed to decide some problems formulated in the language of arithmetic; but the problem so obtained is not of ordinary number theoretic interest.)

These conjectures dominate current research; 1 is supported by work of Scott and Rowbottom, 2 by results of Solovay on the definition of non-measurable sets of real numbers (in the sense of Lebesgue).

Gödel's researches on *constructive foundations*, though less conclusive, constitute some of the principal sources in the subject. Earlier and more clearly then anyone else, he realized the positive (non-restrictive) side of socalled intuitionism, which takes 'constructivity' in its most general sense. About 35 years ago he showed how to reduce classical arithmetic and elementary analysis to intuitionistic principles by means of a suitable 'translation', and more recently (10 years ago) he published a new 'functorial' interpretation of intuitionistic logical operations. The latter is widely used in recent work on constructive foundations of classical analysis.

A more isolated recent result of his is this: the existing rules for intuitionistic logic are incomplete (in contrast to his completeness theorem above for non-constructive logic).

Ever since Gödel's first paper 35 years ago, his influence on all parts of mathematical logic has been, and continues to be, most powerful, and wholly fruitful.

Gödel's reply

Dear Professor Kreisel,

Many thanks for sending me your statement for the Royal Society. In my view it is an excellent presentation of essential points. What I like particularly is that you emphasize the connection between philosophical opinions and precise mathematical questions, a connection which comes out in my publications. There are only a few places where I could suggest improvements in detail.

1. In connection with the refutation of formalism it could be mentioned (before the last sentence) that, by my results, consistency proofs by elementary-combinatorial means (in the sense of Hilbert) are demonstrably impossible.

2. As I showed later (cf. *The Undecidable*, ed. by M. Davis, N.Y., 1965, p. 73) the undecidable proposition is actually an extremely simple theorem concerning *one* diophantine equation.

3. The identification of 'predicative' and 'constructible' seems to me justified only if one can prove some kind of accessibility (from below) of all constructible sets (say, of natural numbers). Is this possible by means of Rowbottom results? If not, the relation between these two concepts should be formulated a little more cautiously; for example, as an expectation (by talking of 'predicative in its most general sense'; better still: 'in a generalized sense').

4. *I* should *not* say that Cohen's work is *merely* a 'refinement' of the theory of constructible sets (even if his work starts out from that theory). I believe you are right that Cohen has said something of this kind (even though I could not say this for certain); but this is in my view unfounded modesty.

5. I have *not* proved the incompleteness of intuitionistic logic. After all, I have not published anything on it, and rather conjectured the opposite. At most one could say that this result was achieved in pursuit of a suggestion on my side.[1]

With best regards, also from my wife.
Yours
Kurt Gödel.

[1] This refers to the proof(s) of incompleteness I published in the '60s. (G.K.)

Further developments

There is a detailed reply (by me) dated May 16, with the motto: Above all I wanted to avoid unnecessary doubts; now one can concentrate on precision.

There is a reply (by Gödel) dated June 9, which goes on forever. The following extracts may be of interest:

1. Finally, I should like to add that, on second reading I have found yet another passage in your statement that does not correspond perfectly to the historical facts. Originally I was *not* convinced that my results were valid for *all possible* formal systems. The passage in question should therefore run roughly as follows:

   He showed that his method is applicable (with the same result) to any formal system satisfying certain very general assumptions (and what he used in the proof turned out to hold for any *formal system as soon as ...).*[2]

2. Of course I do not wish to say that it is *necessary* to mention the result about consistency proofs by elementary combinatorial means. After all, what Hilbert actually meant or expected, is not of great importance in principle. But I wanted to let you know everything significant that crossed my mind.

3. It just occurs to me that the general conditions for my undecidability proof are formulated already on p. 190 of my original paper in the *Monashefte*. That my result were valid for *all possible* formal systems began to be plausible for me (that is, since 1935) only because of the *Remark* printed on p. 83 of *The Undecidable*, by Davis (cf. the footnote on p. 84 about its precise meaning). But I was completely convinced only by Turing's paper.

Warnings

It should not be forgotten that Gödel's final illness had already started in 1968. Consequently, the lack of concentration in his letter is not unexpected. Also, in his depressed state honours (like the election to a Foreign Membership of the Royal Society) were particularly significant for him. So he was unreasonably pleased with me; to an extent I did not foresee in 1966 when I proposed him (an extent that would have disturbed me).

Obviously, I stressed certain aspects of his results; without drawing attention to others (which might have, let us say, confused the issue; even though at the time - with high expectations of logic - his election was a foregone conclusion). However, I myself was not so aware of those 'other aspects'; for example, I had not yet had the surprise of my U.C.L.A. lectures on proof theory in autumn 1968,

[2]This sentence is (of course) in Gödel's English.

where it turned out that the results, which occupied most of the lectures, simply did not fit the general scheme (of *Beweistheorie*) announced at the start.

Chapter 6

Kurt Gödel

Kurt Gödel did not invent mathematical logic; his famous work in the thirties settled questions which had been clearly formulated in the preceding quarter of this century. Despite sensational presentations by crackpots, philosophers and journalists (or even in poems[1]), Gödel's results have not revolutionized the silent majority's conception of mathematics, let alone its practice; much less so than the internal development of the subject since then.

Certainly, those results refuted most elegantly each of the grand foundational 'theories' current at the time, of which Hilbert's (on the place of *formal rules* in mathematical reasoning) and those associated with Frege and Russell (on its reduction to *universal systems* like set theory) were most popular.[2] For obvious reasons, in his original publications Gödel made a point of formulating his work in terms acceptable to these theories, and to stress its bearing on them. But it is fair to say that they were suspect anyway and, less trivially, that they can be refuted more convincingly by simple *constatations* rather than by (his) mathematical theorems (as explained in more detail in Section 7). Further, as so often with very grand schemes, the refutations put nothing comparable in the place of the discredited foundational views which are, quite properly, simply ignored in current practice.

The first principal aim of this chapter is to restate Gödel's main results in the light of present knowledge, and hence independently of those foundational views. This is done in Sections 4, 6 and 9 by reference to two classes of *axiomatic definitions*, first discovered about a century ago, and familiar to anybody with an up-to-date elementary background in mathematics. Peano's and Dedekind's set-theoretic axioms for the natural and real numbers are typical of the broad

[0]Originally published in the *Bibliographical Memoirs of Fellows of the Royal Society*, 26 (1980) 149–224.

[1]For example, by H. M. Enzenberger, set to music by H. W. Henze

[2]Gödel's own and related results also deflate the particular 'anti-formalist' foundations of the time, Poincaré's and Brouwer's constructivist and Zermelo's infinitistic schemes being extreme examples. They are taken up, respectively, in Sections 7, 9 and 11.

class, elementary algebra and (for that matter) computer programs of the narrow class. The relation to the foundational theories is simple: each of the latter wildly overemphasizes the role of one of the two classes of definitions, and so completely misjudges both classes. Gödel's results establish the different potentialities of the two classes of definitions much more dramatically than had been done before. He did not go on to study just where those potentialities are actually useful. This was done by many others who, over the last sixty years, developed and, occasionally, applied the more successful branches of logic: model theory, recursion theory, and set theory (the latter not as a foundational system, but as a specialized part of mathematics).[3]

The second principal aim of this chapter is to substantiate Gödel's own view of the essential ingredient in his early successes, which solved problems directly relevant to principal interests of some of the most eminent mathematicians of this century, including Poincaré, Hilbert, Brouwer and Hermann Weyl. His view differs sharply from the impressions of many mathematical logicians who, over more than sixty years, have looked in Gödel's work for the germs of some exceptionally novel mathematical constructions or for previously unheard-of subtle distinctions, but not very convincingly. Without losing sight of the permanent interest of his work, Gödel repeatedly stressed - at least, during the second half of his life - how little novel mathematics was needed; only attention to some quite commonplace (philosophical) distinctions; in the case of his most famous result: between arithmetic truth on the one hand and derivability by (any given) formal rules on the other. Far from being comfortable about so to speak getting something for nothing, he saw his early successes as special cases of a fruitful general, but neglected scheme:

> By attention to (or, equivalently, analysis of) suitable traditional philosophical notions and issues, adding possibly a touch of precision, one arrives painlessly at appropriate concepts, correct conjectures, and generally easy proofs.[4]

In terms used by Kant:

philosophy analyses and mathematics builds up concepts

Gödel looked for a combination (where Kant saw only a distinction): for a given problem one may have the choice *between* a solution by means of philosophical analysis and easy mathematics *and* one by elaborate or otherwise subtle constructions. The simplest example is a solution by new axioms, discovered and justified

[3] Readers interested in the reaction of the logical community in the thirties to Gödel's results can find a most faithful description in Kleene [1976], and some of Gödel's comments on it in Kleene [1978].

[4] To be compared to the use of physical reasoning for developing mathematics or, on a smaller scale, to the use of geometry in algebra.

by means of philosophical analysis.[5] Evidently, Gödel's scheme goes counter to the wide-spread ideal of *Methodenreinheit* (purity of methods) in mathematics, made famous by Hilbert's successful use of it in geometry. With great determination and much imagination Gödel looked for other areas of knowledge where this kind of analysis would be rewarding, including the natural sciences (where, after all, Einstein had used such analysis so successfully that it remained a kind of ideal of theoretical science for decades). Section 12 covers this material.

It is clearly beyond the scope of this chapter to assess the value of Gödel's scheme in the arsenal of scientific methods, or even to compare it with the opposite (heuristic) view. But enough will be said of the singular state of foundations sixty years ago, heated up by dramatic 'controversies' over almost half a century, and of alternatives in the recent literature, to limit one's expectations.

Sources

Readers are warned that it has not been possible to take full account of the many papers, ranging from over 80 scientific notebooks to some exercise books from his schooldays, which Gödel left to the Institute of Advanced Study at Princeton. The latter, with the support of the N.S.F., made available microfilms of almost 5000 pages (partly in old-fashioned Gabelsberger shorthand), mainly from the very productive years 1938-1945. As a result it was possible to document most of the points I remembered from our conversations over more than 20 years.

But even the small part of his *Nachlass* that I have seen has altered completely my picture of his extraordinarily methodical working habits, about which he had been very reticent. For example, he has left a stack of envelopes full of library chits for books he borrowed and, presumably, read. Another surprising discovery was a bundle of drafts for lectures both on elementary and on advanced logic, written with love and care and relaxed precision, in a style different not only from his publications, but also from his letters and conversations.

Gödel himself was equally reticent about his personal history, but his wife talked more freely about it, usually in his presence. Section 1, which covers such matters, also uses material from a family history of Gödel's mother, written in 1967 by Dr. Rudolf Gödel, his only sibling, a year after her death, and supplemented in 1978. Some points of detail were cleared up by letters from Gödel to his mother (which his brother put at my disposal, and which are now at the Neue Stadtbibliothek of Vienna), and by documents from the Archives of the University of Vienna (which are now available to the public). Evidently, the parts of Section 1 which are based primarily on the memories of members of the family or of myself, will have to be crosschecked; not so much because of exaggerated discretion but because of 'the influence of the observer on the observation', close observers tending to have a lopsided view.

[5] Section 10 describes a specific proposal under the slogan: axioms of infinity.

6.1 Life and Career

Family background

Gödel was born on 28 April 1906 at Brünn in Moravia, then called the Manchester of the Austro-Hungarian Empire, and now Brno (since it became part of Czechoslovakia after World War I). Gödel's father Rudolf, whose family had come from Vienna, was an early 'drop-out', but practical and energetic. He became managing director and part-owner of one of the leading textile firms. He was almost six years older than his wife Marianne, whose father, Gustav Handschuh, had come from the Rhineland, where he had been a poor weaver, to find success at Brünn, also in textiles. The mother had a broad literary education, partly in France. But she was also a competent and imaginative *Hausfrau*, to whom both her children were very much attached. She was brought up as a Protestant, her husband was only formally Catholic, and the children received no religious training. Gödel's older brother has remained unmoved by religion. Gödel himself developed quite early unorthodox theological interests, had a life-long dislike of the Catholic Church, and a soft spot for the new sects, in the New World, of which he spoke often in conversation, and also wrote at some length to his mother, for example in a letter dated 18 March 1961.

Gödel's family cultivated its German national heritage; a bit self-consciously, as was usual among German-speaking minorities of the multi-racial Austro-Hungarian monarchy. Naturally, this continued after World War I, and is beautifully reflected in one of Gödel's essays (written during 1920–21 at the *Staatsrealgymnasium in Brünn mit deutscher Unterrichtssprache*) on the superiority of the austere life led by Teutonic warriors over the decadent habits of civilized Rome. Most of the family friends were very enthusiastic about the successes of Germany under Hitler. Gödel's mother (who apparently had happy memories of her school days at a French *lycée* at Brünn) is said to have been sceptical, almost alone among her friends and neighbours.

Growing up in Brünn and Brno (1906–1923)

Gödel is remembered as a generally happy, but rather timid and touchy child, unusually troubled when his mother left the house or when he lost a game. Around 1914, at the age of eight, concern for his health began to take up more and more of his daily life; the next paragraph gives only a bare outline.

At the age of six Gödel had a painful bout of rheumatic fever, but resumed a normal life after he got better. At eight (pretty evidently after reading about possible complications of the disease, in some medical book or other) he became convinced he had a weak heart. The conviction remained to the end of his life. Occasionally he developed some of the appropriate symptoms; for example, at the end of the sixties, more than 50 years later. He saw a well known heart specialist in New York. When the electrocardiogram and other tests were normal, Gödel felt frustrated - having overlooked that his particular symptoms were perfectly normal for anybody who worried about having a weak heart. This kind of oversight was by no means exceptional in his medical history. Some of Gödel's exaggerated reactions in later life, though surely going back to a natural predisposition, must have been reinforced by the peculiar difficulties of ill health in childhood. Examples of those reactions, ranging from excessive caution both in everyday life and in the presentation of his work, to distrust of the views of others (especially in medical matters) are sprinkled throughout this chapter. The caution and its frustrations go with the childhood coddling and the vicious circle to which the latter leads. The distrust goes with the logical trauma of listening to explanations by doctors and other healthy people (for example, of that vicious circle), especially for a very inquisitive child like Gödel whom the family called 'Mr. Why' (*der Herr Warum*). Be that as it may, the distrust was there, and delayed appropriate treatment of an ulcer in the forties when his life had to be saved by several blood transfusions; in his final years it aggravated the prostate trouble which he called 'weakness of the bladder', well known to be desperately depressing at best.

But most of his life he managed well enough. If preoccupation with his health limited his energies, he was also careful not to waste them, as his diaries show. His powers of concentrated work and sustained interest were evident already at school (as shown by his home work on geometry in one of the exercise books he kept, or his reputation never to have made a mistake in Latin grammar), and continued into the sixties when his wife still spoke of him, affectionately, as a *strammer Bursche*.

Incidentally, he came upon his first romantic interest without much waste of time: she was the daughter of family friends who were frequent visitors. She was regarded as an eccentric beauty. Because she was ten years older his parents objected strongly and successfully, apparently unimpressed by the neat balance between her age and his valetudinarian habits.

Vienna, with two interludes at Princeton (1923–1938)

As Gödel mentioned in conversation, he was originally undecided between mathematics and theoretical physics. The elegance of the three-year lecture cycle by the number theorist Furtwängler, a pupil of Hilbert and one of the founders of class field theory, tipped the balance. Another singular aspect of those lectures (which Gödel did not mention, possibly because of the medical history involved) may have had equal weight. Furtwängler was paralysed from the neck down, and lectured from his wheel chair without notes, while a scribe wrote the proofs on the board. This virtuoso performance was all the more spectacular because Furtwängler, like his cousin the famous conductor, had an exceptionally fine head.

But Gödel's principal teacher was the analyst Hahn, who was actively interested in foundations, and a member of the *Wiener Kreis* (Viennese Circle), a band of positivist philosophers around Schlick, who was shot and killed during a lecture in 1936. The meetings of the *Kreis* were held in a seminar room, off a corridor that led to the department of mathematics, and mathematicians tended to drift in and out of the meetings. Gödel attended more regularly. By a lucky chance it is possible to document what he later remembered as his (negative) reaction; by reference to the record [1931a] of a meeting on foundations organised by the *Kreis*, a few months before he discovered the incompleteness theorem. In it he gives a brilliantly succinct and precise analysis of the inadequacy of consistency as a sufficient condition for sound mathematics, contrary to formalist positivist doctrine. His analysis uses freely, almost ostentatiously just those concepts which are anathema to the doctrine - without a word about the latter, as if it were not worth mentioning. A year later, still only 25, he used similarly elegant tactics in a letter to Zermelo,[6] after the latter's criticism of the incompleteness theorem at the 1931 meeting of the German Mathematical Society (cf. Zermelo [1932]).[7]

In December 1932 Gödel's paper on incompleteness was accepted as *Habilitationsschrift*, as being well above the norm. In March 1933 he was made *Privatdozent*, unpaid lecturer, a title which was abolished in 1938 when Austria became part of Germany. A candidate was required to have either independent means (which was called *reich*) or a job, quaintly reminiscent of the rules for retiring

[6]Of 12 October 1931, reprinted in Grattan-Guinness [1979].

[7]Taussky [1987] contains her view of a meeting between Gödel and Zermelo at Bad Elster in 1931. Whatever (illusion of an) understanding may have been reached at the time, it is certainly no longer expressed in Zermelo's letter of 7th October 1931 to Reinhard Baer (quoted by Taussky). This letter does not, of course, settle the (perhaps more interesting) question of whether Gödel's brilliant exposition in his letter of 12th October 1931 (cf. note 14) was of more help to Zermelo than the apparently cozy walk near Bad Elster.

I know too little of Zermelo's personality, apart from the fact that it is very alien to me, to make further speculations on the details above rewarding. However, there is another side to the matter which can be viewed as an instance of an almost universal temptation.

As shown convincingly in his [1932], and elaborated in [1935], Zermelo had been (or, at least, had become) convinced, quite independently of Gödel's incompleteness theorem, that formal systems were inadequate for understanding logical reasoning (formal systems have recursive - that is, hereditarily finite - rules, and [1932] concerns the logic of the Infinite). Of course, Zermelo's view was shared by the silent majority. But, in contrast to the latter, he said something about his view (in [1935]).

In terms of Walpole's distinction, in Zermelo's place a man of thought would have used Gödel's theorem to strengthen the case against formal systems. They are not merely inadequate in the obvious sense already mentioned, of 'embracing all valid methods of proof', but even with respect to provability (of true Π^0_1 sentences). In contrast Zermelo, apparently a man of feeling, reacted to Gödel's theorem as if any contact with formal systems (even by way of negative metatheorems about them) was bound to be pernicious, like witches and other creatures of the devil.

Gödel himself is of course the perfect counterexample to such fears; cf. p. 118. But it should not be assumed that the man of feeling is statistically all that wrong about his many fellowmen (of feeling).

officers of the Austro- Hungarian army (a genuinely rich wife would do). Gödel's father had left the family comfortably off when he died in 1929, at the age of 54, from a painful abscess on the prostate. The mother moved to Vienna, took a large flat, and shared it with her two sons till 1937 when she returned to her villa in provincial Brno. Rudolf, the elder son, was already an active and successful radiologist without being wholly absorbed in his profession. The mother was enchanted by the theatre in Vienna where her long-standing literary interests were brought to life, and the sons went with her. And if Gödel preferred musicals, as he did all his life, he was very willing to form opinions on Art and Literature, and to defend them energetically, especially when they were unorthodox. - Though his work was quickly recognized in Austria and abroad, at home among his family he always went out of his way to 'hide his light under a bushel', as his brother put it.

After Gödel's first visit to Princeton (1933–34) he had a nervous breakdown. It began with severe anxiety when he got off the boat. (He telephoned his brother from Paris, who almost went to meet him there.) Wagner-Jauregg was called in, a Nobel Prize winner, and at the time perhaps even more famous than Freud, at least in Austria. No indications of psychosis were found. But there were two frustrations, each perhaps sufficient to trigger a breakdown in someone of Gödel's personality. More than twenty years later he still spoke of the frustrations of (tacitly) his bachelor life in Princeton where he had just spent a year. The second stress awaited him in Vienna. At 21, a couple of years before his father died, he met Adele Porkert at a Viennese night spot, *Der Nachfalter*. She had ben briefly married before, and was six years older than Gödel. Once again his parents, especially the father, objected. In fact, Gödel did not marry Adele till 1938.

I visited them quite often in the fifties and sixties. It was a revelation to see him relax in her company. She had little formal education, but a real flair for the *mot juste*, which her somewhat critical mother-in-law eventually noticed too, and a knack for amusing and apparently quite spontaneous twists on a familiar ploy: to invent (at least, at the time) far-fetched grounds for jealousy. On one occasion she painted the Institute for Advanced Studies, which she usually called *Altersversorgungsheim* (home for old-age pensioners), as teeming with pretty girl students who queued up at the office doors of permanent professors. Gödel was very much at ease with her style. But this is not all: in a sense the principal logical theme of this chapter goes back to her banter. She would make fun of his reading matter, for example, on ghosts or demons (but never of pages of logical formulas which have their funny side too, if she only knew). Quite often, the topics she mentioned explicitly fitted perfectly what I had read between the lines in his publications without paying attention; for example, to ghosts and universes with cyclic time considered in [1949] and [1949a], and further discussed in Section 12 below. Since I had noticed the connection spontaneously, presumably showing the pleasure which goes with this kind of *Aha-Erlebnis*, he found it worthwhile expanding on it; in a totally natural style, fully and freely - very much in contrast to his almost staggering responses, logical slaps in the face as it were, when he felt in duty bound to have an opinion on uncongenial matters. As already mentioned on p. 51, his wife's conversations also shed light on his personal life or, at least, suggested how to find out more about it.

Breaking the Austrian connection (1938–39)

Gödel, by and large, had the political views which were standard in his youth, in his immediate surroundings and in large parts of Central Europe. America was the land of opportunity, Germany was efficient, Austria *schlampig*. But granted all this, his aversion, after World War II, to Austrian academic institutions seems out of all proportion, and remained a total puzzle to his family (as documented, for example, by his mother's letter of 28 January 1963 to her brother Karl). He was offered, and refused, sometimes for mind-boggling reasons, membership and later honorary membership of the Academy of Sciences in Vienna, and the highest national medal for science and the arts. (He had no chance to refuse an honorary doctorate of the University of Vienna, since it was awarded posthumously.) He had accepted other honours, and was to accept more. For example, he was delighted to be a Foreign Member of the Royal Society, although England remained *Perfidious Albion* for him. And if the Academy of Sciences of Vienna is not of quite the same level, neither is the American Academy of Arts and Sciences, and he was a member.[8]

The story is not heroic, but it is beautifully coherent. Gödel was a most remarkable logician, he never pretended to be a dashing hero; nor was he impressed by heroes. (He admired General Eisenhower, while his wife was a great fan of General McArthur.)

When Austria became part of Germany in March 1938, he was not made *Dozent neuer Ordnung*, (paid) lecturer of the New Order, in contrast to most university lecturers who had held the title of *Privatdozent*. He was thought to be Jewish. (For the same reason he was once attacked in the street by some rowdies whom his wife chased off with her handbag.) He was convinced that nowhere except in Austria could there be such a *Schlamperei*, such a careless error. As he told me, he left Austria for Princeton (crossing Russia on the Trans-Siberian Railway) at the end of 1939 because he did not wish to be conscripted into the German army. Of course he felt he was not physically fit for military service; but given the evidence he had of *Schlamperei*, the risk was too great.

However, by and large, life went on smoothly for him in Austria during the spring and the summer of 1938: according to his diaries, he worked actively,

[8]He was also made an Honorary Member of the London Mathematical Society in 1967, a Corresponding Fellow of the British Academy in 1972, and was a corresponding member of the Académie des Sciences Morales et Politiques.

read widely, and travelled to Göttingen to lecture on his work in set theory. In autumn, after the Munich agreement, he married. He spent the first term of 1938–39 at the Institute at Princeton, the second at Notre Dame (where he prepared some beautiful lectures). He returned to Austria in spring 1939. In short, his misfortunes in 1938–39 were minor compared not only to what went on more or less quietly around him, but also to the much publicized hardships during popular uprisings (*Volkserhebungen*) of the past, like the French or Russian revolutions.

The fact is that he was bitterly frustrated. Once again, despite great care he had not escaped trouble. Specifically, in the words of the *Dozentenbundsführer* (in a letter of 30 September 1939 concerning Gödel's application of 25 September for a *Dozentur neuer Ordnung*), Gödel was not known ever to have uttered a single word in favour or against the National Socialist movement, although he himself moved in Jewish-liberal circles (and though the letter acknowledges mitigating circumstances, it neither supports nor rejects Gödel's application, which was accepted on 28 June 1940). Incidentally, the *Schlamperei* may have added a touch of insult to injury, if something was still left of the views in his essay on Teutonic warriors mentioned on p. 52: certainly, most of the essays already reflect perfectly the views he held all his life.

A bit more courage or highmindedness might have reduced Gödel's bitterness about his particular predicament. But, as the fate of his mother shows, even those commodities were not enough to ensure a cool head at the time. Till 1944 she stayed in her villa in Brno, openly critical, losing most of her former friends, and worrying her son, Rudolf, who was running the X-ray department of a hospital in Vienna. By 1944 both expected the defeat of Germany. She had had a good offer for her villa, toyed with it, but did nothing despite her almost daily criticisms (of the National Socialist regime): in effect, she did not expect reprisals by the Czechs after the war, not even confiscation of German property, let alone the deportations. Fortunately, she herself moved to Vienna, but not by calculation. She happened to be there with her son, there was a heavy raid, and they simply wanted to stay together. After the war the Austrian government negotiated with the Czechs, and according to the treaty the mother got the usual, inadequate compensation for her villa, one tenth of its assessed value. The fact that the same rate was almost universally applied to confiscation by the Germans was quite irrelevant for Gödel (since, logically, two wrongs do not make a right), and he never got over the injustice to his family. He himself was always most punctilious, and incidentally helped his mother as soon as possible.

The New World: the first 30 years (1939–69)

Gödel was well prepared to like America, given his general views and his particular resentment against Austria and its bureaucracy (in particular, the academic bureaucracy, which he knew well). Almost every letter to his mother between 1946 and 1963 which I have seen contains some variation on this theme. He became a U.S. citizen in 1948. He was especially attached to the Institute, of which he was an ordinary member till 1946, and a permanent member till 1953 when (at the age of 47) he was made professor. In a touching letter, of 25 March, he tells his mother that he would not have any lecturing duties though the salary was even higher than at universities. (He had the illusion that he was expected to have opinions on all details of the Institute business.) He saw a good deal of von Neumann, who is said to have astonished his first wife on their honeymoon in Vienna, in the early thirties, by the long hours he spent with Gödel talking about mathematical logic and foundations.

In the forties, except while weakened by an ulcer (and his own treatment of it, as mentioned on p. 53), Gödel worked with great intensity. A turning point was his wide-ranging essay [1944] on Russell's mathematical logic. It collects together a number of incisive points, most of which are formulated in a more relaxed style in his unpublished notes from the thirties mentioned on p. 51 (and used below). There are also some quite different, and much better known points (reviewed in Section 11), for instance those that have led to the label: Gödel's platonism. He could use [1944] to take stock of his whole logical experience without the slightest trace of self-indulgence: Russell's writings touched on every issue that could conceivably cross anybody's mind. Having thus arrived at his mature (heuristic) views sketched on p. 50, the time was ripe for Gödel to apply them outside the narrow area of mathematical logic too.

The place, the Institute, was right for an excursion into the general theory of relativity. Einstein was there and Gödel, perhaps more than most, was impressed by Einstein's singular success in using philosophical analysis for (presenting) his special theory of relativity; with a bit of luck, 'singular' would allow for repetitions. Einstein was enchanted by Gödel's combination of elegance and precision, and they saw each other constantly till the death of Einstein. It may be difficult to decide how Gödel's work on general relativity (described in Section 12) was influenced by their conversations (as so often when a decision has few consequences, and so, practically speaking, does not matter). At any rate, one can be sure that Gödel would not have brought up the subject before he had something new to say. Gödel's mother was overawed when she heard of the friendship, and began to read about Einstein. In a letter of 8 January 1951, Gödel recommends her not to be afraid of abstractions in Einstein's expositions, and not to try to understand everything at a first reading, but to go about it as she would read a novel.

In the early fifties Gödel's achievements began to be formally recognized: by honorary doctorates at Yale and Harvard, the Einstein Award (split with Schwinger), the Gibbs Lectureship of the American Mathematical Society. In 1955 he was elected to the National Academy of Sciences.

As far as the next 15 years or so are concerned, it is doubtful (and certainly impossible for me to decide) whether my picture is representative of his principal interests; I met him in autumn 1955, and remained in close contact with him till his illness at the end of the sixties. But what I know is sufficient to correct two widespread impressions, namely that:

1. though courteous, he lacked sensibility and warmth

2. his conduct of the Institute business was impenetrable.

In connection with 1, I myself witnessed a degree of understanding (whether intuitive or as a result of reflection) which is exceptional by any standards. Before I met Gödel I was of course impressed by the clarity of purpose shining through every line of his, but not carried away, mainly because it seemed to me (and to Gödel in 1930–31, as he told me later) that at the time it was a matter of months before somebody would stumble on the completeness and incompleteness theorems, his most famous results.[9] Worse still, I was simply put off by his general essays [1944] and [1947] (particulary by the most widely quoted passages, mentioned on p. 58), and I made no secret of the fact. With patience and unerring judgment Gödel led the conversation to points of common interest. In no time I saw for myself the many civilized passages of [1944] and [1947], which are hardly ever quoted. In due course I even went back to the offensive passages, and saw them in a different light, particularly in connection with so-called intuitionistic notions (described in more detail on p. 90).

Later, a different obstacle appeared, as so often when things are going too well. Given common logical interests and (as readers may have guessed on p. 53) a touch of hypochondria also on my part, there were exchanges on those minute reactions, to bugs or drugs, to which doctors will not even pretend to listen. In the after-glow, the conversation occasionally strayed to Gödel's general views on men and events and his all-pervading distrust. Another set of impatient questions: Did he expect me to find, behind his actions, the kind of devious motives he saw in others? Was he not frustrated to let others govern the world, since he knew so well what was good for it? (and almost in the same breath) How well did he know the world, since he was constantly surprised by what happened? Again he, and he alone, helped: apparently without a trace of resentment or even irritation, he avoided general topics (until his illness). At the same time he continued to ask me about my own doings and preferences, with a convincing mix of curiosity and personal sympathy. I remember only one occasion when I reciprocated, one evening when both he and his wife were in particularly good form. Since they so clearly liked being hospitable, why did they not have (other) guests more often? Gödel had noticed that most people showed more excitement in company than they felt, and he found this very tiring. Clearly, at times he needed very few data to reach, painlessly, a very sound conclusion.

[9] In those days I was more impressed by the 'broad sweep' of Hilbert's program, and especially by Herbrand's originality in logic (whose theorem was a much subtler business: it was not even properly understood or used for many years).

In connection with 2, especially in his selection of logicians for temporary membership at the Institute, his practice followed quite simply from his general heuristic principles explained at the beginning of the memoir: he gave preference to applicants whose work used (at least implicitly) or was likely to use philosophical analysis. He tried to judge this by reading their publications repeatedly, but generally not carefully. He seems to have been pretty successful. Besides this 'long-shot', of philosophical analysis, he also encouraged others; for example, the filigree work classifying sets of natural numbers by so-called Turing degrees: he thought it might suggest new ideas in cardinal arithmetic. In the fifties he looked, in vain, for logicians interested in the partition calculus of Erdös and Rado. (Given that the mathematical interest of logic, especially of its elaborations, is marginal, his encouragement of a few long-shots was reasonable anyway, and Gödel never pressed for having a horde of logicians at the Institute.)

Once he had made a selection, he avoided contact with people who were not temperamentally congenial to him; particularly introverted, tongue-tied, and generally affected personalities made him uncomfortable. He was fond of keeping pests at a distance by means of ambiguous remarks reminiscent of de Gaulle (*Messieurs, je vous ai compris*); for example: it would be interesting to see the work in print. He never edited any journal. Presumably, he did not usually give his simple reasons for his selections. After all, he always stressed the conflict between his views and the *Zeitgeist* to which (naturally, without empirical checks) he supposed his colleagues at the Institute to be subject. He was more disappointed than he let on by his occasional failures to persuade them; but not nearly as much as he would have been had he realized that he was battling a *Zeitgeist* from another time, the early thirties; and then not what it was, but what the *Wiener Kreis* would have wanted it to be.

The final years (1969–1978)

The events during this period would have unsettled Gödel at his best. His wife suffered two strokes and a major operation. There were (obviously interrelated) changes for the worse in America and at the Institute, the country and the institution to which he was so much attached. For example, student radicals were making headlines, and (admittedly, less charismatic) professors at the Institute could hope to make, at least, the correspondence columns of the *New York Times*: an issue was bound to present itself, and did. (This appeal to the *Zeitgeist* was not congenial to Gödel.) More subtly, there was a general air of despondency among the large number of able but jobless young mathematicians who were herded together at the Institute, constantly talking to each other, and so reinforcing each other's illusions about clever tactics for getting a job.

But the decisive factor was his own illness, mentioned already on several occasions. This is not the place to give a detailed medical history which, however, will be essential for a correct interpretation of what he said or wrote during those years. The particular character of the self-doubts which go with even mild prostate trouble are well known: usually there is a grain of truth, but magnified out of all proportion. This spoils completely the victim's perspective of his work over the years.[10] Superficially, at least in the early seventies, the changes appeared minor to those who had not known him well. After all, his mind remained nimble; only his exquisite sense of discretion had obviously gone. Perhaps as a result he was more gregarious than before; less formidable, as a perceptive secretary at the Institute later said. Even if this brought him some solace, it did not seem to me to go very deep, and accounts of his close family since his death have more than confirmed this impression. Actually, several of us who knew him well were alarmed already at the end of the sixties: his efforts not to show his depressions were evident, and soon became too much for me to watch. There were some bright spots: the U.S. National Medal of Science in 1974, after an honorary doctorate in 1972 from Rockefeller University, which gave him pleasure. In 1967 he had received one from Amherst College.

Gödel died, sitting in a chair in his hospital room at Princeton, in the afternoon of 14 January 1978.

6.2 Gödel's First Results in Focus

Gödel's first two famous results, which appeared in [1930] and [1931] about sixty years ago, concern *formal rules* or, as we should now say, computer programs. Put simply, [1930] establishes the 'positive' result that Frege's rules for elementary logic (of truth functions and quantifiers) proposed some fifty years earlier, generate exactly the logical truths in the precise mathematical sense corresponding to Leibniz's truths in all possible worlds. [1931] shows that the rules of *Principia Mathematica*, and in fact those of a large class of 'related' systems, do not generate exactly the arithmetic truths (among the formulas of *Principia* built up logically from polynomial equations with integral variables and coefficients). Even without going into refinements of the statements and proofs and leaving more ethereal foundational schemes for later, readers will imagine easily the striking implications of these simple, memorable results.

Historical perspective

100 years ago, [1930] would have had the glamour: simple mechanical rules can be proved *mathematically* to replace logical reasoning (at least its results, not necessarily the details of the process), and logic is about all possible worlds, so to speak, the height of abstraction! At that time, [1931] would merely have ratified the general impression that arithmetic is too difficult to be formalized (another

[10] Except for pp. 105 and 118, Gödel's views in the seventies quoted below correspond to earlier publications, notes or conversations.

word for 'mechanized'); after all, the diophantine equation $x^2 = y^3 + k$ is hard enough.

Today, Frege's rules (even without [1930]) still stand out as the first convincing example of non-numerical data processing by mechanical means. Examples of simple mathematical proofs, as in [1930] and [1931], showing what can or cannot be done 'in principle' by such means, are obviously essential for orientation and, at least occasionally, useful in practice, provided they are used with discretion and imagination. For realistic expectations this should be compared to the use of whole numbers in place of formal rules and of elementary theorems about them, where much skill is needed to find properties studied in higher number theory which are really significant for the bulk of scientific or other uses of whole numbers. It would not be hard to work up a parallel between [1931] and the irrationality of $\sqrt{2}$ in the uses of formal rules and of whole numbers respectively (cf. Chapter 7).

But in between, at the time of [1930] and [1931], the latter had all the glamour. For one thing, *Principia* had claimed to provide great weight of *empirical evidence*, in three heavy volumes, for the possibility of formalizing 'all' of mathematics, and certainly arithmetic. What is more, the claim was widely accepted including even Russell's contention that only empirical evidence, taken from mathematical practice (as codified in texts, etc.), was relevant. [1931] was shocking, especially if one glanced at the proof. *Principia* had left out an obvious type of argument which reflects on its own rules (and implies in a simple way a certain true arithmetic statement that cannot be derived in *Principia* at all): *Principia* had proposed a mathematical model of a certain phenomenon (mathematical practice) and had forgotten to look at the mathematical properties of the model itself! A moment's thought makes [1930] almost as disturbing as [1931] for Russell's doctrine of empirical evidence. What was the difference in the 'degree of confirmation' of the claims of *Principia* as far as logic and arithmetic were concerned? Anyhow, what was the claim? To describe (and perhaps to perpetuate) the defects of current practice, or to find out something about the potentialities of mathematics and the mathematical imagination? And was *Principia* any worse than what, for example, is done in studies of non-mathematical reasoning, by linguists and the like?[11]

Philosophical perspective

Returning to the aims of *Principia*, mathematicians had lapped up the idea of beginning with a formalization of all of mathematics. For example, Bourbaki's treatise starts with a chapter on set theory: not exactly *Principia*, but [1931]

[11]Incidentally, in Section 12 several examples will be given how reflection on [1931] and on its development in mathematics throws light on various arguments in the natural sciences too, the kind of thing one expects from a useful philosophy of science.

applies too. In their manifesto (Bourbaki [1948]) they get round to asking themselves about the point of this enterprise, and conclude that it is 'the least interesting side of the matter', or that formal rules of inference serve for 'logical hygiene' (rarely applied, since the rules are barely quoted later: more like ritual ablutions). If anything, [1930] serves as logical hygiene in giving a logical justification for the choice of the formal rules! Obviously, the notion of *set* is here to stay; but there is not a shred of evidence in Bourbaki [1948] that the ritual of giving formal axioms and rules for sets is of effective use in the later development (more effective than a description of the intended notion; for example, as in Sections 8–10 below).

To anticipate: since 1948, modest but sound answers have been supplied by mathematical logic to the questions implicit in Bourbaki [1948]. The possibility of defining many mathematical notions and problems in elementary terms has found uses, forshadowed in [1930] by the so-called finiteness theorem; and derivations built up by elementary rules are easy to unwind (cf. p. 86 for details). As to [1931], incompleteness results explain quite well why certain questions (for example, about groups) have not been settled yet, though more difficult arguments than those of [1931] are needed. More positively, just because of incompleteness we know more if a theorem can be formally derived by given rules than if it is merely true and, perhaps less obviously, we know more if a (true) theorem can be derived by given rules, but not by a subset of those rules. As always, the discovery of the terms in which this additional knowledge is to be expressed in a principal part of research; successful examples are to be found on p. 74. In short, slowly, the early ritual is becoming a scientific tool.

But also, and this is much more striking, the tools found are pitiful compared to the original expectations associated with mathematical logic. Specifically, Boole [1854] looked for the *laws of thought* in propositional algebra, and Hilbert [1930] thought that he had found the laws in his own favourite rules (a mind boggling exaggeration since, as already mentioned, even the positive result in Gödel [1930] concerned only results, not the details of reasoning, treating the latter as a matter of black boxes). Then there was the retreat to logic as providing a *standard of rigour*, an 'ultimate' criterion for checking proofs, the 'hygiene' which is not applied (in fact, one applies more often interpretations, clever cross checks, to verify formal derivations).

The development of logic since Gödel [1930] and [1931] has moved away from the aims mentioned; in particular, soon after [1931], the emphasis on formal rules for the special purpose of building up derivations and representing proofs was quietly dropped, as reflected in the terminological change from

formal undecidability of a particular problem $P_{\mathcal{F}}$,

(depending on the formal system $\mathcal{F}$ under consideration) used in [1931], to

recursive undecidability of a class of problems (including $P_{\mathcal{F}}$).

A readable account of this matter is in Davis, Matyasevic and Robinson [1976]. More about the whole matter of representing proofs is to be found later in this chapter.

We leave this disturbing side of [1930] and [1931] with the few snippets above. Of course the latter do not convey even approximately the bearing of [1930] and [1931] on the ideas current at the time, let alone on the principal people active in logic. Russell, Hilbert and Brouwer were not narrow specialists: they were fascinated by the turmoil of ideas current during the first three decades of this century, a very special period in the development of science. There was an unbounded confidence in high theory, as already mentioned in Section 1 in connection with Einstein. There was progress with understanding phenomena where, previously, one just did not even know where to begin; and so Kant's odd question how this or that experience was possible at all (*überhaupt möglich*, instead of the ordinary scientific question, what things are like) seemed appealing. And, last but not least, there were extraordinary successes of building up the physical world (or, at least, matter) from a few particles; so why not mathematics and mathematical reasoning from a few primitives (set and membership) and a few rules of inference? Nothing remotely like existing logic is even a candidate for an analysis of mathematics or mathematical reasoning comparable in scope to those successes in the natural sciences.

Accentuating the positive: purity of methods

Gödel's results and even Hilbert's conjectures (which were refuted so simply that they have been described as 'blind spots') appear in a totally different light if we go back to the last century, to what even now are *Aha-Erlebnisse*. Two of them were already mentioned, namely set-theoretic (or: broad axiomatic) definitions of familiar structures by Peano's and Dedekind's axioms, and Frege's formal rules for elementary logic (in the precise sense explained on p. 66). The third is the exposition of geometry in Hilbert [1899], with striking examples of a mathematical scheme for choosing a formalization (in contrast to the business about empirical evidence in *Principia*).

Today the principle of choice is better illustrated by considering the ordered field of the real numbers instead of geometry, passing from *arbitrary* Dedekind cuts to those defined by *elementary* formulas (about ordered fields), and thus to a natural (if not very well-known) axiomatization of so called maximal ordered fields. In the context considered, the reference to arbitrary sets or cuts could really be described (by Hilbert) without exaggeration as a mere *façon de parler*: as far as results (and, at least at the time, also proofs) were concerned, Dedekind's arbitrary cuts gave no more than those defined by elementary means. Hilbert was quite conscious of the obvious relation between this discovery and an age-old ideal of *Methodenreinheit*, as he stressed inthe peroration to Hilbert [1899]; 'age-old' in that it goes back to the time of the Greeks, when Archimedes was criticized for using properties of space to prove theorems about the plane (cf. Knorr [1978]).

For elementary theorems, you use elementary cuts. Number theorists will think of heated but inarticulate arguments about impure methods (analytic number theory at one time, l-adic cohomology now).[12]

From this point of view, Gödel's paper [1930] establishes that logical purity can be achieved in principle, and [1931] that arithmetic purity cannot be achieved; in fact, the result in [1931] is so general that it is quite insensitive to any genuine ambiguities in the notion of purity of method.

Legalistically, Gödel's papers only settle questions about the possibility of purity of method. But inspection of the arguments suggests quite strongly that *the whole ideal of purity of method is suspect, even when it can be achieved*. (As will be seen in Section 10, Gödel turned the ideal upside down, wanting to prove finite combinatorial theorems by use of properties of very large infinite cardinals.)

In any case, today there are plenty of examples in ordinary mathematics where impure methods are employed: the *restriction* to pure methods has to be 'justified' (when it is appropriate at all), at least as much as the use of impure methods. The most familiar reason for restricting methods of proof is the greater generality of the theorems proved, their validity for more cases (of interest). Trivially, where purity can be achieved, the essential difference between pure and impure proofs cannot be analysed in terms of validity. But

> the validity of a theorem (in fact, the validity of a proof) is only a small part of the significant knowledge contained in the proof: it just happens to be the part which is most easily put into words.

And if that part is regarded as the specifically *logical* aspect of proofs, then logic is marginal for understanding the actual phenomena of proofs. As already anticipated on p. 63, in practice (if not in rhetoric) this conclusion has been accepted, and new aims (mentioned there) are pursued.

6.3 Background to [1931]: Axiomatization and Formalization

Although the axiomatic tradition goes back to Euclid, it was changed radically about 100 years ago by two new methods (and aims).

[12] Incidentally, though this was not stressed by Hilbert himself, his later and much more famous consistency program is also a particular case of this search for pure methods: so-called finitist theorems should have finitist proofs (of which old-fashioned school mathematics is typical).

A neat, but purely technical observation of Hilbert was that this aim is assured under suitable conditions if the *formal consistency* of a system $\mathcal{F}$ is proved finitistically, the aim being now restricted to finitist theorems derived in $\mathcal{F}$ itself (cf. p. 75 or, for a more pedantic exposition, the section on Hilbert's second problem.

First, by use of the notion of *set* (which had just become prominent through the work of Cantor) familiar objects together with (what are regarded as) their principal features, could be *defined* axiomatically (as one says: up to isomorphism). The still most famous examples are Peano's axioms for the natural numbers with the successor relation as principal feature, and Dedekind's for the (ordered field of) real numbers; but cf. also Zermelo's axioms in Section 8 for segments of the so-called cumulative hierarchy of sets. This use of axioms (as definitions) distinguishes them from Euclid's, which were not intended to be, and are not definitions unique up to isomorphism, since they are satisfied both by the full (uncountable) plane and by the part consisting of points constructible from two points by means of ruler and compass. In modern terms, the new axioms use a richer, so-called *non-elementary* language; in contrast to Euclid, arbitrary *subsets* of the sets (of numbers) involved are used to state induction and completeness (for Dedekind cuts) in, respectively, Peano's and Dedekind's axioms.

The second new element was introduced by Frege, his famous *formal rules* (of inference). They were intended as an analysis or 'definition' of *logical deduction from axioms* (more precisely, as we realize now, from *elementary* axioms) built up from relations (between the objects of some domain D) by means of the logical operations $\neg$ (not), $\wedge$ (and), $\vee$ (or), $\rightarrow$ (implies), $\forall x$ (for all elements of D), $\exists x$ (for some elements of D). In particular, such elementary axioms do not use the new non- elementary quantifier: for all sub*sets* of D, needed for the definitions in the last paragraph. Systems of elementary axioms together with Frege's rules are called *formalizations*.

Realistically speaking, neither the new definitions nor the new rules were needed for mathematical practice at the time (nor before: the *Disquisitiones* of Gauss would not be improved by starting with Peano's axioms, or by writing the proofs of the law of quadratic reciprocity in Frege's formalism.) But clearly there was a raw interest to the two enrichments of the axiomatic tradition. It fired the imagination of mathematicians and philosophers. Readers can well imagine how the surprisingly compact definitions (in the language of sets and logic) of Peano and Dedekind made them into ideals for all definitions in mathematics, and how Frege's simple rules led to wild exaggerations about the laws of thought, mentioned on p. 63. For all we know, these exaggerations served as a useful body guard, protecting the new interesting methods until their significance was discovered too.

One of the first convincing indications of significant uses is to be found in Hilbert [1899]: the use of non-elementary axiomatizations for a systematic choice of formalizations, already alluded to on p. 66. To be precise, the passage involved was not explicitly formulated by Hilbert, but fits very well his work on the foundations of geometry, where Dedekind cuts turn up as non-elementary axioms of continuity, which explains the connection between Hilbert [1899] and the exposition below.

From non-elementary axiomatizations to formalizations

The passage is best illustrated by the step from Dedekind's axioms to a very natural formalization of elementary real algebra, known in the trade as the theory of *ordered real closed fields*. The principle is to restrict cuts to those defined by elementary formulas about ordered fields, instead of arbitrary cuts; this is expressed by an *infinite* axiom schema corresponding to each formula. No other change is made, since the rest of Dedekind's axioms are elementary anyway. Incidentally, real closed fields were not considered by Hilbert himself, but were stumbled on 'empirically' in the twenties by Artin and Schreier.

Though there are many real closed fields (for example, of all the real numbers and of the algebraic real numbers), every elementary proposition which is true in *one* such field is true in *all* the others. This was established by several logicians around 1930, including Tarski and Herbrand, but also Gödel who, as he mentioned in conversation, did not publish the result when he learnt that Tarski had found it independently. They showed that all elementary formulas F (without free variable) about those fields are *decided*; that is, either F or $\neg F$ is derivable from the axioms by means of Frege's rules. In fact, for each F, a *finite* subset S_F of the infinitely many axioms is determined which is sufficient to decide F. Equivalently, if F is true for the field of real numbers then F follows logically from S_F: the formalization is *complete* (in this sense).[13]

Peano's axioms also illustrate the passage to formalizations, but with an added twist on the choice of 'principal features' of the structure considered. Apart from equality, Peano's axioms mention only one relation (say S) for the *successor*. So, taken literally, the passage leads to the successor axioms and induction restricted to (elementary) properties defined from S alone. Again, this formalization decides every elementary formula (about S); but precious little can be expressed about the natural numbers in this way.

Substantially more is expressed by elementary formulas about *addition* (for example, about congruences). Here the passage starts with Peano's non-elementary axioms together with the usual recursion equations for $+$ (in terms of S and $=$), which define addition implicitly. Again, the resulting formalization decides all its

[13] *Remarks for specialists.* First, logically less complicated cuts are sufficient, namely those defined by (the least zero of) polynomials of odd degree and the (lesser) square root of positive elements.

Secondly, the famous result in Milnor [1958] on division algebras over $\mathcal{R}$ conveys the flavour of the implications of the facts above. Thus the result of Milnor [1958] is true for all real closed fields. But nobody has developed K-theory in that context sufficiently, and the only known proof of the general result uses the transfer principle mentioned above. Again, the fact that there is no division algebra of dimension 16, is expressed by an elementary formula, say F_{16}. By the finiteness principle, for suitable N_{16}, F_{16} holds automatically for all fields in which all polynomials of odd degree $\leq N_{16} + 1$ have a zero, and positive elements have square roots. Incidentally, the least N_{16} is not known; a bound for N_{16} is part of any pure derivation of F_{16} from the formal axioms.

formulas.

Unquestionably, Hilbert expected similar complete formalizations for additional number-theoretic functions defined by recursion equations (for example, *multiplication*). Now the expressive power of the formalism is considerable: every diophantine problem can be stated.

Methodenreinheit: how to test philosophical ideals

The three formalizations in the last subsection for ordered real closed fields and for arithmetic of the successor relation and of addition, fit perfectly Hilbert's ideal of purity (on p. 51): to settle elementary problems, one does not need arbitrary cuts or sets, but only cuts defined by elementary formulas, and only elementary instances of induction, built up logically from the relations used to state the problems.

The *logical* question is to settle to what extent purity of methods can be achieved: in all of mathematics, parts of mathematics, in logic or metamathematics itself. But this leaves open the *philosophical* question whether purity of methods is at all basic (in the sense of fundamental) to mathematical knowledge, the sort of thing one cannot know too much about.

If purity is not basic then work done with this ideal as principal aim will have to be reexamined under the maxim: *dégager les hypothèses utiles*, appropriate to the assessment of tools. The discovery of good uses (as in note 13) becomes a major problem in contrast to the study of fundamental laws, which can be relied upon to have applications.

Defects of ideals are generally seen most clearly in areas where they have been realized, and so the results can be compared both with earlier expectations and with alternatives (which violate the ideal in question). In the cases under discussion (algebraic and number-theoretic purity), plenty of comparisons are available since, with time, impure proofs have become more common in practice, not less. Moreover, and this is often neglected:

1. their actual reliability or 'security' is obviously unaffected by the possibility of pure proofs, if that possibility has not been realized

2. impure methods are not only used heuristically, for discovering conjectures and proofs, but have turned out to be essential for *checking* proofs.

Far from being a mere aberration, the neglect of 1 and 2 is typical of what happens in the kind of intellectual void left by the questions asked in Section 2. First of all, the *unproblematic uses of formalization* (or, generally, purity of methods) have not become widely known; so there is a tendency to thrash about for some uses, and the easiest thing is to cling to dubious doubt which are to be removed by formalization, as in the business of logical hygiene. But also, there is the void created by simply not saying out loud *what (knowledge)* *is gained by impure proofs* (for example, by analytic proofs in number theory: knowledge of relations between the natural numbers and the complex plane or, more fully, between arithmetic and geometric properties). It is precisely this knowledge which provides effective new means of checking proofs: if this conflicts with some ideal of rigour, so much the worse for the ideal (which is being tested).

In short, the whole matter of formalization and purity of method is just much subtler than suggested by generalities about mathematical rigour (however persuasive the latter may be at first glance; cf. Section 13). A corollary to this observation is, of course, that *the significance of Gödel's incompleteness theorem is a subtle business* too. For if we do not restrict ourselves to complete systems even when they *are* available (as in the examples of algebraically or number-theoretically pure methods in the previous subsection), then incompleteness has lost its apparent philosophical sting: since its raw interest is clear, it is a *problem* to analyse its interest(s), philosophical or otherwise.

6.4 The Incompleteness Theorems

Below, reversing the historical order, Gödel's work in [1931] on incompleteness will be presented first, because it requires less background on 'abstract nonsense' (about logical validity) needed for [1930].

Formalization and numerical computation: generalities

Ever since the introduction of Frege's rules, it was evident that numerical computation was a particular case, and (in some ways) even typical of all formal deduction. Thus computations of polynomials with integral coefficients and arguments (≥ 0) are formal deductions from the (elementary) *axioms*:

$$\begin{array}{rclcrcl} n+0 & = & n & \qquad & n+m' & = & (n+m)' \\ n\cdot 0 & = & 0 & \qquad & n\cdot(m') & = & (n\cdot m)+n \end{array}$$

(where $'$ means the successor) by the *rules* of substitution, equating equals to equals. A computation evaluates an expression without variables as a *numeral*: $0, 0', 0'', \ldots$

Computations can be checked mechanically, and so the formalization above is *complete for equations between numerical expressions*, say p_1 and p_2:

> If $p_1 = p_2$ is true (for the usual interpretations of $0,', +, \cdot$) then $p_1 = p_2$ can be derived by the rules above.

If, further, the usual formal rule for existential quantifiers is added, then:

> If a diophantine equation $p_1 = p_2$ in the variables $x_1, \ldots, x_n$ has integral solutions then $\exists x_1 \cdots \exists x_n(p_1 = p_2)$ can be formally derived by the rules mentioned.

Thus the rules are *complete for solvability of diophantine equations*.

More generally, one expects some sort of parallel between

formal derivability and solvability of diophantine equations.

This was stressed early in the century by Hilbert, who saw here a unity between school boy arithmetic and all of (formalized) mathematics. Taken literally, the parallel equates *Hilbert's tenth problem* (to give a general method for deciding whether any diophantine equation has a solution) with deciding whether any formula F has a derivation by means of given formal rules $\mathcal{F}$. This remained, in fact, the source of Hilbert's later conjectures: the pay-off for replacing the allegedly difficult abstract notion of *truth* by the apparently wholly manageable notion of *formal derivability*, was to have been the effective decidability of derivability in properly formalized branches of mathematics (as in the case of real closed ordered fields).

Today we know that the parallel above holds literally, in that there is one diophantine equation (in just 9 variables!) with a parameter f and an effective way of finding a value of f corresponding to any pair $(F, \mathcal{F})$. Hilbert's tenth problem has a negative solution, and his conjectures about the decidability of formal derivability were false.

But at the turn of the century (in fact, up to [1931]) weaker variants of the parallel had not been excluded; for example, that no one formal system 'coded' all formal procedures (for each set of rules $\mathcal{F}$, derivability in $\mathcal{F}$ is formally decidable, but not by a method adequately represented in $\mathcal{F}$). But the price for this possibility would have been high, since the obviously elementary *character* of (verifying) formal derivability in $\mathcal{F}$ would not be reflected in an adequate *definition* for derivability by means of $\mathcal{F}$. Part of the work in [1931] established the definitional 'adequacy' (technically, completeness for formal derivability) of a general class of formal systems, including (Hilbert's) *pure number theory* (the system derived by the passage, in the subsection on p. 67, from Peano's axioms together with the recursion equations for $+$ and $\cdot$).

As a final preliminary, a curious blind spot has to be mentioned. In all the discussions of decidability and completeness (of formalizations) in the first three decades of this century, an obvious connexion was not noticed: *completeness* (for example, of pure number theory) *would yield a decision method* in a 'finite number of steps' for formal derivability (in particular, for Hilbert's tenth problem), as follows. Given a diophantine equation $p_1 = p_2$ in n variables, it is enough to lay out the formal derivations in some ω order, try them out one by one, until a derivation of either $\exists x_1 \cdots \exists x_n (p_1 = p_2)$ or of $\forall x_1 \cdots \forall x_n (p_1 \neq p_2)$ is reached (completeness ensures that this process terminates). This blind spot is a glaring oversight if one means 'finite number of steps' literally, without consideration for the practical value of such a method by trial and error.

Incompleteness of formal systems for number theory and beyond

To fix ideas, the reader may wish to think of pure number theory. In any case no details of the system will be used in the simple sketch below, which supports Gödel's claim (on p. 50) that [1931] did not need new mathematics. In fact, the sketch uses only:

1. The particularly *elementary character of the set of formally derivable formulas* (compared above to the set of solvable diophantine equations); so to speak, the *raison d'être* of formalization itself.

2. Cantor's *diagonal argument* (the class of all sets of natural numbers is not enumerable), here applied to sets and enumerations defined by restricted means.[14]

Proposition 6.4.1 *Let $\mathcal{C}$ be a class of (formulas defining) number-theoretic predicates with one and two arguments, closed under identification of variables and negation. Then there is no (binary) predicate in $\mathcal{C}$ which enumerates all (monadic) predicates in $\mathcal{C}$.*

Proof. The hypothesis means that if $F(n,m)$ is in $\mathcal{C}$, so are $F(m,m)$ and $\neg F(n,m)$.

The conclusion means that no formula $F(n,m)$ in $\mathcal{C}$ has the property that for for each formula $G(m)$ of $\mathcal{C}$ there is a number g such that

$$\forall m[F(g,m) \Leftrightarrow G(m)].$$

A counter example is obtained by taking $\neg F(m,m)$ for $G(m)$, and putting $m = g$. $\square$

In contrast to 6.4.1, there is in general no obstacle to enumerating all *formulas* G (with one variable) by giving them numbers g in such a way that simple syntactic operations are defined by formulas in $\mathcal{C}$; for example, *substitution* σ:

[14] Concisely, the argument uses the Σ_1 *definability of formal derivability*, and the Σ_1 *undefinability of* Π_1 *truth* (by means of straight diagonalization).

More efficient expositions are available in the literature (cf. Chapter D.1 of Barwise [1977]). The interest of the proof below is that it follows Gödel's presentation in his letter to Zermelo (already mentioned on p. 54) rather than his publication [1931], where a relation to the so-called liar paradox is prominent. (In conversation Gödel could not resist the temptation of paradoxical formulations. In publications he dramatized the trauma of ever having been taken in himself by a paradox).

Zermelo's criticism, though clumsily worded, was closely related to 2 (the inadequacy of any formal language for defining - all - sets of natural numbers), but failed to stress 1. So, in particular, it fails to pin-point the difference between the formal systems (of number theory) involved in the incompleteness theorem and those (in the subsection on p. 67) for the field of real numbers or the additive semigroup of natural numbers, which do decide every formula.

> $\sigma(g,n)$ is the number of the formula (without variables) $G(n)$,[15] where G is the formula whose number is g.

More pedantically, since in formal systems $\mathcal{F}$ the number m is represented by the m-th numeral $\overline{m}$ in some standard notation, e.g.

$$0, 1, 1+1, (1+1)+1, \ldots \qquad \text{or} \qquad 0, 0', 0'', 0''', \ldots$$

(as on p. 69),

> $\sigma(\overline{g},\overline{n})$ is the numeral whose value is the number of the formula $G(\overline{n})$, where G is the formula whose number is the value of the numeral $\overline{g}$.

Far from being subtle, the difference is so crude that in ordinary mathematics it would hardly be mentioned. For example, in the case of polynomials $x^n + a$ with numerical a, the *defining* expressions are numbered by an enumeration of the pairs (n, a), which can certainly be done polynomially by: $\frac{1}{2}(n+a+1)(n+1)+n$. But an enumeration of the functions $x^n + a$ *defined* by these expressions, which is a function of triples (x, n, a), cannot be done polynomially. (In the case of functions, we have $=$, where in the case of predicates above we had $\Leftrightarrow$.)

Incidentally, contrary to a widespread misunderstanding, there was nothing particularly novel in Gödel's numbering of formulas or derivations, that is (finite sequences of formulas): this was implicit in Cantor's well-known enumeration of finite sequences of elements taken from an enumerated set.

Theorem 6.4.2 Gödel's First Incompleteness Theorem. *Let $\mathcal{F}$ be a formal system, given with a numbering of its formulas with one or no free variable, where (the value of the numeral) $\overline{n}$ is the number of the formula N. Then $\mathcal{F}$ is incomplete, provided:*

1. *some formula D of $\mathcal{F}$ defines derivability (in $\mathcal{F}$)*

2. *$\mathcal{F}$ is* sound*; that is, for formulas N without free variables,*

$$\textit{if } N \textit{ is derivable then } N \textit{ is true.} \tag{6.1}$$

Proof. If $\mathcal{F}$ were complete, $D[\sigma(n,m)]$ would define an enumeration of the monadic predicates of $\mathcal{F}$. If g is the number of $\neg D[\sigma(m,m)]$ with variable m, then

$$\text{neither } D[\sigma(\overline{g},\overline{g})] \text{ nor } \neg D[\sigma(\overline{g},\overline{g})] \text{ is derivable.}$$

[15] $G(n)$ is obtained by replacing the (only) free variable of G by n.

The false one is not derivable by 2, the other (true) one is not derivable by 1 and the choice of g.

Now, the value of $\sigma(\overline{g}, \overline{g})$ is the number of $\neg D[\sigma(\overline{g}, \overline{g})]$ and the latter is not derivable, but D is assumed to define derivability. So $\neg D[\sigma(\overline{g}, \overline{g})]$ is true, but not derivable. $\square$

At the time there was great interest in weakening the conditions on $\mathcal{F}$ and D, especially 6.1 (that is, to avoid the reference to *truth* of N in favour of *derivability*). Inspection of the argument above leads to:

1. if $D(\overline{n})$ is derivable, then so is N
2. if $\neg D(\overline{n})$ is derivable, then N is not derivable
3. if g' is the numerical value of $\sigma(\overline{g}, \overline{g})$, then either both of $D(\overline{g}')$ and $D[\sigma(\overline{g}, \overline{g})]$ are derivable, or none of them is (and similarly for their negations) .

Then:

- Derivability of $\neg D(\overline{g}')$ and so, by 3, of $\neg D[\sigma(\overline{g}, \overline{g})]$ contradicts 2 for $n = g'$, since G' is $\neg D[\sigma(\overline{g}, \overline{g})]$ itself.
- Derivability of $D(\overline{g}')$ implies, by 1, that G' be derivable, contrary to what just proved.

On p. 77 the conditions above will be further weakened. In particular, in accordance with the basic parallel (on p. 70) between (checking) computations and derivations, the converse of 1 will be used (for relevant N: the *completeness of $\mathcal{F}$ for derivability in $\mathcal{F}$* expressed by D).

Gödel [1931] gave a detailed verification of 1 and 2 for a specific definition D, his $\mathcal{F}$ being (an improved formulation of) the system of *Principia*, which claimed to give a 'complete' formalization of mathematics. But [1931] gave also general conditions on systems $\mathcal{F}$ to which the argument applies, and soon afterwards the analysis of Turing showed that arbitrary formal systems containing a certain minimum of number theory (or of the theory of finite sets) satisfied those general conditions.

By p. 70, today $D(\overline{g}')$ can be replaced by the assertion that a certain diophantine equation has a solution. But technically it was certainly much easier to find $D(\overline{g}')$, an assertion about formal derivability in *Principia* which is undecided in *Principia* (than one in familiar mathematical terms).

Gödel's strategy for going into details, further elaborated in Kleene [1978], avoided controversy.[16] But even without those details, the proof given here, on

[16]Gödel said he felt the heavy formalization was pedagogically useful (at the time) to avoid even superficial resemblance to Finsler's [1926] earlier speculations on incompleteness (which, in fact, gave no hint of sufficient conditions).

the *assumptions* 1 and 2, establishes beyond a shadow of doubt the following inadequacy of (any) formal systems $\mathcal{F}$ for elementary number theory:

> *Either* such a simple notion as formal derivability cannot be defined in $\mathcal{F}$ (in the sense of 1 and 2), *or* $\mathcal{F}$ is incomplete (in the sense that a true formula, namely $\neg D(\overline{g}')$, cannot be derived in $\mathcal{F}$).[17]

Actually, Gödel's own proof (in terms of definability) is so simple that it can be applied to situations which have little in common with formal systems: to sets of axioms which are not recursively enumerated or 'listable', to languages with infinitely long formulas or so-called infinitary rules, and the like (cf. the sections on such matters in Barwise [1977]). Some of these generalizations are in fact needed in connection with the new questions, in the next subsection, which involve a *rethinking of the role of* (necessarily incomplete) *formal systems in mathematics*.

Some lessons from the first incompleteness theorem

The first and principal lesson is related to the questions on p. 63:

> What more can we expect to know from a proof of a theorem by means of (incomplete) rules (for, say, number theory or set theory) than if we merely know that the theorem is true?

And, of course, as a corollary:

> What do we know about a problem if it is not decided by given rules?

At least at the present time, it is not so much the *general* incompleteness theorem for formal systems that has found uses, but incompleteness tied to *specific objects*, like (size of) ordinals in note 17 or (rate of growth of) number-theoretic functions in the examples below.

The second, subsidiary lesson is that, in the cases mentioned, incompleteness of suitable *informal* systems is needed for the most rewarding results (in other words, the generalizations mentioned at the end of the previous subsection).

Incidentally, Part II of [1931] was (at least partly) intended to go into possible criticisms of Part I. It was not (only) illness which stopped Gödel from writing Part II, but also the fact that those expected objections never materialized.

[17] An even more general argument of the same type applies to the case of set theory for any set $\mathcal{A}$ of axioms, not only for formal systems. Suppose $\mathcal{A}$ can be justified at all for the intended meaning (set out in Zermelo [1930] and in the subsection on p. 94 below); that is, for appropriate segments of the cumulative hierarchy, including the segment α. Then:

> *Either* α cannot be defined in set-theoretic language, *or* if D_α is a definition then $\exists x D_\alpha(x)$ is not decided by $\mathcal{A}$.

Of course, $\exists x D_\alpha(x)$ is about the 'abstract nonsense' of sets, while $D(\overline{g}')$ above is about the 'formal nonsense' of derivability.

The points above are illustrated by theorems of the form: some diophantine equation D in, say, 9 variables (by p. 70, the typical case), has infinitely many solutions, i.e.

$$\forall n \exists m_1 \cdots \exists m_9 [m_1^2 + \cdots + m_9^2 > n \ \wedge \ D(m_1, \ldots, m_9) = 0]. \tag{6.2}$$

If 6.2 is *true* then, for each n, such $m_1, \ldots, m_9$ can be computed by a programme which tries out each 9-tuple; in short, recursively. But if 6.2 is *proved* by restricted means (in $\mathcal{F}$), bounds for $m_1^2 + \cdots + m_9^2$ (in terms of n) can be specified. The literature speaks of *provably total recursive functions*, defined by the class $\mathcal{R}_{\mathcal{F}}$ of programmes which can be proved (in $\mathcal{F}$) to terminate. A significant part of proof theory describes the functions defined by $\mathcal{R}_{\mathcal{F}}$ in familiar mathematical terms.

An obvious conclusion is that even if 6.2 is true, but the least $m_1^2 + \cdots + m_9^2$ grows too rapidly with n, then 6.2 cannot be proved by means of $\mathcal{F}$. The converse is not true because, for some n,

$$\forall m_1 \cdots \forall m_9 [m_1^2 + \cdots + m_9^2 > n \ \Rightarrow \ D(m_1, \ldots, m_9) \neq 0] \tag{6.3}$$

may be true, but not derivable in $\mathcal{F}$. A very simple piece of logic shows that the converse does hold if all true propositions of the form 6.3 are added to $\mathcal{F}$. If $\mathcal{F}^+$ is the new (not formal!) system, the class $\mathcal{R}_{\mathcal{F}}^+$ of programs is greater than $\mathcal{R}_{\mathcal{F}}$, but not the class of functions defined. Thus *metamathematical* knowledge of underivability of 6.2 in $\mathcal{F}^+$ gives information about bounds for 6.2 of ordinary *mathematical* interest. Though this connection has been publicized for more than thirty years, the first convincing use was made only more recently (cf. Paris and Harrington [1977] on a problem in combinatorial partition theory). For logically more complex assertions than 6.2, a more sophisticated connection by means of so-called *functional interpretations* is used.

Without exaggeration: the answers above totally reverse the unsophisticated aims of using formal systems for an overview of mathematics; for example, for arranging problems according to the means needed for their solution.

> The new aim is to start with a problem P (one wants to know about), and to look for a bunch $\mathcal{F}_P$ of *relevant* systems (in the arsenal of systems with manageable metamathematical properties).

The metamathematical study (proof-theoretic or model-theoretic) of particular systems $\mathcal{F}$ is here only a preliminary; for example, to get some idea of the sort of problems P' for which $\mathcal{F} \in \mathcal{F}_{P'}$. A different strategy would apply if ever genuinely fundamental systems turned up (for example, related to the laws of thought on p. 63).

Consistency and consistency proofs

We now go into a reformulation of $\neg D(\overline{g}')$ which attracted great attention in the first decade after [1931].

Consistency of $\mathcal{F}$ (for short: $\mathrm{Con}_{\mathcal{F}}$) means that there is no formula F for which both F and $\neg F$ can be derived in $\mathcal{F}$. Since $\neg$ is intended to mean 'not', $\mathcal{F}$ had better be consistent. This is not at issue. Rather,

What use is (mere) consistency?

Once again, diophantine problems are typical. Let p be a polynomial with integral coefficients and $\vec{n} = (n_1, \dots, n_k)$ a list of its variables. If (in the technical jargon of p. 70) $\mathcal{F}$ is complete for solvability of diophantine equations, i.e.

if $\exists\vec{n}(p = 0)$ is true then it is formally derivable in $\mathcal{F}$

then the *significance* of $\mathrm{Con}_{\mathcal{F}}$ is that

if $\forall\vec{n}(p \neq 0)$ is formally derivable in $\mathcal{F}$ then it is true.

For, by completeness for solvability, if $\vec{n}$ were a solution then $p(\vec{n}) = 0$ and hence $\neg\forall\vec{n}(p \neq 0)$ could be derived in $\mathcal{F}$. So $\forall\vec{n}(p \neq 0)$ would be a counterexample to $\mathrm{Con}_{\mathcal{F}}$.

This 'significance' of $\mathrm{Con}_{\mathcal{F}}$ is also relevant to the matter of number-theoretic *purity*. Suppose $\mathrm{Con}_{\mathcal{F}}$ and the completeness of $\mathcal{F}$ for solvability are both proved by 'pure' methods (for example, in Hilbert's pure number theory). Then if $\forall\vec{n}(p \neq 0)$ is derivable in $\mathcal{F}$, it also has a pure proof (with an obvious extension to other preferred methods of proof).

Warning. The consequences above of $\mathrm{Con}_{\mathcal{F}}$ obviously do not extend to formulas $\exists\vec{n}(p = 0)$ since, if $\forall\vec{n}(p \neq 0)$ is formally *undecided* in $\mathcal{F}$, the *false* formula $\exists\vec{n}(p = 0)$ can be added consistently. (Thus, 1 on p. 73 is not assured by $\mathrm{Con}_{\mathcal{F}}$.) Gödel gave that warning in [1931a] in the clearest possible terms, actually before the discovery of incompleteness. But [1931a] made much less of an impression than such *conneries* as:

- In mathematics, consistency ensures existence (of what?).
- An inconsistent system would be dull because every formula G could be derived in it, by use of $\neg F \Rightarrow (F \Rightarrow G)$.

In later 'popular' writings, Gödel always treated such *conneries* respectfully.

Theorem 6.4.3 Gödel's second incompleteness theorem.[18] *Let $\mathcal{F}$ be a formal system demonstrably complete for derivability (defined by D), i.e.*

if N is derivable then so is $D(\overline{n})$ (provably in $\mathcal{F}$).

Then $\mathrm{Con}_{\mathcal{F}}$ *is not derivable in $\mathcal{F}$.*

[18]Described by him, in [1931], as *merkwürdig*: a curiosity.

Proof. Inspection of the proof of the fact that $G' = \neg D[\sigma(\overline{g}, \overline{g}]$ is not derivable (on p. 73) shows that

$$\mathrm{Con}_{\mathcal{F}} \Rightarrow G'$$

is derivable in $\mathcal{F}$, if the latter is *demonstrably* complete for derivability (defined by D), since then 2 on p. 73 follows from $\mathrm{Con}_{\mathcal{F}}$. But G' is not derivable in $\mathcal{F}$, and so neither is $\mathrm{Con}_{\mathcal{F}}$. $\square$

A looser formulation says: $\mathrm{Con}_{\mathcal{F}}$ *cannot be proved* (in $\mathcal{F}$) *if* $\mathrm{Con}_{\mathcal{F}}$ *can be proved to be significant* in the sense above.

To produce *examples of formal systems which do prove their own consistency*, given formal rules $\mathcal{F}$ and a numbering of their derivations we pass to new rules $\mathcal{F}_1$, by adding the following requirement on derivations (with number d and end formula F_d):

> For all pairs of (the finitely many) preceding derivations in $\mathcal{F}$ (that is, $d_1 \leq d$ and $d_2 \leq d$), the end formula of one is not the formal negation of the other (F_{d_1} is not the formula $\neg F_{d_2}$).

Evidently, $\mathrm{Con}_{\mathcal{F}_1}$ is proved in the most elementary way: we stop before an inconsistency turns up. But also: if $\mathcal{F}$ is consistent then $\mathcal{F}$ and $\mathcal{F}_1$ have not only the same theorems, but the same derivations! Only, the procedure for *checking* derivations is more elaborate in $\mathcal{F}_1$.

Also (and this is philosophically interesting), though logical texts rarely consider systems like $\mathcal{F}_1$, the latter mirror quite well, albeit crudely, an essential method used in practice for checking proofs: *comparison with background knowledge* (here represented by d_1 such that $d_1 < d$).

Corollary 6.4.4 An improved version of Gödel's First Incompleteness Theorem. *A formal system $\mathcal{F}$ is incomplete, provided:*

1. *$\mathcal{F}$ is consistent*

2. *$\mathcal{F}$ is complete for derivability, demonstrably in $\mathcal{F}$ and $\mathcal{F}_1$ (w.r.t. D and D_1, respectively).*

Proof. In the notation used on p. 73, if g_1 is the number of $\neg D_1[\sigma(m, m)]$,

$$\text{neither } D(\overline{g}'_1) \text{ nor } \neg D(\overline{g}'_1) \text{ is derivable}$$

(in $\mathcal{F}$ or, equivalently by 1, in $\mathcal{F}_1$).[19] As on p. 73:

- $\neg D_1(\overline{g}'_1)$, and hence $\neg D(\overline{g}'_1)$ is not derivable because condition 2 on p. 73 is ensured by the hypotheses, as mentioned above.

[19]Note that there are $\mathcal{F}$ satisfying 1 and 2 in which $D(\overline{g}')$ *is* derivable!

- If g_2 is the number of $D_1(\overline{g}'_1)$ (and hence G'_1 is the formula $\neg G_2$) then, by $\mathrm{Con}_{\mathcal{F}_1}$ (which is derivable in $\mathcal{F}_1$),

$$D_1(\overline{g}_2) \Rightarrow \neg D_1(\overline{g}'_1)$$

  is derivable in $\mathcal{F}_1$. If G_2 were derivable, so would be $D_1(\overline{g}_2)$ by 2, and then $\neg D_1(\overline{g}'_1)$ would be derivable, contrary to the first part of the proof. $\square$

For further information, cf. Kreisel and Takeuti [1974]. Specifically, there is the matter of so-called *canonical numberings* of formulas and derivations (unique up to formal equivalence), which are perfectly analogous to coding finite sequences of sets by sets (unique up to appropriate equivalences). Also, and this is much more interesting, novel questions arise concerning the second incompleteness theorem for systems which were not known at the time of [1931], for example, so-called *cut-free rules*.

Some lessons from the second incompleteness theorem

As Gödel himself stressed, back in [1931], his second theorem is *irrelevant to any sensible consistency problem*. In any case, if $\mathrm{Con}_{\mathcal{F}}$ is in doubt, why should it be proved in $\mathcal{F}$ (and not in an incomparable system)? Gödel's practice followed his theory:

- His last self-contained publication [1958], which goes back to [1933], was presented as a consistency proof.
- Between 1931 and 1958, as his notebooks at Princeton show, he studied other such proofs, especially Gentzen [1935a] (published posthumously).[20]

Very much in contrast to the break with traditional aims, advocated throughout this chapter, Gödel continued to use traditional terminology. For example:

- the original title of Spector [1962], extending Gödel [1958], did not contain the word 'consistency' (but rather stressed the aspect of provable totality of function(al)s mentioned on p. 75); it was added for the posthumous publication at Gödel's insistence.[21]

He knew only too well the publicity value of this catchword, which (contrary to his own view of the matter) had made his second incompleteness theorem more spectacular than the first.

[20] Gentzen used functionals of lowest type, defined by unfamiliar equations, intended to operate on so-called choice sequences; in other words, with special emphasis on continuity.

Gödel's [1958] used functionals of all finite types, defined more elegantly, but intended to operate on rules (except that the last sentence of [1958] does not fit the intention).

[21] Spector used so-called bar recursion, again for all finite types, for which continuity (in a suitable sense) is again essential.

As to uses of the second incompleteness theorem: above all, it provided the first, much needed *cross check on proposed consistency proofs*. The early literature on the subject (supposed to 'secure' mathematics!) had a particular high density of errors. The most famous are in Ackermann [1924][22] and Herbrand [1930].[23]

Another good use of the second theorem, which however always requires some imagination, can be seen as follows, by reference to the basic significance of $\mathrm{Con}_{\mathcal{F}}$ (on p. 76):

> if $\forall\vec{n}(p \neq 0)$ is derivable in $\mathcal{F}$ from the *false* formula $\neg\mathrm{Con}_{\mathcal{F}}$, then it is derivable in $\mathcal{F}$ itself.

The reason is that $\mathrm{Con}_{\mathcal{F}}$ is not derivable, hence $\mathcal{F}$ together with $\neg\mathrm{Con}_{\mathcal{F}}$ is consistent and $\forall\vec{n}(p \neq 0)$ is true. An easy exercise then shows that $\forall\vec{n}(p \neq 0)$ is also derivable in $\mathcal{F}$ itself (but such a derivation may be more difficult to find).

An even better use relies on the details of consistency *proofs* which derive $\mathrm{Con}_{\mathcal{F}}$ from some 'mathematical', manageable principle, say P. So the *false* formula $\neg P$ is consistent with $\mathcal{F}$ too. For suitably complex P,

> if F is derivable in $\mathcal{F}$ from $\neg P$, then it is derivable in $\mathcal{F}$ itself

even for certain F which are (logically) much more complicated than $\forall\vec{n}(p \neq 0)$.

6.5 Background to [1930]: Elementary Logic in the Twenties

Evidently, to document Gödel's own view (p. 50) on his good use of traditional philosophical notions in [1930], a word on the knowledge about elementary logic which had accumulated before [1930] is needed. For balance, other interesting consequences of that early knowledge, which its authors did not recognize, will be used to illustrate negative effects of (ill digested) traditional philosophical aims.[24]

[22] Pointed out in von Neumann [1927].

[23] Corrected in Dreben, Andrews and Anderaa [1963]. Also discovered, but not quite corrected, by Gödel in the early forties (cf. his *Arbeitshefte IV* and *V* at Princeton). Though in the meantime others have also observed the points of detail made there by Gödel, his sure touch remains exceptional.

[24] An at least comparable important obstacle to progress in the twenties was the emphasis on the false conjecture that logical validity or formal derivability by means of Frege's rules was mechanically decidable. As a result, a fair number of partial (and certainly not very memorable) results cluttered up the literature of the twenties.

As on p. 70, before [1931] it was not realized that a proof of the conjecture would have solved Hilbert's tenth problem (and more, as Gödel stressed in [1931] in his discussion of the matter); for the diophantine equation $p_1 = p_2$ has a solution if and only if $\exists x_1 \cdots \exists x_n(p_1 = p_2)$ follows purely logically from the usual axioms for successor, addition, and multiplication.

Non-categoricity of elementary axioms

One of the best known results had been realized before 1920 (by Loewenheim), and proved very simply (by Skolem): the existence of non- isomorphic models of Euclid's axioms (p. 66), or of real- closed ordered fields (p. 67), is typical of all elementary axioms. In particular:

Theorem 6.5.1 *If each of a countable collection of elementary axioms is true for some structure S, there is a countable (or finite) part S_0 of S they are all true too.*

Thus, if S is uncountable then the axioms are not categorical.[25]

Proof. The idea of the proof (which, incidentally, is the clue to Gödel's work in Section 9) is perfectly illustrated by means of the formula $\forall x \exists y \forall z R(x, y, z)$, where R does not contain quantifiers $(\forall, \exists)$. For if $\forall x \exists y \forall z R(x, y, z)$ holds in S, so does $\forall x \forall z R(x, Y(x), z)$ for some function Y with arguments and values in S. For any element a of S, the set S_0 generated from a by Y[26] will do (where $a \in S_0$, and if $b \in S_0$ also $Y(b) \in S_0$); thus $S_0 = \{a, Y(a), Y(Y(a)), \ldots\}$. Evidently, S_0 is countable or finite. $\square$

Readers probably know, and certainly can easily imagine, the thoughtless conclusions which were drawn from the simple result above. At one extreme, differences between *infinite cardinalities* were rejected as 'meaningless' because such cardinalities cannot be distinguished by elementary properties. At the other extreme, the *elementary formalism* or 'language' was rejected as hopelessly inadequate because it cannot be used to express even such brutal properties as differences in cardinality.

What was overlooked for a remarkably long time, was the *positive* aspect of the result above: from the validity of an elementary formula in all countable structures follows its validity in all uncountable structures too. Without exaggeration: some result of this kind is needed to make the 'abstract nonsense' about *validity in arbitrary structures* (called 'truth in all possible worlds' on p. 61) useful at all; not because of any illegitimacy of the notions involved, but because a formula might fail to be logically valid only because it is false in some odd structure that nobody wants to know about.[27]

Bibliographical remark. Skolem [1922] went even further in reducing the 'abstract nonsense'. Suppose F is an elementary formula with relation symbols $R_1, \ldots, R_m$. Then relations $R_1^F, \ldots, R_m^F$ can be (quite explicitly) defined in pure

[25] This is generalized in 6.6 on p. 84.

[26] If there are also operation, not only relation symbols in R, S_0 is required to be closed for the corresponding operations.

[27] Specialists can think of examples in so-called second order classical logic, and to some extent in intuitionistic propositional logic (where propositions about so-called lawless sequences are needed); cf. p. 90 below.

number theory, with the property: if F_ω is obtained from F by replacing the symbols R_i by R_i^F, and if the quantifiers in F_ω range over the natural numbers ($+$ and $\cdot$ in R_i^F having their usual number-theoretic meaning) then:

> if F is true in any structure at all (or, as one says, if F has a model) then F_ω is true (for the natural numbers).

In other words, logical validity is not only equivalent to validity in countable structures, but to validity in structures defined in this restricted way.[28]

Two formulations of completeness (for logical validity)

One formulation occurs in Hilbert and Ackermann [1928], obviously written by the second co-author. It says just what one would expect. A system of rules $\mathcal{L}$ (for 'logic') formulated in terms of elementary logic is *complete* if, for every elementary formula F,

$$F \text{ is derivable in } \mathcal{L} \text{ provided } F \text{ is logically valid}$$

(that is, true in all structures in which the relation symbols of F are interpreted).[29] Pedantically, one can also consider the converse (called *soundness* of $\mathcal{L}$), which is usually verified by inspection.

It will not have escaped the reader's notice that the matter of completeness is neatly by-passed in the formalizations in the subsection on p. 67, which were Hilbert's principal interest, since they *decide* every proposition: no set of sound rules can do more! (in the sense of generating more theorems). Further (and this certainly did not escape Hilbert's notice!) the completeness in question is formulated *purely formally*: for every elementary formula F (about the structure considered) either F or $\neg F$ is formally derivable. Sure, the reason for being interested in this formal property is that it ensures that all true F are derivable in the formalization. But the wording respects the ideal of *Methodenreinheit* (here applied to formal derivability), and the formulation in Hilbert and Ackermann [1928] violates it.

Soon afterwards, despite the handicap of a recent stroke, Hilbert [1930] tried to correct this violation by a pure version of *completeness of $\mathcal{L}$ modulo a formal system $\mathcal{Z}$ for pure number theory*: for every elementary formula F,

$$(F \text{ is derivable in } \mathcal{L}) \text{ or } (\neg F_\omega \text{ is derivable in } \mathcal{Z}),$$

[28] Skolem himself did not state this result: he noticed only the very marginal improvement that, in contrast to his earlier proof of theorem 6.5.1, the proof of the refined result did not use the axiom of choice.

[29] There is an obvious analogous notion of *completeness for logical consequence* (of F from a set $\mathcal{F}$ of formulas); in the case of a finite set $\mathcal{F} = \{\mathcal{F}_\infty, \ldots, \mathcal{F}_\backslash\}$ this reduces to validity of $F_1 \wedge \cdots \wedge F_n \Rightarrow F$.

where F_ω is obtained from F when the relation symbols R_i of F are replaced by suitable expressions of $\mathcal{Z}$ (for example, the relations R_i^F mentioned at the end of the previous subsection.

Once again the blindspot of p. 70 intervened: Hilbert and others overlooked the fact that completeness in his pure sense would prove the (false) conjecture in note 24, providing a method for deciding in a finite number of steps whether F is derivable in $\mathcal{L}$ (provided of course that $\mathcal{Z}$ is sound). One lays out the formal derivations of the systems $\mathcal{L}$ and $\mathcal{Z}$ in linear order, and tries them out alternately. After a finite number of steps one arrives either at a derivation in $\mathcal{L}$ of F, or at one in $\mathcal{Z}$ of a formula of the form $\neg F_\omega$. Since the conjecture is false, so is Hilbert's pure version of completeness: for *any* (sound) rules $\mathcal{L}'$ for logic and $\mathcal{Z}'$ for (extensions of) number theory.

Herbrand's Theorem

Before [1930] a good deal of formally pure work, concerning transformations of formal derivations, had been done; particularly, the very original work of Herbrand [1930] (already mentioned in note 9). Without exaggeration: as we see things now, part of the interest of this kind of work comes from the fact that it does *not* presuppose completeness of the formal rules (since, for example, a transformation may be particularly efficient if applied to derivations that happen to be built up by an incomplete *sub*set of given, possibly complete rules). But it was hard to see this highly positive side of the matter:

- For one thing, there was no hint of it in Herbrand [1930].
- In addition, there were formal errors (cf. note 23).
- Last but not least, there was the terribly complicated (though correct) formulation of his *Théorème fondamental*, again without a hint of possible uses in ordinary mathematics.[30] Instead, Herbrand used the *Théorème* to get some not at all memorable, partial results on the ill-fated decision problem for elementary logic (cf. note 24).

Concerning completeness: Herbrand [1930] refers to (the possibility of) a proof, but rejects the matter out of hand because the abstract so-called semantic notion of logical validity was not precise enough for Herbrand (to deserve attention).

6.6 The Completeness Theorem

By Section 5, though the completeness problem solved in [1930] had been stated in the twenties, there were mixed feelings about it: it certainly did not fit in with

[30] The irony of the matter is that even the simplest case of the *Théorème*, applied to purely existential formulas, has turned out to be at least as useful as the completeness theorem, especially if one is interested in explicit bounds.

the ideal of purity of method, and at least one formal counterpart to completeness, in Herbrand [1930], simply had more mathematical content.

What Herbrand overlooked was that another step was needed before the average logician or mathematician had enough confidence in the subject to want to look at a monster like his *Théorème Fondamental*. In contrast, almost anybody could understand completeness (or misunderstand it, thinking of it as a confirmation of Hilbert's aim). Being simple and memorable, it helped to put elementary logic 'on the map'.

Enter Gödel

In [1930] Gödel established the impure version of the completeness of Frege's rules in the sense explained above. Translated into the notation used in Section 5, [1930] shows that, for every elementary formula F,

$$\text{(either } F \text{ is derivable by Frege's rules) or } (\neg F_\omega \text{ is true).} \tag{6.4}$$

A comparison with Skolem [1922] documents beyond a shadow of a doubt (for anybody prepared to look at the proof of Skolem's dull result) Gödel's view that all ingredients needed for the proof of the completeness theorem were available in the twenties. But Skolem did not see the relevance of those ingredients (cf. note 28), and Gödel did. Given that those mathematical ingredients were standard anyway, it is a minor matter whether Gödel happened to have seen Skolem's paper (for example, in the mathematics library at Vienna, where Gödel was a voracious reader as a student, sometimes making marginal remarks in shorthand).

At this point it is worth recalling Hilbert's pure version of completeness which was seen to be false above. It differs from 6.4 only in one place:

derivable in $\mathcal{Z}$ in place of *true*.

Yet, the difference is quite essential: derivability of $\neg F_\omega$ can be mechanically verified whenever it holds, truth of $\neg F_\omega$ in general cannot (in contrast to Hilbert's expectations, cf. p. 70).[31]

The Finiteness Theorem

To return to [1930], Gödel noted in *Satz X*, more or less, another fact which is much more often used (in the sense of being directly appealed to) in applications of elementary logic, the so-called *Finiteness Theorem* (for infinite sets $\mathcal{F}$ of formulas):

[31] *Warning*. Inspection of [1930] shows that the assertion

either F is derivable by Frege's rules or $\neg F_\omega$

is not only true, but derivable in $\mathcal{Z}$. But *this* replacement of 'true' by 'derivable in $\mathcal{Z}$' is of no obvious consequence. If anything, it hides the essential differences between the pure and impure versions. A moment's thought shows that it is typical of the ritual of formalization.

If each finite subset of $\mathcal{F}$ has a model, so does $\mathcal{F}$ itself.

As a corollary, the non-categoricity result on p. 80 extends to all sets $\mathcal{F}$ with infinite models, as follows:

For an arbitrary set I (of new constants), $\mathcal{F} \cup \{i \neq j : \text{for distinct } i, j \in I\}$ has a model of cardinality $\geq$ card I.

Actually, Gödel stated the theorem only for countable $\mathcal{F}$ although it holds for arbitrary $\mathcal{F}$, and although he himself formulated corresponding results for uncountable sets of propositional formulas in [1932]. But applications in mathematics where the unrestricted formulation is actually required were discovered only later, as the subject of *model theory* (of elementary logic) developed.[32]

Incidentally, Gödel stated the finiteness theorem in impure terms first, in *Satz IX* of [1930], mixing in formal derivability:

$\mathcal{F}$ has a model, or else some finite subset $\{F_1, \ldots, F_n\}$ is formally inconsistent (that is, $\neg(F_1 \wedge \cdots \wedge F_n)$ can be formally derived).[33]

In retrospect the finiteness theorem is seen to fit in well with the only obvious sense of a formal derivation from an infinite set $\mathcal{F}$, namely, that only a finite subset of $\mathcal{F}$ be used. When asked whether this aspect had led him to the finiteness theorem, Gödel could not remember having been conscious of it at the time; and about [1932] he remembered stating the result first for countable sets, and noting afterwards that the proof nowhere used countability. Realistically speaking, it is of little interest what one is *not* conscious of; in any case, Gödel never claimed to have followed consciously his heuristic principles (on p. 50) at the time of [1930] and [1931], but to have *discovered* later that they apply.[34]

Be that as it may, there is no doubt that Gödel's views fit the later development of logic, in particular the latter's need for non-elementary notions (discussed in the next subsection).

Some lessons from the completeness theorem

The completeness theorem establishes that logically valid *elementary* formulas can be proved by logically *pure* derivations. There is an obvious potential conflict here, in restricting *both* definitions of objects *and* methods of proof (as is done in so- called doctrinaire constructivism): a given problem may have a very simple solution, which may be impossible to establish by the restricted methods of proof. All this is plain horse sense.

[32]Malcev was the first to state (in [1936]) the finiteness theorem for uncountable sets of formulas, and to use it (in [1941]) for interesting results (in group theory).

[33]To use this form in model theory, the completeness theorem *is* needed.

[34]In contrast, his ideas in Sections 9 and 10 were developed after that discovery.

Before knowledge of elementary logic (either of elementary definitions or of formal rules) can become an effective part of our intellectual reflexes, one needs some general orientation on the kind of questions where this knowledge is likely to be relevant; this is easiest by *contrast* (with non-elementary notions).

a) Elementary formulas

Obviously, one needs to know results which hold for elementary, but not for all axioms. *Samples*:

- The *finiteness theorem* above certainly does not apply to Peano's non-elementary axioms together with the infinite set of formulas $\{a \neq n : n \in \omega\}$ (every finite subset of the latter is satisfied by some a in models of Peano's axioms).

- *Non-categoricity* (on p. 80) puts a premium on non-isomorphic structures which share the same elementary properties, but one of them is more manageable. For then, with some imagination, one may find an elementary problem which is difficult for one but not for another, as in the transfer results on real closed fields (p. 67).

- A more delicate strategy, discovered in the last 35 years, involves general operations on structures which preserve elementary properties. With a bit of luck such operations, suitably applied to S and S', may produce isomorphic structures, thereby showing that S and S' have the same elementary properties.

  The (still) best-known application establishes relations between the p-adic fields and the fields of formal power series with integral coefficients modulo p. As long as only very simple questions about p-adics were treated, mathematicians got by with a vague perception of some relation between such fields (and exploited their knowledge of formal power series; for example, Chevalley [1936] in the thirties). For more difficult problems, some 30 years later, the precise relation of elementary equivalence was needed (for example, by Ax and Kochen [1965][35] and Ershov [1965]).

[35]Ax and Kochen make use of ultraproduct constructions, which preserve elementary properties. This fact was first exploited, albeit inadvertently, in Skolem [1933a] to establish the existence of a non-standard model of (all true statements about $+$ and $\cdot$ over) the natural numbers.

Gödel's review [1934] dismissed this result, actually an immediate consequence of his finiteness theorem, but for formally incorrect reasons. Precisely, he proposed to use a combination of his completeness and incompleteness results. He did this immediately after mentioning that Skolem considered models of all true sentences of arithmetic, where there is no place for incompleteness at all. It is not, as is sometimes thought naïvely, that the 'spirit of the times' in the thirties prevented one from considering the set of all true sentences at all: one just did not handle them very well.

Secondly, it must be easy to *recognize* notions which have elementary definitions. This is a delicate matter, especially for the logically perceptive mathematician who has been sold on the idea that *all of mathematics is formalized* in some universal system (say, of set theory). Though, of course, elementary formulas can be formally separated from the others in the universal system, the separation seems artificial, and is less easy to remember than if (following Gödel) non-elementary definitions are understood too, and so can serve for contrast.

At the other extreme, less perceptive mathematicians or logicians are led to apply their knowledge of elementary logic indiscriminately, for example to the universal system itself, generally with disappointing results (according to the principle already quoted on several occasions about what is true in general). Specialists will easily think of such results for those non-standard models which are defined by mere use of the finiteness theorem; others can guess the kind of disappointment involved from the ritual formalization of the impure completeness theorem (in note 31).

In short, as a general rule elementary logic is most rewarding mathematically when applied to structures defined by (sets of) formulas which are elementary as they stand, not merely because they are thought of as expressions in a universal formal system. This includes of course non-elementary notions which are demonstrably equivalent to elementary ones (for example, the notions of order*able* and formally real fields). Up-to-date texts on model theory given general conditions for such equivalences (covering the standard example above). One of the rare exceptions to the general rule is elaborated on p. 107, where it is useful to go back to the definition of a non-elementary notion (free basis of a group) in the universal system, and to apply formal incompleteness properties of that system.

b) Logical inference

On the banal side (and contrary to the false impression mentioned repeatedly), the advantage of a logically pure proof hardly ever lies in greater certainty, the usually shorter impure proof being used for checking.[36] But there is an advantage in *additional information* (for example, bounds as in note 13), which can be read off more easily from pure proofs; in contrast, unwinding of impure proofs, even if it is theoretically possible, tends to pass the point of diminishing returns. Against another widespread misunderstanding, though bounds are *more easily* read off from pure proofs, *better* bounds are liable to be established by means of impure proofs.

The place of pure logical inference *within* impure proofs is more delicate. The issue is general, but most dramatic in the case of purely logical theorems. Modern mathematics provides many examples. Thus the notion of ordered field has an elementary definition, say O, and so an elementary theorem T about such fields

[36]To be pedantic, a logically impure proof of an elementary formula F proves the *validity* Val (F) of F, not F itself.

is a logical truth: $O \Rightarrow T$. But the latter, or rather Val $(O \Rightarrow T)$, is often established by impure proofs, involving the embedding of ordered fields in particular real closed fields, and applying set-theoretic and topological operations to the latter. The heart of the proof is to spot relevant (set-theoretic or topological) properties P of the structures so obtained; only the implication

$$P \Rightarrow \mathrm{Val}\ (O \Rightarrow T) \tag{6.5}$$

is derived purely logically, and often this part of the argument is not mentioned at all (in the phrase of Bourbaki [1948], the derivation of 6.5 is the least interesting side of the matter). Seeing those properties P makes many mathematical proofs, as has often been said, more like *perception* (with all its problems) than a sequence of formal steps.

Reminders (on the use of scientific experience). Though the examples just given of logically impure proofs are commonplace today, they were not known 100 years ago, and still are not known to many authors of logical texts, in whose own experience logically pure proofs have a much greater relative significance (frequency). Thus, except for those with uncommon philosophical talent, their limited experience is not sufficient for a correct sense of proportion on pure and impure proofs in possible mathematical reasoning.

At the other extreme, some of the reservations by philosophers and mathematicians about logic depend equally on defective knowledge of this subject, but with a difference: impure proofs have not been widely advertised, while the best (and, often, only) known claims for the interest of logic are the pretensions about laws of thought or formal rigour mentioned on p. 63. They are considered next, in a partial review which sharpens the general picture painted in Section 2.

6.7 Foundational Bearing of Gödel's First Results

In each case, the logical properties of the schemes themselves will be recalled first, and then they are tested by inspection of scientific experience (in the style of Section 3). The famous foundational schemes of Russell and Hilbert are reviewed briefly, and Brouwer's less famous 'anti-formalist' views are explained and examined. The schemes of the 'anti-formalists' Poincaré and Zermelo are more conveniently discussed in Sections 9 and 11.

Russell's scheme

Russell's aim of a *universal system* for all of mathematics has a clear logical or mathematical sense, and a less obvious empirical sense.

The logical aim cannot be achieved by the first incompleteness theorem, and its philosophical significance is problematic, for reasons given already in the discussion (at the beginning of Section 4) of Hilbert's 'universal' (complete) systems for *branches* of mathematics, which he favoured for the sake of *Methodenreinheit*.

Russell's empirical aim has been achieved, at least for *existing* mathematical practice (by use of current set theory in place of *Principia*); partly by the simple device of restricting practice to a given system. The second incompleteness theorem[37] shows up a defect, a kind of blind spot, of this practice. As mentioned repeatedly, formal independence from such a universal system explains (empirically) why certain well-defined problems have not yet been settled (for example, in odd corners of group theory). But, and these are empirical facts too:

1. such problems are relatively rare

2. (in contrast to formal definitions of, say, Bourbaki's basic structures) the specifically formal axioms and rules of the universal system are barely mentioned in the later development

3. (last but not least) those structures can be applied perfectly well to familiar objects like the natural numbers, which are normally not thought of as defined set-theoretically at all.

By 2 and 2, the two properties characterizing Russell's ideal (of a system which is both formal and universal) are hardly used in practice.

It seems plain (in accordance with p. 87 on the use of scientific experience for refuting foundational schemes) that the conclusions above would be less convincing without our experience with universal systems.

Hilbert's scheme

Hilbert's scheme is a kind of opposite extreme to Russell's empirical aim and, especially, to the doctrine (mentioned on p. 62) that *only* empirical case studies can support universal systems.[38]

The difference is very well expressed by Hilbert's favourite slogan, in Hilbert [1931], which eventually replaced the modest business of purity of method: his aim was a *final solution*[39] of all foundational problems by *purely mathematical means*. Actually, his aim is more modest than it sounds, because of the tacit assumption (which alone makes the aim even remotely plausible) that only those foundational problems which concern proofs of finitist theorems are 'real'. (By

[37] In particular, footnote 48^a of [1931] on the restriction to finite types (in *Principia*).

[38] As a matter of historical curiosity, neither Hilbert nor Russell ever stressed that particular difference between them.

[39] Outside mathematics Hilbert liked big words like 'final solution' (or 'world formula' in relativity theory), rather than little things (like the perihelion of Mercury).

p. 70, the latter are of the same general character as theorems asserting that some diophantine equation is insoluble.) The 'final solution' was to establish the autonomy 'in principle' of the subject (exactly in the same sense as the subsection on p. 67 establishes the autonomy of real algebra).

Despite Hilbert's severe restriction, eloquently criticized in Gödel's [1931a], the first incompleteness theorem is enough to exclude a final solution. To be final, it would have to provide a method which decides every finitist problem, so to speak: here and now (and certainly every diophantine inequality; equivalently by p. 76, its consistency with some formal system which is complete for solvability of diophantine equations). Otherwise, if that system leaves the problem undecided, tomorrow we might think of another system which settles it. The new system would have to be justified, and so on *ad nauseum*.

The second incompleteness theorem is also relevant to Hilbert's scheme, but (by p. 78, and contrary to an almost universal misunderstanding) in a much more subtle way, involving the following fact of experience:

1. For any formal rules or axioms actually used in mathematical practice (in contrast to those experimented with in foundational studies), somebody has an abstract interpretation in mind which establishes their consistency instantaneously.

The second incompleteness theorem refutes an additional conviction (apart from the business of a 'final solution'), formulated by Hilbert, but widely current at the turn of the century:

2. Set-theoretic and other abstract notions constitute a mere *façon de parler*, and thus can be eliminated straightforwardly.

The second theorem pinpoints a particular class of counter examples to 2, since the specific use of abstract notions in the instantaneous consistency proofs of 1 cannot be so eliminated.[40]

Though the use above of the second incompleteness theorem has unquestionable elegance and charm, detailed inspection of scientific experience establishes more. As to 1, abstract notions are essential not only for consistency proofs (which constitute a kind of singularity in mathematical reasoning) but, generally, in algebra and number theory. Moreover, and this is a philosophical defect of the aim of eliminating abstract notions from proofs: when this can be done, essential knowledge contained in the proof is liable to be lost.

As to 2, by the turn of the century there had hardly been time to learn to use set-theoretic methods efficiently: Hilbert's conviction was quite *consistent* with the 'empirical evidence', which is not the same thing as being *supported* by the evidence! In fact, those methods were used particularly cautiously, even though

[40] For systems which prove their own consistency, the corresponding abstract notions cannot be eliminated from the proof of the equivalence between $\mathcal{F}$ and $\mathcal{F}_1$ on p. 77.)

(including theparadoxes) less mistakes had been made with sets than in finitist consistency proofs in the twenties (cf. p. 79 and [∞]).

Finally, at least as a matter of common sense, the foundational problems (about the 'certainty' or 'security' of mathematical knowledge) which Hilbert had in mind do not seem at all promising. After all, though surely not the only reliable means of knowledge, mathematical proofs have long stood out by their certainty. Further analysis of that certainty, in terms of anything remotely resembling existing ideas, is therefore at best a calculated risk, and the more specific aim of increasing that certainty still further *assumes* that the certainty already achieved is not 100%. Here it should be recalled from p. 86 that the unwinding of impure proofs into pure ones, originally presented as eliminating dubious abstract principles, simply yields other information. More generally, preoccupation with certainty is liable to draw attention away from other possibly genuinely problematic and therefore less sterile aspects of proofs.[41]

Brouwer's scheme

Brouwer's intuitionistic doctrine is best known for its *polemical side*, about defects of set-theoretic definitions, and of (Hilbert's problems about) formal rules. The corresponding *positive side* is Brouwer's aim of doing what Russell and Hilbert neglected: to make the mental activity of proofs, not only formal derivations, into the principal subject of foundational studies. An essential (though by no means well known) element of that positive side is a new interpretation of the logical operations, as maps from proofs to proofs.[42] Naturally, for this different interpretation, some of the familiar formal laws of logic fail; what is less well known is that new ones hold.

This positive side was not much stressed by Brouwer himself, whose polemics insisted on a reform of mathematics (or at least of its exposition). Moreover, he

[41]This paragraph obviously conflicts with several old ideas. For example:

1. Contrary to the tradition going back to Descartes, doubts and assertions (including restrictions and extensions of principles of proof) are here treated symmetrically.
2. Contrary to the opening paragraph of Gödel [1944], here logic is not expected to set general norms prior to all science (not even to all mathematics).

Of course, having survived, these old ideas sound plausible enough in the abstract. But they evidently conflict with scientific experience, and inspection of the latter shows up their *obvious* oversights; for example (in the case of 1) doubts can be dubious too, and (in the case of 2) norms valid for literally all imaginable experience are liable to be useless for any particular domain. (This last point is illustrated in the subsection on Russell's schema too, by the weakness of universal systems).

The price for dropping the simple-minded ideas 1 and 2, about the nature of knowledge as one says, is high (cf. the subsection below on the problems involved in a useful representation of proofs).

[42]Readers familiar with the intellectual climate of the first quarter of this century, mentioned already on p. 64, will recognise here the then-privileged place of mind in nature.

never presented so charmingly simple a scheme as those of Russell and Hilbert (or, for that matter, as 1 and 2 in note 41). Nevertheless, what he said is clear enough to be examined in the style of the previous subsections. Once again, results by Gödel and by others profiting from his work, correct wide-spread first impressions (both of intuitionistic doctrinaires and their critics) about the logical properties of Brouwer's scheme.

1. Since Brouwer's doctrine stressed inadequacies of formal systems, the doctrinaires could be expected to err in the opposite direction: not seeing what formal systems could do.

   For example, Heyting [1956] says that no such system can embrace all valid methods of proof. This is true, and in fact made quite specific by Gödel's second theorem: no system 'embraces' methods which use its own validity. (Actually, the idea of 'embracing' the totality of proofs is mind-boggling even when specialized to proofs of the one 'theorem': $0 = 0$).

   But this leaves the question whether a formal system 'embraces' all its valid theorems; in other words, its completeness (naturally, for the intended intuitionistic interpretation). In the fifties and sixties such matters where taken up, and several *positive* results were obtained; as on p. 80, with special care to reduce the (intuitionistic) 'abstract nonsense' (not, however, down to arithmetic, but to the subject of so-called lawless sequences; cf. Troelstra [1977] for an exposition).

   Again, though Brouwer repeatedly objected to formal consistency as a sufficient criterion of soundness, he neither saw its significance (on p. 76) nor pin-pointed its limitations as exactly as Gödel did in [1931a].

2. On the other side of the fence, the critics, perhaps encouraged by Brouwer's dramatic 'contradictions' with ordinary logic ('contradictions' with a different interpretation of the logical operations!) objected to (his) supposedly paralysing restrictions on mathematical practice.

   Gödel was one of the first to expose (in [1933]) the triviality of these particular, still widely believed objections. Since then we have learnt, slowly, to set out the bulk of mathematics quite elegantly by efficient use of intuitionistic methods (to be compared to p. 89, and especially Section 10, on the slow exploitation of specifically set-theoretic methods).

3. Gödel was also one of the first to recognize genuine defects. Specifically, in his early notes on [1958] preserved at Princeton, he pin-pointed a principal defect of Brouwer's logic (which is also not yet widely known). In terms of this chapter: provided the comparison applies, the unwinding of derivations built up by intuitionistic formal rules is of about the same order of complexity as for the corresponding 'usual' systems. For example, in the case

> of theorems (6.2 on p. 75) which show that some diophantine equation has infinitely many solutions, the 'unwinding' consists in computing the n-th solution of the equation (in some given ordering).

In short, by 1–3, the original impressions (of all concerned) about the *logical* properties of Brouwer's scheme were about as wrong as those of Russell and Hilbert about their schemes. (Naturally, the corrections of the more famous errors have also become more famous.) But given [1930] and [1931], which show how easy it is to correct that kind of error, it was a foregone conclusion that the logical properties of Brouwer's scheme would be straightened out sooner or later.

In contrast, a philosophical assessment of Brouwer's scheme is more delicate. Since he proposed a *reform* (not analysis) of ordinary mathematics, experience of the latter is not enough. Instead it is necessary to apply a basic lesson from general scientific experience, on the *choice of data* needed to represent relevant features of the principal objects of study. As already mentioned, Brouwer's scheme was to study the mental activity of proofs. His polemics certainly show up the superficial character of known representations (not only in formal systems, but in ordinary texts with diagrams and all the rest). But he has no satisfactory answer to the question:

> What better scheme is there than the known representations of proofs?

According to Brouwer [1948a], he proposed to explore (his) deepest consciousness, presumably to arrive at ultimate reasons, as others have chased final causes. This sort of pretentiousness is of course suspect, because it generally goes with simple-mindedness. But here it is possible to be more precise.

The proposal errs by ignoring the basic lesson alluded to above, as follows. Reports from (his) deepest consciousness may be *quite enough for us to recognize the* (mental) *object involved, but useless for its theoretical study*. Perhaps to be compared to reports on the shape and colour of minerals or plants in natural history; going into 'deepest' consciousness then corresponds to a meticulous description of nuances in shape, and shades of colour. True, such data are amply sufficient for recognizing the (mental or physical) object meant; but they are not adequate for a theory. Thus, in the case of minerals, rough knowledge of the molecular structure tells us much more about their physically significant properties than do very precise superficial data. In the case of proofs, a similar improvement would be expected from even a crude idea(lization) of the memory structures involved. As matters stand today, Brouwer's aim was shortsighted; for though the others neglected the potentially interesting topic of (actual) proofs altogether, what he had to add to the subject added too little, and stopped him from looking for results which are independent of our ignorance (about proofs).

Evidently, this ignorance concerns a theoretical analysis since, practically speaking, we know a great deal about proofs, using them constantly as tools. It is precisely in such circumstances that only a really substantial theoretical advance has a chance of competing with unanalysed practical knowledge or perceptive *aperçus*; as a corollary, simple-minded schemes are then intellectually especially unsatisfactory. But though unsatisfactory (and this has been a general lesson of this chapter) a study of such schemes can be fruitful; the (mathematical) uses of incompleteness properties, or of elementary formulas and their model theory (in the subsections on pp. 74, 78 and 84) developed form Gödels' studies of Russell's and Hilbert's simple-minded schemes. A price paid for the philosophical weakness of these schemes was the imagination needed to find the reinterpretations which lead to applications of Gödel's results.

From foundations to technology

The need for such reinterpretations is not particularly unusual in the sciences, especially if the work in question was originally used to refute a theory. But the frequency can be expected to be particularly high in the case of those foundational schemes or theories which, in line with Kant's view mentioned on p. 64, make 'possibilities in principle' their primary object. For, given that preoccupation, they will be satisfied with answers that *simulate* striking properties of the (mathematical) phenomena under study. This cannot be expected to tell us much about the phenomena themselves. But it will lead to technological progress, provided (as is natural) the answers are formulated in familiar (say, mechanical) terms. For then there is a chance that the effects which originally struck us can be achieved by those familiar means too, perhaps even more economically than by the things originally considered. *Achieving a given effect, rather than understanding a given (natural) phenomenon, distinguishes technology from science.* Evidently, the word 'technology' is suggested by the relation (on p. 62) between Frege's rules for logic which had the ethereal purpose of analysing deduction, and the application of computers to non-numerical data. But the word also applies quite well to the mathematical uses of Gödel's results.

The parallel with technology applies also to the relative difficulty of discovering foundational results (which usually correspond to first impressions) and effective uses which, by above, require imagination. In contrast, in the case of (what are normally called) fundamental sciences the applications look after themselves. Incidentally, the relative difficulty of those two kinds of discoveries is badly obscured by the slogan of 'pre-established harmony', so popular among logicians from Leibniz to Hilbert.

6.8 Background to [1938]: Zermelo's Set Theory

The material in Sections 8, 9 and 10 will appeal most to either mathematicians or historians of mathematics. The reason was given already on p. 50: modern set theory has turned out to be of some interest when regarded as a specialized branch of mathematics. Its original appeal as a foundational system has turned out to be deceptive, as argued in Bourbaki [1948], with a few exceptions (in corners of advanced mathematics) discussed above and illustrated on p. 107 below.

Sets before the twenties

Older readers may still remember the embarrassing level of traditional 'debates' about sets and their properties.

- At one extreme there was the fixation on the paradoxes; despite the fact that, for example, the most famous version, due to Russell, has a perfect parallel in arithmetic if one assumes that there is a greatest integer.

- Even more thoughtless were the sweeping generalities stirred up by those paradoxes.

  The most innocent *connerie* was the idea that, somehow, axiomatization would be a safeguard, as if there were no inconsistent formal systems (like Frege's).

  The most pretentious was the appeal to a general theory of knowledge (along the general lines familiar from Section 2 and, especially, p. 92 on the matter of proofs). In the particular case of sets, the stress was on definitions (rather than proofs) from which sets were supposed to be 'constructed', to be compared to the then-current business of sense data from which physical objects are 'constructed'; all this despite the fact that (corresponding to p. 91 on proofs of $0 = 0$) any *one* set has a truly mind-boggling 'totality' of definitions, and that sense data tend to fall to pieces on a closer look while (most) objects do not.

By the end of the twenties, at least some mathematicians had become sufficiently familiar with the vague mixture of things called 'sets' to decide which objects they wanted to talk about, instead of relying on accepted usage or on its (premature) codification in formal axioms. Some basic distinctions had been made, to be compared to distinctions between natural, rational, real or complex numbers: without such distinctions, properties of $+$ and $\times$, which are common to all of those numbers, are trivial for any one kind (and 'paradoxes' result if one puts together properties which are of interest for different kinds).

Fat hierarchies of sets

In particular, in Zermelo [1930] there is a lucid description of what is nowadays called the *cumulative hierarchy of sets*; that is, sets generated by iterating the *power set operation* $\mathcal{P}$:

$$\mathcal{P}(x) \text{ is the collection of } all \text{ subsets of } x$$

(more precisely, iterating $\mathcal{P}$ transfinitely up to some stage α).

The simplest particular cases of the hierarchies of sets described in Zermelo [1930] are V_ω and $V_{\omega+\omega}$ where

$$\begin{array}{lcl} V_0 & = & \emptyset \quad \text{(the empty set)} \\ V_{\alpha+1} & = & \mathcal{P}(V_\alpha) \\ V_\omega & = & \bigcup_{n\in\omega} V_n \\ V_{\omega+\omega} & = & \bigcup_{n\in\omega} V_{\omega+n}. \end{array}$$

More simply, also for α beyond $\omega+\omega$, and without distinguishing between successor and limit ordinals α:

$$V_\alpha = \bigcup_{\beta<\alpha} \mathcal{P}(V_\beta).$$

V_ω or, more pedantically, the structure $(V_\omega, \in_\omega)$ satisfies the familiar axioms of set theory *without* the axiom of infinity; $V_{\omega+\omega}$ satisfies those of Zermelo [1908], from which the formal system called 'Zermelo's axioms' in the current literature is derived.[43]

Zermelo [1930] introduced an additional parameter, an arbitrary collection U (for *Urelemente*) of distinct atoms without any elements, and a corresponding hierarchy $V_\alpha(U)$ with $V_0(U) = U$. This is useful, and used in practice; for example, in number theory the natural numbers are thought of as atoms. But the $V_\alpha(\emptyset)$, called V_α above, are good enough for the present purpose.

Non-elementary axiomatizations

Zermelo [1930] contains also the *non-elementary axiomatization* (mentioned on p. 66) for all segments of that hierarchy up to so-called *Grenzzahlen* α, also called *inaccessible cardinals*. Thus, in contrast to Peano's or Dedekind's axioms, Zermelo's are not categorical, but determine a family. A trivial modification yields categorical axioms for such specific segments as the first (where $\alpha = \omega$) or the next (the first uncountable inaccessible).

In terms of p. 66 the 'principal feature' of the hierarchies in Zermelo [1930] is the (binary) membership relation $\in$. There are three non-elementary axioms (the remaining ones being elementary):

1. *Well-foundedness of* $\in$ for arbitrary predicates P:[44]

$$\exists u P(u) \;\Rightarrow\; \exists x[P(x) \wedge (\forall y \in x)\neg P(y)].$$

[43] Interested readers are advised to stop a moment, and actually verify a couple of axioms. Remember that V_ω is the collection of hereditarily finite sets and that, for $n \geq 1$, $V_{\omega+\omega}$ is the closure of $V_{\omega+n}$ under power set and subset formation.

[44] Its contrapositive is called $\in$-*induction*.

This holds for all V_α, since they are 'built up from below': if $u \in C_\alpha$, there is a least $\beta \leq \alpha$ for which some $x \in V_\beta$ and $P(x)$.

2. *Comprehension*, again for arbitrary predicates P:

$$\forall x \exists y \forall z (z \in y \;\Leftrightarrow\; [z \in x \wedge P(z)]).$$

This too holds for all V_α: if $x \in V_\alpha$, $x \subseteq V_\beta$ for some $\beta < \alpha$ and hence also $y \subseteq V_\beta$; but since $V_\alpha \supseteq \mathcal{P}(V_\beta)$, $y \in V_\alpha$ too.

3. *Replacement* for arbitrary *functional* relations R (or, equivalently, predicates of ordered pairs):

If the domain of R is restricted to a set $\in V_\alpha$, so is the range.

This holds for $\alpha = \omega$ and, if one wishes to be pedantic, also for $\alpha = 2$. It does not hold for $\alpha = \omega + \omega$, etc.

A principal result of Zermelo [1930] is this: granted the rest of the axioms, replacement holds only for V_α where α is *strongly inaccessible*, that is card $V_\alpha =$ card α.[45]

Moreover, and this is the non-elementary *axiomatization of the family of V_α where α is strongly inaccessible*:

if any structure (D, E) with domain D and the binary relation E on D satisfies the axioms of Zermelo [1930], then (D, E) is isomorphic to some $(V_\alpha, \in)$ where V_α is in that family.

Digression on the passage to formalizations of set theory

In contrast to those derived from Peano's or Dedekind's axioms (by the passage on p. 67), Zermelo's axioms are better known than the non-elementary axiomatizations, and are familiar from current texts on logic or from introductions to mathematical texts. As always, the three (infinite) schemata[46] of formal axioms arise form the three non-elementary axioms. But there are some curiosities.

[45] Equivalently, in terms of Cantor's cardinal arithmetic: for $\beta, \gamma, \beta_\delta$ (all $< \alpha$),

$$\beta^\gamma < \alpha \quad \text{and} \quad (\sum_{\delta<\gamma} \beta_\delta) < \alpha.$$

[46] Such schemata can be finitely generated by introducing a second type of variable, for predicates (usually denoted by capitals X), a new binary relation symbol η ($x \eta X$: X applies to x), and axioms for the X corresponding to the (finitely many) syntactic rules for building up formulas. In the particular case of set theory, those X are called 'classes' and, as Gödel observed in [1940], $\in$ can be conflated with η in a natural way.

For example, non-elementary comprehension implies not only the schema for formulas F with a single free variable z, but also

$$\forall u_1 \cdots \forall u_n \forall x \exists y \forall z (z \in y \;\Leftrightarrow\; [z \in x \wedge F(u_1, \ldots, u_n)]).$$

This schema (with 'parameters' $u_1, \ldots, u_n$) is not formally derivable from the other one. (This has a parallel in the case of Dedekind's axioms but not, as an easy exercise shows, in the case of Peano's).

More interestingly, the formalizations are satisfied by V_α for suitable *accessible* α too. The proof is similar to the one of 6.5.1, but the results is incomparable:

1. the structures V_α involved are not countable, since already $V_{\omega+1}$ has the cardinal of the continuum

2. not only the (elementary) logical symbols retain their standard meaning; for example, $\mathcal{P}$ does too.

Typically, these simple facts were mentioned relatively late, in Montague and Vaught [1959], and are still not prominent in the literature.

Gödel was the first to find really striking differences between the non-elementary axiomatizations and their formalization. With one proviso, those differences are very well illustrated by the differences between the full Euclidean plane (or, more simply, our geometric imagination) and its 'thinned' part constructible by use of ruler and compass (already mentioned on p. 66); both the full and the thin plane satisfy Euclid's elementary axioms, but only the former satisfies the non-elementary continuity axiom. By and large, the geometrically most obvious properties are easier to verify for the Euclidean plane, even when they hold for the thin part too (for example, the existence of a regular polygon of 17 sides). Also, Euclid's axioms do not decide every elementary formula: already the Greeks asked questions which have a different answer for the full plane and its thin part.

The proviso is connected with the logical form of questions common in geometry and in set theory. The former are often purely existential, and so a solution for the thin plane is automatically a (refined) solution for the full plane, as in the case of a regular polygon of 17 sides. In set theory the logical form is more complicated, and so solutions to formally the same problem will be incomparable; in the case of the plane, the set $\{(x, 0) : (x^3 - 2)x = 0\}$ consists of one point in the thin part ($x = 0$), but of two points in the full plane.

The much more heavily publicized comparison with the parallel axiom is wholly irrelevant to Gödel's contribution, since it has nothing to do with the difference between elementary and non-elementary axioms (the parallel axiom is undecided by the remaining axioms of Euclid *together* with full continuity).[47]

We now return to the principal topic.

[47]The business of the parallel axiom corresponds quite well to relatively easy analyses of the non-elementary axioms in Zermelo [1930], sufficiently illustrated by the easy incompleteness argument for set theory in note 17.

Consequences of the non-elementary axioms

As an immediate pay-off, the familiar axioms are seen to be valid by inspection; and the more 'elementary' the axiom, the more segments α satisfy it.

Secondly, a point which Gödel like to stress (for example, in [1947]), Frege's formulation

$$\exists x \forall y [y \in x \;\Leftrightarrow\; P(y)]$$

is *obviously* false for all α when $P(y)$ is $y = y$; for example, for the segment up to ω (of hereditarily finite sets), all the infinitely many objects y in that segment satisfy $P(y)$, but every set x in the segment is finite.

One of the easiest consequences is the *axiom of choice*; for example, in the form:

> for any set x of disjoint (unordered) pairs $\{u, v\}$, there is a set y intersecting each pair of x in exactly one element.

Theorem 6.8.1 *The axiom of choice is true for each V_α.*

Proof. If $x \in V_\alpha$, all the $\{u, v\}$, and hence u and v, are in V_β for some $\beta < \alpha$. So $y \subseteq V_\beta$, and hence $y \in V_\alpha$. $\Box$

In other words, for the fat (or 'full') hierarchy, the axiom of choice is quite evident. The fact that this axiom was used tacitly till Zermelo [1904] should be compared to similar uses in geometry of axioms for *order* which were not listed by Euclid: such tacit uses do not cast doubt on the soundness of the axioms (for the intended meaning), though possibly on the competence of the axiomatizers.

Cantor's *continuum hypothesis CH* asserts, in effect:

> any subset of $V_{\omega+1}$ (which is in one-one correspondence with the real numbers) is either in one-one correspondence with $V_{\omega+1}$ itself or with a member of $V_{\omega+1}$ (equivalently, countable or finite).

Theorem 6.8.2 *The continuum hypothesis is decided by the non-elementary axioms.*

Proof. CH concerns only elements of $V_{\omega+4}$. Now, for the natural definition (say, $\mathcal{C}_{\omega+4}$) of $V_{\omega+4}$, the non-elementary axioms are obviously categorical (even without replacement). That is, if the formula $\mathcal{C}_{\omega+4}$ holds in any model (D, E) of the axioms for an object c in D, then (c, E_c) is isomorphic to $(V_{\omega+4}, \in_{\omega+4})$, where $E_c = E$ is the restriction of E to $c \times c$. $\Box$

Thus CH is decided just as, say, the prime pair conjecture[48] is decided by Peano's axioms; only we don't know which way.

The so-called *generalized continuum hypothesis* GCH is obtained from CH when ω is replaced by arbitrary infinite α.[49] A moment's reflection on GCH conveys a feeling for the content of non-elementary decidability, because GCH is not *obviously* decided by the non-elementary axioms. It is not if, for example, GCH is true for all infinite α less than the first *un*countable strongly inaccessible cardinal, but not for all. Then GCH is true in the smallest segment V_α which satisfies those axioms, but not in all such segments.

Bibliographical remarks

Zermelo [1930] has made little impression. For one thing, though the non- elementary character of the axioms is prominent enough, there is no hint of such easy but memorable consequences as those listed in the last paragraph. Perhaps more significantly, two basic points were slurred over.

First of all, the reader was not prepared for the striking effect of adding the replacement axiom (on the ordinals α for which the axioms are satisfied by V_α). As early as 1931 Gödel alluded to some reservations, evidently on this score, in his correspondence with Zermelo; the latter did not take them up, and Gödel repeated them throughout the thirties in his notes for lectures and courses. Those reservations go well with the fact that the axiom was a late-comer, having been introduced in the twenties by Fraenkel (in a restricted form, for definition by transfinite recursion) and in Skolem [1922] (for formal reasons), but was first properly used only by von Neumann [1928]. There it replaces the power set axiom for developing a good part of then-current set theory; but, above all, it is used for a *canonical well ordering* (by $\in$) of what are now the standard (set theoretic) 'numerals' for ordinal numbers, in which all well orderings can be embedded. As we see things now, von Neumann's work suggests a *thinning* of the hierarchy V_α. For any so-called *regular cardinal* ρ (for example $\aleph_1$, in Cantor's notation for the first uncountable), let

$$V_\alpha^\rho = \bigcup_{\beta<\alpha} \mathcal{P}^\rho(V_\beta^\rho),$$

where $\mathcal{P}^\rho(x)$ is the set of all those subsets of x which have cardinal $< \rho$. Then, for $\alpha \geq \rho$, $V_\alpha^\rho = V_\rho^\rho$ and V_ρ^ρ satisfies the non-elementary axiom of replacement

[48]Being purely universal, Fermat's conjecture is a less suitable analogue because (by the discussion on p. 79) if the conjecture were proved consistent with number theory, the same methods would prove the conjecture itself. Basically: if $\forall\vec{x}(p \neq 0)$ is consistent, then its negation $\exists\vec{x}(p = 0)$ is not derivable, and hence (being existential) is false. Then $\forall\vec{x}(p \neq 0)$ is true.

[49]The restriction to infinite α is needed, since $\alpha = 0$ and $\alpha = 1$ are the only finite α which would satisfy the unrestricted version.

(but generally not the power set axiom). Familiarity with such V_ρ^ρ is a useful preliminary for really effective use of replacement.

In this way one also comes to see the principal open problem presented by the hierarchy V_α: not the innocuous power set operation, but the *number of its iterations*[50] (in fact, the problem remains open and is a principal subject of Section 10). But for the majority of potential readers of Zermelo [1930] at the time, the operation $\mathcal{P}$ *was* problematic, the key words being 'vicious circle principle' or 'impredicativity'. Consequently, quantification over arbitrary predicates, so essential to non-elementary axioms, seemed to be an evasion of the problem on the part of Zermelo. Perhaps it was; in any case, all he did was to repeat his would-be telling terminology of *definite Eigenschaften* (predicates) which had been ineffective since Zermelo [1908] (where it was first introduced). The irony is that he never seems to have spotted the crucial ambiguity between:

1. definite in the sense of well-defined, perhaps even: decidable (as in Gödel [1931]: *entscheidungsdefinit*)

2. having a definite extension (as is implicit in Cantor's explanation of sets as: varieties which can be grasped as a unity, varieties being defined by predicates).

As to 2, this is ensured by the restriction of the comprehension axiom to: $z \in x$ (in Gödel's terms in [1947], sets y are *sets-of-something*: of x's). But for the majority of readers hung up on the business of definitions (or on *predicativity*, in the jargon of the day), only sense 1 was natural. Certainly, Zermelo himself had made great progress in the twenty five odd years before Zermelo [1930] appeared, but not enough to find the *mot juste* sufficient to remove that hang-up.

A much easier method towards this end would have been to look at the alternatives which predicativist critics had offered; specifically, what have come to be known as *ramified hierarchies*. (Of course, they were originally intended as hierarchies of definitions, while nowadays we look at the sets so defined). The literature ranged from Poincaré's reflections on the matter to hoary details in *Principia*, and to particular examples in Weyl [1918]. Zermelo himself may have had too little confidence in Poincaré's predicativist philosophy to look at those alternatives; for one thing, he had had bad experience of Poincaré's reflections on the mechanical theory of heat, which are criticized in Zermelo [1896].[51]

[50]This is quite parallel to the 'problem' presented by V_ω if one 'believes' only in finite sets x: for each such x, $\mathcal{P}(x)$ is not problematic, but the notion of 'arbitrary' (finite) iteration is.

[51]Incidentally, that paper has also made little impression on mathematicians, although it contains the first really elegant proof of the *recurrence theorem* for dynamical systems.

6.9 Constructible Sets

Gödel introduced a formal variant of the predicativist ramifications at the end of Section 8. Earlier attempts were much clumsier; as so often, but perhaps exceptionally so here (or, on p. 90, in Brouwer's logic) since (by p. 94) it goes against the grain to think of definitions instead of the objects defined (or of proofs instead of the theorems proved). Even so, in retrospect the complications appear pretty marginal, mainly because of pointless restrictions: to sets of integers (instead of abstract sets), or to so-called simple (instead of cumulative) types.

First step: a restricted power set

In any case, if the original intentions of ramified *definitions* were to be formalized after Zermelo [1930], the natural scheme was to use as definitions: formulas of the elementary language of set theory with (the usual meaning of) its one principal symbol $\in$, and to let quantifiers range over a given (in applications: 'previously' defined) set x. Then a formula $F(z, u_1, \dots, u_n)$ defines the subset $\{z : F(z, u_1, \dots, u_n)\}$ of x (where $u_1, \dots, u_n$ are elements of x), and one writes $\mathcal{P}^-(x)$ for the collection of all sets defined in this way.[52]

The ramified hierarchy is defined (by ordinal recursion) as:

$$L_\alpha = \bigcup_{\beta<\alpha} \mathcal{P}^-(L_\beta).$$

Then:

- For $\alpha \leq \omega$, $L_\alpha = V_\alpha$.
- For $\alpha > \omega$, in sharp contrast, $L_{\alpha+1}$ has the same cardinal as L_α (and if the 'parameters' $u_1, \dots, u_n$ were omitted, $\mathcal{P}^-(x)$ would even be countable for any x).

The hierarchy is *ramified* in the sense that (at each stage new definitions of any one set are introduced, but also) new subsets of L_α appear beyond $\alpha + 1$, in contrast to V_α. For example, when $\alpha = \omega$, new sets of integers appear in $L_{\omega+n+1} - L_{\omega+n}$ for each $n < \omega$. An immediate consequence is that, in general, the comprehension schema is not satisfied by the sets $\in L_\alpha$ (certainly not for $\omega < \alpha \leq \omega + \omega$).

By the turn of the century it was not unusual to begin analysis with the principle of the least upper bound: in logical terms, the comprehension schema was used at the very start. This seemed desperate if, like Russell in *Principia*,

[52] By note 46, it is clear how to replace $\mathcal{P}^-$ by a finite number of operations with a more algebraic look, at the price of slowing up the growth of the hierarchy below. This was done by Gödel in [1940], using seven operations. Below, following his original exposition in [1938], $\mathcal{P}^-$ itself will be used.

one wanted a (theory of some) ramified hierarchy to provide a 'universal' system for mathematics. He introduced the so-called *reducibility axiom* which says, in effect, in the case of subsets of L_ω, that no new ones appear in L_α for $\alpha > \omega + 1$. These tactics seemed equally desperate, especially coming from a philosopher who had compared the advantages of the axiomatic method to those of stealing over honest toil; more so than Zermelo's mildly evasive business of 'definite predicates' on p. 100.

Second step: the number of iterations

Already back in 1931, Gödel concentrated on another weakness of the ramified hierarchy in *Principia*: it stopped at $\omega+\omega$, for no good reason. More specifically, footnote 48^a of [1931] points out that the consistency of (the appropriate formal theory of) $L_{\omega+\omega}$ can be proved in $L_{\omega+\omega+1}$.

It might be added that, by 1930, transfinite definitions (for example, of the real closure of an arbitrary ordered field) were common in mathematics. And usually there was a very good reason for stopping at some stage α: either when no now objects are introduced after α, or when the objects accumulated by that stage satisfy some clearly stated (closure) condition.

Footnote 48^a was essentially negative, containing no hint, even remotely satisfactory for predicativist aims, where to stop (beyond $\omega + \omega$). Gödel's decision was not to stop the hierarchy L_α at all. More formally (with Zermelo [1930] as background) he did not stop before κ, the first uncountable inaccessible ordinal (defined on p. 96).

Gödel then *discovered a problem* for which the decision of where to stop the hierarchy was irrelevant.

Constructible sets: reculer pour mieux sauter

To answer the broad question:

What does one want to know about L_κ?

a good start is:

Which of the usual axioms of set theory are satisfied by L_κ?

Several are verified immediately. For example, *extensionality* or *pairing*. But also the non-elementary axiom of *well foundedness* of $\in$, simply because $(L_\kappa, \in_\kappa)$ is the restriction of $(V_\kappa, \in_\kappa)$ to $L_\kappa \times L_\kappa$. There are also pleasant surprises, the first one being:

Theorem 6.9.1 *L_κ is closed under the power set operation.*

Proof. For example, let $x = L_\omega$. There are $2^{\aleph_0}$ subsets of L_ω, so $\leq 2^{\aleph_0}$ such subsets in L_κ. Suppose they appear in

$$A_\kappa = \{\alpha : \alpha < \kappa \text{ and } L_{\alpha+1} - L_\alpha \neq \emptyset\}.$$

Since κ is regular and $> 2^{\aleph_0}$, A_κ has an upper bound $\alpha_\kappa < \kappa$. Then the set y defined (in L_{α_κ}) by the formula '$z \subseteq L_\omega$' is in $L_{\alpha_\kappa+1}$.[53] □

The proof above is certainly not difficult, once one has understood that y is to be the set of subsets of L_ω that occur in L_κ, and not $\mathcal{P}(L_\omega)$ itself. Of course, the proof by inspection that, for limit numbers α, V_α satisfies the power set axiom, is even simpler (cf. note 43).

The *replacement* property, which implies *comprehension* (by taking characteristic functions), presents *the* new aspect. There is no evidence at all that the *non-elementary* version is satisfied in L_κ, but:

Theorem 6.9.2 *The formal schema of replacement is satisfied in L_κ.*

Proof. By induction on the logical complexity of the (elementary) formulas defining the functional relation involved, and use of familiar closure properties of sets of ordinals.[54] □

The *axiom of choice* also holds, and again the proof is a little more involved than mere inspection (in 6.8.1). But it also gives more:

Theorem 6.9.3 *There is an explicit definition of a well ordering of L_κ.*

Proof. By recursion on $\alpha < \kappa$. Suppose the elements u of L_α are well ordered by $<_\alpha$, which induces a well ordering of finite sequences $\vec{u}$ of L_α (also written $<_\alpha$). As usual, elementary formulas F are numbered. For $x \in L_{\alpha+1} - L_\alpha$, let F_x be the first formula which defines x (in L_α):

$$x = \{z : F_x(z, \vec{u}\} \text{ for suitable } \vec{u} \text{ in } L_\alpha.$$

Let $\vec{u}_x$ be the first such $\vec{u}$. Then the elements of $L_{\alpha+1} - L_\alpha$ are ordered lexicographically according to $(F_x, \vec{u}_x)$. It turns out that the (natural) definition of $<_\kappa$ uses only quantification over elements of L_κ.[55] □

Using the parallel on p. 97, we can fairly say that the definition of the well ordering $<_\kappa$ of L_κ is easier than Gauss's construction of a regular polygon with 17 sides.

[53]Clearly, the argument applies not only to κ, but to any cardinal β which is regular and $> 2^{\aleph_0}$. As it stands, the argument leaves open whether, for such β, new subsets of L_ω appear in $L_\kappa - L_\beta$ (cf. p. 106).

[54]Cf. the few lines on p. 456 of Barwise [1977] needed for a full proof.

[55]Incidentally, similar care is needed in checking that the (natural) *definition $\mathcal{L}$ of constructibility* is invariant (that is, it defines L_κ both when its quantifiers range over L_κ and over V_κ), and that $\forall x \mathcal{L}(x)$ holds in L_κ (that is often written: $V = L$).

Bibliographical remarks

In keeping with his reservations (mentioned on p. 99), Gödel first tried to do without the replacement property, and to describe the constructible hierarchy L_α only for $\alpha <$ card $V_{\omega+\omega}$; in particular, without using von Neumann's canonical well ordering (p. 99). Instead, well orderings had to be defined (painfully) in $V_{\omega+\omega}$.[56] The simplification achieved by using von Neumann's notations (for which *higher types* are needed) provided the second memorable lesson (after footnote 48^a of [1931]) in Gödel's education on the virtues of transfinite iteration.

The titles of [1938], [1939] and [1940] about *consistency* properties of formal set theories do not even mention the notion of constructible set, although he considered the use of that notion as his most significant contribution in the area.[57] Actually, his choice of titles involved him in painful details: it had to be verified that the properties of V_κ used to establish facts about L_κ were formally derivable from the axioms listed in the theories considered. Gödel's strategy of going into details avoided controversy at the time, as in note 16. But it also left the (false) impression that the most urgent (if not the only fruitful) problem was to complement his work by establishing the consistency of the *negations* of the propositions he had established for the constructible sets; in other words, to show their *formal independence*. This turned out to be of a different order of difficulty (cf. p. 108).

Gödel himself paid a price for his cautious tactics. For example, in footnote 2 of [1947], he recognized the absurdity of stressing the consistency of the axiom of choice since (by 6.8.1) it is as easily seen to be true for the hierarchy V_α as the other axioms.

A more startling oversight (corrected in [1964]) occurs in [1947]. There he assumed the formal independence of the continuum hypothesis, and played with the idea that CH should be judged by its arithmetic 'fruits' (that is, its arithmetic consequences). Certainly, mere consistency leaves open the possibility that CH has new (even false) arithmetic consequences; but a glance at his own definition of L (in particular, at $V_\omega = L_\omega$) 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.[58]

[56] According to conversations in the fifties and sixties, Gödel originally tried to define L as an inner model for (arbitrary models of) Zermelo's set theory, where nothing like the 'standard' well-orderings of von Neumann's ordinals by $\in$ are available. Suitable well-orderings have to be defined as, for example, in modern expositions of the constructible analytical hierarchy (of sets of natural numbers), and Gödel had no taste for such exercises. His notes for various lectures at Notre Dame in the thirties (left to the Institute of Advanced Study at Princeton) confirm these conversations.

[57] This was not a late afterthought (for example, in his comments reported in Kleene [1978]), but is already stressed in his notes for lectures in the thirties.

[58] The opposite view was publicized for nearly a decade, before convincing though temporary use was made of it by Ax and Kochen [1965] in their proof of the decidability of the theory of p-adic fields (by means of the CH).

For a realistic view of Gödel's heuristic ideas on p. 150, two more points are relevant:

- He himself missed several interesting results by giving attention only to the theorem stated, not to the details of its proof.

  This concerns less formal errors (for example, at the end of [1933a]), but certainly his review of Skolem [1933a] (mentioned in note 35).

  Returning to L: according to Gödel's notes, not he, but Ulam, steeped in the Polish tradition of descriptive set theory, noticed that the definition of the well-ordering (6.9.3) of subsets of ω was so simple that it supplied a non-measurable PCA set of real numbers (when all objects involved are taken from L).

- Conversely as it were, Gödel tended to be uncritical of logically exciting claims; for example, regarding non-standard models (in [1974], admittedly written in the seventies; cf. p. 61). He attributed the scepticism of number theorists to broad prejudice, mysteriously connected with the recursive undecidability of Hilbert's tenth problem. In fact, later developments more than justify the suspicions of the majority.[59]

GCH: a variant of the axiom of reducibility

We now come to Gödel's principal discovery about the (thin) hierarchy L_α. To understand the issue, it is necessary to recall p. 98 on the continuum hypothesis (in its intended sense, that is, for the fat hierarchy V_α), and the objects involved: *elements* of $\mathcal{P}(\omega)$, *subsets* X of $\mathcal{P}(\omega)$, *mappings* of X onto $\mathcal{P}(\omega)$ or ω. These objects are in $V_{\omega+4}$. In contrast, when the continuum hypothesis is meant for the thin hierarchy L_α, the corresponding objects can occur at levels far beyond $\omega+4$ (by note 53). Also, inasmuch as there are liable to be more of all these objects in L_α than L_β for $\beta < \alpha$ (and more in $V_{\omega+4}$ than in L_κ), the truth or falsity of the continuum hypothesis may well be sensitive to the length of the segments L_α (and to the kind of hierarchy) considered.

A corresponding sensitivity is found in a more familiar formulation of the CH in terms of *cardinal arithmetic*: card $\mathcal{P}(\omega)$ is the first cardinal $>\omega$. This is the least ordinal which is *not* in one-one correspondence with the set ω by a map in the stock of sets considered. The ordinal is denoted by ω_1 if all maps (of ω onto initial segments of the ordinals) are considered, and by $\omega_1^{L_\alpha}$ if only all such maps in L_α are considered. Evidently, $\omega_1^{L_\alpha}$ is liable to be $<\omega_1$ even though L_α and V_α

[59]The best-known claim for non- standard models is in Robinson and Roquette [1975] (where incidentally arbitrary ones are used, in no way tailored to their problem). But the only novelty is their use of a (known) generalization of Roth's theorem to arbitrary number fields, which has nothing to do with non-standard models.

have the same ordinals.[60] In fancy language, already used in note 55: while the property of *being an ordinal* is invariant or absolute, the property (of ordinals) of *being a cardinal* $> \omega$ is not.

This point is often overlooked in the (popular) 'debate' on the CH, where the *orderliness* of the ordinals (in V_κ or L_κ) is contrasted with the *mess* of $\mathcal{P}(\omega)$ (in V_κ): a similar mess is involved in the collection of maps (in V_κ) of ω onto initial segments of the ordinals. It does not seem at all surprising that we have not (yet) decided whether the two 'messes' match. This fact is perfectly consistent with the non-elementary decidability of CH 6.8.2; after all, we don't even know how to match up the surely less 'messy' set of prime pairs with the ordinals $< \omega$, though the matter is certainly decided by Peano's non-elementary axioms (as mentioned on p. 99).

Returning to the GCH: by a quite simple use of 6.5.1 (cf. also p. 97) Gödel established the GCH for L_κ, and by careful formalization in [1938] even a little more:

> for any model (D, E) of formal set theory, the 'inner model' defined by the condition $\mathcal{L}$ (in note 55) always satisfies the GCH.

This result is a consequence of the following more delicate property of the constructible hierarchy (where 'cardinal' refers to constructible maps):

> *Reducibility for cardinals.* If α^+ is the first cardinal $> \alpha$, then all subsets of L_α which occur in the hierarchy at all are already in L_α.[61]

A proof, in less than a page, can be found on pp. 465–466 of Barwise [1977]. It is remarkably similar to some early expositions by Gödel, especially in his notes for general lectures (for example, to the American Mathematical Society in December 1938). As he mentioned in conversation, the idea of the sort of argument involved occurred to him when he learnt Skolem's proof (6.5.1) as a student.

Other properties of L

Few other memorable properties of L were discovered in the 30 years after [1938], until the so-called *Souslin hypothesis* was shown to be false for L in Jensen [1972]: there is a dense ordering in L without end points, which is complete (for cuts in L) and any set (in L) of non-overlapping intervals is countable in L; but the ordering is not order isomorphic to the real numbers (of L).[62] As is clear from

[60] Cf. the example on p. 97, where the least integer n which satisfies $(\forall x > n)[(x^3 - 2)x \neq 0]$ for all x constructible by ruler and compass, is less than the least integer with that property for all real numbers x.

[61] This is a sharpening of 6.9.1 (closure under the power set operation).

[62] Incidentally, Gödel's notes (for example, in *Arbeitsheft XI*, 47-54) contain material on Souslin's hypothesis and the related matter of Aronszajn trees but, apparently, nothing in relation to L.

its title, Jensen [1972] concerns the details of L_α also for ordinals α which are *not* cardinals. Such concern would have appeared marginal (*Kleinarbeit*[63]) even only ten years earlier, since the significance of such L_α was first established in the sixties in so-called generalized recursion theory (cf. Chapter C.5 of Barwise [1977]).

The discoveries in Jensen [1972] about L were used by Shelah to solve a purely *algebraic-sounding* problem about certain (abelian) groups satisfying a condition W (for 'Whitehead'). Provided G is countable, the only case that arises in the (topological) context where the condition W was first introduced, they have a free basis. The problem, explained in detail in the very readable exposition Eklof [1976], was whether

all groups satisfying W have a free basis. (6.6)

Though the word 'set' is not mentioned in 6.6, the stock of sets considered is clearly liable to be relevant (as in problems on cardinals on p. 105). The more sets, the more groups (satisfying W): the (universal) proposition 6.6 is more difficult to satisfy. But for any given group G: the more sets, the more subsets of G, and so the better the chance of there being a basis (in the stock considered). Shelah established 6.6 for *constructible* groups (and bases), by essential use of Jensen [1972]. The example illustrates two points of general interest.

First, in terms of current mathematical jargon: how easy is it to guess whether some phenomenon in group theory is set-theoretical (as one speaks of, say, gravitational phenomena)? That is, whether knowledge of set theory is relevant or even decisive. Flash judgment does not seem reliable. For example, in the superficially similar case of the (particular, but also uncountable) abelian group G_0 of *bounded sequences of integers with pointwise addition*, the constructible part of G_0 has a (constructible) free basis, simply by use of the continuum hypothesis, as observed in Specker [1950]; but a quite different, much more informative proof in Nobeling [1968] establishes the result (and more) for G_0 itself.

Secondly: how useful is it to eliminate (some particular) specifically set-theoretical restrictions? (in the case above, of uncountable groups satisfying W: to constructible sets). As in similar cases, the answer will have to wait till group theorists are familiar enough with such groups or their constructible parts to have a chance of spotting whatever uses such objects may have; in other words, until St. Thomas's *adaequatio rei intellectu* applies (to those group theorists).

Fundamentalists in set theory follow the simple rule that all 'hypotheses' not known to be satisfied by (suitable) segments of the *full* hierarchy should be eliminated. But, at least in terms of the guiding parallel on p. 97, this simplicity is spurious. For example: in number theory, the full Euclidean plane or the field

[63] Gödel's *Arbeitshefte* contain some attractive *Kleinarbeit* too. For example, *XV*, 11-13 or *XVI*, 38-40 on the axiom of extensionality; but this is superseded by the much more thorough analysis in Scott [1962].

of all complex numbers is not always most relevant; in fancy language, it may be more rewarding to embed the numbers in some subfield with suitable properties (which we happen to know). So, if set theory is ever to become significant for number theory (if ω or V_ω is to be embedded in some variant of the full hierarchy, with the axioms of formal set theory - or of a subsystem! - playing the role of the axioms for fields above) a prerequisite is that we should *know* something about that variant. As matters stand today, the constructible hierarchy has at least as good a chance of being useful as the full hierarchy from which it is extracted: after all, we know literally more about the L_α than the V_α.

Formal independence results

We now return to a topic, already broached on p. 104, which has led to a sophisticated arsenal of 'subhierarchies'. We already provided general orientation on the topic, in particular (on p. 93), on the imagination needed to *discover* uses. As so often in such cases (cf. p. 66), as it were to protect the results in question, a body guard of exaggerations has developed. For example, the *connerie* that problems about a specific structure like $V_{\omega+\omega}$ are 'meaningless' when they do not happen to be decided by the sort of properties so far codified in axioms. (This *connerie* is involved in regarding the CH as 'settled' by its formal independence.)

In view of the last subsection, the body guard is now superfluous. Since Cohen [1963] a great number of subhierarchies have been introduced to establish the (formal) independence of most propositions mentioned in the last few pages, including Souslin's hypothesis and 6.6.[64] For the purposes of this chapter, there is no need to enter into details. But a general outline of this successful work is relevant in relation to both Gödel's heuristic views on p. 50, and to his own (early) results contained in his notes at Princeton.

For one thing he observed some simple *conditional independence results*, rediscovered (in one way or another) in the fifties; cf. Hajnal [1956], Lévy [1957], and Shoenfield [1959]. Specifically: suppose that $V = L$ (in note 55) is independent of the remaining axioms (*without* the axiom of choice) and some set A whose members are constructible, is not constructible. Then a new hierarchy $L_\alpha[A]$ defined by putting $L_0[A] = \{A\}$ (in place of: $L_0 = \emptyset$), satisfies (for suitable α) the usual axioms *including* the axiom of choice, but not $V = L$. Thus $V = L$ *is also independent of the axiom of choice*.[65] In view of the rediscoveries of such extensions it is fair to say that not the general principle, but the discovery of particular sets A needed for specific (absolute) independence results presented the principal difficulty.

Around March 1942 (in *Arbeitshefte XIV-XVI*) Gödel made extensive notes for proving the formal *independence of the axiom of choice* (for sets of pairs of

[64]Readers may wish to verify that the former is decided by the non-elementary axioms, while 6.6 is not (at least: not obviously; cf. the discussion of CH and GCH after 6.8.2).

[65]As in note 53, there are obvious possibilities of refinement.

integers), and hence of $V = L$. The general idea goes back to [1933b], on so-called modal logic and its topological models. With present experience it is not too difficult to complete the proof. But something essential (in Gödel's words, in conversation: a *method*) had been missing; cf. also his letter of 1st May 1968 on p. 46 of Chapter 5, where he corrected the description of Cohen [1963] as being a 'refinement' of his work. Gödel had just as much admiration for the later reformulation of Cohen [1963] in terms of so-called Boolean-valued models (which are more obviously related to his ideas in 1942). Again, not the 'broad principles' involved in that later work, but their appropriate use, constituted the progress. For example, Church [1953] explicitly considered Boolean-valued models for propositional logic and, as explained in some detail in Scott's introduction to Bell [1977], the notion of forcing (though not the catchy name) had so to speak forced itself on several people who toyed with set-theoretic models of intuitionistic logic in the late fifties.

Some logical and foundational lessons

The development of set theory followed a pattern which seems to be often successful at the *beginning* of research. After some experimentation in the general area of experience under investigation, one selects objects in the area which seem to lend themselves to theory; this presupposes of course that some non-trivial facts are known about those objects.

In the general area of sets (and definitions), the most successful selections were those of the *full cumulative hierarchy* (generated by the power set operation) and, at the other extreme, of the ordinals (generated by iterating the operation: $x \mapsto x \cup \{x\}$). In the latter case, some functions have to be added since not much can be expressed by (elementary formulas built up from the order relation) $\in$ alone; cf. p. 67 concerning the successor or, for that matter, the order relation on the finite ordinals.

The non-trivial facts known at the start are the familiar axioms. Later the classes of objects selected are restricted or enriched to realize structures with additional properties; in the case of sets, the full hierarchy is restricted, ordinals (or constructibles) are enriched (cf. $L_\alpha[A]$ above).

Gödel took a lively interest when, in the fifties and sixties, the area of experience adumbrated by Brouwer in his writings on choice sequences began to be studied according to the pattern above. A particular kind of such sequences, mentioned already on p. 91, turned out to have a simple theory: *lawless sequences*. The term is due to Gödel who objected to their original name: absolutely free, their principal property being that no restriction may be imposed on them beyond a finite number of values (a restriction on restrictions reminiscent of anti-trust laws which are intended to ensure a free market, but nevertheless are felt to be a shade short of absolute freedom). Compounds of lawless sequences, later called 'projections', have played much the same role as compounds of L and A to form

$L[A]$. Parts of the story are to be found in Troelstra [1977].

Gödel's interest is significant for a correct estimate of what has come to be known as his *platonism*. He never questioned the possibility of a *part* of mathematics which is intended to be about our own 'constructions' or choices. Thus, once objects of this sort (in particular, lawless sequences) had been described, the search for non- trivial facts about them was, for him, just as well-determined a project as his own search for axioms of set theory (and of course easier, since the pioneers in set theory had already discovered the more obvious interesting properties). But he did not regard that part as at all useful for mathematics itself, let alone as the whole of legitimate mathematics.

Gödel himself was less interested in the general pattern above than in the use of (his) more specific experience in set theory for other parts of logic. As early as 1936 (commenting in his note book on a report by Bernays of Gentzen's lecture to a philosophical congress in Paris) he felt that the actual details of his proof of reducibility on p. 106 should be useful for a *consistency proof of analysis*; and nearly 40 years later, in [1972], he repeated this impression, though less explicitly. Evidently, the idea was that the 'collapse' should not stop at countable substructures, but should somehow go on to (suitable families of) finite orderings. Even if successful, this idea would no doubt have to be supplemented by the difficult step from foundations to technology in Section 7, presumably, by the discovery of a significant problem in set theory itself which is solved by use of that idea.

In accordance with his heuristic views, Gödel took little notice of what seems to be the principal *foundational* lesson to be learnt from the work described in Section 9 (in particular, on formal independence results on p. 108): the contrast between research at an early and at an advanced stage of a subject, well illustrated by the difference in meaning of 'axiom'. For example, Martin's axiom (in Chapter B.6 of Barwise [1977]) and Jensen's $\Diamond$ (in Chapter B.5) were discovered by inspecting elaborate *proofs* (like many axioms of current mathematical practice), and are not meant to be seen by inspecting familiar *objects* (like V_α in the subsection on p. 98).

However, in the brilliant programmatic lecture [1946], Gödel derived a foundational lesson in the traditional sense of 'foundations', from his own work on *definability*.[66] The most obvious inadequacy of any formal system for analysing even approximately the possibilities of definitions follows from *diagonalization*, as recognized already by Poincaré. In [1946] Gödel pointed out how the use of ordinals in the constructible hierarchy prevents diagonalization, and thus provides a class of definitions with better closure properties. In conversation he mentioned

[66]The lecture contains a second lesson of this sort, on (higher) infinite cardinals, which belongs to Section 10.

Incidentally, though the topics of these lessons are so to speak at opposite poles of the early 'debate' on sets, the inadequacies of formal systems are central to both.

that for a while he thought of it as exhausting *all* definitions.[67] But he soon noticed that, for example, once one understands (not only L_α, but) the V_α, quantification over sets in V_α is also meaningful, and so he arrived at the notion of *ordinal definability* in [1946]. Without being dogmatic or even particularly specific about a more realistic candidate for an idea(lization) of the possibilities of humanly intelligible definitions, Gödel felt that L provided at least an idea for such an idealization.

He also mentioned, in passing, that (so to speak, at the lower end of the spectrum) the familiar class of computer programs (for recursive functions) escaped diagonalization too, but for a different reason: only the larger category of programmes for *partial* functions has a (partial) recursive enumeration. Similar ideas on definitions were pursued in the fifties and early sixties (but without reference to [1946], which appeared only later), arriving at *subclasses* of the class of recursive definitions because now definitions were required to be justified by appropriate proofs. This was achieved by restricting the ordinal logics in Turing [1939] by a so-called autonomy condition: before an ordinal was introduced, it had to be (formally) proved to be one. Here diagonalization was prevented, even though everything in sight was recursively enumerated, since only proper segments of the system are justified according to the scheme adopted. The claim was that, in this way, one had a (simultaneous) characterization of certain *informal notions of proof and definition*.

Not surprisingly, whatever its formal merits, the weaknesses of such a characterization are similar to those pointed out (in the subsection on p. 90) in Brouwer's attempts to make proofs into a principal object of study. In fact, with all the additional detail in front of one, the criticism can go further. It concerns *growth* (and here it does not matter if one means, literally, growth of neurological connections or simply of understanding). Evidently, the introduction of hierarchies is reminiscent of growth; but the specific laws of growth by the particular hierarchies have no visible counterpart in experience. The weakness of those characterizations is not a matter of principle, for example, a conflict with empiricist methodology; (successful) rational mechanics and (unsuccessful) hydrodynamics of ideal fluids are not one bit less *a priori* than those hierarchies. The difference is that so-called reasonable assumptions about our reason are just much wider off the mark than our ideas about 'rational' behaviour of the planets.[68]

Be that as it may, the silent majority of logicians has not taken up the part of [1946] on definability, and provability; certainly less than the other ideas in [1946] on axioms of infinity, which provide a beautiful illustration of Gödel's heuristic views on p. 50. Fortunately for the purpose of *testing* those views concretely, in

[67] In fact, 'L' stood for 'law'. Cf. also [1946]: 'constructible' means definable.

[68] There is a charming description in the Preface to Dedekind [1888] of 'rational' ideas about *reading*: spelling out words is reading in slow motion, with the logical corollary that, for literal certainty, one *ought* to slow down (and accept the literal text including printer's errors, rather than an obviously intended meaning.

the last 35 years also other candidates for new axioms of set theory have come up: they concern infinite two-person games and winning strategies for one of the players.[69] The work on those other axioms has been as different, both in style and content, from Gödel's [1946] as can be imagined. So, for contrast, it will be briefly described in the next section.

6.10 Gödel's Program: Axioms of Infinity

The axioms in question are intended to hold for suitable segments V_α of the fat cumulative hierarchy. Evidently, this aim makes sense only for those who know a basic minimum about that hierarchy. Gödel tried to convey this minimum knowledge in three publications in the forties, in terms varying according to (what he considered to be) his audience: for philosophers in [1944], sophisticated mathematicians in [1946], school masters in [1947]. Instead of speaking of a 'minimum knowledge', he spoke of the 'reality' of the V_α (as will be described in more detail in Section 11). In any case, here we accept the aim. But before one gets to specific problems, there are at least two further broad preliminary questions.

First of all,

What do we naïvely want to know about V_α?

Gödel concentrated on Cantor's continuum problem; that is, whether the CH is true or false (which, by the proof of 6.8.2, concerns only $V_{\omega+4}$). Believing it to be false, he regarded a refutation of the GCH as an easier first step. He took the formal undecidability by means of current axioms for granted, and so new axioms had to be discovered.

By p. 108, there is also the more delicate matter:

> Is it at all rewarding to study the V_α further? Or are there variants of V_α which are (perhaps) less easy to describe, but more manageable (by use of what we already know of the V_α): should one *reculer pour mieux sauter*?

Gödel did not encourage the interest of the naïve question to be questioned by the others. Instead he gave in [1946] a beautifully plausible account of likely ways to find new axioms; in other words, of continuing the process which has led to the currently used axioms.

Enriching the language of formal set theories

The most obvious loss in the passage (on p. 67) from non-elementary axiomatizations to formalizations is that not arbitrary predicates, but only those defined by

[69] Incidentally, Zermelo [1912] is one of the first papers on such 'determinate' games.

elementary formulas are used. So the most obvious step is to write down (new) axioms with the aid of some of those lost predicates.

This is easy, contrary to a wide-spread misunderstanding (generated by the idea that there is something 'universal' about the usual systems of set theory). Specifically, in terms of numberings n of formulas N (on p. 71): the predicate T of natural numbers (called *truth definition* by Tarski)

$$T(n) \text{ if and only if } N$$

is not definable explicitly (by diagonalization), but it has an obvious implicit definition (by recursion on the number of logical operators in N), also known in the literature as 'Tarski's adequacy conditions'. The definition involves, as an auxiliary, an enumeration $S(n, x)$ of all monadic predicates defined by formulas N with one free variable x, any finite sequence of sets being coded by one set. Now, given a formalization $\mathcal{F}$ of set theory, let $\mathcal{F}^+$ be obtained by adding the relation symbol S to the formalism and the implicit definition as a new axiom, the axiom schemata of $\mathcal{F}$ being extended to all formulas in the enlarged formalism. Then $\mathcal{F}^+$ is stronger than $\mathcal{F}$: for example, the consistency of $\mathcal{F}$ can be proved in $\mathcal{F}^+$. In short, one of the inadequacies of formal languages (on p. 110) is that *not all implicitly definable predicates are explicitly definable*.

Those with special interest in geometry would think of extending the language of set theory by symbols for geometric relations, and the axioms by propositions expressing geometric properties of those relations (with the proviso that all sets of real numbers or subsets of $V_{\omega+1}$ considered, represent geometrically meaningful figures).

Gödel had a different idea, going back to footnote 48^a of [1931], and his other fruitful contacts with higher types mentioned on p. 104.

Axioms of infinity

He pointed out in [1946] that, for current formalizations $\mathcal{F}$ of set theory, the extension $\mathcal{F}^+$ above can be replaced by *an axiom $I_\mathcal{F}$ in the usual language of set theory*, where $I_\mathcal{F}$ is seen to be valid by the same considerations as $\mathcal{F}$, and all theorems of $\mathcal{F}^+$ in the usual language can be derived from $I_\mathcal{F}$ in $\mathcal{F}$.

In fact, $I_\mathcal{F}$ is nothing else but the proposition used in footnote 48^a in [1931] some 15 years earlier, the existence of the least V_α for which $\mathcal{F}$ is obviously valid (cf. note 17). The argument is standard: though no enumeration of the monadic predicates definable in formal set theory can be explicitly defined, there is such an enumeration, say $S(n, x, y)$, for the predicates $N_y(x)$ defined as:

$$x \in y \text{ and the quantifiers in } N \text{ are restricted to range over } y.$$

Gödel concluded that, if such a modest use of higher types (actually, more than) replaced the most natural alternative (extension by enlarging the language of set theory), then a little more imagination would do miracles. Of course, nothing in mathematical practice gives even a hint of any more imaginative use, unless (following Gödel [1964]) one regards analytic number theory as an instance of passing to type $\omega + 1$ in order to solve problems about V_ω. But then there are plenty of open problems in mathematical practice: so why stick to its tradition?

To summarize his programme as it were, Gödel proposed to *solve every problem by use of a suitable axioms of infinity*. Naturally, he was not specific about the term, but the idea was clear enough: a new axiom of infinity is to be satisfied by some V_α, but only for α much bigger than β if V_β satisfies the already established axioms (the bigger the α the better). And, certainly, the evidence for Gödel's program was not worse than the evidence (mentioned on p. 68) which Hilbert had for his programme (of *Methodenreinheit*).

There can be no question of summarizing here the massive work done on Gödel's program over the last 45 years (cf. for example Kanamori and Magidor [1978]). But two directions of such work, firmly established by the end of the fifties, are worth noting specially.

First, a new[70] *style* of axiom was discovered, known to logicians under the name of (Lévy's) *reflection principle*, and to mathematicians as (Grothendieck's) *axiom of universes*. The general idea is that all properties of V_α, stated in some given (elementary or non-elementary) language, should also be satisfied by some *element* x of V_α (either simultaneously, or by an x depending on the property considered). The idea corresponds clearly to the (intended) unending character of the hierarchy V_α. Already the simplest case of a non-elementary language (so-called Π^1_1-*reflection*) ensures that V_α is closed under all earlier schemes (for example, in Mahlo [1912]) of building up the hierarchy 'from below' (cf. Bernays [1961]). Evidently, except for properties stated in the most primitive language, reflection principles are *not* satisfied by V_ω.

The second line of work goes in the opposite direction, as it were: some simple set-theoretical property $P(\alpha)$ about α and its power set $\mathcal{P}(\alpha)$, which holds for $\alpha = \omega$ (and, usually trivially, for $\alpha = 2$) is asserted or, at least studied, for certain $\alpha > \omega$.

- One typical example, going back to work in Poland in the thirties, is the existence of a two-valued measure on $\mathcal{P}(\alpha)$ which is additive for subsets of $\mathcal{P}(\alpha)$ of cardinal $< \alpha$ (α is then called a *measurable cardinal*).

- Another typical example is derived from the partition theorem in Ramsey [1928].

By the early sixties, any $\alpha > \omega$ which has the properties P considered was known to be larger than all familiar cardinals (for example, for any such α there are

[70]'New', even though the early instances are formally derivable in current set theory $\mathcal{F}$ or from $I_{\mathcal{F}}$.

α strongly inaccessible cardinals $< \alpha$). Far from being disturbing (for Gödel's programme), this knowledge is a prerequisite if $\exists \alpha P(\alpha)$ is to deserve the name *axiom of infinity* at all! After all, one wants here cardinals α which differ from ω 'as much as' ω differs from, say, 2. However, with remarkably few exceptions, the particular properties P that people have stumbled on are very poorly understood; mostly, one does not know if they are satisfied by any $\alpha > \omega$ at all, nor even whether superficially similar P are not satisfied. (A notable exception is the property called *weak compactness*, which follows from Π^1_1-reflection.)[71]

The first genuine implications of axioms of infinity for questions outside cardinal arithmetic were discovered in the early sixties: *if $\alpha > \omega$ and α is measurable then some subset of $V_{\alpha+1}$ (and even of V_ω) is not constructible* (as shown by Scott and, respectively, Rowbottom); in fact, the constructible subsets of V_ω are then countable. Whatever the defects of the particular property of measurability may be, one sees from the proofs how an arithmetic question might be settled by 'looking down on ω from above'.

For people more familiar with another sense of 'L' (for Lebesgue measure rather than for constructible sets), an implication discovered later by Solovay is more instructive: *if there is a measurable cardinal $\alpha > \omega$, then every PCA set of real numbers is L-measurable*. The required coverings (by open sets) of such a set and of its complement are defined by use of the assumed measure on α.

However, Gödel's particular candidate CH is left (demonstrably) undecided by any (consistent) axiom of infinity so far proposed. Indeed, it is fair to say that the only memorable result[72] on *cardinal exponentiation* discovered in the last eighty years (by Silver) can be proved by methods not too different from those current at the turn of the century: *if α is of cofinality $> \omega$ and $< \alpha$ and, for all $\beta < \alpha$, $2^{\omega_\beta} = \omega_{\beta+1}$, then $2^{\omega_\alpha} = \omega_{\alpha+1}$* (a nice proof is on pp. 388–389 of Barwise [1977]).

Before taking stock of work done on Gödel's program (and, particularly, of his heuristic views), the quite different direction of research mentioned on p. 112, has to be summarized.

Axioms of determinacy

In the fifties, when the theory of games was popular, certain so-called infinite games attracted special attention on Poland, where infinitistic generalizations had been popular for a quarter of a century.

Suppose G is a set of sequences of natural numbers. Two players choose

[71] Incidentally, though Gödel [1946] was published only in the sixties, work on those new axioms (especially of the second kind) could have profited from Gödel's presentation in 1946, since Tarski and Erdös were principal contributors: the former was present at the lecture, the latter was in contact with Gödel.

[72] Before the quite recent results of Shelah on the same subject, which similarly do not require any axiom of infinity.

alternately natural numbers x_{2n+1} and x_{2n+2} for $n = 0, 1, \ldots$ If $\vec{x}$ stands for the sequence $x_1, x_2, \ldots$, the first player has *won* if $\vec{x} \in G$. A *winning strategy* for that player is, by definition, a function f_1 (with finite sequences as arguments and with numerical values) such that, for all choices x_{2n} (where $n \geq 1$) the sequence

$$f_1(\langle\ \rangle), x_2, \ldots, f_1(\langle x_2, x_4, \ldots, x_{2n}\rangle), x_{2n+2}, \ldots\ \in G,$$

where $\langle\ \rangle$ is the empty sequence. Similarly, a winning strategy for the other player is a function f_2 such that, for all x_{2n+1} (where $n \geq 1$),

$$x_1, f_2(\langle x_1\rangle), \ldots, x_{2n+1}, f_2(\langle x_1, \ldots, x_{2n+1}\rangle), x_{2n+3}, \ldots\ \notin G.$$

Another way of writing these conditions is

$$\exists x_1 \forall x_2 \cdots \exists x_{2n+1} \forall x_{2n+2} \cdots (\vec{x} \in G) \tag{6.7}$$

and, respectively,

$$\forall x_1 \exists x_2 \cdots \forall x_{2n+1} \exists x_{2n+2} \cdots (\vec{x} \notin G). \tag{6.8}$$

A kind of dual to such 'games' without an end are 'games' without a beginning, where the winning conditions for the two players are

$$\cdots \forall x_{2n+2} \exists x_{2n+1} \cdots \forall x_2 \exists x_1 (\vec{x} \in G) \tag{6.9}$$

and, respectively,

$$\cdots \exists x_{2n+2} \forall x_{2n+1} \cdots \exists x_2 \forall x_1 (\vec{x} \notin G) \tag{6.10}$$

(and the distinguished player is now the one with the last move).

Steinhaus, later in collaboration with Mycielski, experimented with the unpromising proposition, called *axiom of determinacy*: $6.7 \vee 6.8$ (formulated for all sets G of sequences: the x_n are arbitrary sets, not only natural numbers). The proposition is unpromising because the general idea behind it is nothing else but an extension of the well-known law for *negating finite sequences of quantifiers*. But a glance at its proof shows that it uses, in an obviously essential way, the fact that finite sequences have a beginning and an end (in particular, this is needed for the two basic properties of negation: the laws of contradiction and of the excluded middle). In fact, $\neg(6.7 \wedge 6.8)$ holds, but not necessarily $6.7 \vee 6.8$.[73]

For about two decades many articles (of uneven quality) were published on determinacy, and certainly none that is remotely comparable in distinction to Gödel's [1946]. But finally, Martin [1975] proved that *all Borel games are determinate*; that is, $6.7 \vee 6.8$ holds if G is a Borel set. This is not only of interest to the subject of infinite games, but easily *the most convincing contribution to*

[73]Incidentally, by Galvin and Prikry [1976], neither $6.9 \vee 6.10$ nor $\neg(6.9 \wedge 6.10)$ need hold.

Gödel's program so far. More specifically, the proof proceeds by transfinite recursion on the countable ordinals α ($< \omega_1$, the first uncountable ordinal), where α is the number of applications of Borel operations (complementation, projection, countable unions) used to generate the set G_α. In a very transparent way, the determinacy of $G_{\alpha+1}$, a set of sequences of elements of X, is derived from the corresponding result for a suitable set G_α of sequences of elements of $\mathcal{P}(X)$. In set-theoretic terminology, the proof uses V_{ω_1} to establish the determinacy of all Borel sets of sequences of natural numbers (a proposition about the quite familiar, low level $V_{\omega+2}$ of the cumulative hierarchy).

Even if not for Gödel, for mathematical practice the assumption of V_{ω_1} is an 'axiom of infinity'. Of course, its interest is established up to the hilt by the *particular* proof of Martin [1975],[74] since the transfinite iteration of the power set operation is seen to be useful by inspection. But more is true, by work going back to Friedman: at least for the usual formulations of set theory, the determinacy of sets G_α cannot be formally derived at all without use of V_β, where β is of the same order of magnitude as α. As on p. 79, this negative result can be given a positive twist: for any $\alpha < \omega_1$, *if a proposition of suitably simple* (syntactic) *structure is derived from assuming V_α, it can also be derived from Borel determinacy*.

A similar positive twist can be given to the results in the literature on the consistency of stronger axioms of infinity than V_{ω_1} relative to the assumption that larger classes of sets than Borel sets are determinate. In short, *determinacy is an alternative to Gödel's program*.

Superficially, the relation between axioms of determinacy and the use of V_α for large α is quite similar to that between the axiom of choice and the use of transfinite recursion in measure theory at the turn of the century. As long as only consequences of a simple syntactic structure are formulated, the details of the definitions in the proofs by transfinite recursion are lost in the statement of the theorems (just as the details of the winning strategy defined in Martin [1975] are lost if only simple consequences of Borel determinacy are considered). It remains to be seen whether somebody *discovers* problems, also in measure theory, for which those details are relevant.

[74] In addition, the (original) proof has the virtue of making a very convincing use of (a simple instance of) the very attractive so-called *priority argument*, which was discovered in recursion theory nearly twenty years earlier, but applied only to the somewhat teratological subject of degrees of undecidability of recursively enumerable sets.

Maritn has subsequently published (in [1985]) a version of the same proof that does not use the priority method.

6.11 Gödel's foundational views: balancing the account

As already mentioned on several occasions, in his publications Gödel used traditional terminology; for example, about *conflicting views* of 'realist' or 'idealist' philosophies.

In conversation, at least with me, he was ready to treat them more like different *branches* of the subject, the former concentrating on the things considered, the latter on the processes of acquiring knowledge about these objects or about the processes. Naturally, for a given question, a 'conflict' remains:

> Which branch studies the aspects relevant to solving that question?

with obvious parallels in mathematics and the natural sciences.

Gödel rejected only so-called positivist philosophy which (at least for logic) is distinctly negative, since it accepts (as arbitrary 'conventions', or as 'facts of our natural history') phenomena which the other branches see as problematic, or at least as capable of a rewarding analysis.

Successes: mixing the realist and idealist traditions

In mathematics, the idealist tradition is involved (in one form or another) in *constructivist* foundations which stress the use of definitions and proofs in the process of acquiring mathematical knowledge (cf. Section 7 and the subsection on p. 109). In particular, Poincaré stressed definitions, Hilbert and Brouwer stressed proofs.[75] As already mentioned, Gödel solved problems which either had been formulated explicitly by these three famous constructivists or, at least, are patently relevant to their foundational schemes.

Gödel himself stressed (most clearly in his letters reprinted in Wang [1974]) that his results are best understood in terms of notions from the realist tradition which were rejected or simply ignored in the constructivist schemes, such as: logical validity, arithmetic truth, various fat or thin hierarchies. Gödel's analysis was adopted in this chapter.

But another reason for his success was, obviously, his familiarity with the subjects derived form the constuctivist programs: formal systems, intuitionistic logic, ramified hierarchies.

By (yet another!) fortunate coincidence, the relative importance of the two elements in Gödel's successes can be illustrated by the case of Zermelo (cf. note 7), who had an equally staunch realist *Weltanschauung*: so much so that he simply refused to look at the tainted subjects! Thus, the stated reason for his outburst in Zermelo [1932] (against Gödel's incompleteness results) was that Gödel considered

[75] Incidentally, contrary to an almost universal misunderstanding, Hilbert's finitist proofs are much more restricted than intuitionistic ones.

formal systems at all, establishing their inadequacy instead of dismissing them as obviously inadequate. (An unstated reason could have been that detailed work on any subject is liable to create a vested interest in it, and a reluctance to look at alternatives.)

At that time the still little-known Zermelo [1935] was in preparation, sketching what we should now call *infinitary systems*, with infinitely long formulas and infinite proof figures, intended to represent the meaning of propositions and the structure of mathematical thought adequately: in short, an alternative to formal systems. Whatever his conscious motives may have been, Zermelo's instincts to protect his alternative were more than justified: he did not get beyond his intentions! What he actually said about those infinitistic representations was not only, trivially, formulated in finite terms, but (and this is the critical defect) already fully expressed in *current* systems of set theory, as implied by Gödel's analysis on p. 113 (though Zermelo [1935] is not mentioned in Gödel [1946] at all).[76]

Gödel's program is nothing else but the first genuine proposal for *implementing* those realist intentions, by deriving from them new axioms; to be compared to deriving mathematical laws from a physical conception (or physical 'picture'), Maxwell's derivation of his equations from Faraday's picture being the standard example.

Gödel's program involves quite different problems from those he had solved earlier: for one thing, there was no idealist bias to be corrected by injecting suitable realist elements. He was treading new ground (though surely not 'rushing in', unlike the gamblers on infinite games on p. 115).

Neglected problems: beyond naïve idealism

Evidently (as already illustrated by the opening of Section 10, in the particular case of sets) recognizing some phenomena as a (legitimate) subject of research is necessary, but by no means sufficient for progress with understanding them. To put it paradoxically: once generalized doubts about them have been removed and some simple useful properties have been noted (here: doubts about infinite sets, and axioms codifying some obvious properties of the V_α), *the principal problem is selection*; selection of *objects* (among those recognized) which lend themselves to theory by something like available means, and selection of *properties* which have implications for such a theory. Such selection involves, besides the phenomena, just those processes which are the business of the idealist branch of philosophy (its sophisticated part, as it were). When reasons for new axioms (that is, matters of evidence) are at issue (as in Gödel's program), questions belonging to

[76] It is also not very well known that so-called fully analysed infinite proof figures were considered in intuitionistic mathematics in the twenties; cf. Brouwer [1927]. Again, analysis in the sixties of the properties actually stated about those figures (cf. p. 109) shows that they do not go very far either.

sophisticated idealism must be expected to become important, or even dominant.

Here it is to be emphasized that the bulk of the constructivist literature at best ignores sophisticated questions of selection, but more generally dismisses them (as a matter of some vague kind of 'convenience', or of sacrosanct 'personal taste'). Instead, that literature assumes miracles from the slogan about mathematics being 'our own construction', as opposed to some 'external reality'. This assumption leads to what might be called *naïve idealism*, which is no less widespread than naïve realism (though the name is not). It is naïve on at least two counts, besides those already listed in the remarks on proofs and definitions in Section 7 and in the subsection on p. 109:

- First, it forgets the problems arising in those parts of mathematics which are simply intended to be about our own constructions (for example, about computation rules); parts in which realist questions of a correct representation (definition) of a previously understood notion, or of the mathematical structure of some external phenomenon, do not arise at all.

  Nevertheless there remains the problem of recognizing whether a construction does or does not have some property: it is no simpler to decide if a diophantine equation has a solution when this problem is interpreted purely computationally than when one thinks of the natural numbers as properties of (extensions of) concepts.

- Secondly (and, if anything, this assumption is even more naïve), the constructivist literature regards as particularly fundamental those parts of (mental) experience of which we are most acutely aware; for example, in the case of definitions and proofs, principal attention is given to the slow early stages in the learning process (not to our predispositions, nor to our reasoning after that elementary knowledge has become part of our intellectual reflexes).

  Parallels in naïve natural philosophy are obvious; for example, whenever the visible part is simply assumed to be decisive (not only, trivially, for our view of the world in the literal sense of the term, but also for scientific understanding).

This naïve part of the idealist tradition, which has of course completely overshadowed its sophisticated part, is viewed with great scepticism by the silent majority (whose objections, as expected, are not very articulate).

Naturally, Gödel too had strong reservations about naïve idealism, though he would not apply the term 'naïve' to any part of traditional philosophy. But, at least in mathematics, he never seems to have faced squarely the problems raised in sophisticated idealism. This omission is only too obvious from his *obiter dicta* on evaluating the evidence for new axioms.[77] If I had not known him personally, I should have dismissed those *dicta* as another 'body guard', to protect his program until the time was ripe for progress. As it is, the level of his discussions in [1944] and [1947] troubles me. It is not much above that of the 'debates' on the paradoxes mentioned in the subsection on p. 94; more troubling still: it is utterly different from what I remember of our conversations (up to his illness at the end of the sixties), more than 20 years after [1944] and [1947] were written. Perhaps others, less involved than I am, will one day read the masses of his notes in Princeton, and fit that troublesome material into a more interesting picture of Gödel.

Bibliographical remarks

Both in [1944] and [1947] Gödel tries to use the *parallel between mathematical and physical objects* to support his program: without reference to specific examples, but simply as 'realities independent of ourselves'. This is doubly suspect. Trivially (again by the opening of Section 10), this makes his program only a candidate for research, without the slightest hint of its chances of success: some physical phenomena are far from having a satisfactory theory. Less trivially, there is no emphasis on the fact that mathematical notions enter into the description of the simplest physical phenomena on a par with other notions, not to speak of physical laws: for example, objects have chromatic and arithmetic properties (a yellow table with four legs). So one is left without an issue at all (quite apart from the fact that the methods needed for studying very different kinds of physical objects differ markedly among themselves).

The proposal in [1947] (already mentioned on p. 104) to judge new axioms by *deciding (demonstrably) formally undecided propositions* conflicts with the previous point, vague as the latter may be. For, in judging new scientific hypotheses, essential use is made of consequences which are tested *independently* (for example, observationally). A more convincing parallel involves (the use of new axioms for) *new proofs of old theorems* (as on p. 521 of [1947], implicitly taken back on p. 271 of [1964] where such uses, so-called weak extensions, are described as sterile).[78]

Gödel developed a remarkable obsession with *mere cardinality*. Thus [1947]

[77] A foretaste was given on p. 104, and the full flavour will be conveyed by the samples cited in the next subsection.

[78] *For the record.* Today, some 40 years later, the general level of derivative literature on assessing new axioms is even more embarrassing. For example, in two of the otherwise most brilliant expositions in Barwise [1977].

First, on p. 344, a cardinality principle is announced: 'Thus we see, the more problems a new axiom settles, the less reason we have for believing that the axiom is true'.

Secondly, on pp. 813–814, in connection with instances of determinacy which are known *not* to hold for L_α, their validity for V_α is regarded as plausible because Borel determinacy holds (but also for L_α if $\alpha > \omega_1$!), and because the consequences for descriptive set theory are coherent and pleasant (so to speak: fat is beautiful).

suggests in effect that the most fundamental problem about the continuum is to decide whether the continuum hypothesis CH is true or false (for V_α where $\alpha \geq \omega + 4$), as if one did not want to know the geometry of the continuum just as much. Obviously, Gödel wanted to forestall the inevitable *conneries* which the expected proof of the formal independence of the CH was to produce. He was doubly unsuccessful:

- First, even some 30 years later, Martin [1976] questions whether the CH has a definite truth value for the intended meaning at all because, despite many attempts (by looking at many variants of ordinary set theory), the CH has not yet been decided: as if there were not infinitely many false starts, perhaps due to a systematic oversight, for any problem.

- But also, and this is not at all a matter of mere *conneries*, Gödel's exaggeration gives no hint of the kind of implications which make those independence results significant (apart from the technical uses on p. 104): obviously not by casting doubts on the precision of the continuum problem, but on its 'fundamental' character, its interest.

  Specifically, because of those results, the CH may be true even if some perfectly straightforward subset X of V_{ω_1} can be mapped onto V_ω or V_{ω_1} only by a very odd map, and the CH may be false simply because some quite odd X cannot be mapped onto V_ω or V_{ω_1} (at all).

  In terms of the subsection on p. 80 (explaining the unexpected usefulness of the general notion of logical validity), the CH *lacks stability* (with respect to perturbation of the domain of sets X involved). Without such stability, the problems of sophisticated idealism become decisive: which sets X and which maps do we want to know about?

  Examples were considered by Cantor and Brouwer more than eighty years ago: the former showed that the CH does hold for *closed* sets X; the latter considered (in [1908a]) geometrically meaningful, *topological* maps, when the CH is false even for quite simple X.[79]

Still in connection with *geometrical properties*, Gödel notes in [1947] that not all sets of points are geometrically significant, but calls certain consequences of the CH 'paradoxical': at most a conflict with geometric impressions is involved, but not more than in the case of several well-known consequences of the axiom of choice (and, by p. 104, he had recognized the validity of that axiom for the V_α).[80]

---

[79] For spaces of choice sequences with the usual topology, all maps are automatically continuous: then the CH is obviously false.

[80] Actually, even without the axiom of choice one gets geometrically meaningless results when $\forall x \exists! y A(x, y)$ holds but the unique function f which satisfies: $\forall x A[x, f(x)]$ is highly discontinuous (for example, characteristic functions of sets defined by the comprehension principle applied to logically complicated predicates).

The points above troubled me, as already mentioned on p. 59, before I met Gödel personally. The following gem occurred to him later, concerning weakly and strongly inaccessible cardinals:[81]

> for *finite* cardinals, the two properties are not equivalent: 1 (in contrast to: 2) is only weakly, not strongly inaccessible (since $0^0 = 1$).

From this Gödel concluded that the GCH was implausible, since it implies that all weakly inaccessible *infinite* cardinals are also strongly inaccessible. Before publication in [1964], Gödel told me his discovery. He added, with the expectant look he always had when he thought he was saying something particularly naughty, that I surely regarded it all as no more than a play on words (*Wortspiel*). I still remember my pleasure (and his) when a totally ambiguous comment occurred to me. No similar banter is to be found in [1964] itself.

6.12 Philosophy: Speculations and Reflection

The three published samples of Gödel's speculations on spectacular topic are about time travel, minds and machines, and the origin of life on earth.

- The first was developed by him in considerable detail in [1949], [1949a] and [1950].
- The second was a principal topic of our conversations in the sixties.
- The third happens to bring out particularly clearly, by contrast with Crick and Orgel [1973], the single most distinctive point in Gödel's heuristic views: his preference for using general qualitative (rather than specific 'empirical') data, in accordance with the ideals of traditional philosophy.

On present evidence Gödel's contributions to the topics above are not conclusive, and certainly not comparable to his successes in the previous sections. It is beyond the scope of this chapter to go into the many differences involved, in the nature of the topics, the stages of development, the attention Gödel gave to them, and so forth. Nevertheless, quite apart from a pleasing freshness and wit, the special twists he gave to his speculations provide striking illustrations, in the quite different area of natural science, for the lessons learnt from reflecting on his incompleteness theorems in the foundations of mathematics (cf. the subsections on pp. 74 and 78).

[81] The latter are defined in note 45; the former need not be closed under exponentiation.

General Theory of Relativity

Gödel's early interest in the subject and his close contacts with Einstein were described in Section 1.

Gödel's writings on the General Theory of Relativity[82] were not extensive (consisting of the three quite short articles [1949], [1949a] and [1950]), but they were highly original and, in the long run, quite influential. In these articles he described a family of cosmological solutions of Einsteins' equations that possessed a number of novel features.

Most striking among these was the presence of *closed timelike curves* in his original non-expanding model. Thus, in this model, it would be possible in principle for an observer to travel into his own past. While for the majority of physicists, this feature might be regarded as a sufficient criterion to rule the model out as 'physically unrealistic', Gödel appears to have taken a contrary view. Indeed, in [1949a] he computed, in a footnote, the amount of fuel required for the execution of such a journey and, finding this to be absurdly large, concluded that his model could not be ruled out as contradicting experience. (He did not, however, consider the vastly 'cheaper' but equally paradoxical possibility of an observer merely sending a signal into his own past.) In the modern theory of global general relativity (for example, in Hawking and Ellis [1974]) it has been found necessary to examine the various types of 'pathology' that can exist in space-time models even when these features might be regarded as sufficient to rule out the models as 'physically unreasonable'. Thus, this original Gödel model has provided an interesting and significant example of a space-time precisely because of this 'unphysical' feature. Indeed, the Gödel model was the first simply-connected such example, the closed timelike curves being therefore 'essential' in the sense that they cannot be removed by passing to a covering space. The model is interesting also for a more philosophical reason. It shows that a concept of time that is globally quite different from that seemingly implied by our normal experiences cannot be ruled out merely on the basis of the known local physical laws, once some of the ideas of general relativity are taken into account.

A second feature possessed by Gödel's models is that *the matter in them rotates* relative to the local inertial frames. Thus, the models show that at least one form of 'Mach's principle' is not a consequence of general relativity (contrary to what Einstein had originally hoped). Gödel also proposed expanding rotating models (without closed timelike curves) which could be serious candidates for the actual large-scale structure of the universe. Gödel's demonstration of the existence of apparently realistic models in which there is a relative rotation between inertial frames and distant matter led to the speculation among cosmologists that such a feature might also be detectable in the actual universe. However, very low observational limits can now be placed on this hypothetical rotation (apparently

[82] The following account of Gödel's publications in the area is due to R. Penrose.

$< 10^{-16}s^{-1}$).

A third feature of the models is that they possess *spatial homogeneity but not isotropy*. Gödel appears to have been the first to study such models and to introduce the appropriate non-holonomic frame techniques (largely unfamiliar to relativists of the time) for their detailed analysis. However, much of his work in this area remained unpublished and had to be rediscovered by others. The study of spatially homogeneous models has become an important part of theoretical cosmology in more recent years (cf. Heckmann and Schucking [1962]). Gödel was concerned only with space-times filled with incoherent matter. A corresponding analysis of empty space-times was made by Taub [1951] shortly afterwards.[83]

Thus the direct physical interest of Gödel's papers is limited, in accordance with Einstein's comment on [1949a] in the same volume. (Gödel's papers appeared in the middle of a long period during which the general theory made little progress.) But Gödel's work served as a cross check on mathematical conjectures and proofs in the modern global theory of relativity. This is the first of the striking parallels to his incompleteness theorems promised at the beginning of this section; in particular, to the use (on p. 78) of his second incompleteness theorem as a cross check on proposed consistency proofs (though, as mentioned there, the direct foundational interest of that theorem is quite limited).

Bibliographical remark. In a long typed essay at Princeton, Gödel expanded [1949a] in the style of academic philosophy, using [1949] to interpret Kant's ideas on time. Perhaps closer study will show what more is gained from this pedantic attention to Kant's elaborations than from the simple idea of ghosts which, by p. 55, had long been in Gödel's thoughts (while, as he says in the essay, he never had much sympathy for Kant's general philosophy). In any case, though the typescript dates from the fifties, Gödel did not publish it; but he put it among the items to be published after his death, on lists he made in the seventies (especially on days when he though he was going to die).

Non-mechanical laws of nature

Throughout his life Gödel looked for good reasons which would justify the most spectacular conclusion that has been drawn from his first incompleteness theorem: *minds are not (Turing) machines*. In other words, going back to p. 63, the laws of though are not mechanical[84] (that is, cannot be programmed even on an idealized computer).

The popular reasons are quite inconclusive. Certainly, by (Matyasevic's improvement of) the incompleteness theorem, those minds which can settle all diophantine problems are not machines; but we have not found any evidence of such

[83] Here ends Penrose's account.

[84] 'Mechanical' should not be confused with 'mechanistic', in the sense of deterministic; the usual probabilistic laws are mechanically computable.

minds. Nor there is the slightest hint of any computer programs which stimulate (even in outline) actual proof search; not even for solving problems which do have a mechanical decision procedure (for example, propositional algebra).[85]

In conversation Gödel brought up one of his favourite twists:

> Either mind is not mechanical, or mathematics (in fact, arithmetic) is not our own construction.

The tacit assumption here, one of those reasonable assumptions about our reason (cf. nore 68), is that we can decide all properties of our own constructions. Gödel remained unsympathetic to the admittedly tasteless comparison with our physiological productions, which can have painfully unexpected properties.

His reaction was quite different to another objection I made in the early sixties, expressed by the question: *Is mechanics mechanical?* More formally:

> Are the laws of current physics mechanical in the sense that, according to current theory, every analogue computer can be simulated by a digital computer (with the same probability of error)?

It is certainly not evident that celestial mechanics is mechanical; in particular, that collision properties of n-body configurations are mechanically decidable, even in finite time intervals.[86] Other candidates for non-mechanical laws came from statistical mechanics of co-operative phenomena (such as boiling).

The question above expresses an objection; for if some laws of ordinary matter are non-mechanical, then the notion of machine is not adequate 'in principle' to separate mind and matter. Gödel was at first tempted to dismiss the question, by the familiar *petitio principii* of supposing that only mechanical laws are precise (for a non-mechanical mind?), but he stopped himself in the middle of the sentence (I believe, the only time in all our conversations). Afterwards, he took an active interest in the search for non-mechanical laws, both in physics and in the part of logic which studies specifically mental constructions.[87] (In the latter the *petitio principii* to be avoided is the requirement that those constructions must be represented by a mechanical procedure.)

The question above is not yet settled. Here the parallel (promised at the beginning of this section) to the lessons learnt (on pp. 74 and 78) from the incompleteness theorem concerns the evaluation of the 'empirical evidence' provided by existing solutions in mechanics.

- Certainly, the bulk are mechanical; just as the bulk of ordinary mathematics is easily formalized in *Principia* (despite its incompleteness).

[85] Cf. the use of impure methods stressed at the end of Sections 2 and 3.

[86] A little care is needed in the formulation: the data should not be points, but neighbourhoods in phase space.

[87] According to Wang [1974], in the seventies Gödel seems to have gone back to his original twist; but his arguments for supposing that physical laws must be mechanical have a, for him, strangely positivistic flavour.

- If there are mechanically undecidable problems in some parts of mechanics, they may have been discounted by now (replaced by more tractable questions); just as number-theoretic practice concentrated on more rewarding problems about diophantine equations, long before the negative solution of Hilbert's tenth problem so to speak ratified the practice.

- Corresponding to the positive aspects emphasized in Davis, Matijasevic and Robinson [1976], in mechanics one would hope to have a new kind of analogue computer.

- Last but not least, the mere existence of some non-mechanical laws of nature, just as the mere existence of some formally undecided problems in mathematics, does not settle their significance in the sense of their *frequency* in different branches of science. Incompleteness phenomena are, on present evidence, much more significant for set theory than for arithmetic (tacitly, for the questions that strike us as interesting). it certainly cannot be excluded that, similarly, the phenomena of consciousness (that strike us) follow non-mechanical laws as a rule (in contrast to the phenomena of ordinary physics, at least those on which physical theory concentrates).

Bibliographical remarks. Komar [1964], which was overlooked in our conversations, points out that non-mechanical laws arise in those parts of physics where theoretically admissible states σ are represented by so-called primitive recursive sequences s of natural numbers (given by a description of the experimental set up), and some observable relation R between σ_1 and σ_2 corresponds to: s_1 and s_2 differ infinitely often. But Komar [1964] is inconclusive, since the theories considered are not shown to permit arbitrary primitive recursive s (or enough for R to be non-mechanical).

More recently, on p. 59 of Browder [1976], Arnold mentioned other candidates for non-mechanical laws in statistical mechanics involving vector fields given by polynomials with rational coefficients; but, like the n-body problems on p. 126, Arnold's seem to need neighbourhoods instead of discrete coefficients.[88]

Pour El and Richards [1981] have found computable initial data (in dimensions > 1) for which the wave equation has a unique, but non-computable solution. It remains to be seen if a physical system, perhaps by use of lasers, can realize these solutions with the kind of probability of error expected in the execution of a computer program (by a digital computer).

[88]Mechanical undecidability is usually easier to establish in the latter case. Occasionally there are neat theoretical reasons, even in classical mechanics, for discrete data; for example, Newton's for the exponent -2 in his law of gravitation.

Chemical evolution of living organisms on earth

Though Gödel's published comment on this topic (on p. 326 of Wang [1974]) is very brief, it is worth mentioning since it fits in with his general views expressed in many conversations. His particular conjecture was that the probability of a living organism developing in geological time as a result of random chemical operations was vanishingly small. The initial distribution of matter is assumed to be random, and nothing is said about the significant features of either the chemical reactions or of living organisms which would be used in the calculation. Evidently, he hoped that only 'basic' knowledge, for example, schoolboy chemistry would be needed, since (as he often said) the use of specific detail could not be convincing in Big Questions.

As it happens, Crick and Orgel [1973] present a perfect example of so to speak the opposite heuristic view. Briefly, they use two 'very' specific details about molybdenum: it is rare in our part of the universe, and it occurs in living organisms. So it seems a foregone conclusion that, with these two additional hypotheses, Gödel's conjecture holds, and should be easy to prove formally. But also, while Gödel's conjecture (as formulated) gives no hint at all of any positive theory about the origin of life on earth, Crick and Orgel [1973] inevitably looked for a source in regions of the universe where molybdenum is more plentiful and where chemical evolution could have succeeded, free from terrestrial constraints. (After that they followed *Genesis*: like Jehovah those extra-terrestrial beings set about populating the Earth, in their fashion.)

Their speculations have not settled the origin of life on earth. But their use of 'specific detail' about molybdenum provides a neat parallel to one of the lessons on incompleteness on p. 74. The aim of Gödel [1931] and its improvements in the thirties was independence of subject matter, for as broad a class of formal systems as possible. The so far most successful applications of incompleteness involve 'specific' properties, such as the size of ordinals in set theory or the rate of growth of number-theoretic functions, and above all the informal notion of arithmetic truth (at least, for diophantine problems).[89]

General interests (and a contrast)

Judged by the amount of space in Gödel's note books dealing with *general philosophy* and *theology* (including demonology), these subjects occupied a great deal of his attention ever since his student days. They were rarely touched in our conversations, since there was not enough common interest. However, during the 15 years or so when I saw a great deal of him, he would occasionally quote pas-

[89] *Remark on another spectacular topic.* The literature on hidden variables in the quantum theory contains several impossibility proofs which are also *incomparable*, without stressing this fact. By and large philosophers and logicians try to avoid specific details of the theory, and prefer to use (familiar) properties studied in logic or probability theory.

sages from his preferred reading at the time: Kant, then the slow-paced Husserl, then so to speak the opposite extremes, Fichte and Schelling. The quotations were not at all well-known and, at least for me, very perceptive. Given Gödel's methodical habits mentioned on p. 51, he may well have kept a record of these and similar passages. The publication of such an anthology is likely to produce a minor revolution in philosophy: if we came to associate Hegel or Husserl with a dozen crisp and memorable ideas, we could cherish them as much as (what we know of) Heraclitus.

Gödel's conversations on the general topics above, at least until his illness, had the light touch and exquisite discretion noted already elsewhere in this chapter, in contrast to the impression left by some of his more popular writings (p. 59). In this respect there is a striking parallel to the difference between the letters of Archimedes and his public image which has him look for a fulcrum in outer space to move the earth (as might be expected from some kind of misfit, ill at ease on this planet).

Incidentally, if Gödel's work is to be compared to that of one of the ancients, Archimedes is a better choice than Aristotle (who invented logic, but proved little about it). Archimedes did not invent mechanics, as Gödel did not invent logic. But both of them changed their subjects profoundly, by work with an almost unsurpassable ratio of interest of the results to effort (as seen in Sections 4 and 6 above, or in the laws of the lever).

6.13 Foundations and the Common Understanding

As promised at the outset, this chapter has described Gödel's contribution to our present understanding of formal and (non-elementary) axiomatic notions; in particular: logical validity in Section 6, arithmetic truth (or equivalently, in fancy language: consequence from Peano's non-elementary axioms) in Section 4, and truth for segments of the cumulative hierarchy of sets in Section 9. Those notions had been neglected by most logicians for a quarter of a century before Gödel's famous results put them back into circulation, by establishing memorable relations between them and formal notions. Apart from any heuristic value which non-elementary notions may have had for Gödel's own discoveries, they continue to be essential (even for an effective use of elementary logic itself; cf. the subsection on p. 84). The interest of Gödel's contribution is in no way diminished by the checkered development of the subject since then: by the efforts needed to discover rewarding applications, the limitations of Gödel's general program to apply traditional philosophical notions more broadly (in Sections 11 and 12), and not even by the endless refinements of his work which have gone far beyond the point of diminishing returns.

Also as promised at the outset, the refutations of the best known foundational schemes of this century by use of Gödel's results were compared with an alternative critique, by inspecting (later) mathematical experience. The comparison is familiar from so-called *purely mathematical* and *experimental* refutations of theories in the natural sciences:

- The former involve conflicts with very familiar facts, so to speak with the bare minimum expected of the theories; a standard example is Galileo's refutation of his (first) proposal that the velocity of a freely falling body is proportional to the distance covered (and so a body at rest would never start to fall at all, contrary to very familiar experience).

- So-called experimental tests of theories, even of those presented as abstractly as Newton's or Einstein's theories of gravitation, generally require a high level of unfamiliar extensions of ordinary experience (and, if positive, the tests supersede mathematical refutations of competing theories).

In the case of foundational schemes: Gödel's results provide mathematical refutations, while details of mathematical experience are used to pin-point less obvious defects of the schemes.

As a corollary (already asserted on p. 49): since the silent majority has the experience needed for the alternative critique, Gödel's results could not be expected to affect significantly the conception (let alone the practice) of that majority. Sooner or later, it would discount foundational ideals; either ignoring them altogether, or putting them in their place by reference to experience. For the same reason, the majority has no need for a pedantic formulation of the ideals themselves.

The presentation above leaves out of account a side of Gödel's contribution to foundations (in fact, of the subject of foundations itself, which is literally of the highest interest), for two principal reasons:

1. First of all, the scientific experience needed for the alternative critique has not been, and cannot be absorbed at all widely. Some philosophers (including Wittgenstein) have attempted something like that alternative, using only examples from quite elementary mathematics. This was unconvincing. It left a nagging doubt whether the examples were representative. More formally, as shown in logic (cf. Section 7), large parts of mathematics can be set out in accordance with conflicting foundational schemes, usually quite elegantly after some practice with the style involved. So, quite objectively, elementary experience is not enough for a decision between such schemes, let along against (all of) them.

2. Secondly, foundational questions occur to us when we know little; as little as a school boy in his teens, or even as little as the Greeks 2300 years ago. At this stage of experience the familiar foundational answers or schemes have

> a great attraction (in keeping with the objective fact, mentioned in 1, that limited experience does not decide against them). In such circumstances, it is rare indeed that anything significant (let alone conclusive) can be done using only a mild extension of familiar experience.

Gödel's results which are relevant here are significant, and can be *fully* understood with a minimum of background (especially those in Section 4 and 6 use no more additional knowledge than the elementary parts of logic available in the twenties, practically no more than needed to state the foundational schemes in mathematical terms[90]). So these results have an exceptional value, measured by the simplest criteria of all: the size and probable duration of the market for his contributions (or, equivalently, measured by the particular kind of fame which Schopenhauer analysed in Chapter 4 of his *Aphorismen der Lebensweisheit*).

This value of Gödel's results is of course quite separate from their value (for foundations or for science) at a more developed stage, perhaps to be compared to those elements which are valuable or even vital at an early stage of evolution, and less rewarding or even superfluous later. With one difference: in the case of the evolution of knowledge, each generation starts off at an early stage. Besides, for all of us there are areas about which we know little, and have first impressions analogous to foundational schemes. Gödel's results on the famous schemes of Russell and Hilbert, at least when looked at in the way just described, give one confidence in the possibility of analysing other schemes of this sort instead of simply suppressing them (and the analyses of other foundational schemes mentioned in Sections 7 and 9, support this confidence).

Sub specie aeternitatis, or at least as long as our age of intellectual affluence lasts, the value to the common understanding described above may well be seen as the most extraordinary part of Gödel's contributions, memorable as their scientific uses (reported in Sections 4, 6 and 9) undoubtedly are.

[90]His notes to later reprints or translations give, with loving care, the most economical formulations.

Chapter 7

Gödel's Early Works

Gödel became instantaneously famous after a couple of articles on mathematical logic at the ages of 24 and 25, proving the completeness of the (formal) rules for predicate logic ([1930]) and the incompleteness of formal systems for arithmetic ([1931]). They raise two principal questions, which could not have been answered at the time and arise as follows.

Gödel himself presented his work (as stressed in the introductions of his two papers) by reference to the sensational *formalist thesis*:

all mathematics is like doing sums

(in a sense explained in more detail below). As with other extravagant theses, programs, ideals, or what have you, the first question is:

Why not simply ignore such things?

Next, since work on extravagances rarely suggests convincing improvements, there is a second question:

What might be done with the (here, mathematical) tools used?

At this point one might agonize about ignoring these two questions in turn. Instead we'll ask them first about another thesis.

The Pythagorean thesis

The following is usually attributed to Pythagoras, more than 2500 years ago:

(natural) number is the measure of all things.

[0]This chapter is based on 'Gödel's Collected Works, Volume I', *Notre Dame Journal of Formal Logic*, 29 (1988) 160–181.

It is not, because no natural number or ratio between such numbers measures the diagonal of the unit square, which has length $\sqrt{2}$. The tools used in this refutation are the geometric theorem of Pythagoras about right-angled triangles, and an arithmetic theorem about the (so-called diophantine) equation $n^2 = 2m^2$, which has no solution in natural numbers n and m.

This refutation has not suggested any improvements, say, in the form of another measure for all things. On the contrary, being mundane enough to remove any sense of awe inspired by the Pythagorean thesis, the refutation raises questions about *assumptions behind the thesis*. For example,

> What would be so wonderful if the thesis were true?

After all, even where only rational numbers are 'needed' (as in limits for experimental errors), others are used (e.g., the interval for a measured length of some diagonal of a unit square is often given in the form $\sqrt{2} \pm \epsilon$, with rational ϵ and thus irrational endpoints). More generally,

> Is it sensible to demand *one* measure for *all* things, rather than relatively *few* measures for relatively *many* things?

So to speak in the opposite direction:

> Are there phenomena, possibly far-removed from everyday experience (of Pythagoras or even of ourselves) that do lend themselves to one, as it were, fundamental measure?

and so forth.

Similar questions come up about the thesis refuted by Gödel's work, but also analogues to the sociological fact that a few bands of the faithful continue to pursue the Pythagorean thesis (in numerology, or in reductions of mathematics to arithmetic).

As to the tools mentioned earlier, it would be unrealistic to try to be precise about cause and effect in the last 2500 years. (Did the yodeler or the echo trigger the avalanche?) As somebody said, such matters tend to be difficult just because they have so few consequences. Be that as it may, the tools used remain *memorable samples*.

Summary

Readers of Gödel's papers should not expect similarly colloquial language in the rest of the chapter. Partly this is a matter of temperament. But also, those of us who know the detailed analyses made in the meantime can now judge which familiar ideas are typical enough to illustrate a particular general issue reliably; occasional uses below of erudite language will thus serve as reminders of those analyses (for example: instead of 'doing sums' there will be 'formal procedures', i.e. computations according to the logical idea(lization) of the perfect computer).

But before we can even try to answer questions about tools used in Gödel's results, we'd better know (about) them. They are reviewed in Section 2, while Section 1 provides the needed background. We then return to the two questions in Section 3.

7.1 Background: Doing Sums Formally

We briefly describe formal numerical computations, as a quite typical example of those formal procedures which Gödel's early results refer to.

Numerical equations

The *alphabet* consists of the following symbols:

$$0 \quad 1 \quad + \quad \cdot \quad (\quad) \quad = .$$

The *terms* are defined inductively as follows:

- 0 and 1 are terms
- if t_1 and t_2 are terms, so are $(t_1 + t_2)$ and $(t_1 \cdot t_2)$.

An *equation* is an expression of the following form (where, here and below, t_1 and t_2 stand for arbitrary terms):

$$t_1 = t_2.$$

The *axioms* are:

- $1 = (0 + 1)$
- $(t_1 + 0) = t_1$
- $(t_1 + (t_2 + 1)) = ((t_1 + t_2) + 1)$
- $(t_1 \cdot 0) = 0$
- $(t_1 \cdot (t_2 + 1)) = ((t_1 \cdot t_2) + t_1)$.

A *derivation* is a finite partially (or, for ordinary writing, linearly) ordered sequence of equations E, such that E is *either* an axiom *or* obtained from preceding equations by the rule of *substitution of equals for equals*:

> if both $t_1 = t_2$ and $t_3 = t_4$, one can substitute t_2 for one or more occurrences of t_1 in $t_3 = t_4$.

An equation E is a formal *theorem* if there is a derivation whose last equation is E.

Exercises 7.1.1 a) *$t = t$ is a theorem.* (Hint: $(t+0) = t$ is an axioms. Substitute t for $(t+0)$ in $(t+0) = t$.)

b) *If $t_1 = t_2$ is a theorem, so is $t_2 = t_1$.* (Hint: Substitute t_2 for the left t_1 in $t_1 = t_1$.)

The *numerals* are defined inductively as follows:

- 0 is a numeral
- if t is a numeral, so is $(t+1)$.

Thus the numerals are

$$0 \qquad (0+1) \qquad ((0+1)+1) \qquad \cdots$$

and the n-th numeral is indicated by $\overline{n}$.

Exercise 7.1.2 *For each term t there is a numeral $|t|$, such that $t = |t|$ is a formal theorem.* (Hint: 0 is itself a numeral. $1 = (0+1)$ is an axiom. Note that for numerals $|t_1|$ and $|t_2|$, $(|t_1| + |t_2|)$ and $(|t_1| \cdot |t_2|)$ can be proved to be equal to numerals, and follow the buildup of terms.)

The symbols of the alphabet can be given their usual arithmetic meaning (parentheses being part of the notation for addition and multiplication). In that case, we can talk of an equation being *true* or *false*. The words *completeness* and *soundness* of formal rules (used in Section 3) mean here that exactly the true equations $t_1 = t_2$ are formal theorems.

Exercise 7.1.3 *The rules are sound and complete for equality of terms.* (Hint: The axioms are true and the rules preserve truth, hence all formal theorems are true. But $t_1 = t_2$ is true only if $|t_1|$ and $|t_2|$ are identical, in which case $|t_1| = |t_2|$ is an instance of $t = t$, and thus a formal theorem.)

Like other formal rules, those given above leave a choice in their order of application. When computing a numerical value, knowledge of arithmetic properties helps in an efficient choice; for example, knowing that $|(0 \cdot t)| = 0$ helps in computing the value of $(0 \cdot t)$.

Exercise 7.1.4 *For each term t,* $(0 \cdot (t+1)) = (0 \cdot t)$. (Hint: for $t_1 = 0$, $(t_1 \cdot (t+1)) = ((t_1 \cdot t) + t_1)$ becomes $(0 \cdot (t+1)) = ((0 \cdot t) + 0)$, which becomes $(0 \cdot (t+1)) = (0 \cdot t)$ when $(0 \cdot t)$ is substituted for t_1 in $(t_1 + 0) = t_1$.)

Thus, when computing $(0 \cdot (t+1)) = 0$, do not compute t first (as suggested by exercise 7.1.2), but use $(0 \cdot (t+1)) = (0 \cdot t)$ as the first step.

Numerical inequalities

The formal system above can be supplemented as follows, to deal not only with equalities of terms, but also with inequalities:

- The *alphabet* is enriched by the new symbol '$\neq$'.
- An *inequality* is an expression of the form $t_1 \neq t_2$, where t_1 and t_2 are terms in the previous sense.
- For any term t, the following is a new *axiom*: $(t+1) \neq 0$.
- There are two new *rules*, one for substitution:

  if $t_1 = t_2$ and $t_3 \neq t_4$, one can substitute t_2 for one or more occurrences of t_1 in $t_3 \neq t_4$,

  and one for deductions:

  $$\text{if } (t_1+1) \neq (t_2+1) \text{ then } t_1 \neq t_2.$$[1]

Exercise 7.1.5 *The new rules are sound and complete for inequality of terms.* (Hint: If $t_1 \neq t_2$ is true then one term, say $|t_2|$, of the pair of numerals $|t_1|$ and $|t_2|$ is a proper part of the other. Let $|t|$ be (the numeral of) the difference. Then $|t| \neq 0$ is an axiom, and $|t_1| \neq |t_2|$ is inferred from $|t| \neq 0$ by ($|t_2|$ applications of) the new deduction rule.)

Diophantine questions

They have been mentioned already (on p. 133) in connection with $\sqrt{2}$, and concern equations between *polynomials* with numerical coefficients; in the notation above, such polynomials are the terms generated by adding the 'variables' $x_1, x_2, \ldots$ to the 'constants' 0 and 1.

Diophantine questions ask whether or not an equation between polynomials has a solution by natural numbers:

- For *positive* answers, the rules for equality given above are enough: if $x_1, x_2, \ldots$ are (the numerals of) solutions, this fact is verified by computation.
- But there are simply no formal rules at all that are (correct and) enough for all *negative answers*, so-called diophantine inequalities (recall $n^2 \neq 2m^2$ in connection with $\sqrt{2}$).[2]

[1]This corresponds to the *cancellation rule* for equations:

$$\text{if } (t_1+1) = (t_2+1) \text{ then } t_1 = t_2,$$

which is superfluous in the sense that the rules for equations are complete without it.

[2]This is a negative response to Hilbert's demand for such rules (in his tenth problem).

Thus, even if only logical relations are considered, *solving diophantine equations is not like doing sums.*

There is a plausible parallel between diophantine questions and metamathematical questions about derivability by formal rules:[3]

- If a formula *is* derivable, this fact is verified by (nonnumerical) computation with formal objects.
- *Un*derivability is usually established by use of specific properties of the formula considered, as illustrated by non-Euclidean models in the case of Euclid's fifth postulate.

Thinking about sums: facts of experience

Building up derivations (for example, by following the rules above mechanically) is more exhausting than, say, *reflecting on shortcuts* and thus, realistically speaking, more liable to error.

This fact is dismissed in foundations as human weakness, and thus as irrelevant to logic. Perhaps; but if so, the broader philosophical topic of human data processing just does not have a (primarily) logical character. In particular, those 'weaknesses' may be of the essence in determining the extent to which human data processing is *not discrete*.[4]

7.2 Two Twists by Gödel on Cantor's Results

A formal counterpart to Cantor's coding arguments

Cantor's enumerations of pairs and of finite sequences of objects in an enumerated set are familiar. For example, the rationals are enumerated as pairs of integers, and the algebraic numbers by the sequence of coefficients of their primitive equations. Formal objects like those of Section 1 are sequences (of letters in the alphabet of the system used), and can thus be enumerated too.

Cantor himself did not pay attention to the numerical properties and relations that correspond to those for numbered sequences. But given a numbering it is

[3] Specialists know a precise sense of this parallel from work on Hilbert's tenth problem (cf. note 2).

[4] In *Philosophical Investigations*, Wittgenstein queries about *following* (mechanical) *rules correctly*. His wording is purely logical: it concerns the idea of a correct application of a given rule (like Kant, *Kritik der reinen Vernunft*, A 132–133), and possibly irreconcilable conflicts over different interpretations of that idea in certain imagined situations.

Actually, computational errors do occur; with the difference that their presence is recognized (even if no erroneous step is located!), and so those (imagined) conflicts are rare. However, it is to be noted that the errors are of a kind that would not at all be expected to occur frequently in wholly discrete data processing.

often, in practice, a matter of routine to write them down.[5] Here 'in practice' means that any of the familiar arithmetic operations, like exponentiation, are used; in any case, polynomials alone (the subject of Section 1) would not be enough.[6]

The following twist by Gödel involves something new, even if it is little more than remembering (the possibility of) incompleteness; in other words, a difference between truth and formal provability.[7] It concerns two kinds of formal representation of properties of numbers (of formal objects).

First, let T be the property of being the number of a *formal theorem* of the system considered, and D[8] be a formula with one free variable.

Definition 7.2.1 *D is said to* **represent** *T numerically if, for each number n,*

$$T(n) \text{ is true} \quad \Leftrightarrow \quad D(\overline{n}) \text{ is a formal theorem.} \tag{7.1}$$

Here, as in Section 1, $\overline{n}$ is the numeral of n.

Exercise 7.2.2 *Verify that in an inconsistent system, where every formula is a formal theorem, only one property is representable.* (Hint: Consider the property which holds for all n.)

Let $\neg$ be formal negation. For a *complete* (formal) system, where either $D(\overline{n})$ or $\neg D(\overline{n})$ is a formal theorem, it follows that

$$T(n) \text{ is false} \quad \Leftrightarrow \quad \neg D(\overline{n}) \text{ is a formal theorem.} \tag{7.2}$$

---

[5]For example, it is in the case of the operation τ defined as follows:

> $\tau(w, v)$ is the result of substituting (the sequence) v for some chosen element in w,

a version of which is considered below.

[6]*Numberings* (of words of a numbered alphabet and of other syntactic objects) are used traditionally, and appropriately if some proposition about numbers is to be shown formally independent.

But if the rhetoric about set-theoretic foundations were taken literally, one would consider systems for sets, and code (that is, represent) their syntactic objects by means of *hereditarily finite sets*, as is done in some elementary texts. For the rhetoric mentioned the 'identification' of symbols with sets is a matter of course, and the representation of sequences of sets by sets is familiar. Viewed this way, it is quite lopsided to present *arithmetization* as a most central component (let alone novelty) in Gödel's proofs of incompleteness.

Arithmetization or, more precisely, ingenious variants have *become* central for such delicate later developments as reducing the number of variables in a 'universal' diophantine equation.

[7]Actually, Gödel's proofs apply also to suitable sets (of axioms) that are not formal (or, equivalently, recursively enumerable). Basically, the set of consequences should be representable (it is if, for example, all true $\forall$-sentences are added to formal arithmetic).

[8]Pedantically, D_T should be used here (since D depends on T).

To underline the point, the literature uses a new word for representations that also satisfy 7.2: originally, *entscheidungsdefinit*; today, more often, *invariant definitions*.

The idea is extended to *sequences* $\vec{P}$ of properties $P_0, P_2, \ldots$

Definition 7.2.3 *A formula F with two free variables represents $\vec{P} = \{P_0, P_1, \ldots\}$ if, for all m and n,*

$$P_m(n) \text{ is true } \Leftrightarrow F(\overline{m}, \overline{n}) \text{ is a formal theorem.}$$

Evidently, even if each member of a sequence has a representation, the whole sequence need not be representable by any formula of the system (for example, in Section 1, each polynomial is represented, but no sequence that includes all the polynomials).

This warning serves as a foil to *representing a sequence of all representable properties* by use of D and the substitution operation σ:[9]

> $\sigma(\overline{m}, \overline{n})$ is the numeral[10] of the formula $M(\overline{n})$, where M is the formula whose number is m.

If P_m is the property represented by the formula with number m then

$$P_m(n) \text{ is true } \Leftrightarrow D[\sigma(\overline{m}, \overline{n})] \text{ is a formal theorem.}$$

If the system considered is complete, then the representation is a definition and $D[\sigma(\overline{x}, \overline{y})]$ would *define* an enumeration of all representable properties (a notion familiar from Cantor's cardinal arithmetic).

A formal counterpart to Cantor's diagonal arguments

As usual, two properties of numbers are called *different* if some number has one of the properties, but not the other.

Proposition 7.2.4 *For every sequence $\vec{P}$ of properties there is a property $\vec{P}_d$, depending of course on $\vec{P}$, that is different from each property of the given sequence.*

Proof. If $\vec{P} = \{P_0, P_1, \ldots\}$ let

$$\vec{P}_d(n) \text{ be true } \Leftrightarrow P_n(n) \text{ is false.} \tag{7.3}$$

[9] The operation σ is a variant of τ in note 5.

[10] In the rest of this section, for convenience, when we say 'numeral' of a formula we mean 'numeral whose numerical value is the number' of a formula.

Consider any P_n. It differs from $\vec{P}_d$; specifically, at the argument n which, by definition, has the property $\vec{P}_d$ if and only if it does not have the property P_n.[11]
$\square$

For a long time, the principal use of the passage (that is, the operation) from $\vec{P}$ to $\vec{P}_d$ was made in cardinal arithmetic, Cantor's pet. Specifically, in contrast to (say) the set of all algebraic numbers, and like the set of all real numbers, *the set of all sets of natural numbers is not enumerable*. An enumeration is nothing else but a sequence, and it would not include the corresponding diagonal set 7.3.

Long before the representation of all representable sets (by Gödel, cf. the end of the previous subsection) Cantor's argument caused malaise, and people thrashed about for ways of expressing this malaise. There were those dubious doubts about the existence of (the) uncountable sets mentioned above, but also talk about the language in which they are defined, although cardinal arithmetic is not restricted to sets that happen to be specified in any particular language.

Similar words (but with quite a different meaning!) get a point in the twist from cardinal arithmetic to formal representations. Gödel's twist[12] on Cantor's diagonal construction is applied to the sequence represented by $D[\sigma(\overline{x},\overline{y})]$.[13]

Theorem 7.2.5 *If the property of being a formal theorem (of T) is representable in T, then T is incomplete.*

Proof. As already noted, if T were complete then $D[\sigma(\overline{x},\overline{y}]$ would *define* an enumeration of all representable properties.

The diagonal set of that sequence would be defined by $\neg D[\sigma(\overline{x},\overline{x})]$, which is of the form

$$D[\sim \sigma(\overline{x},\overline{x})],$$

where $\sim \overline{n}$ is the numeral of the negation of the formula whose numeral is $\overline{n}$.

Now, the formula

$$G(x) = D[\sim \sigma(\overline{x},\overline{x})]$$

[11]The literature sometimes speaks of *self-reference* here. This is literally true, since an argument where $\vec{P}_d$ differs from P_n is the subscript of P_n itself (after all, for most function terms their evaluation at the argument n refers to n).

Psychoanalysts may speculate on the fact of experience that the word has clouded the critical judgment of many, but the practice itself is harmless.

[12]The twist is only implicit in Gödel's famous paper [1931]. It is explicit in Turing's work [1936], but also in Gödel's letter to Zermelo (of 1931), which can be found in Grattan-Guiness [1979].

[13]This is a *representation of a sequence* of all representable properties (of the system considered) in the sense above or, equivalently, an *enumeration* of them for $m = 0, 1, \ldots$ It is perhaps *satisfaisant pour l'esprit* that Kleene called such enumerations, which are indeed central to *in*completeness, *complete* (for recursively enumerable sets); 'universal' is more usual now.

In the language of functions, they correspond to partial recursive enumerations of all partial recursive functions. Here the difference between *total* and *partial* functions corresponds to the difference between truth and provability, mentioned above.

has a number g. By the diagonal construction, this purported representation of the diagonal set is certainly not a definition for $x = g$, since (by 7.3),

$$D[\sim \sigma(\overline{g}, \overline{g})] \text{ is true } \;\Leftrightarrow\; G(\overline{g}), \text{ i.e. } \neg D[\sim \sigma(\overline{g}, \overline{g})], \text{ is false.}$$

Hence T is incomplete. $\square$

Readers familiar with the literature are warned that, instead of

$$G(\overline{g}) = D[\sim \sigma(\overline{g}, \overline{g})] \text{ with } g \text{ number of } G(x) = D[\sim \sigma(\overline{x}, \overline{x})],$$

usually the formula

$$G_0(\overline{g}_0) = \neg D[\sigma(\overline{g}_0, \overline{g}_0)] \text{ with } g_0 \text{ number of } G_0(x) = \neg D[\sigma(\overline{x}, \overline{x})]$$

is used.

Interpretations

For the rest of this section T is a system as in 7.2.5, that is in which the property of being a formal theorem is representable (by D).

Since the proof above is a little slick, it pays to *interpret* the formulas used; in particular, in terms of two related metamathematical properties of formal systems:

- *soundness*: for every formula F, if F is a formal theorem then F is true;
- *consistency*: for every formula F, F and $\neg F$ are not both formal theorems.

We start first with a *literal interpretation*. Since D represents the property of being a formal theorem, $G(\overline{g}) = D[\sim \sigma(\overline{g}, \overline{g})]$ expresses

'the formula with numeral $\sim \sigma(\overline{g}, \overline{g})$ is a formal theorem'.

But $G(\overline{g})$ has numeral $\sigma(\overline{g}, \overline{g})$, and thus the formula with numeral $\sim \sigma(\overline{g}, \overline{g})$ is $\neg G(\overline{g})$. Then

$G(\overline{g})$ expresses '$\neg G(\overline{g})$ is a formal theorem',

and

$\neg G(\overline{g})$ expresses '$\neg G(\overline{g})$ is not a formal theorem'.[14]

Theorem 7.2.6 *If T is consistent, neither $G(\overline{g})$ nor $\neg G(\overline{g})$ is a formal theorem.*

[14] Similarly, for the formula G_0 considered above, one has that

$G_0(\overline{g}_0)$ expresses '$G_0(\overline{g}_0)$ is not a formal theorem'.

Proof. Suppose $G(\overline{g}) = D[\sim \sigma(\overline{g}, \overline{g})]$ is a formal theorem. Then, by the right-to-left direction of 7.1, (the formula with numeral $\sim \sigma(\overline{g}, \overline{g})$, i.e.) $\neg G(\overline{g})$ is also a formal theorem.

Conversely, suppose $\neg G(\overline{g})$ is a formal theorem. Then, by the left-to-right direction of 7.1, ($D[\sim \sigma(\overline{g}, \overline{g})]$, i.e.) $G(\overline{g})$ is also a formal theorem.

Thus, if one of $G(\overline{g})$ and $\neg G(\overline{g})$ is a formal theorem, so is the other, and the system is inconsistent. □

In particular, $G(\overline{g})$ and $\neg G(\overline{g})$ are witnesses of the incompleteness of T.

Theorem 7.2.7 *$\neg G(\overline{g})$ is an instance of soundness (precisely, for $\neg G(\overline{g})$ itself).*

Proof. We have already seen that $G(\overline{g})$ means '$\neg G(\overline{g})$ is a formal theorem'. As usual, $\neg G(\overline{g})$ means '$\neg G(\overline{g})$ is true'. Thus, the instance of soundness for $\neg G(\overline{g})$:

$$\text{if } \neg G(\overline{g}) \text{ is a formal theorem then } \neg G(\overline{g}) \text{ is true}$$

reduces to

$$G(\overline{g}) \Rightarrow \neg G(\overline{g}),$$

which is equivalent to $\neg G(\overline{g})$. □

As a corollary, (even) the particular case of the soundness principle applied only to $\neg G(\overline{g})$ is not a formal theorem of the systems considered in 7.2.5.

Theorem 7.2.8 *If the system is complete for $G(\overline{g})$, i.e.*

$$\textit{if } G(\overline{g}) \textit{ is true then } G(\overline{g}) \textit{ is a formal theorem,}$$

then $\neg G(\overline{g})$ is an instance of consistency (precisely, for $G(\overline{g})$ and $\neg G(\overline{g})$ themselves).

Proof. The completeness assumption reduces to

$$G(\overline{g}) \quad \Rightarrow \quad D[\sigma(\overline{g}, \overline{g})]. \tag{7.4}$$

Consistency for $G(\overline{g})$ and $\neg G(\overline{g})$ is equivalent to

$$\text{if } G(\overline{g}) \text{ is a formal theorem then } \neg G(\overline{g}) \text{ is not a formal theorem,}$$

which reduces to

$$D[\sigma(\overline{g}, \overline{g})] \ \Rightarrow \ \neg D[\sim \sigma(\overline{g}, \overline{g})],$$

and hence to

$$D[\sigma(\overline{g}, \overline{g})] \quad \Rightarrow \quad \neg G(\overline{g}). \tag{7.5}$$

By 7.4 and 7.5,

$$G(\overline{g}) \ \Rightarrow \ \neg G(\overline{g}),$$

which is equivalent to $\neg G(\overline{g})$. Thus consistency implies $\neg G(\overline{g})$.

Conversely, $\neg G(\overline{g})$ is equivalent to $\neg D[\sim \sigma(\overline{g}, \overline{g})]$, and hence it trivially implies consistency for $G(\overline{g})$ and $\neg G(\overline{g})$ in the form given above. $\square$

In terms of Section 1, here two special properties are used: that of $G(\overline{g})$ being *like* solvability of a diophantine equation; and of course of the system considered, which must prove of itself that verification by computation is possible.

Corollary 7.2.9 *If a consistent system is provably complete for $G(\overline{g})$, then consistency (for $G(\overline{g})$ and $\neg G(\overline{g})$) is not a formal theorem.*

Proof. The hypothesis means that

$$G(\overline{g}) \ \rightarrow \ D[\sigma(\overline{g}, \overline{g})]$$

is a formal theorem.

The conclusion means that

$$D[\sigma(\overline{g}, \overline{g})] \ \rightarrow \ \neg D[\sim \sigma(\overline{g}, \overline{g})]$$

is *not* a formal theorem. If it were, by the proof above $\neg G(\overline{g})$ would be a formal theorem too, contradicting 7.2.6. $\square$

In particular, for systems S as in 7.2.5 with the *additional* hypothesis of 7.2.9, the consistency of S is not a formal theorem, and so *it is consistent to assume* (that is, add to S the axiom) *that S is inconsistent*.

It should be recalled that many current systems satisfy the additional condition that, demonstrably, each inconsistency implies an arbitrary formula. With this *further* condition,[15] *it is consistent to assume that every proposition is a formal theorem*.

Finally, it is worth noting that the (second) relation above, with the soundness principle, makes consistency - for the special systems considered - more than a purely necessary, so to speak negative virtue (i.e., the absence of particularly crass errors like contradictions).

Discussion

We have conflicting requirements for a scientifically successful *pursuit* of mathematical logic, and for a philosophically adequate *examination* of foundational claims like Hilbert's thesis. For the former it is most rewarding to pursue the details of the proofs above, including the details of D, thus establishing the potential of formalization. But for a proper assessment of the intended thesis it is essential to realize how few details are needed (for its refutation)!

[15]Some current systems (for example, cut-free ones) do *not* satisfy this additional condition.

> *Either* the various representations are not available in the system considered (and the system fails because it lacks expressive power), *or* the system is incomplete (and so fails because it lacks deductive) power.

This refutation is, perhaps, little more than a wisecrack. If so, the punishment fits the crime, as it were: the formalist thesis is so badly wrong that a refutation is so *un*demanding.

This kind of thing is familiar in memorable foundations, and not only in the case of refutations. Thus Tarski gave an impeccable translation of

'snow is white' is true

by

snow is white

(used in 7.2.7). Whatever problems there may be here, they concern the optical properties of snow, not the general notion of truth. This shows, for example, that Pilate had no reason to stay for an answer (in any case, it does not seem to be recorded whether he asked his question with bated breath or a shrug-and-a-wink).

This use of mathematics (in particular, mathematical logic) is a refrain of the whole chapter. Being mundane enough to remove any sense of awe inspired by those Big Words, it *corrects* our view of them.

Digression on representations (optional)

For the results up to 7.2.7, the notion of numerical representation for *properties* is perfectly adequate. Specifically, the results stated hold for all representations (of the particular property in question), even though the latter need not be formally equivalent.[16]

More generally (for example, for 7.2.9), attention is required not only by the representation of properties, but also of *propositions*. Here, as so often, the best guide for progress comes from broad mathematical experience rather than

[16] For example, if σ represents the substitution operation, so does $\sigma_1 = \sigma + 0$, but the formulas $D[\sim \sigma(\overline{x}, \overline{x})]$ and $D[\sim \sigma_1(\overline{x}, \overline{x})]$ have different numbers g and g_1. However, both $D[\sim \sigma(\overline{g}, \overline{g})]$ and $D[\sim \sigma_1(\overline{g}_1, \overline{g}_1)]$ satisfy the condition for so-called *Gödel sentences* S:

$$S \Leftrightarrow D(\sim \overline{s}),$$

where s is the number of the formula S.

In (literate) English the popular use of the definite article ('the' Gödel sentence) requires some equivalence relation connecting all formulas that satisfy the condition above. For broad classes of common systems all Gödel sentences are in fact formally equivalent, for others they are not or not known to be equivalent. (Specialists will think here of so-called Rosser systems.)

Similar questions can be raised about 'this' in 'This sentence is false'.

from the recent logical literature; specifically, from the introduction of (algebraic) coordinates in (synthetic) geometry:[17]

- Codes (arithmetic or set-theoretic) correspond to the *coordinates*;
- syntactic objects (either thought of as finite words of an enumerated alphabet, or as finite trees) correspond to *geometric points*;
- properties of the objects (such as terms, formulas, derivations) and relations (say, between derivations and their last formula) correspond to *geometric relations* (for example, collinearity).

At least one (historical) difference is to be noted, especially with respect to the delicate cases of projective or desarguean geometry, which would be called *weak* systems in contemporary logical jargon (compared to the geometry of the full Euclidean plane).

- The geometric axioms came first, and the corresponding algebraic ideas (of skew fields, fields etc.) afterwards.
- In contrast, axioms for arithmetic (including weak subsystems of familiar formal arithmetic) came first, while formal systems for syntax (concatenation theory, or theories of finite trees) are still not very familiar; especially, the choice of axioms satisfied by the syntactic properties and relations above is usually left implicit.

---

[17]Readers familiar with *interpretations* in the sense of Tarski (that is, uniform definitions of a model for one theory in another) may wish to compare the introduction of coordinates to the interpretation of the particular systems of geometry considered in the corresponding algebraic systems; to be quite precise, systems for vectors over the coordinate space (projective geometry corresponds to skew fields, Pappus to commutative fields, etc.). But over and above an interpretation, coordinates provide an *embedding*, which induces the (definable algebraic) relations that interpret the so-called nonlogical constants of the geometry considered. All this is well known.

Hilbert went further in his *Foundations of Geometry*, and interpreted the algebraic systems in the corresponding geometry too. Specifically, he defined ternary relations A and M in the language of projective geometry, and proved for them geometrically the laws of *addition* and *multiplication* that hold in the corresponding algebra, with the same (formulas) A and M for all the extensions of the projective axioms and the axioms for skew fields (and the same embedding, the identity, for the planes and the algebras in question). This *tour de force* is a high spot in the tradition to which so-called reverse mathematics belongs.

In the present subsection coding of syntax is meant as interpreting a suitable formal theory of syntax (i.e. of finite words or finite trees, with certain inductive definitions of such syntactic properties as: being terms, proofs, provable) in generally weak systems of arithmetic. The reverse direction corresponding to Hilbert's *tour de force* does not seem to have been investigated. Roughly speaking, it looks for an arithmetic 'structure' in - the language of - syntax (even though so far this has no more been needed in coding than the *tour de force* was needed in geometry).

Strictly speaking, a *relative* interpretation is involved since the embedding is not *onto* (that is, not all natural numbers are codes of some syntactic object).

> Put differently, it is left open which *data* determine the formal rules considered for generating terms, etc.[18]

There is also a difference in jargon. The mathematical tradition does not have the refrain, introduced in this section, about the difference between truth and *provability*, but about truth (in, say, the Euclidean plane) and *validity* (for all projective or desarguean planes): the definitions involved in the introduction of coordinates are *uniform* (for all projective or desarguean planes considered).

There are other differences too. For example: in provability logic one has iterated codes (of codes, etc.), but practically never in geometry (coordinates of coordinates). In the opposite direction: the literature on (Gödel) numberings is rarely explicit about the 'structure' on finite sequences (to be represented).

Readers will recall two elements in the introduction of coordinates. First, there is the matter of determining (algebraic) coordinates for points *uniquely up to transformations* of a suitable kind (usually, some transformation group). Secondly, for a given choice of coordinates it is shown that algebraic relations that satisfy the geometric axioms must be of a certain form (for example, a relation satisfying the axioms for collinearity in two dimensions will be linear in the algebraic sense).

The notion of *canonical representation* (of properties of numbers of formal objects), follows the model just recalled.[19] In particular, given a (canonically defined) coding of the syntactic objects, such canonical representations of syntactic properties are arithmetic predicates satisfying (demonstrably) the inductive definitions induced by the Post production rules for those properties. Any two canonical representations of a given property are then demonstrably equivalent, provided of course the system of arithmetic considered contains enough induction.

The corresponding *uniqueness condition* satisfied by the (canonical) codings themselves is adequately illustrated by the humble matter of (surjective) *pairing* π with its left and right inverses λ and ρ, satisfying

$$z = \pi[\lambda(z), \rho(z)]$$

---

[18] If Post production rules are meant, the corresponding additional axioms of the concatenation theory have the form of so-called elementary inductive definitions (including the principle of proof by induction, which expresses that the least solution of the inductive definition is meant).

[19] For the record, I was not conscious of the close relation between canonical representations and the introduction of coordinates when I introduced the former some thirty years ago (though I had learnt the latter from Hilbert's *Foundations of Geometry* twenty years earlier; cf. the footnote on p. 261 of Hodge and Pedoe [1947]).

Incidentally, the classical literature on the foundations of geometry was far more sophisticated than the (sloppy) literature on *natural* representations, especially common in connection with consistency statements. It did *not assume* that what happened to be familiar from ordinary analytic geometry since Descartes would also be natural for, say, all projective planes. On the contrary, it investigated *whether* familiar techniques were adequate in the case of unfamiliar, so to speak nonstandard (projective) planes; 'adequate' for its elegant (though half-forgotten) representation theorems, with their explicitly stated adequacy conditions.

for all z.[20] If (π', λ', ρ') is another such pairing, then the transformation from z to z' is given by:

$$z' = \pi'[\lambda(z), \rho(z)].$$

This uniqueness up to definable isomorphism here corresponds to uniqueness within the geometric transformation group in the case of coordinates.

Before leaving the subject of canonical representations, it is as well to recognize a *wide-spread malaise* about details of coding, often felt to be boring. This is a 'subjective correlative' of the fact that, generally, the results actually proved (about a particular coding or by use of it) are also valid for any other coding that may come to mind. In other words, the details are introduced for some ritual of 'precision', which draws attention away from the more demanding questions of what one is being precise about, and why.

- At an *elementary* stage there are two principal strategies for progress.

  One, exemplified by the canonical representations above (and categorical axioms in another sphere), is to formulate some kind of maximal requirement, enough for any developments in sight.

  At an opposite extreme one looks at minimal requirements for each of the more prominent results.

- At a *later* stage, brute power may be used (as in the 'ingenious variants' alluded to in note 6) and, as always, there is the delicate job of discovering relatively few requirements adequate in relatively many situations.[21] (All this may be hard, but it is not boring.)

---

[20]The weaker condition on (not necessarily surjective) pairings that used to be quoted, namely that

$$\pi(x_1, y_1) = \pi(x_2, y_2) \;\Rightarrow\; (x_1 = x_2 \wedge y_1 = y_2),$$

would of course not be enough for the following transformation.

[21]A good example is provided by the popular and successful *modal language of provability logic*. It realizes the idea of 'relatively many situations' by its established expressive power. A moment's thought shows that the full panoply of canonical requirements is liable to be excessive here: they are enough to cover also propositions about *proofs* (i.e., derivations), while the language is restricted to prov*ability* statements.

For example, the latter require only conditions on D which ensure that statements in the modal language are equivalent when D is substituted for $\Box$. Because of the iteration of $\Box$, it is not enough that D be a representation in the sense used in the text.

As experts surely know but do not publicize, if modus ponens or, pedantically,

$$\Box(p \to q) \;\to\; (\Box p \to \Box q)$$

is derivable for the representation D used, the required condition simplifies. Instead they talk of modus ponens being 'natural', ignoring both the *discovery* that cut-free rules are useful, and a *philosophical* observation (cf. note 32) about the general weakness of the formal picture for understanding the phenomena of proof.

Systems that prove their own consistency

The 'deviant' (self-referential) sentence introduced by Rosser (to eliminate Gödel's condition of ω-consistency for a system S) is simply a Gödel sentence (in the sense of note 16) for the canonical representation of the system S_R defined as follows:

> d is a derivation of A in S_R if d is a derivation of A in S, and there is no derivation d_1 of $\neg A$ before d (i.e. with $d_1 < d$).

I do not know whether, in general, S_R proves its own consistency (of course, formulated canonically), but S_R^+ certainly does, where it is required in addition that

$$\text{if } A = \neg B, \text{ there is no derivation } d_2 \text{ of } B \text{ before } d \text{ either.}$$

Note that the passage from S to S_R^+ is (not only a simple but) a recursive operation.[22] Incidentally, and without forgetting limitations of *all* formal systems, S_R^+ *possesses a feature of actual experience with proofs that is not present in more usual systems*: the results are cross- checked against background knowledge.

For consistent S, S and S_R^+ have not only the same set of theorems but the same proofs; only the procedures for checking the latter are different. Thus the (still common) formulation of Gödel's second theorem 'for all sufficiently strong systems' is not at all sloppy, but just sadly ignorant (in particular, of anything like S_R^+ above).

A civilized formulation (modulo minimal conditions) is in the next result, sharpened and expressed in modal language (where $\Box$ is substituted for D, and completeness is restricted to provability statements).

Theorem 7.2.10 *Either consistency or completeness for Σ_1^0 statements is not (internally) derivable.*

Proof (following Jeroslow). Suppose

1. $G \leftrightarrow \Box\neg G$[23]
2. $\Box\neg G \to \Box\Box\neg G$
3. $\neg(\Box G \wedge \Box\neg G)$

are all formal theorems. They are if, respectively:

1. G is a Gödel sentence

[22] Here meant in the usual strong sense, not merely the sense of having recursive values at recursive arguments.

[23] This a modal counterpart of so-called literal Gödel sentences (cf. note 16), constructed earlier in this section.

2. completeness for Σ_1^0 statements (in particular $\neg G$) is internally derivable

3. consistenct for Σ_1^0 statements is internally derivable.

Then

$$\Box\neg G \leftrightarrow \Box\neg\Box\neg G$$

by 1 (substituting the right-hand-side for G on both sides). In particular,

$$\Box\neg G \rightarrow \Box\neg\Box\neg G.$$

Together with 2, we have

$$\Box\neg G \rightarrow \Box(\Box\neg G) \wedge \Box\neg(\Box\neg G)$$

and, by 1 again,

$$\Box\neg G \rightarrow \Box G \wedge \Box\neg G.$$

But the conclusion is the negation of 3, so ($\neg\Box\neg G$ and, by 1) $\neg G$ is derivable. This is impossible, since G is a Gödel sentence. $\Box$

Corollary 7.2.11 *Ordinary cut-free systems do not prove their own consistency.*

Proof. The proof above does not use closure under *modus ponens*, and ordinary cut-free systems are complete with respect to Σ_1^0 statements. $\Box$

Notice that here consistency is taken in a sensible form, not merely in Hilbert's coy version '$\neg\Box\perp$', which implies usual consistency in the presence of *modus ponens*. In cut-free systems, $\neg\Box\perp$ *is* provable in the system itself.

Finally, to repeat what cannot be repeated too often: *without completeness with respect to Σ_1^0 statements, consistency is very pale indeed.* Indeed, it does *not* even ensure the truth of (proved) Π_1^0 statements (the formula $G(\overline{g})$ in 7.2.5 is Σ_1^0, and so $\neg G(\overline{g})$ is Π_1^0).

7.3 Back to the Two Questions

Formal (or, equivalently, mechanical) aspects of mathematics: what are they, and what are they supposed to do?

These questions, but not the specific answers below, come from the title of Dedekind [1888].

Computations with $0, 1, +, \times$ (as done in elementary school) are quite typical of formal procedures. Section 1 has sketched both the drill involved, and shortcuts resulting from (humanly inevitable) reflection on it.

Today, computations on an electronic computer are familiar. They are even better examples, with one proviso: *not* too much attention to details either of the hardware or of the wetware (computer jargon for 'intellect'), since the logical idea of the computing machine corresponds only to Simple Simon's image of computing Man.

Both kinds of examples convey very well what formal procedures *are*, but they do not provide an effective background for Gödel's (best known) result. This states some odd things that even an ideal computer can*not* do, which are wholly overshadowed by the many things that even actual computers *can* do; realistically speaking, far beyond Leibniz's dream.[24]

The extravagant formalist thesis provides, of course, a much more glamorous background. The first of the details about that thesis (promised at the outset) is a *distinction*, between:

- thought processes in mathematics
- mathematical thoughts (or, more simply, results).

The pioneers, especially Frege, concentrated on the logical relations between results; without any claims on the full range of the mathematical imagination, or even dismissing questions about the latter as irrelevant (to mathematics).[25] Brief indications of this distinction, applied to addition and multiplication, are in Section 1.

What is at issue is an understanding or theory of *reasoning* (at least, in mathematics). And the claim of the formalist thesis is that the formal elements reflect all that is significant. Thus a theory takes the form of a *formalization*, which consists of: a formal language with a basic alphabet, a formal grammar, and formal rules of derivation. Section 1 provides a sample, but the formalization of elementary (predicate) logic, which appeared more than one hundred years

[24]Even if not the oft-quoted part about settling moral or legal squabbles mechanically. It seems to me it should not be too hard to program expert systems that generate, statistically, more or less the usual judgments. But the market is likely to be limited. Lawyers would not be enthusiastic, for obvious reasons; nor those, among the accused and litigants, who are not satisfied with statistical justice (even when they accept the two cherished principles of the uniqueness of each individual and of equality before the law, which would make it wise to be satisfied).

[25]In contrast to the pioneers a century ago, around the middle of our century Turing proposed a test for 'identifying' thinking behavior by its not being distinguishable from human performance. Naturally, distinctions by reference to certain aspects (so-called results) were meant; comparable in the case of *artificial locomotion* to 'identifying' walking with roller skating as long as one starts and finishes together. For artificial locomotion viewed as a branch of engineering, which lives on achieving given results (i.e., tasks) by novel processes, Turing's proposal is a matter of course. It becomes remarkable if it is viewed as contributing somehow to elucidating the processes in data processing: by putting them into black boxes, as it were, or ignoring them in other ways. Needless to say, that 'test' has become popular among the vulgar, like the most vulgar uses of Ockham's razor (where the fact that something is not needed to explain a particular bunch of phenomena is interpreted to show that it does not exist). For the record, that 'test' is the only indiscretion of Turing that I both know of and have found at all disturbing.

ago, remains more impressive; above all, because of the expressive power of its simple notation (vocabulary and grammar). This became relatively soon part of mathematical (and even general) culture; much more so than its rules of inference, inevitably reminiscent of processes.

Let there be no mistake: even the formalization for doing sums is amazingly simple, compared to the phenomena that present themselves naturally; specifically, in the many nuances of natural (written, and especially spoken) mathematical language(s). The price to pay for this simplicity is a malaise: the formal elements constitute a *very pale picture* of even the two elementary parts of mathematics in the last paragraph, let alone a broad mathematical experience.

One consolation, albeit overlooked by Goethe himself, is implicit in 1.2037 of *Faust I*: *all* theory is grey. So paleness by itself is not a defect of any theory, logical or not. More substantial encouragement comes from the mechanical picture of the physical world around us in terms of point masses and their motion in space-time (which leaves out colors and shapes, not to speak of chemical composition). This picture, hardly less pale than the formal picture above, has not only been most successful in its domain, but remained for centuries a model for theoretical understanding. After Section 2, readers may pursue the parallel with the formal picture(s) of the world of mathematics a step further: not only are the objects of mechanics here represented, but also spatio-temporal relations between them (cf. Gödel's twist representing properties of Cantor's numerical representations of words).

Memorable formalizations were proposed in the two decades around the turn of this century:

- by Whitehead and Russell for *all of mathematics* (in the three volumes of *Principia Mathematica*);
- by Hilbert for *various branches* of it (beginning, in the *Foundations of Geometry*, with elementary geometry), with respect for both the venerable ideal of purity of method, and for the mathematical tradition of concise exposition.

But for the sequel, and probably *sub specie aeternitatis*, the following difference between the two styles is much more consequential:

- In his so-called metamathematics Hilbert paid attention to *global mathematical* properties of the formal pictures, such as completeness (taken up in the next paragraph).
- *Principia* did not; in the tradition of natural history, which is content with a compact description of data that happen to catch our attention (by mathematical formulas, when it uses mathematics at all).

Admittedly, this side of metamathematics got lost in Hilbert's later rhetoric, for example, about real and ideal elements in the tradition of Ockham's razor (already mentioned in note 25).

A specific principal requirement on formal pictures is this. For propositions P represented by p in the formalization considered, p should be a formal theorem if and only if P is true. When P is about some particular notion or 'structure', then

either P is true or its negation is.

The corresponding mathematical property of the formalization is that

either p is a formal theorem or the formal negation of p is.

This is called *formal completeness* (cf. p. 136).[26]

Perhaps it is worth adding that, even when P is about some specific structure, formal completeness would not be generally even plausible, if thought *processes* were the main object of study. There is simply not a shred of evidence that every problem of, say, Higher Arithmetic is solvable in any even remotely realistic sense, let alone that we should want to look at every problem (in current formalizations).

The incompleteness theorem: a literal refutation

Gödel's (most famous) incompleteness theorem was originally stated for *Principia* and related systems; in fact, for the parts that serve to represent arithmetic properties. For each such system T, some true proposition of arithmetic is not a formal theorem (of T), where the proposition depends upon T. Thus, *doing higher arithmetic is not like doing sums* (cf. Section 1 on diophantine equations).

As we saw in 7.2.7, the particular true proposition obtained by Gödel has a simple interpretation: it is (the so-called arithmetization of) an instance of the soundness principle.

For our ordinary view of (the logical relations between results in) mathematics, this principle is a minimal consequence of understanding the rules of T at all (and, it may be added, rules actually used are understood; at least, enough for this consequence).

Viewed this way, T would simply be said not to prove the particular property of itself expressed in the principle, though T proves many things (also about

[26]In general, the restriction to particular notions is needed, as seen in the case of pure logic where (as Leibniz put it) truth in *all* possible worlds is meant. So P is not logically true if it is false in *some* such world. But the negation of P is logically true only if P holds in *no* possible world. Thus neither P nor its negation need be logically true, and so formal completeness is not required here.

Nor is it required when P is about (say) sets, before we have made up our minds on the particular kind of set to be considered; for some P, neither P need be true for all kinds of sets contemplated, nor its negation.

itself). Thus *the formalist thesis is refuted according to the letter*, since one of its explicitly formulated (cl)aims (completeness for arithmetic) is not realized.

The completeness theorem: a Pyrrhic victory

For grand theses, and especially for ideals (here, of theoretical understanding), a broader sense of 'refutation' is philosophically appropriate:

> when the ideal is realized, the realization is found to be unsatisfactory.

To spell it out: the realization is found to lack (or have) certain properties that, as can *now* be seen, had been tacitly assumed to go with (or against) the ideal. In this way, inspection of a realization can identify *tacit assumptions behind the ideal*; which constitutes a refutation in the broader, second sense (by a Pyrrhic victory, as it were).

An example of such a refutation of the formalist thesis is Gödel's other very well-known early result: the *completeness* of Frege's formalization for elementary predicate logic, but now in the sense that for each formula F, *either* F is a formal theorem *or* F is not true in all possible worlds.

Viewed dispassionately, this result does not at all give a privileged place to Frege's rules. It just shows that they are sufficient-in-principle (as Kant liked to say)[27] or, more soberly, sufficient to generate the logically true formulas in the formalization.

On the contrary, by using the concept of logical truth, the result draws attention to quite different possibilities of proving logical theorems; specifically, the possibility of drawing on knowledge (if not of all possible worlds, at least) of many corners of our world; in other words, the possibility of proving logical theorems by so-called *logically impure methods*.[28]

In point of fact such impure methods have been used increasingly since the formalization of pure logic more than one hundred years ago, and especially since Gödel's result sixty years ago; incidentally, this is often done by people totally ignorant of the formal rules (so that it is wide open in which way the conviction carried by their proofs can be realistically related to the formalization).

---

[27]The ethereal business of *possibilities-in-principle* has been most prominent in foundation; not only in the writings of Kant, but also of logicians like Russell, who described *Principia* as 'a parenthesis in the refutation of Kant'. Many of Kant's observations on reasoning apply impeccably to actual phenomena, but are false if interpreted as *needs-in-principle*. Thus, contrary to Kant, appeal to geometric experience (especially, visualization, also called *Anschauung*) is not needed in principle for mathematical deductions; but it continues to be used widely, and to good effect. More specifically, Euclidean geometry does not have the privileged place that Kant, taken literally, gave to it (nor, of course, for physical space near massive bodies; and it is not the geometry of visual space either); but, to this day, mathematicians continue to think in Euclidean terms also when defining non-Euclidean spaces.

[28]Specialists will think here of theorems that are or can be formulated in Frege's logical language (for example, about ordered or real closed fields), but are proved by topological methods.

Sensationalism and utilitarianism

Before turning to more positive aspects of the two refutations above by the incompleteness and completeness theorems, a couple of comments seem in order.

The first is general. The proofs of both theorems are pearls of logic. But *not both results can be sensational!* If the completeness of a formalization for mere elementary logic is a sensation, then incompleteness of Higher Arithmetic is not.[29] Historical counterfactuals aside, one hundred years ago it would have been fitting to give pride of place to the completeness theorem, bolstering up the then-tentative project of formalization, with the incompleteness theorem ratifying (formally, as it were) the then general distrust of that project.

The second comment concerns the open secret that, outside mathematical logic, *neither of these two results turns up in the ordinary mathematical literature*. There is no conspiracy against them. As stated, they just have not found a use; fittingly, in view of the fact that they are tailor-made for (refuting or supporting, no matter) a *refuted thesis*.

This is the situation considered at the outset, with all its problems; in particular, the one of finding some sober use for the tools employed. Especially when, as in the present case, the tools are of obvious 'raw' interest, the principal obstacle to solving those problems is *blindness* to them; in particular, the illusion that pretty mathematics must 'somehow' have already solved those problems. *Tractatus* 6.21, about mathematics not expressing any thought, is surely literally false. But, equally surely, *mathematics very often leaves questions open that require more (demanding) thought than the mathematical solution*.

Shifts of emphasis: what more do we know from formalization?

Obviously this question does not even arise if the formalist thesis is elevated, as indeed has been done, to the doctrine that *there is nothing (precise) besides formalization*.

The question is less innocent than it may look. For effective contributions to some particular area the additional knowledge (in answer to: 'What more …?') must be expressed in terms used in that area, and is thus liable to require familiarity with it. The following two general points have a larger market.

First, there is the matter of *choosing a formal system* rationally. By Gödel's incompleteness theorem, the only general idea for reducing the arbitrariness of

[29] For example, a couple of years before Gödel's discovery of incompleteness Siegel [1929] and Weil [1928] discounted, in effect though in different terms, the possibility of a complete formal theory for diophantine equations. In fact, their equations had just two variables, respectively for integral and rational solutions. Weil's equations were even only of degree four.

The question of whether or not there is a complete formal theory for these special cases is still open. Incidentally, the two papers referred to were immediately famous.

such a choice (namely, completeness) is not even 'in principle' available for systems of arithmetic. One of the current favorites for answering the question above provides a better (and sometimes even practical) idea for a choice, based on the *rate of growth* of those functions that can be shown in T to solve suitable[30] problems P (provided one wants to know about such things). Here, P *comes first* and T *is a tool*, not a 'foundation'.

The second point concerns *complete formalizations*, such as (Frege's) logical rules. To repeat what cannot be repeated too often: unless formalization is required as a matter of doctrine, the question 'What more ...?' is just as hot here as for incomplete systems. The answer is simply: *a new description of the object involved*; in the case above, of the logically true propositions (in Frege's language). For effective knowledge *this description competes then with others*, such as the 'impure' kinds in the last section. Logicians think here of the competition between descriptions in model-theoretic and diverse proof-theoretic terms.

There is an obvious parallel here with the Pythagorean thesis, specifically with the use of irrationals even when they could be avoided-in-principle. But a more significant element of the parallel is this: the proofs of completeness and incompleteness are mundane enough to remove any sense of awe inspired by the refuted formalist thesis, and leave us free to examine some *assumptions behind* the latter.

What is so wonderful about formalization?

People have been thrashing about for an answer. Only one will be considered here. It is the tacit assumption of some ethereal need (here, satisfied by formalization) for an *ultimate norm of precision*; a tacit assumption popular not only in the foundations of mathematics, but almost throughout all Western culture.

Presumably according to the books they read in their teens, particular authors writing on such norms refer to:

- the finiteness of formal objects
- their spatio-temporal character
- more generally: their public character
- going the whole hog: the idea that the thought processes themselves are formal (i.e., mechanical in the sense of the perfect computer), in which case only formal rules can be unambiguous.

[30]'Suitable' means in practice that P has the form $\forall\exists$, and that incompleteness applies even if all true purely universal propositions are added to T as axioms. The bounds are established by so-called consistency proofs for those (necessarily incomplete) T.

All this would carry little weight without the master assumption that there are *genuine doubts about the precision or reliability of principles currently in use*. (Viewed this way, all those little paradoxes are a godsend for this assumption.)

Now, if the principles in question are in fact 100% reliable, then doubts about them are dubious (which doubts can be, just as much as can assertions). The privileged place given to doubts (for example, by Descartes) looks very much like other pious conventions that are only too familiar. But it pays to be more specific.

To put first things first, there can be perfectly proper doubts, even about principles. Thus, not so long ago, (antecedents of) those now current about sets were problematic; the problems were solved by saying out loud which sets were meant (and *not* by putting principles into formal dress which, after all, Frege did in his *Grundgesetze*).[31] Though (occasionally) obviously problematic principles are investigated, most often those proper doubts are about (the probability of) *incorrect applications of correct principles*. In this case preoccupation with reliability-in-principle (that is, reliability of principles) simply *distracts from the dominant factor* (here, dominant source of error).

Is this factor a foundational concern? According to a principal tradition of the subject, it is not. In this case *the topic of reliability has been discovered not to be primarily foundational* (by the way, not necessarily a comedown; see the discussion on p. 143). A discovery of the nonfoundational character of a topic (here, of reliability) is to be compared to discoveries in the natural sciences; for example, of the gravitational or magnetic character of some phenomenon, say, near the surface of the earth (in other words, whether the dominant force, if there is *one*, is gravity or the magnetic field of the earth). Similarly, in note 28 concerning 'impure' methods, problems stated in logical language but solved by topological methods are thereby discovered to have topological character.

As in other scientific experience, research has produced ways both to cope with such delicate points as mixtures and, above all, ways to recognize when *enough is enough*: that is, enough for using such characterizations without looking for new evidence, or even referring to the old. (Our existing knowledge of those characterizations is thus treated as *a priori*.) Given the level of generality of all this, it applies of course to norms as well; here of precision, but surely also in practical life. For the literal sense of 'philosophy', recognizing when enough is enough has always been a central concern. Thus in the *Metaphysics* (Γ, 4, 1006a, 6–9) Aristotle relied on good breeding. The following couple of points help too:

- *There is extended experience*, which may but need not confirm ordinary practice.

  For example, as already explained, experience has rehabilitated our ordinary 'norms' (or, more simply, checks) in the case of logical reasoning, as opposed

[31] Cf. note 26 on completeness.

to the demands of so-called formal reasoning.

In contrast, after the discovery in the last century of so-called abstract reasoning we have never looked back. Specifically, we express in axiomatic terms what we feel to be essential about an argument. For example: in elementary number theory, analysis in terms of abstract finite groups is used; we do not isolate only, say, the 'numerical content'.

- Particularly significant for the present subsection: (a good deal of) *the foundational literature obscures the consequences* of such extended experience, including logical experience.

  Thus, it wails about 'far-reaching reductions' (of mathematics to formal systems) that are (allegedly) lost by the refutation of the formalist thesis, apparently without remembering the fact that those would-be magic reductions are *already available* in substantial areas (such as elementary geometry or logic) but *not used*: least of all, as a norm of precision.

Readers familiar with foundational (bad) habits surely know much more along these lines. But the snippets above are enough to show that *the topic of reliability is not a rewarding market for formalization*, and certainly less so than the answers above to the question 'What more ... ?'.

This suggests also a genuinely philosophical conclusion: it is simply not plausible that this particular topic will get very far without *closer attention to thought processes*, which (by p. 151) are neglected in the pale formal picture.[32]

Short answers to the initial questions

Gödel's results and the refutations they provided of the formalist thesis were involved in locating and examining assumptions behind the latter, such as dubious doubts or doctrinaire norms of precision. Admittedly, those assumptions can be and were questioned before settling or even formulating the thesis. But the refutation and, particularly, its elementary character (stressed in Section 2) can help establish *a sense of proportion for the examination*: above all, by eliminating undue worry about not having grasped the full inwardness of the thesis.

As to the *tools*, sketched in Section 2, the situation is not too different from that of the Pythagorean thesis; except that 60, not 2500, years have passed: this

[32] *Warning.* The heading 'length of proofs' of Gödel [1936] suggests consequences for understanding thought processes; but the formal picture is just too pale to support such colorful interpretations! A short description of a long formal proof, especially with underlining of its memorable parts, is in fact easier to process than a relatively short formal proof.

In a similar vein, the process of discovering a *new* axiom to prove some given result is often not only less demanding than discovering a proof of it from *given* axioms, but has a similar flavor. Though such a proof by given axioms can in principle be found mechanically by trial and error, in practice it is not so found (cf. note 27).

affects both the choice of problems and the methods of solution.[33] But Gödel's proofs remain *memorable samples* in the second subject.

For the record, I still find his direct and self-assured style in the early works appealing, compared to the extremes of pedantry and sloppiness rampant at the time.[34] The so-called substance, of proofs and results, has been superseded, in accordance with Buffon's *ces choses sont hors de l'homme* (those things - meaning results - are impersonal), which preceded his much better known pronouncement (on August 25, 1753): *le style est l'homme*.

So much for logic. But logicians do not live by logic alone, not even intellectually. To me, a principal reward for refuting (rather than ignoring) extravagant theses is in that broader area, as follows.

Refutations: For a better quality of life

Material pollution presents a similar choice. In the austere 50's, those complaining of noise or other pollution were told to ignore it. Progressives, always characterized more by temperament than by specific views, had only contempt for the (to them, obviously absurd) Air Force general played by George C. Scott in the film *Dr. Strangelove*, who worried about chemicals in drinking water. And (for all I know) they may have a point, if selection by resistance to pollution helps the species flourish in our cold, unfriendly universe. But for some (of us) it was not so easy to ignore the pollution.

Similarly, (we) logically sensitive souls do not so easily ignore the logical atrocities in (Hilbert's) presentations of the formalist thesis, especially when research stagnated and the claims inflated. The pollution was all around, spread by a band of the faithful who found the presentation so congenial that it matters little whether it helped form or 'only' consolidate their views.

Here are a couple of samples:

1. The laws of thought are mechanical, and *non ignorabimus*.

Actually, the idea was that those laws were already formulated in Hilbert's systems, and that we shall (want to) know the answers to all the problems formulated there.

As a kind of fallback, Hilbert had a weaker (would-be cute) meaning for *non ignorabimus*:

2. It is consistent to assume that every problem can be solved.

---

[33] Thus there is a greater difference between the proofs of Mordell's conjecture (a recent highlight) and the irrationality of $\sqrt{2}$ in the theory of diophantine equations than between, say, 'double' diagonalization in so-called priority arguments of Recursion Theory and simple diagonalization in Section 2.

[34] And also, for that matter, to some constipated parts of the editorial material in the volume (Gödel [1986]) collecting those works.

In other words, to assume for every proposition P that either P is a formal theorem (of the 'foundational' system) or its negation is.

Gödel's results refute 1 and 2 conclusively and most elegantly. For 1, this is clear without further analysis. For 2, a corollary is needed that is valid only for a more special class of systems, but certainly for all current at the time (cf. the discussion after 7.2.9):

3. It is even consistent to assume that every proposition can be proved,

from which 2 follows trivially. We breathe more freely.[35]

This relief is hardly necessary in practice; not, for example, for the robust among us who ignored the formalist thesis in their work even if they gushed about it in private. Philosophically, the relief is *only* a palliative: it just distracts from the source of the pollution, the assumptions behind the thesis. But (as seen by the subsection on p. 155) cleaning this up is a costly business, requiring more capital (scientific experience) and labor (sustained reflection). Palliatives have a wider market.

To pursue the parallel a little further, but also to end on an irreverent note, it may be remembered that similar technologies can create as well as clean up pollution, and that quite often the same firm manufactures both types. In the 50's, Gödel himself (as so often in this chapter, in effect, not in these words) modified 1 above by adding *nil nisi externum* (to *non ignorabimus*: what he actually asserted was that we know everything about our own constructions). Then his incompleteness result implies the hot news:

> Either mind is not mechanical, or the natural numbers[36] are not our own construction

For the record, I don't choke at that emission, but almost cherish it, together with memories of many conversations with him, spiced with spontaneous twists of a similar flavor.

But all in all it's fair to say that elementary metamathematics and, particularly, *Gödel's contributions are good value*; at least for those who, for one reason or another, have learned about them. What seems to me wide open is *how effective those contributions are for conveying that which is of general interest in*

[35] Hilbert continued to repeat his mantra even after Gödel's incompleteness paper, and never saw that 2 follows from 3. This is not merely a personal, but a metaphysical affront!

What is so offensive here is that Hilbert had been repeating his 'grand' pronouncement with the conviction that 2, being so deep, could never be refuted; obviously never dreaming that it could actually be proved! And trivially, to boot.

And it would be absurd to split hairs by saying that 2 is not proved by the means of proof for P considered, since Hilbert's many false conjectures are quite implausible once the relevant difference in methods of proof is remembered at all.

This is the kind of stuff of which metaphysical anger is made.

[36] Nor, *pace* Kronecker, the reals.

them; in particular, that which is genuinely generally accessible (effective compared to metaphors from more widely available knowledge, as in the Pythagorean thesis). The matter is wide open because my skepticism seems to be shared by others in the trade, and so existing (unsuccessful) attempts at popular exposition have been perpetrated by the uninformed,[37] in line with the proverb:

fools rush in, where angels fear to tread.

[37] At least in my view, Smullyan's *What is the name of this book?* does not belong here at all. (For one thing, the author is well informed.) It contains a remarkable collection of puzzles, puns, and other *jeux d'esprit* in which their logical aspects may fairly be said to be dominant. They are understood by use of propositional (or, at most, propositional provability) logic. But, realistically speaking, this recreational corner of experience seems to me to be of very specialized interest; for example, more so than the broader matters in the last chapter of his book, with more jokes, but without much (relation to any earlier) point. True, Smullyan's fancies are no further from ordinary linguistic experience than, say, Galileo's bags of feathers falling behind leaden spheres are from ordinary mechanical phenomena. But they bring (at least to my mind) a jolly wake for a defunct two-thousand-year-old tradition, that of the Liar, rather than a first step to higher things like celestial mechanics.

Chapter 8

Gödel's Later Works

After his completeness and incompleteness theorems at the ages of 24 and 25, Gödel's fame was consolidated by his work in set theory, on Cantor's continuum hypothesis, published when he was 32. Gödel himself stressed, in the titles and introductions of those papers, the connection of his results with the foundational views and controversies in the first quarter of this century. Briefly, Gödel's results refuted or deflated those grand schemes, without however putting nothing comparable in the place of the soundly discredited and, in practice, (properly) ignored foundational views.

The mathematical development of his work over the last sixty years have involved (in effect, if not by intention) a search for local uses of the general notions and ideas, whose significance had previously been so grossly misjudged in foundations.

Gödel instead devoted a good part of the remaining two thirds of his working life (till his illness, at the end of the sixties) to use his scientific successes in a quite different direction, of interest not only to mathematics but also to the natural sciences (and, of course, also independent of the discredited views). Specifically, in retrospect, he saw his successes as special cases of a general scheme, to be applied to suitable traditional philosophical notions and issues: by making them precise, one arrives painlessly at fruitful concepts, correct conjectures, and generally easy proofs.

Naturally, Gödel's later reflections are not presented as formal theorems but in essays, relatively short notes, and programmatic lectures. Far from being the result of some kind of premature senility, at the age of 36, Gödel's new interests were simply ahead of his contemporaries: without exaggeration, more so than in the thirties. In particular, he saw the need for some analysis of his successes to forestall a useless kind of 'revisionist' history.

The two trends in Gödel's writings can be related, respectively, to (the following) two traditions.

---

[0]Originally published in *Notre Dame Journal of Formal Logic*, 31 (1990) , as 'Gödel's Collected Works, Volume II'.

Metamathematics

Gödel's work from the first half of his life (till his mid thirties) is squarely in one tradition, going back to Hilbert's *Foundations of Geometry*. Through the latter not only (young Gödel's kind of) mathematical logic, but generally the axiomatic method in its modern sense was put on the map.

Hilbert's agenda at the turn of the century[1] was:

> to replace the often tedious literary forms of logic chopping in the foundations of mathematics by those of mathematical logic, with emphasis on the idea(l)s of *consistency*, *completeness* and *decidability*.

These household words were applied by Hilbert to formal objects, defined independently of any further interpretation.

The scheme recalled (the best of) rational mechanics beginning in the 17th century, which both replaced logic chopping concerning matter and motion, and gave scope to the armchair (applied) mathematician. Results in the literary forms of mathematical logic were expected to 'speak for themselves' too.

Gödel's contributions to this line of business ([1930], [1931], [1938]) remain (among) its most memorable successes.

Logic chopping

The later part of Gödel's work (from [1944] on) belongs to an older tradition, variously known as logic chopping or exact philosophy (in the academic sense of this word); the latter, in turn, derives from the heroic perennials familiar from philosophy in its more popular sense.

Gödel's agenda for this period became:

What is lacking (in the earlier work)?

Specifically, 'lacking' if logic is to be a science prior to all others, which contains the ideas and principles underlying all sciences (cf. the opening of [1944]). In less academic terms, logic is to be *a seed from which the tree of knowledge grows*, and the logical order of priority is the corresponding (tree) ordering.

Roughly speaking, Gödel saw the best prospects for this idea(l) in going back to the older tradition; in particular, to elements that are (as it were) prematurely disregarded in Hilbert's scheme. Gödel presented such elements at various levels of sophistication, both in mathematical and other literary forms.

Our agenda for this chapter will be:

to balance the account.

[1]Not to be confused with his later program, in which so-called *finitist* parts of (what had come to be called) metamathematics were privileged (with the usual consequences of such practice).

More precisely:

- The *internal coherence* of Gödel's view is emphasized, with some formal consequences of those neglected elements from the last 40 years.
- There are (and this is, of course, of broader interest) reminders of *genuine alternatives* to that heroic perennial of knowledge growing like a tree from a seed;[2] made particularly memorable, I believe, by *contrast* with (Gödel's) pursuits of that idea.[3]

The first six sections of this chapter deal with isolated significant topics prompted, or at least suggested, by Gödel's later works. We then conclude, in Section 7, by going back to the begining of Chapter 7 about 'principal questions, which could not have been answered at the time', but now applied to the entirety of Gödel's work.

8.1 How Adequate Are Those Would-Be Fundamental Metamathematical Notions? ([1938], [1939], [1940])

The adequacy meant here is a common-or-garden variety; viewed

- neither as a mere matter of principle (of being ever or never adequate), since this is not in doubt
- nor (of course) according to the heroic ideal of logic on Gödel's agenda, since this is on trial here.

We use as a sample the particular notion of (relative) *consistency* prominent in Gödel's own titles for his work on the continuum hypothesis (cf. [1938], [1939], [1940]). The details are familiar enough to rely on the following reminders:

- First, it is a common place that the work in question is more adequately described by *different labels*; for example, 'inner model constructions', 'absoluteness' (albeit relative to the ordinals), or 'conservation'. Pedantically, results (stated in such different terms) are obtained as corollaries to the work by means of general logical theorems, which correspond to so-called abstract nonsense in current mathematical jargon. Then, contributions of the work to effective knowledge are discovered to follow from those different descriptions, but not from Gödel's titles.

[2]Pedantically, provided the mind is permitted to use its natural capacity for processing data in parallel (not only 'systematically').

[3]Here, as elsewhere, the general idea of this theme (on alternatives to the ideal of a tree of knowledge) is abstractly so familiar as to be banal. But also, as elsewhere, it is in conflict with venerable ideals; for example, of a 'definitive evaluation'. More information on that conflict (where and how it arises) is in the chapter, especially in Section 7.

- Secondly, and this keeps the first point topical: the later literature (for example, relating axioms of infinity and determinacy) continues to rely on relative consistency as if it were an adequate description.

This alone would be enough to illustrate vividly Gödel's reservations about following Hilbert's ideal (of course, not despite, but because of the fact that they conflict with Gödel's own practice in his 'salad days'). A closer look underlines the point as follows.

Finite axiomatizations. If relative consistency (and the metamathematical methods of proof used) were the first order of business, and such alternatives as (inner) model constructions only a means, then finite axiomatizations would have compelling consequences. For example, mere validity of those constructions ensures a relative consistency proof by quite elementary methods, while (in general) this is not so (even) for r.e. systems of axioms. But the trade dealing in relative consistency results does not generally regard finite axiomatizations as so privileged.

The points above, about labels and formal incongruities, would be as irrelevant as the name 'rose' in horticulture, if mathematical logic could be viewed as a similar trade with an established market. Such a view may well apply, by now, to certain parts of logic and their markets (dealing with finitely generated groups, finite fields and a few more). But those markets had only just been discovered at the time of Gödel's agenda, and certainly are not prominent in it (nor in this chapter).

Almost as a corollary, there is a *positive side* to all this (at least for those readers who are looking for object lessons from experience with problematic aspects of the logical enterprises). This need not be a parochial interest, since those aspects are (as it were) chemically pure specimens of those found throughout the *commerce of ideas*.

Digression on the commerce of ideas (optional)

Throughout this chapter, the metaphor of 'commerce of ideas' serves to underline aspects (in the 'knowledge business'[4]) central for our agenda; more fully, aspects which are neglected by venerable ideas, but correspond to household words in ordinary commerce. *Samples*:

- Above all, there is so-called *marketing*; but not confined to door-to-door salesmanship (though this too has its parallels in the commerce of ideas). It requires the *discovery of markets* or (at least) gaps in the market, for products over and above their (mere) legitimacy (that is, absence of fraud).

  In our commerce, legitimacy corresponds to the likes of existence, truth or consistency.

[4]A literal translation of Kant's *Vernunftgeschäft* (A 724), but with a twist in meaning.

- In a somewhat different vein: there are *different priorities* for those who labour (say, on the shop floor, or on arranging shop windows) and, more broadly, for trade unions and management of corporations. These, in turn, combine against those concerned with *correcting traditions* common to the 'supply side' (for example, by philosophies of mergers and unbundling).

  More topically, failures of so-called *Big Science* often recall those of big business conducted according to idea(l)s familiar from (and well established in) experience of craftsmen and village grocers.

These simple home truths, concerning the emphasis on different aspects, are in conflict with erudite, traditional (here, economic) *isms*, for which certain aspects alone are privileged.

For effective use of those home truths, one should realize that generally, and especially for our agenda, the appropriate pay-off in the commerce of ideas is *not monetary* (which is relatively prominent in ordinary commerce[5]). A venerable alternative payoff, just *knowledge*, is memorably described in Goethe's letter of 18 June 1795 to Humboldt (by reference to his experience with science; according to him, in contrast with works of art).

Concerning my own use of the parallel (not only in this chapter), I have nothing to say about changing the world (here, of our commerce), only about interpreting it.[6]

8.2 Logic Chopping: Elementary Samples ([1944])

Whatever the literary defects of the essay [1944] on Russell's mathematical logic, the topic itself was (and remains) perfect for anybody who has anything to add to the logical literature.

[5]But recall the striking exceptions in the pretty theory(!) by Kreps under the heading 'bounded rationality' (tacitly, in Simple Simon's sense of 'rationality').

[6]These words come from one of Marx's (many) popular dicta:

> The philosophers have only interpreted the world in various ways; the point is to change it.

This is not my point. I do not see that I know enough (or even that enough is to be known) about *predicting* the world (or history) to make Marx's philosophy above even remotely plausible (I mean, globally: one's backyard is another matter, as Candide might have added.)

For the record, ever since my teens I have viewed that dictum (of course, not as the single most distinctive, but) the shrillest element in (what I have learnt of) Marxist thought. Viewed this way, and contrary to Engels' quote from Marx:

All I know is that I am not a Marxist,

not only he was very much a Marxist, but so are many so-called anti- Marxists and specially ex-Marxists.

Passing thought. Temperament is involved in my particular brand of anti-Marxism.

Russell had an exceptional talent for formulating memorably almost any thought that could cross anybody's mind, and used it freely. So he has left us plenty of pegs (one of them being his paradox), on which to hang (actually, often salutary) additions; obviously, to consolidate, not to introduce broad ideas. In the present case the broad idea involved is of course this:

> logical results, which do not speak (well) for themselves, may do so (better) when supplemented by some traditional logic chopping.

After nearly a century of experience it is fair to say that some of the items thrown up by people thrashing about for something to say about that paradox, are more rewarding here than the paradox itself;[7] let alone, Gödel's oft-quoted, would-be dramatic comment:

> By analyzing the paradoxes to which Cantor's set theory had led, [Russell] freed them from all mathematical technicalities, thus bringing to light the amazing[8] fact that our logical intuitions ... are self-contradictory.

---

[7]In the body of this section, the focus will be the *logical* interest of the paradoxes (more fully, the interest of their logical aspects). For contrast, here and in notes 8 and 10 the focus is different: on alternatives to *assuming* that the interest of the paradoxes is primarily related to their logical aspects.

For one thing, paradoxes (in other words, refutations of familiar, implicit or explicit, more or less thoughtless assumptions) are commonplace; especially, around (ab)uses of the definite article, as in Russell's 'the class of ... ', but also in 'the greatest integer'. (In the latter case, 'the' is misplaced both for the usual sense of 'integer' and for, say, integers mod p, when the usual order is incompatible with the ring operations.) In short, the logical interest is dubious.

Secondly, inasmuch as the kind of logic around the paradoxes is typical, *logic just isn't mathematics*. More soberly, notes 8 and 10 below draw attention to other aspects of phenomena around the paradoxes. So it is (just) simple-minded to *assume* that their mathematical aspects are rewarding (let alone, decisive). This twist is two-fold:

- For one thing, it reminds of how marginal (especially, higher) mathematics is; not only here, but generally in the broad commerce of ideas.
- Secondly, and on this score unions and management close rank across the board (in the world of academic disciplines), *traditions of a trade are not sacrosanct*; specifically, not those of academic philosophy, which claims to know the extent to which the *raw* potential of a commodity is enhanced by its own resources (here, potential of the paradoxes).

[8]What is regarded as *amazing* is obviously, at least partly, a matter of temperament. Less obviously, this applies to 'our' intuitions, for example when a solitary temperament does not get an opportunity to compare personal impressions with wider experience. (Of course, dimwitted people have little chance of benefitting from such experience even if they have a different temperament.)

Once again, an abuse of the definite article (here: in 'the problem of paradoxes', when the latter have, obviously, many different aspects) helps to bring in tacit assumptions about 'the' solution. For example, that amazing facts *must* have great inwardness; never mind, whether as sources of profound wisdom or as the work of demons (or in another quarter, witches, so to speak as a matter of sexual preference). The tacit assumption is that they are *not* simply blind

Frege had written down the following axiom scheme for (his idea of) predicates and classes: for every predicate P,

$$\exists y \forall z [z \in y \;\Leftrightarrow\; P(z)].$$

For $P(z) = \neg(z \in z)$ this becomes a special case of

$$\exists y \forall z [R(z, y) \;\Leftrightarrow\; \neg R(z, z)],$$

which is impossible for arbitrary (binary) relations R, since it implies (when $z = y$)

$$\exists y [R(y, y) \;\Leftrightarrow\; \neg R(y, y)].$$

Frege's scheme, which had been totally ignored in scientific trades of the commerce of ideas before Russell's paradox, gained a little notoriety by it. More high-minded (and less experienced) traditions assumed (as it were, as a matter of course) that there was a specific error (to understand) in Frege's scheme; so to speak, 'the' root cause of its failure (in some tree of ignorance). For example, the following *demons*:

infinity, self-reference, impredicativity.

But are they really demons?[9] And, whatever the answer(s), just where are they in the result above?

---

spots.

Needless to say, matters of temperament and background are often difficult to disentangle (since one may not wish or be able to manipulate either). But in some cases such aspects of background as general education are more visible, and thus easier to document and convey, even to solitary temperaments. 'Convey', not discover, since experience of men and events (not available to solitary temperaments) is needed to use the historical record sensibly, or even to have an inkling of possible snags.

In the case of Gödel's 'amazing fact', there is the historical record of Cantor's and specially Frege's complaints about being ignored. In other words, for philistine intuitions topics like 'being' (Gödel mentions also 'truth', 'concept', 'class') weren't even candidates for (mathematical) study.

[9]For one of the parochial concerns in this chapter (the growth of knowledge of sets), infinity, self-reference and impredicativity are *straws clutched at* by people thrashing about for something to say after Russell's paradox. Preoccupied with the latter (and, generally, with Frege's naughty axiom) they neglected, by and large, more rewarding questions about sets; for example, about the replacement axiom and (some of) those called axioms of infinity (cf. section 4).

Obviously, straws have properties too, and so can be perfectly legitimate objects of (precise) study (as in predicativity on p. 172). Clutchers differ in style and power, and so they too can be subjects of musings (for example, in note 10 on Weyl).

For wider concerns, one of the more glamorous candidates for a straw among heroic perennials is the following (allegedly fundamental) 'opposition':

Objective and subjective (knowledge).

Just as in the case of sets above, (scientific) experience presents (in fact, many) questions *around* the glamour issues above. Certainly, for the broad philosophy (in the popular sense) of this

- *infinity*
  Gödel does not go into this. Readers should recall 'the' barber (in some hypothetical finite village), with its embarrassingly blatant abuse of the definite article.[10]

- *self-reference*[11]

---

chapter, the focus on some opposition between objective and subjective knowledge, which is of course high-minded, is above all simple-minded (and probably even below any threshold of informed discussion).

The general idea of clutching at straws is perfectly commonplace. What is stressed here is that this idea applies to (and may be even adequate to specify principal errors in) would-be sophisticated enterprises.

Disclaimers (again in terms of a refrain of this chapter). The idea of *clutching at straws* is not thought of (here) as a seed from which a tree of knowledge grows. So, as matter of practical politics, it (is of course a legitimate topic of analysis or what have you, but) is not *assumed* to be a rewarding object of recondite or otherwise extended study. Metaphors for alternatives abound:

- At one extreme there are *pegs* (with luck, not straws) on which diverse items (of knowledge) are assembled, which would otherwise float in thin air, and be inaccessible.
- At another there are (mathematical) *attractors* in chaotic dynamics; so to speak, steady states presenting a more or less adequate idea(l) in turbulent surroundings.

[10]For the record, Weyl did not remember this in his comment on the paradox in his review [1946].

Incidentally, Weyl had a choice between the literary forms of ordinary mathematics (of which he was a master), of mathematical logic, of metaphors and similes, and others from the ordinary literary tradition. He chose the latter, calling Gödel's essay [1944] 'the work of a pointillist'. This simply does not fit what he wanted to say. A pointillist worth the name uses points, which are individually of no weight, to produce the impression of a recognizable object with global features. But then Weyl goes on to say, of course in different terms, that he had not derived any global idea from Gödel's points. (By what said above, he had not perceived their individual weight either.)

An alternative available at the time was to use details of the constructible hierarchy, if only for a metaphor (cf. footnote 11 of Gödel [1947]): Weyl's only (but well-known) 'intervention' in the foundational debate was his emphasis on the first level of ramified analysis, obviously related to the ramified 'theory' of types up to ω considered by Whitehead and Russell. Gödel's essay [1944] emphasizes the (close) connection to the constructible hierarchy.

[11]*Reminder* (with a shift of emphasis away from the 'self' in 'self reference'). As stressed in the subsection on p. 144 of Chapter 7, contemporary mathematics provides a good deal of effective knowledge on *representing* one kind of thing by another (so-called choice of data). Such representations are then used to *refer* to those other things.

As in the general topic of representation (including reference), it goes without saying that its mental aspects (for example, intentions involved in reference) strike the mind's eye most vividly. But it also goes without saying that this fact does not guarantee that those aspects lend themselves well to theory (tacitly, as always: by anything remotely like current means). Of course, we know a lot about them; it's just not theoretical knowledge.

For reassurance: inanities about reference (in so-called theories of meaning), which 'identify' it either with *mental aspects*, or the *software* (that is, mathematical aspects of representations), or with the *wetware* (data processing in the brain) will fall into place in Section 7. It would be

In the result above, this is the specialization of z to y. Gödel ridicules the assumption that self reference is a demon, by a more subdued reminder:

> there doubtlessly exist sentences referring to a totality of sentences to which they themselves belong as, e.g., the sentence: 'Every sentence (of a given language) contains at least one relation word'.

For the record, (by temperament and) by contrast with my earlier experience of traditional logic chopping, I (continue to) find this example of the genre compelling.

- *impredicativity*
  This is of course related to the broad topic of self-reference since, pedantries aside, it is about defining an object by reference to a totality containing it. So what?

  Of course, such definitions do not fit the metaphor of knowledge growing like a tree: *any cycle vitiates any* (tree) *order*. For our agenda, where that metaphor is on trial, it is an open (and main) question whether this conflict is evidence for or against it (and such other items on Gödel's agenda as his logical order(s) of priority). But, for those of us grateful for small mercies, it is relief that he raises at least the less demanding question:

  > Just where is there any impredicativity in Russell's paradox?

  The predicate used ($\neg z \in z$) contains no quantifiers, which could be said to 'refer' to a totality. Gödel focuses on the *range* of the variable z itself; in other words, something left implicit in the logical notation.[12]

  Here too Gödel's (compelling) point, which does not use formal constructions, fits the academic tradition of philosophy; as, for example, in Kant's *aperçu* (A713):

  > philosophy analyses and mathematics builds up concepts.[13]

A sound perspective (here, on logic chopping) requires the following elementary distinction, and above all attention to its neglected consequences.

The samples above correct errors; both in the traditional literature and on the would-be 'purely' mathematical side. It is a common place that it would be merely

premature to agonize over them here.

[12]Incidentally, Cantor's criticism in his review [1885] of Frege's *Grundlagen*, more than 15 years before Russell's paradox, also focuses on the indefinite range, albeit in different (medieval) terms.

[13]The *aperçu* is, as so often, useful provided only it is *not* taken literally! There are plenty of mathematical analyses of concepts, and there were some at his time, too.

high-minded (and thus liable to be simple-minded) to assume that, in some given area of knowledge, the correction of errors *must* contribute rather than simply distract; more generally, that extended logic chopping must (help to) contribute (in A716 and A718 Kant ridiculed the assumption by the example of geometry and analysing the concept of triangle). One simply may do better by making a fresh start. But it should also be added (and this is illustrated throughout the chapter) that errors are not automatically corrected by contributions (in the area considered): the latter do not generally 'speak for themselves'; not enough for the corrections in question.

8.3 Absolutes: a Top Priority in the Logical Tradition ([1946])

In [1946] Gödel emphasizes absoluteness of such venerable notions as *definability* and *provability*. This is in sharp(est) contrast with Hilbert's scheme, which applies them only to some formal system or 'language' (leaving open, or at most paraphrased, the adequacy of such choices; cf. p. 258 of Chapter 11 in the case of completeness).[14]

Gödel introduces the topic by reference to (mechanical) *computability* and to its analysis by Turing. Quite explicitly, but so innocently that this too sounds absolute, he assumes that the (mathematical) concept of *recursiveness* itself derives its importance from its absolute character; that is, the independence of this definition (of computability) from any particular formalism:

> In all other cases treated previously, such as demonstrability or definability, one has been able to define them only relative to a given language ... For the concept of computability, however, ... the situation is different.

Now, computability evidently involves both definability and provability (by routine verification). So, after Turing's success, Gödel proposes to go the whole hog and analyse the absolute idea(l) of the two perennials above:

> This, I think, should encourage one to expect the same thing to be possible also in other cases (such as demonstrability or definability).

Incidentally, Gödel's flourish in [1946] about Turing's analysis being a first in human history requires a *correction*: classical propositional logic too is (not only deductively, but also) functionally complete; in other words, the adequacy of its

[14]Model theory, central to mathematical logic (today, not 45 years ago), also concerns aspects ignored in Hilbert's scheme (models or structures), but not those stressed by Gödel. Thus model-theoretic definability is generally relative to a language, and hence *not absolute* (in the sense used by Gödel here).

formalism is established. For our agenda, the correction is less innocent than it may sound; actually on two scores:

- it takes some of the glamour out of any logic chopping that may be used in establishing such absoluteness
- it shows that doctrinaire (formalist) objections to Gödel's proposal (namely, that the notions involved are 'essentially' relative to some formalism) are below any threshold of informed discussion.

Once grasped, these insights (may help to) shift the emphasis to (naturally, more demanding) matters above threshold. For example, incompatibilities between the aspects (of definability and of provability) required by the logical orders of priority and by effective knowledge. But this is quite another (and outside the logical trade) only too familiar story.

To end on a positive note, here are a couple of *cheerful news* (from [1946]):

- For one thing, at least for suitable variants of absoluteness and suitably adjusted expectations (cf. the next subsection), Gödel's remarks on definability (say, of sets of natural numbers) have been checked. For example, for such variants it has been proved that only countably many are so definable (and, naturally, there is no enumeration that is so defined).
- The second cheerful item is Gödel's own blithe disregard for his own idea(l) of absoluteness when he goes on to muse about (the possibility of a complete set of what are now called) axioms of infinity. Here, quite cheerfully, completeness for the *ordinary language* of set theory is meant:

  > 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.

  It is a matter of temperament whether, like (the older) Gödel, you like to 'aim for the stars'. Ever since my teens I have been told that, in this way, 'you may hit the moon' (and even long before there were astronauts, I wondered whether this was really good advice to those who have a chance of actually going to the moon).

  Be that as it may, a later generation had a few successes with a few axioms of infinity; in accordance with some, but certainly not all elements prominent in Gödel's musings.

Autobiographical remarks: absoluteness scaled down

Though I saw Gödel [1946] only in the mid 60's (when it was published), the general drift was clear enough from his conversations 10 years earlier.[15] At the time I was equally ill at ease with popular claims for *and* against general, so-called epistemological notions (cf. the opening of [1946]). But also I had no idea how to formulate the malaise; not even, for example, in terms of the obvious incompatibility between the logical order of priority and orders discovered to fit the facts of experience better (which is a refrain of this chapter). So I could not (and of course did not) take Gödel's words literally. However, they struck several chords:[16]

1. *Finitist provability*

   This too is 'absolute', for those so benighted that other proofs are inaccessible to them. Now, footnote 2 of Kreisel [1951], which I had quite forgotten by the time I met Gödel (and nobody including him had challenged), is spoilt by a blind spot. I had assumed there that a (satisfactory) definition of finitist provability should also be *established* finitistically; in other words, that it should be satisfactory to a finitist, too. Now, whatever malaise I had with Gödel absolutes taken literally, what he said about them was enough to remove my blind spot.

   His (admittedly arresting) terminology jarred with the view I took of finitism. Given my temperament, it would not do (for me) to sanctify benighted shortcomings by terms like 'absolute'. So instead I used the (admittedly colourless) word 'informal'; for example, in my [1960] (cf. footnote 4 of Gödel [1958]).

2. *Predicative provability*

   As described elsewhere, by a fluke I came across Kleene's papers on hyperarithmetic predicates at about the same time. Of course, there was nothing wrong with them. But, to me, they became more rewarding when related to the traditional literature on predicativity. Later I noticed a more specific use for analysing different proofs of Cantor-Bendixson, which I had learned in my teens in Littlewood's lectures, the only course I liked at Cambridge. (There was a corresponding shift of emphasis in the questions asked about hyperarithmetic objects.)

---

[15] As usual, he did not breathe a word about his earlier (here, oral) publication (cf. Chapter 10 [∞]).

[16] Since Gödel thought of the isms in 1–3 below as 'opposed' to realism, *his* realism could have led to my interest in the others only out of perverseness (which I do not wish to exclude; but then: Who am I to judge my unconscious motives?). However, I was totally conscious of the fact that what he had to say not only revived my earlier interests (mentioned in Kreisel [1951]), but consolidated them.

3. *Intuitionistic provability*
   As in 1, in the course of conversations with Gödel about his system T I came to see some merits in this idea (which were indeed enormous, compared to my earlier expectations!). In contrast to the topic of sets, here was lots of virgin territory, starting with the possibility of completeness theorems without concocted semantics. In Gödel's terms, one had now absolute results on intuitionistic provability (albeit only about propositions stated in familiar formal languages); 'positive' ones for propositional logic, 'negative' ones for predicate logic. Most memorably, at least for me, results on new propositional operators served as a foil to the functional completeness of classical propositional logic (a specimen of absoluteness ignored by Gödel, as mentioned above).

Nothing in 1–3 had shown (to my satisfaction) that the informal notions considered are particularly suited to describe the facts of (here, mathematical) experience. I was aware of this risk at the time [∞], and others noted my awareness (cf. the introduction to Benacerraf and Putnam [1964], but dropped from the second edition).

Once again, it is a matter of temperament whether correcting widespread misconceptions (here, about the scientific potential of traditional informal notions of proof) is or is not adequate pay-off; to be compared to correcting the idea that butter (which tastes good) is bad for you, as opposed to discovering that seal fat (which stinks) is good for you. (In both cases you have to look at those fats.)

8.4 Selected Thoughts About Sets

They come mainly from scattered footnotes in Gödel's piece on Cantor's continuum hypothesis ([1947] and [1964]). But first a few words on Gödel's own emphasis are in order, lest the shift below cause unnecessary malaise; 'unnecessary' today, not 40 years ago, when I at least knew nothing simple enough about it (or set-theoretic foundations generally) to dispel (anybody's) malaise.[17]

[17] Gödel's essays in the 40's had blinded me so completely by their (to me still atrocious) opening fanfares that, 40 years ago, I declined von Neumann's proposal (via the geophysicist Bullard, for whom I had done some work) to visit the Institute at Princeton for contact with Gödel. The visit would have interfered with some plans for frivolity, which was (as I saw things at the time) more rewarding.

Fortunately, actually at the *ICM* at Amsterdam in 1954, I was reassured by a friend, whose views I had found compelling for more than 10 years and who had personal knowledge of all parties concerned.

When I met Gödel in 1955 I learnt to see - what still appear to me - gems in those essays (admittedly, perhaps brighter than they are, against those 'philosophical' fanfares as a foil). But above all I very soon discovered to my delight (and, admittedly, again possibly all the greater by contrast with expectations) that, in practice, he took a very catholic view of his pet ism:

Gödel's own perspective and our agenda

His first order of business is the question whether the concepts used to state the continuum hypothesis CH are well-defined; his answer is in terms of 'some well-determined reality' (philosophical realism):

> in this reality 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.

There are many, albeit partial, results which he relates to his answer. But, as for everything else in the world, this is not (and certainly does not remain) the only compelling emphasis; in fact, not even remotely so. (By the same token, the constructible sets also constitute *some* well-determined reality.)

For the record, 40 years ago my (and my chums') malaise with Gödel's piece was simply compounded by all that heavy breathing about 'reality'; recalling *Hamlet*:

> The lady doth protest too much, me thinks.

Today I can be more explicit; partly by reference to specific logical discoveries in the meantime: they show that the results about the continuum, established (and rewarding) in geometry, are even logically independent of CH.[18] Probably more convincingly, at least for those with broad research experience, there are *general reminders*: when a question is of little consequence, that is it has few consequences of interest in the area considered (like CH above in geometry), it is likely to be both difficult *and* unrewarding to decide.

In terms of (erudite) isms, the assumption implicit in Gödel's emphasis is tantamount to the most vulgar form of pragmatism: what exists is useful (here, in the commerce of ideas).

But just as in the last section, here too there is *cheerful news*. Despite his unpromising general perspective, Gödel's piece has some almost equally memorable points of obviously lasting use. But before going into them, some warning and reminders (int the next two subsections) are salutary.

> You trust in God (represented by philosophical isms) and keep your powder dry
> (by being both realistic and constructive, in the popular senses of these words).

[18]More precisely, the single most significant discovery about the continuum, made early in this century and consolidated since then, is the following.

First, the set of points and maps prominent in general (set-theoretic) topology are unrewarding geometrically except, as always, for a few general (and therefore simple) facts.

Secondly, the opposite is true of those prominent in the (incidentally, relatively few) branches of contemporary topology. Furthermore, the properties studied in the latter are (generally demonstrably) insensitive to the cardinal of the continuum (as in the case of homotopy).

Terminology

Instead of Gödel's 'set in the sense of arbitrary multitude' or 'extension of definable property' (in footnote 2 of [1947], quoted below), corresponding more familiar words are used below; in particular, sets in (segments of) the cumulative hierarchy V_α generated by the power set operation, constructible sets in L_α, etc. Generally, 'kind of set' is used without agonizing whether the kinds involved are restrictions of some (more) general kind. In terms of the refrain about the growth of knowledge (here, knowledge of sets): without agonizing whether such a general kind functions like a seed, or is the result of (growth by) accretion.

In any case, throughout this section the emphasis is on sets in suitable V_α's.[19] In terms of the properties listed among axioms of familiar set theories, all α or all limit ordinals α may be 'suitable', depending on which V_α have the property in question. For example: all V_α satisfy the axiom of union; for each limit ordinal α, V_α is closed under pairing and, of course, under its generating (power set) operation.

Some home truths, half truths and untruths

Above all, in the first place, V_α for particular α are meant. Where conditions (and ways) have been spotted for extending results to 'all' ordinals, it is sensible to do so. It is not sensible merely because one 'wants' to (cf. Dirac [1978] on sensible mathematics). A *rough* parallel is a child's experience with finite α on the one hand, combined with general (that is, indefinite) properties on the other; pedantically, of indefinite extension.

But parallels between V_ω and V_α (for specific $\alpha > \omega$) generally reach the point of diminishing returns quite soon; much sooner, as it were, than those between, say, $\mathcal{Z}$ and the rings of algebraic integers in number fields (cf. the Appendix). *Samples*:

- In V_ω (what since Cantor are called) cardinals and ordinals satisfy the same arithmetic laws, but not beyond.

  For example, for cardinals a: $2^a > a$ holds generally, while $2^a > a^n$ $(1 < n < \omega)$ is a consequence for all $a \geq \omega$, but not for all $a < \omega$.

  The generalized continuum hypothesis GCH (2^a is the successor of a) is false in V_ω, except for $a = 0$ and $a = 1$, but it would be as pathetic to draw any conclusion from this about V_α for $\alpha > \omega$ as from the fact that 1 is weakly, but not strongly inaccessible.

[19] The emphasis has an obvious parallel in the case of *number* (in place of *set*), for example when axioms for rings or fields are explained by reference to familiar or otherwise easily described kinds of numbers. As with any emphasis, something is lost; for example, we have no parallel here for Conway [1976]'s numbers-large-and-small.

- V_ω (like everything else in the world) has many descriptions, all of which can be generalized (again, like everything else in mathematics); but not necessarily to V_α for $\alpha > \omega$.

  Thus V_ω is generated from $\emptyset$ by: $x, y \mapsto x \cup \{y\}$. In old-fashioned terms: this satisfies *l'esprit fin* (of algebra, or its infinitistic analogue in L), in contrast to *l'esprit géometrique* of the (impredicative) power set operation. Thus ω has at least two descriptions, as the closure ordinal of each of those generating operations (of course, without the usual accumulation at limits).

For the metaphor of a tree of knowledge there is, as always, some (logical) order of *evidence* or at least of *justification*. But its use or, for that matter, niggling about it simply distracts from effective knowledge of those phenomena in familiar experience that have the labels above (and are genuinely in demand in the commerce of ideas). Specifically, the logical order imposed on descriptions (where one is privileged as a definition, and the others are deduced) produces artifacts; certainly with respect to the historical order. For example, the Greeks managed well with their geometric (impredicative) descriptions.

The reminders above, including the last one about effective knowledge (serve to) correct errors. They are not contributions to effective knowledge (of V_α for $\alpha > \omega$). But, compared to the logic chopping in earlier sections, they rely more on mathematical constructions than on other thoughts. (The latter are meant as in *Tractatus* 6.21, about mathematics not expressing any thought.[20])

Axiom of choice

Footnote 2 of [1947] recalls that

> this axiom is, in the present state of knowledge, exactly as well-founded as the system of the other axioms ... It is exactly as evident as the other axioms for sets in the sense of arbitrary multitudes and, as for sets in the sense of extensions of definable properties; it also is demonstrable for those concepts of definability for which, in the present state of knowledge, it is possible to prove the other axioms.

Thus, the axiom is valid for the V_α's (the 'arbitrary multitudes'); actually, at each α, at least in its multiplicative version. As for the flourish about 'evidence', the axiom is (realistically) *more* evident than, say, replacement (see below). Historically, it was used freely; to be compared to axioms for *order*, which are used in Euclid, but not stated there either.

Secondly, Gödel notes that the axiom also holds for L; in other words, for sets defined from the ordinals by familiar logical operations and accumulation. This flourish misses a couple of opportunities:

[20] *Tractatus* is (best regarded as) an ode to propositional calculus, when its otherwise irritating exaggerations become perfectly acceptable instances of poetic license.

- *Logical hygiene* concerning the sets (say, in $V_{\omega+1}$) which are definable in the absolute sense adumbrated in [1946] (where this topic is presented as a first order of business).

  Now, whatever doubts there may be about the scientific sterility of this sense, there is no doubt that it is among the first thoughts that cross anybody's mind. For those sets the axiom of choice is quite dubious. Viewed this way, the fuss about it is not merely thoughtless and thus simply embarrassing (as it is in the bulk of the literature).

- *Reminder* concerning the contemporary sense of the word *axiom* (for the motto: *dégager les hypothèses utiles*).

  Suppose the property P satisfies some general conditions which ensure $\exists xP$ by the axiom of choice, and (the lemma, as it were) $\exists xP$ is enough to infer Q. Then, given a proof of Q from $P[x/f]$ with some more or less elaborate definition of (a choice function) f, the motto requires *either* the use of the axiom of choice *or* some rewarding strengthening of Q which follows from $P[x/f]$, but not from $\exists xP$ alone. The traditional preoccupation with the validity of choice for airy-fairy notions of sets distracts from the sensible use of 'axiom' above.

  Viewed this way, well known conservation results applying to Q of suitable logical form are in accordance with the motto; at least, when utility for formal derivability is meant.

Comprehension

Footnotes 12-14 of [1947] are about (what is there called) the operation 'set of x's':

> The operation 'set of x's' cannot be defined satisfactorily (at least in the present state of knowledge), ... but as opposed to the concept of set in general (if considered as primitive) we have a clear notion of this operation.

For the record, Gödel did not object to the use which I make below of those footnotes, and made in many conversations with him.

In a nut shell, the use is (not, of course, to agonize over paradoxes but) to find *a memorable interpretation of the literature*: on Frege's oversight (which Cantor called '*unglücklich*', an unfortunate idea)

$$\exists y \forall z[z \in y \leftrightarrow P(z)];$$

and on Zermelo's

$$\forall x \exists y \forall z[z \in y \leftrightarrow z \in x \wedge P(z)],$$

which has superseded it. Mere (mathematical) survival certainly does not depend on reviving Frege's fossil in the evolution of ideas. The (cl)aim is merely about a *suitable* place for it in a museum of them, if we can afford it in our age of intellectual affluence.

Zermelo's clause '$z \in x$' gets a memorable interpretation in Gödel's thought of (here, y being) a set of x's. Actually, in conversation I went the whole hog, and interpreted '$z \in x \wedge P(z)$' as a *property of* x's, too. On the other hand, rightly or wrongly, the relation to Cantor's sense of *definite* (applied to extensions of properties, here, to elements $z \in x$) was to me too obvious to require mentioning it (to Gödel); as opposed to: undetermined for some particular z. After all, in connection with impredicativity in Section 2, he himself had drawn attention to the (indefinite) range of the membership relation used in Russell's paradox. And if this relation were undetermined in set theory, what on earth should be determined there? Be that as it may, comprehension holds for all V_α.

Reminder on the other kind of property (not 'determined without arbitrariness'). Suppose a quantified first order formula of ordinary set theory is presented as a definition of P for some given $x \in V_\alpha$, but without stating the range V_β of the quantifiers in the formula. In general, the property so defined does depend on β, and if the choice of β is regarded as arbitrary, the property is indeed not determined without arbitrariness.

Replacement and (uncountable) strongly inaccessible cardinals

By Zermelo's [1930] and for infinite α, V_α satisfies replacement for arbitrary (in other words, second-order) functional relations if and only if α is strongly inaccessible; naturally, well defined relations are meant.[21]

Now, the replacement axiom differs in many down-to-earth respects from the others; in particular, from choice and comprehension:

- For one thing, *it is not true for all* α; in fact, not for any α ($> \omega$) usually encountered in the commerce of ideas (where it is, realistically speaking, an axiom of infinity).
- For another, *its first-order and second-order versions differ* in a more brutal way than that illustrated in the *Reminder* at the end of the previous subsection. There simply are accessible α such that the former version holds for V_α (but, of course, not the latter).

As to Gödel's own perceptions of evidence, they required education over several years after his correspondence with Zermelo in 1931. There he alluded to his

[21] Incidentally, in [1930] Zermelo uses the homely words 'definite property', but without would-be erudite explanations, to be compared to the use of 'finite' without any ritual of set-theoretic definitions.

malaise with Zermelo [1930], but could not (politely) pursue the matter since Zermelo did not take him up on it. At least according to what Gödel told me, he was ill at ease with replacement.

This malaise had a practical consequence for him in his work on the constructible hierarchy. He had the general idea for his proof of GCH for L as a student, and even lectured on L in 1936. But with his malaise he hesitated to use von Neumann's ordinals of the required high types; and, without them, tiresome definitions of well-orderings are needed up to ω_ω in $V_{\omega+\omega}$ (cf. [1939]). This delayed the final publication for a couple more years; until he had satisfied himself about inaccessibles.

For the record, I once asked him for his later thought(s) on these things. His answer:

> those inaccessibles are implicit in the concept of ordinal

was to me a reminder:

> the series of ordinals is (conceived as) absolutely unending.

I did not have sufficiently specific questions to have a chance of getting much from any specific answers (he might have). So I did not pursue the matter at the time. Only later did I see in the many Pyrrhic victories of set theory in general, and especially of set-theoretic foundations, a *fertile subject of cultural interest*; too late to become sufficiently steeped in it for genuine interest (let alone, contributions).

Addendum on replacement (optional)

Much less than the above is enough for the following *correction of errors*; actually, at two extremes:

- So to speak, on the negative side: not every instance (of replacement) that holds for a particular V_α, provides support for the axiom itself.

  For example, $\alpha = \omega$ does not; not surprisingly, after what was said in the *Samples* on p. 175 about (other) parallels with ω. If $\alpha = \omega$ then, for arbitrary functional relations restricted to V_α, domain and range are finite. Thus, if the range consists of ordinals, it actually contains its supremum. But accessibility of α can be problematic only if the supremum of a set (of x's $< \alpha$) may exceed all elements (and be $= \alpha$).

  Strong accessibility involves the exponential function, which is not well understood for $\alpha \geq \omega$ (in contrast to $\alpha < \omega$, from which fact the literature both on computation and on certain non-standard models of arithmetic distracts). In simplest terms, in the present case, parallels with V_α when $\alpha > \omega$ are spoilt by incomparably greater experience with V_ω; including the impredicative knowledge of facts about V_ω, which are described by use of the idea *finite*.

- At an opposite extreme is the superstition that replacement is suspect merely because it is a kind of (infinitistic) 'closure' condition, and hence tainted by Frege's 'closure' condition on properties and sets: $P \mapsto \{z : P(z)\}$.

  It is familiar that Frege thus obliterates the traditional distinction between properties and sets, which applies the membership relation only to the latter; Cantor spoke of 'arbitrary varieties' in contrast to those 'grasped as unities'.[22]

Remarks on what may be lacking. First, what is *not* lacking is progress in our understanding of inaccessibles; specifically, in the last quarter of this century:

- Thus Solovay's use of them for defining kinds of sets (that is, models of set theory), for which every set of reals is L-measurable is simply qualitatively more substantial than any earlier material.

- Even more directly pertinent is Shelah's demonstration, as it were in the opposite direction: how knowledge of L-measures for certain projective sets provides new descriptions of ordinals, which establish the latter to be inaccessible in L; to be compared to impredicative knowledge about ω used in the discussion of V_ω above.

Secondly, at least with respect to demands in the market of which I am representative, what is certainly lacking is an analogue to Gödel's *felicitous expression*

[22]Pertinent, but less familiar are the following *reminders*.

Some 2500 years ago, in *Physics* (III, 6, 206b 33-35), Aristotle made a distinction, admittedly in clumsy terms, under the heading: the infinite.

> It is exactly the opposite to what is [tacitly, sometimes or, perhaps, usually] said to be: not what has nothing outside, but what has always something outside [each of its parts].

Applied to strong inaccessibility, viewed as an axiom of infinity, this emphasis on 'parts' corresponds to the use of *strict* order:

$$\text{If } \beta < \alpha \text{ and } \gamma < \alpha \text{ then } \beta^\gamma < \alpha.$$

This would be obviously false with $\leq$ in place of $<$. (Frege's condition does not have a corresponding bound.)

Some 100 years ago much attention was given to a particular class of closure conditions, called 'from below' or 'inductive'. These do admit equality, with a *least* fix point (and others 'outside' it).

To repeat what cannot be repeated too often: the reminders above correct elementary (that is, brutal) errors (here, dubious doubts) about inaccessibles, but are not enough for contributions.

'set of x's' for a *thought* that supplied what was previously lacking in connection with comprehension.[23]

Thirdly, this last and similar defects may be connected with a lack of a sensible perspective; specifically, in the emphasis on the (logical) need for axioms of infinity for new results, which distracts from the market for their *uses as a better bargain* (in competition with possibly already existing proofs):

- Cantor's cardinality argument for the mere existence of transcendental numbers remains a better bargain than the specific transcendental $\sum 10^{-n!}$, which Liouville had produced 10 years earlier (with more work, for a very limited market).

- When Martin gave his proof of Borel determinacy by use of (what is realistically) an axiom of infinity, the emphasis was *either* on the theorem itself *or* on the logical need of the axiom for proving it; overlooking the *main* novelty of Martin's product: the use of higher cardinals for something that at least remotely resembles some things in demand by the market in question.

8.5 Intuitionistic Logic: Hitting a (Little) Moon First, and Then Dreaming of the Stars

By our agenda, this kind of logic is viewed here like the kinds of set in the last section: without agonizing whether any general notion of logic is a seed for some tree of knowledge of which our kind is a (sturdy) branch, or whether any such notion has grown by accretion from components, among which our kind remains a visible entity. For one thing, future research may have something to say about these two options. For another, (premature) agonies about such options are traditional, and our agenda requires emphasis on alternatives to the traditions involved.

Background

First, there is another kind of logic, which is meant for propositions:

- without incompletely defined terms

- with the property that they are either true or false.

These aspects were considered by Aristotle; quite explicitly, as necessary to make *propositions* rewarding objects for study. Since the latter was his trade, those

[23]Incidentally, when work on axioms of infinity began some 30 years ago I was (obviously wrongly) convinced that, *at a minimum*, it would lead to an adequate expression for a correspondingly adequate thought in connection with inaccessibles.

aspects were indeed essential. The sanctimonious expression 'of the essence' is apt if trade interests are sacred (whatever the actual usage of that expression may be).

Now, it is a simple fact of life that mere *truth* is often a distraction; for example, for the accused who (knows he) *is* guilty, provided only the law happens to consider him innocent unless *proved* guilty. His business is not truth, but an *irrefutable* defense; formally, double negation is weaker than truth.

Intuitionistic logic, which ignores truth (in favour of evidence) altogether, provides a coherent scheme for a corresponding interpretation of the logical particles. It therefore also provides a literary form for underlining that first order of business for the accused.

As always, it is a separate question how (if at all) theoretical elaborations of intuitionistic logic contribute here. At any rate, as always, it would be a philosophical mistake to assume that all (effective) thought must be theoretical. Though I can see the broad interest of the rhetorical aspects touched above, I am not sufficiently familiar with them to be interested; let alone, to report on details.

The second reminder is more parochial (but less than higher set theory in the last section). In mathematics too, we generally know more about propositions than whether they are true or false (and sometimes want a vehicle for expressing such additional knowledge); most often, *how* they depend on parameters (so-called functional dependence). This is obvious in the case of $\exists xP$ and $p \vee q$, with some (explicit or implicit) parameter; but as a moment's thought shows, also with $p \to q$. There are at least two options:

1. One is to make the dependence explicit, when the logical particle involved is simply eliminated.

2. Another one is to adapt the logical laws, and thus ensure that (logically proved) theorems are subject to *suitable* dependencies. Logicians are familiar with such dependencies from various kinds of reducibility in recursion theory.

Gödel's scheme

In a lecture on 15 April 1941, Gödel asked:

In what sense is intuitionistic logic constructive?

Here he meant, roughly, option 2 above; to be precise, he had to explain what dependencies are meant in the case of logically compound expressions like $(p \to q) \to r$, which do not occur in ordinary mathematical thought (but do occur in formal systems). For simple expressions $p \to q$ his option involves different (familiar) reducibilities, according to the logical forms of p and q; roughly, one-to-one reducibilities if they are in prenex form without alternation of quantifiers (in other words, Skolem functions of a suitable prenex form of $q \vee \neg p$ are used), but not generally.

Gödel meant 'constructive' in its usual mathematical sense; with emphasis on the (functional) dependencies, as above (less on the means of showing that the functions do what they are supposed to do). In the example above, of $\exists x P$ with parameter a and an explicit function X (of a), this would require a proof of $P[X(a)]$ (for the range of a considered).

Corresponding to iterated implications, as in the example above and for which intuitionistic logic is notorious, Gödel had operations of higher (finite) type. His answer to his question was an impeccable interpretation of Heyting's arithmetic, a most familiar system of intuitionistic logic. The definitions of the operations used have a very familiar look: except for the higher types (and the corresponding conventions of a typed λ-calculus), the whole scheme looks just like primitive recursive arithmetic. In terms he used 10 years earlier, on incompleteness of the familiar system of *Principia*, the interpretation is easily seen to be typical of 'related systems', and so enough for the general picture below.

A pyrrhic success

The two outstanding facts here are that Gödel's scheme:

- provides definitions for the operations in the mainstream of constructive mathematics (cf. the background above)
- it has not contributed to that mainstream.

Incompleteness results distract from both those facts.

Now, with both the scheme and that mathematical experience before one, one can also see *what is lacking* in the former; specifically, concerning its 'functional dependencies'.

> By experience, *separation* of some (albeit relatively few) different kinds is necessary (for significant results), while the scheme concerns what they have in *common*.

Refinements (for example, according to the syntactic form of the definitions) produce (of course, precise) results that are not significant for mathematics; in contrast to such classifications as algebraic or topological dependencies. Now, this property of the scheme has a perfect parallel in the common consent among experienced mathematicians about set-theoretic foundations (as being the 'least interesting side' of the business); of course, also where the latter are amply complete. As a corollary, this *neglect of significance* in classifications is not peculiar to any particular foundational scheme (or even ism, with which it may be connected), but is part of the foundational ideal (in other words, of the metaphor of knowledge growing like a tree).

Gödel himself does not touch this foundational side. But, starting in his first paper [1958] (in contrast to the lecture in 1941), and especially in his later additions, his (cl)aims concern quite different aspects of his scheme: relations to such traditional idea(l)s as *reductive proof* (in footnote h of [1972]). It is not particularly hard to elaborate such matters (as in the subsection on p. 172 for other traditional notions of proof), and thus 'formulate the philosophical gain achieved', when the 'gain' is measured by the canons of academic epistemology. This leaves open what gain, if any, there is for a more realistic view (of knowledge). Be that as it may, all this shows vividly that, for good or ill, Gödel's attachment to his so-called Platonism did not keep his hands off other isms.[24]

8.6 Cosmology and Some (Even) More Ethereal Ologies ([1949], [1949a], [1950])

This section has to do with Gödel's three short papers around a previously neglected type of cosmological solution of Einstein's equations for gravitation. For readers with a general mathematical education, the knowledge of differential geometry needed here is no more demanding than the logic in some of Gödel's other papers.

In [1949] Gödel lists some properties of his solution, in particular: there exist closed time-like lines (though every world line of matter is an open line of infinite length). In [1949a] Gödel treats the solution quite solemnly, as a contribution to effective knowledge; of the nature (as one says) of time.

For the present chapter, this is below threshold before looking at alternatives. For example, the following:

[24] Abstractly, there is a staggering contrast between Gödel's:

1. acuteness and imagination in seeing and exploiting logical (and other mathematical) aspects of quite hackneyed idea(l)s
2. blithe disregard for general scientific experience where the idea in question has proved sterile or false.

An extreme case is the idea of God being a mathematician, a (hackneyed) way of saying that spectacular mathematics must be (a guarantee for) 'truth'; of a particular interpretation to boot. The example of the theory of a complex variable, which is also the theory of ideal liquids in two dimensions, is often quoted (the ideal of reductive proofs above is a candidate, too; specifically, Gödel swalled the reductive interpretation of Gentzen's results, mainly because of their obvious mathematical wit).

But (and this too is a fact of experience, not a mere possibility) Gödel is by no means unique in combining 1 and 2. After all, 2 is a simple attachment to ideas learnt in one's teens, when one often really has too little scientific experience to correct them convincingly (cf. *A brief history of time* by Hawking, who wrote an editorial note for a couple of items in the collection Gödel [1989], with many memorable examples of combining 1 and 2).

Analysis of language

This dread idea of the academic tradition is here applied to the language of (theoretical) sciences, in particular, the mathematics of differential geometry. Viewed this way, Gödel's solution unquestionably corrects the neglect of solutions of his type. So what?

- *Singularities*.
  Gödel's type has none; admittedly, this fact is not listed in [1949]. But also, at least since the 60's there has been interest (in particular, by Penrose and later also Hawking) in the matter of singularities in solutions of Einstein's equations. Whatever else may be in doubt, if something is to be *proved* about classes of solutions with singularities, Gödel's type has to be *excluded*. If nothing else, it is a *complement* to (later) results about such classes.[25]

  In contrast, it would be simple-minded to assume that it must have (had) so-called *heuristic value*; 'simple-minded' by overlooking, for example, the possibility that the type was excluded tacitly.

- *Combinations* with (equations for) other aspects, besides gravitation, of the phenomena considered.
  This matter is prominent in the common-or-garden varieties of science; for example, at the beginning of Newton's *Principia*. The development of rational mechanics in the 17th and 18th century produced many examples of such combinations (with his equations for gravity, where actually solving the combined equations is another story altogether). Attention to problems arising from such combinations is one particularly striking difference between practice in the common-or-garden varieties of science (or, in fact, thought generally), and in those would-be all-encompassing schemes, which (cl)aim to leave out nothing (with which to combine aspects privileged in them). Foundational schemes are of course chemically pure specimens of this idea(l).

  It is by now a common place that relativistic (requirements on) equations are hard to combine with others. Dirac's equation for the electron, respecting *special* relativity, became correspondingly famous. Gödel's solution presents itself as a new type of *candidate for object lessons* on such

[25] Downmarket Hawking's *imaginary time*, used in his answer to questions about where the universe 'comes from' (the universe just 'is'), gives the flavour of such complements; more cheaply, since it is familiar that his switch turns hyperbolic into elliptic equations, which are much tamer.

For the record, the later work has been supplemented by yet another type of solution without singularities which satisfies most conditions prominent in earlier singularitiy theorems (but not, for example, the convergence of light rays); cf. Senovilla [1990] or the breezy account on p. 201 in *Nature* (17.V.1990).

> combinations; at least by analogy with my own experience, as in the following *digression*.
>
> The assumption is that the 'ends of time' are irrelevant in at least some situations where the combination in question is a main problem. Given that not much is known about them, what options (as always, if any) are there? One, of course, is to leave those 'ends dangling'; another is to 'glue them together'. This has nothing to do with conventionalism or any other ism. In either case some key element may become visible (that is, some memorable obstacle to combinations) and with luck means of removing it, which then opens the way in other situations, too. Incidentally, the obstacle may be (later seen as) a blind spot.[26]

The points above are - meant to be - *alternatives to the solemn tradition* (and, in fact, anathema for it), especially the motto:

nothing but the universe is good enough for us.

For the record, the price paid for this motto seems (to me) fair enough: it generally spawns work that, at best, corrects errors rather than contributes to effective knowledge. Exceptions exist (often by Big Science, cf. p. 165), but they tend to be quite costly by any realistic account(ing).

Sundry tit-bits: old and new

Aristotle's primary (measure of) time is cyclic (cf. *Physics*, VIII, 265a 15). His student Eudemus mused about time being cyclic too. They do not seem to have agonized over effects possibly preceding causes (for a local direction, as in (4) in [1949]); perhaps not surprisingly. For one thing, Aristotle has little to say about temporal causes; for another, neither of them discusses the possibility of going back in time. Still, I had hoped that some sophist at the time had objected: What happens *if* you go back and *kill your father when he was a baby*? Better

[26] Obviously, nobody (in his senses) with my limited experience here would be tempted to pontificate about cosmology. But there are the following *easy parallels*.

Jets in rapid motion, producing cavities, often disintegrate into turbulence (about which not much is known, and the details of which pretty obviously have little to do with the general motion). Experience has shown that, for a successful theory of those broad aspects, the 'ends' of the jets are sent off into a different Riemann surface.

But also, still in connection with rapid motion with forces that are very much greater than gravity, the combination with gravity may be *dramatic in permitting a new type of solution*. Specifically boundary conditions, which now determine a solution, have none - for equations - without gravity; for example, an infinite jet with a free surface (at the top) deflected by an infinite plate partially immersed (below that surface).

Incidentally, there is a recent claim by Motz and Motz [1990] about a similar story for equations of the photon.

still: And what if East ever met West? Alas, even if anything of such pastimes among sophists is known, it has not come my way.[27]

The next and last tit-bit is a little different. It is about *markets for literature* concerning the universe:

- As far as ordinary commerce (say, of the book trade) is concerned, the facts are striking, and I certainly have nothing to add by way of interpretation.

- The commerce of ideas is a different matter. At one time venerable theologians constituted an, as it were, captive market. Times have surely changed but, perhaps, not as much as suggested by Carl Sagan in his preface to *A brief history of time*. He considers pre-school children who ask:

  Where do we come from?

  and envisages an answer in cosmological terms; perhaps:

  the Big Bang.[28]

  I have heard it said that the idea of quite ordinary bangs, which do not even move the earth for those involved, also finds a market among pre- school children.

Other ethereal ologies

It is time to return to Gödel. Related topics came up in our conversations, as I report below. His musings were not as coarse (in either sense of the word) as the last subsection, but they did not consist of solemn logic chopping or erudite references to the ancients either. He just had a general interest in ethereal ologies; for example: *theology* itself, but also *pneumatology* (not only of the Holy Ghost, but of ordinary ghosts too), and *demonology*.

Gödel's (broad) interest in these matters is common enough; what I find most *satisfaisant pour l'esprit* is the parallel to experience with the logical 'demons' of Section 2 (infinity, self-reference, impredicativity): a kind of 'clutching at straws' [∞]. As could be expected from the broad philosophy of this chapter, this shifts the emphasis to different aspects, pursued in the next subsection. But first some

[27]Fortunately, the question above is in the introductory note to [1949] (in Gödel [1989]) by Hawking, solemnly shown in the introductory note to [1949a] by Stein to be a less decisive objection than his brasher co-editor clearly assumes. To be precise, Stein considers a variant, where you go back and, more simply, *murder your own former self*; pedantically, you 'murder' your 'spirit' (since, as mentioned, world lines of matter are open and of infinite length). It's all good, clean fun for us more cheerful readers, whatever the editorial intentions may have been.

[28]Unless 'the universe just is', like the Creator according to theologians when asked: Who created the Creator?

Anecdotes. When I read the following passage on time travel in Gödel's article [1949a], it struck me as attempting to provide a kind of explanation why one rarely sees familiar ghosts (that is, of the recent past; cf. p. 124 of Chapter 6):

> It is possible in these worlds ... to travel into the near past of those places where [one] has himself lived. There [one] would find a person who would be himself at some earlier period of his life ... But the velocities which would be necessary in order to complete the voyage in a reasonable length of time are far beyond everything that can be expected ever to become a practical possibility.

I happened to tell him about it in the presence of his wife, who then spoke mockingly and at length about his life long interests in ghosts (which, according to her, was shared by Viennese washerwomen), and about the many books he had read on the subject (I remember using this to change the subject, by noting that those washerwomen surely did not rely on books).

Gödel became very expansive on the need for a great deal of basic agreement if conversations are to be fruitful (but read on). I did not need much persuading, since I have always applied this 'non-missionary' view even to my writings. Here are a few additional points.

Gödel's faith in the wisdom of the ancients (including ghosts, demons, deities and so forth) was not shocking to me. It reminded me of a lecture by Lord Keynes at Trinity College, Cambridge. Keynes had bought in Ireland a ship trunk full of Newton's papers, and spoke most memorably of Newton's attempts to deduce all sorts of things from the number of the beast (in the *Apocrypha*). I tried to convey my impressions of Keynes' style to Gödel, but doubt that I succeeded.

Without sharing Gödel's faith at all, I did not reject it out of hand; at least, not abstractly. Obviously, unusual skills were needed to make good use of it. Common place objections to that kind of faith seemed to me weaker than the faith itself; perhaps, comparable to my distaste for the usual objections to informal rigour in the analysis of intended meanings, even granted that the latter may be unsuitable for their intended purposes. Besides, in human affairs the wisdom of the ancients is widely accepted; in effect, if not in these words.

But the agreement quoted above did not go very deep. On one occasion, I think out of the blue, Gödel brought up the familiar asymmetry of the universe: so many more events are unpleasant. From this he concluded the existence of demons. I don't remember, if ever I knew, what came over me to talk about an evenly distributed universe, but my being so often below par that I could not exploit pleasant opportunities. Evidently, I had forgotten to consider the role of demons in my being below par. At any rate, in his gentle way, Gödel attributed my blunder to lack of interest, not demons. We never talked about the subject again.

There clearly was something to Gödel's view; at least, if the loaded expression 'wisdom of the ancients' is replaced by, say, 'naive ideas'. Gödel's own favourite refrain was:

> if one could be so successful [with such ideas] as he was, one must expect marvels if one tried harder.

During his life I never felt quite comfortable about the whole business. Only when I came to the end of the original version of Chapter 6 did I put my reservations into words:

> Perhaps those ideas are good to remove blindspots, and then they are wonderfully efficient. But it still bothers me that the law of diminishing returns seems to apply to them so very soon.

Digression on logical aspects of theology (for intellectually cheerful readers)

The general idea follows from a refrain of this chapter:

> Though everything has logical aspects, they will be most visible where they are not overshadowed by other (more rewarding) aspects.
>
> Furthermore, by experience, neglect of logical aspects can be occasionally costly (in the commerce of ideas, too).

The occasional observation is that there seems to be a gap in the market for *supplementing the literary forms of mathematical logic*, which are used for making particularly elementary logical properties memorable (enough to be remembered when they present themselves). The theological literature is one (re)source. And a good bargain too,[29] since many possess knowledge of it already.

In the following examples, the emphasis is on particularly crass logical errors in (familiar) theology, and on how to use them (as it were) as vaccines for immunity in more delicate situations (where related errors occur).

1. Various ontological arguments, preferably in Latin (which has no articles at all), concerning 'the perfect Being' illustrate abuses of *the definite article*, which are not covered by Russell's 'the present king of France'.[30]

[29]Naturally, not for those who are determined either to remain committed to the solemn tradition, or to stay away from it altogether.

For the rest of us, it pays to know something of the conventions of such literature. What to do with such knowledge may depend on temperament: whether we want to interpret or change the world (of this sector in our commerce); in the latter case, whether by merger, take-over or unbundling.

[30]Both 'the present king of France is bold' and 'the present king of France is not bold' are false on Russell's analysis, which conflicts with (a first reading of) Tarski's 'adequacy' condition for truth: $T(\neg p) \leftrightarrow \neg T(p)$ (cf. Gödel [1944]).

For example, by the tradition of theology (of their day), critics of those arguments could not assume that there was no such being. Nevertheless, they were able to make their point *without any ritual of formal 'paradoxes'*.

The word 'ritual' is meant to underline (not only the obvious possibility, but) the fact of experience that formal contradictions are neither the only defects of reasoning, nor particularly instructive. Specifically, while Cantor's review [1885] of Frege's *Grundlagen* specified a convincing defect in Frege's naughty axiom, Russell's paradox continued to attract attention to itself (and not, for example, to the definite article in: 'the class of all classes not belonging to themselves').

2. One of the properties required of 'the perfect Being' in 1 above is omnipotence or (in terms of Section 3) *absolute* power. Cusanus had some formally very simple *closure conditions* on his idea(l) of omnipotence: not only (the power to create) an immovable material object (in his case, a stone), but also the power to move all material objects (including stones). Incidentally, it appears that Cusanus *wanted* a Creator with such absolute power.

   Now, not only did Frege impose closure properties (on such logical objects as predicates and classes) which have a *prima facie* similarly absolute flavour (and so are suspect *if* Cusanus is remembered). But, more than 50 years after Russell's paradox, the faithful (in a then-new sect devoted to categories) blithely *wanted* a category of all categories (without any of the many qualifications which present themselves after a moment's thought).

Of course, this treatment conflicts with (theological) conventions, as follows.

First of all, the *style* is a breech of good manners in solemn circles; most simply, by its lack of respect for what is *holy*. It is of course a matter of routine to avoid it, *if* one wants to do so.

Secondly, the solemn tradition assumes that that *shift of emphasis* risks a permanent loss by distraction (from the full inwardness of *higher aspects*); a kind of mirror image to the refrain in this chapter about distractions from effective knowledge (by clutching at straws). In fact, that assumed risk is often presented as involving a loss of effective knowledge, too (especially by Gödel, albeit in an exceptionally innocent manner; cf. the end of [1944]). This overlooks at least two snags:

- In such complex situations it is simple-minded to assume that relations of cause and effect are appropriate at all, and even more to rely on flash judgement.

- It is equally simple-minded not to balance the account of such (assumed) gains against the cost (of futile pursuits of solemn idea(l)s).

This is the 'theoretical' side. More significantly:

- What do we know of the probability that the risk in question materializes?

Even for interpreting the solemn tradition it is, by experience, good policy to correct for the lopsidedness of its preoccupation which that risk introduces. As before, changing it would be a more difficult matter, requiring the practical skills of unbundlers. Their policy is to give their targets rope: they are a better bargain if they have pursued their idea(l)s further, when it is easier to see (that is, cheaper in the commerce of ideas) what to keep.

8.7 What Was Lacking (60, 40, or Even 20 Years Ago)?

Pedantically, 'lacking' concerning the main refrain of this chapter about alternatives to the idea(l) of a tree of knowledge growing from (logical) seeds.

60 years ago, there was just *lack of logical experience*; in particular, of what else to do (with knowledge of logic; besides growing and trimming trees in logical foundations).

40 years ago, one had *elementary results* which, at least when used with discretion (and in sometimes imaginative combinations with more specific knowledge), contributed effectively:

- The most striking example remains mechanical computation; naturally, more for what can be done with it than for its limitations.
- Malcev had published other such combinations about 50 years ago.[31]

But, certainly, those of my chums who took any interest in logic at all had a view of scientific knowledge that was still dominated by expositions of the relativity and quantum theories according to the ideal of a tree of knowledge (or, in Dirac's *tour de force*, almost along a single branch!). Not even Bourbaki's scheme (of *relatively few* basic structures to be used for *very many* combinations) was presented (by its authors) or recognized (by us) as an alternative to the foundational ideal. Besides, it was not familiar enough (to us at the time) for its effectiveness to be seen (by us).

20 years ago alternatives to that idea(l) had become spectacularly visible through *scientific experience*:

- Molecular biology, full of brilliant thoughts, just isn't a theory according to the stone-age ideal.

[31] For the record, nobody had drawn my attention to them even 40 years ago. I had found for myself some significant combinations with proof theory.

- Big science (naturally, by p. 165, when used with skills closer to big business than those in demand by crafts and guilds) had combined successfully (as it were, in parallel) ideas, people and technical apparatus.[32]

Without exaggeration, this experience corrected a parochial idea of *understanding*. New possibilities were established, to be compared to the discovery of new kinds of *solution* (in mathematics). But applied to logic, at least for most, this broader view merely shifted the emphasis *away* from the foundational ideal. One tied up loose ends, by solving (clearly formulated) familiar problems, and combined logic with more substantial mathematics (in commercial terms: with richer resources).

During the last 20 years, even to the outsider, spectacular *events in the commerce of material goods and services* have shifted the emphasis to aspects that have obvious parallels in the commerce of ideas. In particular, (genuine) *new markets* (beyond, as always, new bandwagons) have become prominent which, previously, were genuinely marginal or simply ignored by piety towards traditions of the trade (usually shared by management and unions alike). One example (cf. p. 158 of Chapter 7) is suggested by the discovery that attention to pollution can make (not only for legitimate, but) rewarding business. Here it should be added that it is even more promising in the commerce of ideas. To use (again) the words of Marx, but with opposite emphasis:

> in contrast to the case of material pollution, a change in interpretation is sometimes enough to change the world (of ideas).

It would not do to end on this note of smugness, as if events in the last 60 years belonged to the best of all possible worlds (of ideas). It would be a missed opportunity (at least for interpreting the past, even if we forget it when it is needed in the future) not to mention the following reminder:

> Just think of any (thing that strikes the literal or the mind's eye as an) object. Every relation to anything else (not only to what strikes the eye as a part) is a property of that object.

When this home truth (repeated in different words for more than 2500 years) is remembered, the *broad* ideal of a tree of knowledge is seen as little more than a blind spot; even when only knowledge of the object above is meant. The same applies to related ideals; for example, of a *complete description* (as a seed for that tree).

This ideal is *not logically defective*, since it is realized impeccably in Peano's or Dedekind's axioms: they relate the objects considered (the number series generated by the successor, and the field $\mathcal{R}$) to all objects in the universe of sets. But,

[32]For the record, 40 years ago it had seemed to me that people would merely get in each other's way in such enterprises.

by experience, the ideal is *sterile* here;[33] specifically, compared to the alternative of focusing on *suitable incomplete descriptions* (that is, abstractions).

The literary forms of mathematical logic are well suited for *illustrating* diverse possibilities by memorable examples without necessarily contributing to effective knowledge, too. For instance: examples of abstractions *can* be concocted, of which (as Plato's translators say) $\mathcal{R}$ partakes, but this is best proved by use of Dedekind's axioms and recondite properties of sets. (There are corners in the market for highly touted independence proofs, which establish a *logical* need for such properties.) There is also the separate fact of experience that such examples are *not encountered* often.

The metaphor of a commerce of ideas would be very weak indeed if trades supplying such properties did not advertize their concoctions, and others not dealing in them did not huff and puff. Those others need not be established trades; often they are vendors of other seeds, peddling theirs under labels like 'cybernetics' or 'information theory' (instead of the more venerable logical variety).

8.8 Appendix. A View of Non-Standard Analysis ([1974])

Gödel's view is presented in a little more than half a page, after a lecture by Abraham Robinson at Princeton in 1974. The aspect of Gödel's remarks most pertinent for our agenda, and particularly Section 7, is the following.

The *contrast between traditional logical ideal(l)s* (generally and, in particular, applied to mathematics) *and scientific experience* has become evident, especially in the second half of this century. It becomes particularly vivid in Gödel's remarks, which focus on the (in 1974 even) narrow(er) area of non-standard analysis and arithmetic. The contrast concerns both the interpretation of the results available, and expectations of the future; so to speak, concerning the centre(s) of gravity of a growing body (here, of analysis). By our refrain, such expectations will differ if knowledge is taken to grow like a seed from a tree or by accretion.

Gödel was singularly well equipped at the time to present that contrast (in effect, not necessarily on purpose): he had been out of touch with developments in mathematics in the preceding quarter of the century, and he had had practice in presenting logical ideals since the forties. Allowance should be made for strong language; for example, about non-standard analysis being 'the analysis of the future'. But it is certainly no stronger than the fanfares in his essays in the 40s.

Of course there is also (in terms used in note 24) a staggering contrast between, at least, the literal meaning of what he preaches here and his practice in his

[33] For any remotely realistic sense of 'sterile', the qualification 'here' is *obviously* necessary. Euclid's presentation of (his) knowledge of geometry in the form of a tree is not only (eternally!) legitimate: it also had a market among educated Greeks of his time, and had a good run as a blue-chip-investment in the commerce of ideas.

metamathematical papers. But it is no greater than the contrast, for example, between Hilbert's peroration about purity of method at the end of his *Foundations of Geometry* and his practice in ordinary mathematics.

The following three samples will do.

The tree (of knowledge) of numbers

On Gödel's view this tree becomes a single branch, filling gaps (from $\mathcal{Z}$ to $\mathcal{R}$, with $\mathcal{C}$ regarded as a minor excrescence):

> Arithmetic starts with the integers and proceeds by successively enlarging the number system by rational and negative numbers, irrationals numbers, etc. But the next quite natural step after the reals, namely the introduction of infinitesimals, has simply been omitted.

Non-standard reals are then the next step, if not the holy grail.

This forgets differences between them and other non-archimedean fields, which are not even mentioned by Gödel (though certainly significant for effective knowledge). But, by mathematical experience, other omissions are more serious.

For one thing, Gödel blithely disregards the risk of a point of diminishing returns in the pursuit of any holy grail, here of filling gaps.

More specifically: certainly, as far as number theory goes, branches that are totally overlooked by Gödel are at least as prominent in mathematical experience; for example, finite fields or p-adics (the former, like $\mathcal{C}$, differ in not having an ordering compatible with $+$ and $\times$). Only the *logical* order of priority puts these objects low down.

Ordering theorems by logical implication

Gödel tacitly applies this to theorems (proved at the time by non-standards methods) about invariant subspaces for (polynomially compact) operators, and disregards the quality of the 'improvement' of Robinson's result:[34]

> Non-standard analysis frequently simplifies substantially the proofs, not only of elementary theorems, but also of deep results. This is true, e.g., also for the proof of the existence of invariant subspaces for compact operators, disregarding the improvement of the result.

Gödel also refers to 'other' cases, of which there were few at the time. One of them involved non-standard notions in the statement itself, since it connects function fields of basic arithmetic with number fields in (suitable) non-standard

[34] Tacitly, in Lomonosov [1973]; of course, even without knowing the meaning of the words used, the later result by Lomonosov is seen to imply Robinson's, but its quality requires closer attention.

models of arithmetic. So, on the surface, it resembles Higman's gem which (also) connects logical notions and ordinary mathematics (now, for recursion theory and the theory of finitely presented groups). But closer inspection of that other case was summed up by a perceptive mathematician as follows:

> No wonder you get such a connection, if you call a lot of strange objects 'non-standard integers'.

In other words, as matters stand at the time, knowledge of function fields told you quite a lot about non-standard models, with precious little in return. The assumption that one day the balance of trade *must* be reversed is, by experience and as in ordinary commerce, touching (at least in the young).

Of course, literally, the connection constitutes new knowledge since it cannot even be stated without non-standard concepts. It is a *new* truth, and by the logical tradition this has priority over choosing *among* truths.

Concrete numerical problems

In Gödel's words:

> compared to the enormous development of abstract mathematics, the solution of concrete numerical problems [like Fermat's conjecture] was left far behind

(in then-contemporary mathematics). By implication, and again in accordance with the logical meaning of the words used, *concrete* and *abstract* are seen to be in (would-be fundamental) opposition.

Partly, plain ignorance was involved. In 1970 Baker received a Fields Medal; exaggerating very little, for bounds, in terms of $k \in \mathcal{Z}$, on $x, y \in \mathcal{Z}$ that satisfy $x^2 = y^3 + k$.

Partly, it was thoughtlessness: what could be more 'concrete' than finite fields? Especially, since metamathematical properties like consistency and other Π^0_1 properties are concrete in Gödel's sense.

Compared to these oversights it is a minor detail that, even today (with many memorable contributions, of which one will come up below) Robinson's *logical*[35] versions of non-standard arithmetic and analysis have not been used to solve such concrete numerical problems. On the contrary, such well-known results as Tchebotarov's theorem have been used imaginatively for work on non-standard models.

[35]'Logical' must be stressed, since otherwise number fields are also non-standard 'versions' of $\mathcal{Q}$. The single most striking difference is, of course, that logical versions express the analogy involved in terms of logical classifications (of the properties preserved).

Disclaimer

Ever since the 60s I not only had no qualms about non-standard methods, but was in the market for information about them; however, not for such traditional reasons as presented by Gödel.

For example, I realized (and this was comfirmed by people familiar with the subject) that the problem about invariant subspaces actually solved by Robinson was simply *not a contribution* to effective knowledge in that area. But, to me, it *illustrated* vividly a particular potential of his method: an efficient representation by a single (infinite) nonstandard element of iterated limiting processes. Quite apart from literary talent, I could not possibly have expressed this thought as compellingly as van den Dries and Wilkie in the 80's in the introduction to their [1984]. They had realized this potential by realizing that it contributed to an unquestionably substantial piece of knowledge: Gromov's theorem on finitely generated groups of polynomial growth (a different matter from polynomial compactness of operators).

A report on impressions of non-standard analysis in a different quarter, in the late 60's

Given my own impression described above, also of Robinson's other work, it was natural for me to look for support of my proposal to have him elected a Fellow of the Royal Society; not even a Foreign Member, since he had British nationality.

It had been already established that there was then no 'prejudice' against logic. By statute, initial support has to come from within the Society. On the fact of it this seemed easy (to me). Robinson had not worked on, say, large cardinals or Turing degrees, on which the people involved simply could not be expected to have *informed* views (and I for one did not expect support for my proposal on the basis of uniformed views).

Robinson's (cl)aims were stated in ordinary mathematical terms; in several books, at length, and with a good deal of repetition. His style may not have been everybody's cup of tea, but then mine (though different) isn't either. By leafing through his publications it is certainly possible to get the gist of it, provided at least a few results catch one's attention. But the general response to my proposal was very cool.

Probably, I had underestimated the extent to which would-be advertizing of Robinson's invariant subspace theorem had become known. It has to be admitted that it involved a quite staggering lack of understanding of the subject in question. Be that as it may, this theorem was brought up, and nobody (in the trade) would want to 'disregard the improvement'. I remember the shift of emphasis in the *Disclaimer* above (to illustrating potential) so well, because *this* was acceptable. But it was pointed out that this point was not to be found in *his* books.

Somebody with a temperament different from mine might have pursued the matter successfully. As I saw it, it was the Society's loss, not Robinson's. By chance, in a review of Robinson's book on metamathematics [1955] in the mid fifties, I took the same view of lack of interest in logic: it was their loss, not ours.

For the record, Gödel took a completely different view. Not only his personal, but his professional loyalty was very strong. I have too little interest in such matters to speculate sensibly in what way, if any, this played a role in his remarks; let alone, whether anybody in the audience later helped in the election of Robinson to the U.S. National Academy of Sciences. If it did, one could learn a lesson about the way this world works (in terms used repeatedly in this chapter):

> It can be useful, even if it is not logically necessary, to be able to say with a straight face: 'Non-standard analysis, in some version or other, will be the analysis of the future'.

Chapter 9

Gödel's Last Remarks on The Undecidability Results

Gödel's three remarks in [1972a] occupy barely 2 pages, and each of them contains:

- (more or less explicitly) some *simple point* which, for more than 50 years, has proved to have (what is called in current mathematical jargon) foundational significance, but is not prominent in the literature;
- (more or less implicitly) some *mind-boggling assumption* which has been instead prominent in the heroic tradition.

The word 'undecidability' applies:

1. In the first two remarks, to *particular formulas* and their (formal) *independence* (as in Gödel's [1931])
2. in the third remark, to *classes of formulas* and their (recursive) *unsolvability* (as in Turing [1936]).

The mathematics used in 1 and 2, such as diagonalization, is (exceptionally simple and) closely related; but the choice of notions and problems in further elaborations is very different. Also, Gödel's pearl does not turn up in the ordinary mathematical literature (cf. Chapter 7), while Turing's twist has an established place (for example, through Higman's theorem, in the subject of finitely generated groups).

The titles of the sections below are Gödel's own (for his remarks).

[0] Originally published in *Notre Dame Journal of Formal Logic*, 31 (1990) , as Appendix I to 'Gödel's Collected Works, Volume II'.

9.1 The Best and Most General Version of the Unprovability of Consistency in the Same System

With this would-be dramatic title, the remark itself cannot help being *plus serieux*. Another (catchy?) title would have been:

> How I never had the courage of my convictions expressed at Königsberg in 1930 about consistency as an adequacy criterion.

Indeed, Gödel pointed out in [1931a] (tacitly, even in the particular case of formal systems for arithmetic) that, at best,[1] consistency is adequate for ensuring the validity of (what we now call) formally proved Π^0_1 sentences. In current terminology and notation, this restricted adequacy is:

- Π^0_1-*reflection*: for all formulas $F \in \Pi^0_1$, $(\Box F) \Rightarrow F$.

Gödel's terminology in [1972a] is 'outer consistency'.[2]

Gödel incompleteness paper [1931] provides a formula G which *is* $\neg\Box G$. Pedantically, this is done for the particular system of *Principia Mathematica*, a particular coding of syntactic objects, and particular definitions of the syntactic relations involved. But never mind for the moment the flourish about 'the most general version' (with its innocent disregard of the most elementary conventions about the definite article).

Theorem 9.1.1 *If:*

1. *the system is consistent*
2. $\Box G$ *is provable if* G *is*

then $\Box G \to G$ *is unprovable.*

Proof. Since G *is* $\neg\Box G$,

$$\Box G \to \neg G$$

[1]'At best' because some additional condition, such as (what we now call) Σ^0_1-*completeness* (cf. note 2), is (obviously) required.

[2]If we define:

- Σ^0_1-*consistency*: for all formulas $F \in \Sigma^0_1$, $\neg(\Box F \wedge \Box\neg F)$
- Σ^0_1-*completeness*: for all formulas $F \in \Sigma^0_1$, $F \Rightarrow \Box F$.

then Π^0_1-*reflection follows from* Σ^0_1-*completeness and* Σ^0_1-*consistency*, as follows.

If F is Π^0_1, then F is equivalent to $\neg F'$, where F' is Σ^0_1. Suppose $\Box F$, i.e. $\Box\neg F'$. By Σ^0_1-consistency, $\neg\Box F'$ holds. Then, by (the contrapositive of) Σ^0_1-completeness, $\neg F'$ (i.e. F) holds. Thus $(\Box F) \Rightarrow F$.

is provable in the system. So, if

$$\Box G \to G$$

were also provable, we would deduce $\neg\Box G$.

Again, since $\neg\Box G$ *is* G, G itself would be provable. Now, by 2 (a minimum condition, used in the incompleteness theorem), we would deduce $\Box G$.

Thus $\neg\Box G$ and $\Box G$ would both be provable, and the system considered would be inconsistent, contradicting 1. $\Box$

Corollary 9.1.2 *(A particular instance of)* Π^0_1*-reflection is unprovable.*[3]

Proof. Gödel's formula $\Box G$ is in Σ^0_1 form, and thus G (which is the negation of $\Box G$) is, provably in the system, equivalent to a Π^0_1 formula. Thus $\Box G \to G$ is an instance of Π^0_1-reflection. $\Box$

This is all; it would have been perfectly compelling in the early 30's. By scientific experience, and contrary to the teenage idea(l) implicit in the flourish about 'the most general version', it would have been premature then to try and establish *suitable* generalizations; in other words, relatively few that cover relatively many formal systems (in broad experience). If one tries, as people did, one ends up with inanities about 'natural' systems (cf. note 32 of Chapter 7 on the *obvious* poverty of all formal systems considered; poor for representing the phenomena from mathematical experience).

Points to note today

First, by reference to experience with (pretty) *provability logic*. This uses heavily not only both:

- closure under modus ponens
- Σ^0_1-completeness (cf. note 2) for $\Box G$

but also the provability of those properties of the system in itself. (It is an open secret that the contemporary trade of provability logic would be out of a job without these two properties.) The particular (unprovability) argument above, advocated by Gödel [1972a], *uses neither property*.

Secondly, this need not be the end of the story (at least, not for those prepared to learn from scientific experience): emphasis is shifted to the significance, if any, of *formal systems that do not have those venerable properties* (and thus to the significance of avoiding the latter for a particular result). Cut-free systems are (now, not in 1931!) familiar enough, the other kind less so (but cf. the subsection

[3] By note 2, it follows that Σ^0_1-consistency and Σ^0_1-completeness are not both provable.

on p. 148 of Chapter 7). Let there be no mistake: this shift is *in conflict* with teenage ideals (especially, of a complete description for the essence of proof).

Thirdly, in terms of the metaphor on pollution (used in Chapters 7 and 8), since gushing about (Hilbert's banner) consistency is still around, one may wish to give it attention and, perhaps, thereby develop immunity. One option is to dot the i's and cross the t's by:

- considering different formulations of consistency (by Hilbert, who was never tired of emphasizing that they are equivalent; tacitly, for systems S that *do* have these two properties);
- listing other S for which the different formulations are not equivalent (a salutary preparation for Hilbert's many, tacit or even glib, assumptions in this area).

All this was done in the 50's and 60's, but (in Gödel's words in [1972a]) 'it has not received sufficient notice'.

Gödel's wilder side: legalistic (debating) points

In [1972a] his concern is to 'refute' Hilbert (on the latter's terms) in a court that insists on the letter of the law and relies on precedents:[4]

> it [would be] necessary to prove this 'outer' consistency of S [that is, Π^0_1-reflection] ... in order to 'justify' the transfinite axioms of a system S in the sense of Hilbert's program.

His concern fits the (to the modern reader strange) prominence in the remark of *provability in his equation calculus*, say $\Box_0$, and of the idea(l) of *primitive recursion* (in, of all things, a would- be most general version!).

Now, Hilbert had a thing about (formal) derivability. The harmless little word 'true' was banned, and not even applied to purely numerical propositions $P(n)$. Gödel respects this little whim by a rephrasing, permitted by the following theorem of *Principia*: for numerals $\overline{n}$,

$$(\forall n \in \omega)[P(n) \Leftrightarrow \Box_0 \overline{P(\overline{n})}].$$

Also, in the mid 20's Hilbert committed himself to Ackermann's function as *finitist* (meaning: privileged for the cheerful tradition, legitimate for the other). Now, Ackermann's function enumerates all primitive recursive functions, and so any attempt to make do with less would be teratological; even for that 'most general version' since, tacitly, it is meant for the refutation in question.

[4]For the record, I could not bring myself to do this, but I found the spectacle of Gödel at it simply enchanting.

Gödel's style

As mentioned earlier, read with ordinary horse sense, Gödel's remark is easy enough to follow. But Gödel's refusal to use the literary forms of mathematical logic for the sake of very compressed solemn language interprets Ezra Pound's recipe for great literature (as being 'simply language charged with meaning to the utmost possible degree') too innocently. I am not persuaded that the simple thoughts in Gödel's remark are rewarding subjects for great literature at all; let alone, in the particular traditional language favoured by Gödel. That's the way the cookie crumbles.

9.2 Another Version of the First Undecidability Theorem

The wording of this remark, including the title, may be a little strange; at least, if Gödel's poor health at the time (1972) is disregarded. But the thoughts (I read into or) in them have been long familiar to me: the simplest since before I first met him, the wilder ones from our conversations (even in his good old days).

Ever since the 50's a little cottage industry has been busy classifying formal systems (pedantically, the corresponding sets of theorems) according to their degrees of unsolvability (as in Turing's meaning of 'undecidability').

But, also since that time, this classification has been known to be insignificant (in the ordinary sense of statistics) for many parts of logic, which were prominent then (and have become more so). For example, in modern terminology and notation:

- Σ_n-*induction*: induction for all formulas $F \in \Sigma_n^0$

for different n: their theories are all of degree $\mathbf{0}'$, but differ even with respect to their Π_1^0-theorems (prominent in the last section). This was known then for $n > 2$, now even for $n \geq 1$.

Last but not least, these differences were established by (more efficient use of) the ideas in Gödel's own proof of the first undecidability result. Pedantically, this would not be called 'another version' of the latter. But a suitable version, which implies those differences as corollaries, is easily formulated by routine use of the literary forms of mathematical logic, and proved by means of those ideas:

> in order to solve all problems of Golbach type [that is, Π_1^0] of a certain degree of complication k one needs a system of axioms whose degree of complication ... is $\geq k$ (where the degree of complication is measured by the number of symbolds necessary to formulate the problem (or the system of axioms) ...).

Gödel's wilder side: understanding

The remainders above emphasize the fact that sometimes classification by degree is useful, sometimes by inclusion. They distract from the (more demanding) thoughts needed to determine what significance, if any, either classification has in a particular area of experience.

Lack of this kind of thought (in the sense of *Tractatus* 6.21) spoils the wilder sides of Gödel's second remark on such matters as *understanding* (here, of mathematical concepts and axioms about them):

> all present day mathematics can be derived from a handful of rather simple axioms about a very few primitive terms. Therefore ... the few simple axioms being used today will have to be supplemented by a great number of new ones or by axioms of great complication. It may be doubted whether evident axioms in such great number (or of such great complexity) can exist at all ... But ... more (and ever more complicated) axioms appear during the development of mathematics.

First, readers of Chapter 7 should recall note 32, but also p. 152 on *what* one wants to understand, as opposed to Gödel's preoccupation with the mere 'existence of mathematical yes or no questions' with some heroic, traditional property; like being 'undecidable for the human mind'.

Secondly, less specifically, the general idea of Gödel's (second) remark is commonplace for ordinary mathematical experience: focusing on formal deductions from given axioms is by no means obviously a first step *toward* a realistic view of mathematical reasoning. Its formal aspects are inadequate not only as far as discovery is concerned, but also understanding (including checking); both of theorems and of proofs. Thus *formally unnecessary methods can be essential*; for example, for reliability by cross checks (of numerical calculations by use of general theorems; in higher mathematics, of deductions from axioms for real closed fields, which are formally complete, by means of topological methods, for which there is no similarly complete formalization).

Thirdly, as so often: though all this is commonplace, it is also in conflict with heroic perennials (for example, of 'purity of method').

Gödel's remark compounds the conflict by uncritical (and possibly even premature) precision about the matter above. It pays no explicit attention to what, if anything, (its) precise meanings for words like 'understanding' contribute by way of effective knowledge.

Gödel's wilder side: complexity and abstraction

To fix ideas, the samples below concern the systems of Σ_n-induction mentioned earlier.

Gödel trots out familiar (logical) parameters, but now meant as *measures*. Thus n itself is for (degree of) *abstraction* (in other words, determined by counting quantifiers[5]): the number of symbols occurring in a formal object for (the complexity of) understanding the thought (theorem or proof) represented by the formal object. Naturally, there is no limit to elaborating this numerology.

But what we know of those matters (never mind the many things we don't know) is enough to show that the measures chosen are at cross purposes to the meaning(s) for which abstraction *does* contribute; for example, the passage from $\mathcal{Q}$ to abstract fields.

This is not all.

A solemn assumption and one alternative

The pious *assumption* is that views should be established or defended on terms set by the opponents. Here the commonplace view from mathematical experience is meant, and its opponents are proponents of strong *AI* (who were called 'formalists' 100 years ago). The assumption is enshrined in Turing's test, which requires mental capacities to be measured solely by (the sets of) formal *results* obtained, not *processes* (cf. Chapter 7). Accordingly, (pious) defences of the ordinary view attempt to rely on incompleteness properties. The weakness of any such attempt spawns (valid) objections, which then attract attention to that weakness and away from the strength of the (commonplace) conclusion. Except for the commotion produced in this way, the net result is a step back.

Now, certainly, a most obvious *alternative* is to look at (if you like, just the conscious aspects of) mental processes of human computers, where the *results* generally *do* agree with those of the electronic variety! (one simply looks at the execution of formal rules such as substituting one formal expression for another). Human computers have additional resources; including, for example, appropriate (so-called *ad hoc*) interpretations of formal symbols. If one wants to know about the (biological) resources available, one will be well advised to look at them; not merely at possibilities of replacing them for (sufficiently) similar results by suitable software engineering.

By Sections 6 and 7 of Chapter 8, readers must expect many different things to come to mind at this stage (and unless they are very unlucky, all of them more rewarding than logical straws such as the 'measures' above):

- Those with a classical background will think of Aristotle advice in *Metaphysics* (Γ, 5, 1009a 16–22) on how to treat opponents who object mechanically; perhaps fittingly, if they are proponents of mechanical resoning.

- Those used to ologies (cf. Section 6 of Chapter 8) may have to be remainded of the common-or-garden varieties of science, where there are lots of familiar

[5]Another favourite parameter relates abstraction to *higher types*; not 'type' in the ordinary sense (of 'sort' or 'kind'), but as in 'functions of higher type' or in axioms of infinity.

things to look at (such as conscious aspects of mental processes mentioned above). Some of them will come up again in Section 3.

Anecdote (about Gödel's good old days)

Gödel's own attachment to those (literally superficial) measures came up in our conversations first at a luncheon at Pennington, in a cottage which I shared with Dana Scott. (The occasion was Friedberg's visit to Princeton after solving Post's problem.)

Gödel talked with obvious warmth about the little note [1936] he had published around 20 years earlier (whose general idea is perfectly clear from the reminders above). There was an obvious, so to speak solemn, contrast between his faith in those formal(ist) parameters and his reservations (in particular, at Königsberg) about Hilbert's aims (in his program). But this presented itself to me (then and now) as a kind of aberration or blind spot.

The temperamental side, his *attachment*,[6] has remained for me much more vivid, partly because of a coincidence. Less than a year before that luncheon I had learnt from him his interpretation of Heyting's arithmetic. For me a principal attraction of this was *as a change* from the no-counterexample-interpretation, with which I had been familiar for barely 10 years (and had used effectively and repeatedly in the meantime).

9.3 A Philosophical Error in Turing's Work

By experience with the academic sense of 'philosophical', the title gives fair warning: an elementary blindspot is meant which (like other, even literally superficial, errors) can have profound consequences. By that same experience, Gödel must be expected to be solemn where Turing was (by ordinary standards, not those at Cambridge at the time) particularly breezy.

Turing proposed, in what most readers of his paper [1936] would have regarded as an aside, a (compactness) argument to 'establish' that, at any given moment, there can be only finitely many states of mind. His rather quaint hypothesis was that otherwise there would be confusion.[7]

Granted that at any stage there are only finitely many (conscious or unconscious) states of mind (or, in some equally rough sense, of the brain), Gödel notices that this leaves open *how* any sequence of subsequent states continues or

[6]Incidentally, he remained attached not only to his own discoveries, but also to knowledge he had got the hard way (for example, from uncongenial literature).

[7]For the record, I know at least one mind that gave Turing the opportunity (a few years later) to correct his, let us say, ideal of the human mind, and to remember (from his Tripos questions on ideal fluids) that idealizations need not be even first steps towards understanding the phenomena meant.

(so to speak) grows; especially if the capacity of the mind (which he calls 'faculty' in the supplement to [1964]) grows, of course, over and above its memory:

> what Turing disregards completely is the fact that *mind, in its use, is not static, but constantly developing* ... Therefore, although at each stage the number [of *distinguishable states of mind*] may be *finite*, ... [it] may *converge toward infinity* ...

As a *correction* to Turing's breezy (strong) *AI* Gödel's reminder remains compelling, however weak his *attempts at contributions* to effective knowledge of that faculty may be; cf. note 32 of Chapter 7 on doubts about assuming that the discovery of (logically) new axioms is a very rewarding function of that (mathematical) faculty.

As matters stand today, Turing's focus on those functions which the human mind has *in common* with (conventional) machines (but perhaps less so with the minds of animals) has been more rewarding; tacitly, as objects of theoretical understanding.

Thoughts by association with the topic of finiteness

The additional background here is the, by now, standard material on *recursiveness* applied in various parts of mathematics that serve as the languages of theoretical sciences. As a corollary, so to speak for the *analysis of language(s)* of this kind, the notion of *mechanistic theory* presents itself; for such a theory, all its (scientifically interpreted) aspects are recursive (for example, solutions of partial differential equations).

Despite pitfalls in interpretation (for example, of so-called initial conditions, cf. Kreisel [1982]), the general idea of this mathematical property of theories is clear enough. A problem comes from the ordinary separation between observational knowledge and its theoretical interpretation(s):

- on the one hand, data of the observational kind are (hereditarily) finitely described
- on the other hand, any such (necessarily finite) set of data is recursive.

Evidently, only the most coarse-minded would conclude from this that the mathematical property above (of being mechanistic) is without *any* scientific significance. An obvious question is:

where, if anywhere, is such a significance?

In other words, recursiveness is an *infinitistic* property, and so its interpretation is more demanding (in imagination).

Vast experience in classical physics (in particular, *PDE*) can provide some direction:

- Some *infinitesimal* conditions (on solutions being once or twice differentiable) are, often demonstrably, mathematically significant; most simply, for admitting or excluding a particular PDE as (even) a candidate for a theory of the phenomena (pedantically, of those aspects of them which are) considered. But again, every observational set of data is consistent with those conditions and also with their negation.

- More recently, these infinitesimal conditions have been *discovered* to have significance; roughly, when the PDE of continuum mechanics are viewed as (classical) *limits* of theories for molecular or quantum phenomena and the like (cf. Berry [199?]). Not surprisingly, this requires particular attention to matters that used to be brushed aside by claiming that we understand the classical phenomena 'in principle'. Also not surprisingly, the implications at a miscroscale of the infinitesimal conditions are, as one says, qualitative (or 'matters of principle').

Viewed this way, and in the absence of any corresponding micro *theory* of the mind or brain, pedantic elaborations concerning recursiveness are premature; as it were, a confusion in kind, not only in degree.

Autobiographical digression on effective rules

The topic came up repeatedly in conversations with Gödel during the sixties, but was not pursued. Whatever the reasons may be, our interests were certainly very different:

- He was preoccupied with giving a conclusive proof of the difference between (effectiveness for) mind and matter (cf. p. 126 of Chapter 6).

- My principal (conscious) interest was to give vent to metaphysical anger. Specifically, both extremes (in current jargon, wide-eyed enthusiasts of AI or digital intelligence, and indignant critics) never give a thought to the possibility that we already know *enough to refute* their rhetoric; but, presumably, *not enough to settle* any significant issue. In other words, we may not have reached the kind of *threshold* for genuine progress.

During a stay at the Institute for Advanced Study in the sixties, I prepared the survey article [1965] (about which I spoke very little to Gödel, since he regarded such activities as a waste of time). Footnote 29 in it mentions the possibility that, at least, statistical mechanics may be demonstrably non-mechanical.

I returned to the topic (of rules or, equivalently, inputs that are effective for physical systems for which a theory is available) in [1970] and [1974]. The former excludes some plausible candidates for systems that have non-recursive outputs for some recursive inputs, by adding physically relevant conditions to a simple-minded formulation;[8] specifically, in *percolation problems*.

Since then, also physicists have shown an interest; for example, Geroch and Hartle [19]. The logically more conscientious work on the topic above, by Pour El and Richards, is however spoilt by lack of respect for the physical meaning; cf. Kreisel [1982] on their abuse of the notion 'initial value'. On the positive side, their work has improved my understanding in two ways:

1. On the formal side, the boring lists at the beginning of [1982] (of recursive operators that do and do not necessarily have recursive values for recursive arguments) can be subsumed compactly in terms of *boundedness*; cf. Pour El and Richards [1983], and Turing's 'philosophical error'.

2. More generally, Pour El and Richards [1981] refutes the idea (I had for some time, without mentioning it in print) that people working on a problem (say, the 3 body problem) might make a tacit assumption: sound solutions must be recursive in (what are regarded as) the initial data. In this way they would simply miss non-recursive solutions.

   Pour El and Richards show that this was not the case with Kirchhoff's solution of the wave equation.[9]

Before examining the particular rules for the perfect (intuitionistic) mathematician, I looked at another area of formal work with the potential of being relevant to the matter at issue: the consistency of (suitable formalizations of) *Church's Thesis CT* with intuitionistically interpreted systems, like the theory of species.

As might be expected, this was more to Gödel's taste. Actually, even an inconsistency would only show that *CT* cannot be *proved* by methods in the system considered (it would not necessarily furnish a rule that is effective for the perfect mathematician, but defines a non-recursive function).

Before I dropped this subject, I summarized my experience in [1972]. Here are a few points from it:

- If a system has the $\exists$-*property* and is *formal* (that is, the set of theorems is r.e.), then for any theorem $\forall x \exists y R(x, y)$ there is a recursive function f such that, for each $n \in \omega$, $R(\overline{n}, \overline{f(n)})$, is derivable.

[8]A Markov process with unique asymptotic behaviour, the additional requirement being *non-vanishing* probability of that behaviour.

[9]Non-recursive solutions are often indeed unsatisfactory as they stand. But, once recognized, they may be explicitly *excluded* for physical reasons, or they may suggest *new questions* that have more manageable solutions.

Cf. the famous example in number theory, where there was only the *suspicion* that no simple method decides whether a binary diophantine equation has *some* integral solution. So Siegel asked instead if it has *infinitely many*, and decided that question.

So, if only formal systems are regarded as 'precise', the $\exists$- property excludes the most direct refutation of Church's thesis indicated above; that is, by a proof of $\forall x \exists y R(x,y)$ that is convincing for the perfect mathematician, while no recursive f satisfies $R(\overline{n}, \overline{f(n)})$ for all $n \in \omega$.

Obviously, it would be a *petitio principii*, in connection with effective rules, to *assume* that only formal theories are precise.

- The idea of a *creative subject* proceeding in an ω-series of steps, and the 'axioms' for the property $\vdash_n A$ (A is proved at stage n) implying Kripke's principle are found to be implausible. Roughly speaking, ω is too easily grasped.

- An opposite extreme, as it were, is the idea of *transfinite progressions* (originally called 'ordinal logic' by Turing). In fact, the title of Kreisel [1972], which refers to this topic, is related to conversations with Gödel that are not mentioned in the paper.

  In the early seventies, at a very difficult time in his life, he made several very long, mildly frantic transcontinental telephone calls (from Princeton to Los Altos Hills, where I lived at the time) about the present topic. He felt sure that, for *every* path through $\mathcal{O}$, *some* non-recursive predicate is computable (on any transfinite progression of the kind studied by Feferman). Gödel regarded this as *a refutation of CT*.

  It was not necessary to go into his idea, since it was practically apparent from Feferman and Spector [1962] that Gödel was simply wrong: no non-recursive predicate can be computed on any Π^1_1-path. He insisted that this should be published, and I simply did not have the heart to refuse. He agreed that, since the proof was so close to published material, and the result so inconclusive, it would be out of place to describe it as a refutation of a conjecture of his.

  I had done such a thing in [1958], partly through carelessness, partly for reasons explained on p. 265 of Chapter 11. Actually, Gödels' offhand habit (of making conjectures to visiting members at the Institute, who would then spend hours doubting their own, often completely trivial, proofs or refutations of those 'conjectures') was only brought home to me a few years later, by reports of a victim who spent a year at the Institute.

  My compromise, some 15 years later, was to do what he wanted in (the short) Part I of [1972], with an appropriate title, but use Part II for what I wanted to do myself.

Both Kreisel [1972] and Gödel [1972a] mention Turing's errors (or better, tacit assumptions). Of course, they are trivial compared to Turing's contribution: his focus on rules effective for computing machines had raised the level of the discussion not only far above the drivel about equivalent definitions of recursiveness, but also above Gödel's discovery of absoluteness of the latter (in [1946]). Here are a few additional details:

1. Gödel had surely noticed the *petitio principii* mentioned above (of assuming that only formal rules are precise) that struck me so forcibly, but I am not sure (since, at least in my experience, he never mentioned his independent discoveries, let alone priorities); cf. Kreisel [1972] for references to other instances of that *petitio*.

2. In [1972a] Gödel is particularly critical of Turing's assumption that we do not have enough room in our heads (as it were) for our mental processes to be governed by an unbounded operator. This neglects growth of our (tacitly, finite) intellectual equipment. In other words, Turing's and Gödel's remarks may be regarded as *jeux d'esprit* corresponding to the sober distinction between (the behaviour of) bounded and unbounded operators; cf. 1 on p. 208.

3. When Gödel mentioned 2 to me in the 60's, I brought up the more radical objection, mentioned in Kreisel [1972]: the operators themselves may be non-recursive. His (I believe considered, not offhand) view was that we know so little about the details that only very simple assumptions can be convincing. But he rejected the thought that we may know too little for *anything convincing*. He had no gratitude for small mercies like 1 and 2 on p. 208.

In the mid 70's (of course, after the aberration Kreisel [1972]) Gödel drew my attention to Hilbert's musings in lectures, not in print, on (what is called in Chapter 8) a market for his kind of *axiomatic analysis*, especially of elementary geometry. He proposed it as a substitute for pursuing sterile (cl)aims of contributing knowledge of the nature (or, in modern jargon, the concept) of space; for example, by sterile traditional logic chopping.

At the time Hilbert did not (and could not very well) estimate the (obvious) *risk*: of such axiomatic exercises distracting attention from more rewarding aspects of (something more or less like) space.

But by another fluke, which strikes the (mind's) eye if you keep it open, Hilbert himself had a chance to reassure us about that risk. He was among the first to take up Einstein's *shift of emphasis*: from space and time separately to space-time. (Nothing could be further removed from all that axiomatic analysis, some of which no doubt prettier than Einstein's mathematics.)

In short, for those with a modicum of negative capability, the axiomatic analysis is harmless enough; as somebody put it: this side of the pale.

For the record, I have always regarded exercises around Church's Thesis as similar substitutes. But before learning of Hilbert's musings, it was tedious to talk about this. Now I just tell the story above.

Chapter 10

Gödel's Excursions into Intuitionistic Logic

The topic of intuitionistic logic spans Gödel's working life, and was the main subject of our frequent conversations over some 15 years. It was neglected in Chapter 6, where effective *contributions* to scientific knowledge were given priority over *reflections* about the latter. In that chapter formal systems and sets were stressed, with some attention to various kinds of definability, since they have turned out to be more effective scientific tools than the intuitionistic notions; even for purposes that the latter have been claimed to serve.

Specifically, in connection with *constructivity*, the principal element of intuitionistic logic (of applying this requirement to proofs, and not only to definitions) is highly dubious; cf. Kreisel and MacIntyre [1982], and the following quotation from Gödel [1964]:

In fact, in the 40's Gödel emphasized *definitions*; both in connection with the ramified hierarchy (of the constructible sets) and with logic-free, essentially equational systems discussed in Section 5 (incorporating his successful experience with higher types, cf. p. 104 of Chapter 6).

In contrast, in the early 30's, as developed in Section 2, Gödel looked at typically intuitionistic, logical aspects of *constructivity*, and (as in other areas) returned to these interests in the 70's, albeit in a different style (not the early free display of sound intellectual reflexes, but reverent attachment to the tradition of academic philosophy).

It turns out that intuitionistic logic generally, and Gödel's contributions to it just mentioned can be used very effectively for a part of knowledge that is outside science, and incidentally quite close to the popular meaning of 'philosophy' (cf. Kreisel [1986] for more details). Roughly speaking, as elaborated in Section 1

[0]Originally published in *Gödel remembered*, Weingartner and Schmettered eds., Bibliopolis, 1987, pp. 77–120.

below, here the primary object is not merely to *pursue* traditional ideologies, and problems fundamental to them, but to *examine* them. And intuitionistic logic provides so to speak chemically pure specimens of such ideologies.

The text below is organized accordingly.

Advice to readers. It is probably best to leaf through the text below, letting the eye pause on titles, and to make selections for further reading. It is in the nature of the case that the range of interests, even within intuitionistic logic, of an exceptionally gifted person like Gödel will be exceptionally broad, and so selection is advisable.

10.1 Background and a Manifesto

Looking back at the early 30's, when Gödel did his best known work on intuitionistic logic, the first order of business was to help *clean up the logical pollution* spread by Brouwer and his epigones (comparable to Hilbert's logical atrocities with his claims for consistency). Gödel's distant, sometimes almost offhand (*hochnäsig*) style still seems fitting (cf. [1931a] on consistency). Time, if not Gödel's style, has dispelled the pollution, and it is appropriate to look at another side of intuitionistic logic

To put first things first, intuitionistic logic is most memorable as a reaction to an earlier, also would-be revolutionary [∞] enterprise: so-called *logistic foundations*. Among other things, the latter were said (for example, by Russell) to have exhibited the concepts implicit in our ordinary logical reasoning. Granted, logistic (that is, truth functional) logic may be better (for example, for reasoning well) than the logical notions of natural language. But they are certainly different! For example:

- In the propositional domain, $(p \to q) \vee (q \to p)$ is not a law on any ordinary reading of $\to$ and $\vee$.
- In large domains of natural mathematical language, $\exists$ is taken to mean that an instance can be defined (tacitly, in terms used in that domain).

  For better or for worse, $\exists$ is not regarded as an abbreviation for $\neg\forall\neg$.

Without exaggeration, intuitionistic logic wanted to approximate natural usage better, and succeeded.

So much is clear, though not often stressed. Much more interesting for the kind of philosophy adumbrated already (and explained in Kreisel [1986]) are early impressions (*alias* convictions) about intuitionistic logic. They were particularly *off the mark* where all sides *agreed*; specifically, about:

- some *intrinsic complexity* of intuitionistic logic

- (what we should now call) its *proof-theoretic weakness*.[1]

The faithful were prepared to pay this price for the sake of the Truth; critics saw this as a reason for ignoring intuitionistic logic, generally without investigating further (and, as we shall see in Section 2, some of those who did investigate, did not pause to reflect on the implications). For an examination of ideologies, as at the outset of this Chapter, systematic oversights are just relevant as any successes.

Implications for formal languages

The implications (or, as one says in mathematics, the corollaries) are obtained from formal investigations by one-liners. Nevertheless, experience shows that very often the investigations are not demanding (being done independently by different people), and it takes years before the implications are noted. In Hilbert's terms, here the latter are the building (*Bau*), and the formal work is the scaffolding (*Gerüst*).

As a logical implication of the principal point above, where intuitionistic logic is presented as a better approximation to (the logical features of a dialect, as it were, of) natural mathematical language, intuitionistic logic becomes a standard of reference for studies of other parts of other natural languages.

Generally, a look at the sophistication and elegance of intuitionistic logic during the last half century allows one to see that *threshold* where its study began to touch essentials; both regarding its relevance and its limitations. In the light of intuitionistic logic, other linguistic studies can be viewed realistically, and not as alleged pioneer work (for which only quite lax standards would be appropriate).

More particularly, the main limitation of intuitionistic logic is not, contrary to what is often claimed, the 'internal' difficulty of a formalization or of finding adequate (intended or contrived) meanings, but the superiority (again, for intended or discovered purposes) of *paraphrases*. (This is quite consistent with intuitionistic logic being elegant and satisfying, a virtue of all successful *jeux d'esprit*.)

Last but not least, intuitionistic logic provides a lively reminder of the fiasco of *natural history*, which studies phenomena that strike our untutored attention; in contrast to the now dominant *scientific tradition*, which relies heavily on 'artificial' constraints imposed by experiments. Here phenomena are isolated that lend themselves to rewarding study (as always, by something like available means).

[1] Incidentally, history repeated here the experience with *ramified types* during the first decade of this century, when all sides agreed on complexity (of course), and on the axiom of reducibility as its chief embarrassment. The latter turned out to be true for cardinal levels; cf. p. 105 of Chapter 6.

The last point above is developed at some length in Section 3, after a closer look at some classes of propositions (and operations on them) for which intuitionistic (but not classical) logic holds. In terms used above, they are kinds of propositions that strike our untutored attention, and incidentally have done so since Aristotle.[2]

Conversely, as it were, general considerations on natural languages throw light (in an obvious, but neglected way) on the so-called *creative subject*, a central element of the intuitionistic enterprise (see Section 4).

Intuitionistic rhetoric: some impressions and reminiscences

The presentation above of intuitionistic logic (as a modest correction of syntactic - that is, purely external - defects in the logistic scheme) may be sounder than the popular intuitionistic rhetoric, but it is less memorable. The orthodox rhetoric associates the difference between classical and intuitionistic logic with general philosophical 'positions' on the nature of mathematics (or of the world itself, if you want to go the whole hog): the key words are *objective* and *subjective*. They are familiar (in fact, hackneyed), and hence all the more memorable (even if one does not know too well what one remembers).

a) Diverse reactions to the rhetoric

One extreme is Bourbaki's, reflecting undoubtedly the view of the silent majority: intuitionistic logic is an *historical curiosity* (tacitly, to be ignored). This is a view of the rhetoric, and not of the details of intuitionistic logic, because these never got known.

An opposite extreme is the dramatic *thrill of a conflict*, even though already Georg Lichtengerb put the orthodox 'issue' above in perspective in his *Aphorismen*.[3] But also the deeper thrill of indignation is to be remembered here, which was triggereed by the 'menace' of intuitionistic logic.

[2]However, the classification cuts across such familiar grammatical categories as declarative sentences, etc.

[3]*Kantische Philosophie ist die gewiss wahre Betrachtung, dass wir ja auch so gut etwas sind als die Gegenstände ausser uns. Wenn also etwas auf uns wirkt, so hängt die Wirkung nicht allein von den wirkenden Dingen, sondern auch von dem ab, auf welches gewirkt wird ... Kantischer Geist ... die Verhältnisse unseres Wesens ... gegen die Dinge [,die wir] ausser uns [nennen,] ausfindig zu machen; das heisst, die Verhältnisse des Subjektiven gegen das Objektive zu bestimmen. Dieses ist freilich immer der Zweck aller gründlichen Naturforscher gewesen, aber die Frage ist, ob sie es je so wahrhaft philosophisch angefangen haben wie Herr Kant.*

If the parts in square brackets are kept one merely has relations between different parts of so-called subjective experience; not an equally clean separation. Whatever else *wahrhaft philosophisch* may mean, it is pretty certain that no natural scientist has succeeded in spreading out Kant's *Bethrachtung* over 700 pages. In particular, nobody before him has succeeded in giving comparable weight to Kant's reminder, a matter not to be despised.

At another opposite extreme there are studies that can be seen as *investigating the rhetoric*, even if the authors had different principal or at least additional interests (cf. the rest of this section). For example, for philosophy in the sense above, one may wish to put the rhetoric in its place; or one may wish to see if anything of any sober interest can be extracted from the rhetoric at all (as I tried from 1958–1963), or even something wonderful (as vendors of sheaf models seem to suggest).

b) Gödel's overt reactions

Whatever Gödel's research interests may have been, the styles of his early and later presentations differ sharply, as already mentioned at the beginning of this Chapter. As elaborated in Section 2, the early notes are concise and cavalier, apparently scoffing at the antics of the rhetoric. Later on, even where he disagreed, his comments on any kind of traditional philosophical concerns were respectful to the point of reverence.[4]

Though later Gödel used crude, hackneyed formulations that had proved to have popular appeal (and had put me off), in his very early writings he was more austere. For example, in the introduction to his dissertation [1929] he scoffed at the *Grundlagenstreit* (which Einstein had called a cat-and-mouse game), and soon afterwards, in [1931a], he treated (Hilbert's) claims that consistency was a sufficient condition for soundness similarly.

Towards the end of his life he is quoted to have said:

> intuitionistic logic is bad for mathematics, but important for foundations.

Before 1970 he never made such unbusinesslike remarks to me (at least, not in logic). The following twist is in line with this chapter:

> Intuitionistic logic has not so far proved to be a useful *tool* in the arsenal of mathematics, though it has been a quite rewarding object of (meta)mathematical analysis (cf. p. 220).
>
> On the other hand, it is a gold mine for foundations in the sense of Section 1; that is, for *examining* extreme ideologies.

[4] In the introduction to his dissertation he used the notion of *validity* as a matter of course, and later *truth* of arithmetic statements. This is very different from a non-constructivist *position*, which makes an issue out of accepting those notions; as indeed Gödel himself did in his later popular writings. (Chapter 11 contains more about Gödel's 'non-constructivism'.)

For the record it may be mentioned that, when Gödel said the word 'philosophy' with a trace of awe in his voice, his wife reminded him of his habit, back in Vienna, of stressing that he was a mathematician (incidentally, imitating his voice quite successfully). Actually, this habit would have served a good purpose of keeping philosophical pests at a distance; cf. p. 60 of Chapter 6.

This was not Gödel's sense; above all, not towards the end of his life, when he confided to me that he expected foundations (and philosophy generally) to tell us what the world is really like; as if science did not have this goal, too.

c) Reactions to Brouwer's style

At least in the 50's Brouwer's personal style, of haranguing his audience for hours, did not suit Gödel at all. Gödel complained about having to play the host; also, if I remember correctly, in one of his letters to his mother. Gödel was utterly bored by Brouwer, unlike several logicians and mathematicians who, being dry themselves, were buoyed by Brouwer's probably genuine exuberance (cf. p. 59 of Chapter 6 on Gödel's reaction to exuberance).

I never asked if he attended Brouwer's lectures at Vienna at the end of the twenties (as he presumably did). If he did, his reaction would certainly not have been very different.[5]

I do not suggest that these personal matters could be decisive. But if one's confidence in an enterprise like intuitionistic logic is shaky to start with, the performance of its chief exponent can give one the final push. I have a relevant anecdote myself.

My first encounters with Brouwer's style were in the late 40's, at the first of his lectures at Cambridge, and then at his (invited) lectures at University College in London. I was utterly bored by his exaggerations, and asked him after a lecture if he meant all he said. He quoted George Bernard Shaw on having to exaggerate to make an impression, in a style that made me feel he had used the quotation repeatedly. I pointed out, as innocently as I could, that Shaw had not promised him that he'd make a good impression. Incidentally, Brouwer was not amused. Apparently, he did not like to be interrupted anyway; fittingly, for a good solipsist.

Now, I certainly was sceptical of intuitionistic logic before I ever met Brouwer. For example, I am on record stressing the appeal of the more radical restriction to quantifier-free (in particular, finitist) schemes, if and when it is appropriate to be constructive in mathematics at all. And I had never taken seriously the principal preoccupation of foundations alluded to above, of exhibiting the logical laws implicit in ordinary reasoning. (Philosophers, who have this preoccupation and are interested in constructive aspects, feel obliged to be interested in intuitionistic logic simply because logical words occur in that reasoning.)

Sure, there wasn't much for Brouwer to spoil in my case. But I do remember that a phrase I used quite often in later writings, occurred to me during one of Brouwer's lectures: those iterated implications make my head spin (as they still do; just like higher types, their counterpart in logic-free mathematics).[6]

[5] Gödel (incidentally, like Brouwer himself) did not change his tastes, and was proud of it; he called any change of taste: *Mangel an beständigen Gefühlen.*

[6] The formal theory of those things, even in ramified set theory, is however quite elegant (as

10.2 Early Metamathematics of Systems for Intuitionistic Logic ([1932a], [1933], [1933b])

In the early 30's Gödel published a few notes on intuitionistic logic. As is to be expected, the results are superseded; incidentally, some of them were not demanding even at the time, inasmuch as they were found by others independently.

The style is refreshingly concise, almost offhand; in obvious contrast to the then prevalent heavy rhetoric of intuitionistic logic.

Less trivially, the notes are widely accessible; they (incidentally, like most of Gödel's publications) require little mathematical background. The price is high: the notes provide little indication of any domains where the result might by really relevant.

His comments, right or wrong, are almost uniformly rewarding as a reliable record of first impressions (on intuitionistic logic).

Below, most results are stated for the *propositional* part, which illustrates very well many properties of intuitionistic logic (cf. the very successful Chapter 1 of Chang and Keisler [1973] for classical logic). In intuitionistic logic the propositional part becomes particularly rewarding when its *quantifiers* are included.[7]

Negative fragments ([1933])

The *negative fragment* consists of the operators $\neg$ and $\wedge$ (with $\forall$ in predicate logic) applied to negated atomic formulas. By [1933], *the same formulas of this fragment are derivable in* (the usual systems of) *classical and intuitionistic logic*. Further, there is a quite efficient transformation of any classical derivation into a (generally different) intuitionistic one (with the same end formula).

Since the negative fragment is a so-called reduction class for the full *classical* fragment, the latter is thus embedded in intuitionistic logic (preserving most relations prominent in metamathematics). However, interpretation and scope of this kind of embedding are a delicate matter. Above all, there is the question:

> What is gained by having an intuitionistic (rather than only a classical) proof of a negative formula?

Evidently, granted intuitionistic ideology, the answer is trivial: now one has a *valid* proof. For philosophy in the sense of Section 1, this answer is worse then useless. It stops one from even looking for a convincing answer; for example, in terms of functional interpretations of formal derivations.[8]

---

long as one does not think of instances).

[7] Propositional quantifiers bring little in the classical case. At least, generally; there are some exceptions in so-called complexity theory.

[8] Derivations in intuitionistic logic are usually realized by operations that are continuous in some suitable sense, and so the translation ensures that an arbitrary realization can be replaced

Gödel's early notes do not touch the question above. But they raise quite a number of less delicate points that are still of interest. As already noted, Gödel's own answers usually reflect only first impressions; not only his, but (as shown, for example, by Skolem's review [1933] of Gödel [1933]) also by distinguished contemporaries.

a) Differences in meaning

Concerning *differences in meaning* between classical and intuitionistic operators, Gödel thought they were mild (*gering abweichende Interpretation*). This is like saying that the notions of countable and uncountable structures differ mildly because the same first order formulas are valid classically for both classes of structures.

If one wants to dismiss intuitionistic logic, one has to find a less hackneyed (metamathematical) property than conservation of classical logic over intuitionistic logic for the negative fragment. Incidentally, later Gödel became supersensitive about differences in meaning (cf. Chapter 12).

b) Proof-theoretical strength

Concerning the scope of his result, Gödel said at the end of his note that it might *fail* for so-called *impredicative systems*. This is doubly wrong.

First, the proof extends almost verbatim to the *theory of species* and, with a little care in choosing (among classically equivalent) formulations, also to *set theory*.

More subtly, as Gödel had observed himself, the result extends to formal *classical number theory*, though the latter isn't all that predicative either (at least, in the strict sense). Specifically, an object may be defined by a quantified property A (that is, for which $\exists! x A$ is derivable) without there being a numeral $\overline{n}$ for which $A[x/\overline{n}]$ is derivable (where numerals are typical of definitions 'independent of the totality of all natural numbers'). This is a corollary to the incompleteness phenomena.

But much more significant is the following oversight. Gödel had swallowed the then (and, incidentally, still) widespread superstition, mentioned in Section 1, about intuitionistic logic lacking proof-theoretic strength. Accordingly, he never noticed that this superstition was refuted by his embedding! It was a bewitchment, but not primarily by (clumsy?) language.

c) (Weak) completeness

As a corollary of the embedding:

---

by a continuous one. Cf. first-order formulas about real closed fields, where any realization can be 'replaced' by an algebraic one.

- propositional logic is *complete* for the negative fragment
- predicate logic is *weakly complete* (that is, the double negation of completeness holds).

The proof needs nothing beyond the embedding, except the realization that only very special kinds of propositions and predicates are relevant (cf. Chapter 4).

Yet, the corollary was not observed for a quarter of a century after Gödel's note on the embedding. This is fitting for somebody who, like young Gödel, regards the intended intuitionistic meaning as illegitimate or at least as sterile, and does not want to sanctify it by proving completeness for it.[9]

Beyond the negative fragment (optional)

On the principle:

What do they know of England who only England know?

the heart of intuitionistic logic, as already noted in Section 1, is outside the negative fragment: even intuitionistic rhetoric is dominated by talk about $\exists$ together with $\vee$.

a) Disjunction and existence properties

By now it is fairly generally recognized that these properties are not (or, at least, not generally) required by intuitionistic validity.

On the formal side, there are systems that do not have them; some were manufactured for the purpose, some introduced for other purposes were discovered not to have them.[10]

More instructive are the following reminders:

---

[9]Twenty-five years later Gödel pointed out that *Markov's principle* (in note 12, for $x \in N$ and A primitive recursive, possibly with additional parameters) *implies the incompleteness of Heyting's predicate calculus*.

[10]Friedman has shown that *all formal extensions of HA with the numerical disjunction property, also have the existence property*; the converse being obvious, since

$$(A \vee B) \leftrightarrow \exists x[(x = 0 \rightarrow A) \wedge (x \neq 0 \rightarrow B)].$$

Gödel eventually submitted Friedman's paper for publication, but only after worrying whether the result was really completely general. He had to be reminded of the (good) reasons for his worry! They go back to the widespread belief that the properties in question are needed for intuitionistic validity. Given this blindspot, there was suspicion that some tacit assumption had slipped into Friedman's proof and restricted the systems from the start, thus trivializing the result. In fact, when the paper was published, several outsiders were ill at ease about the paper just because of the blindspot. The paper is of particular interest precisely because those properties are not needed for intuitionistic validity; in short, it is not a mere curiosity, contrary to the impression conveyed by Nerode and Harrington [1984].

- In the case of *logic*, where propositional and predicate symbols are interpreted as variables (when a formula is called 'valid'), $A \vee B$ is prima facie comparable to $a = 0 \vee a \neq 0$ in number theory. The latter is true for all a, but neither $\forall a(a = 0)$ nor $\forall a(a \neq 0)$ are.

  The obvious difference is that, in the case of general validity (tacitly, for a wild stock of propositions or of domains and predicates in predicate logic) too little is known about these things to use their structure in associating effectively either A or B with each value of the variables. In that case, either A can be asserted outright or B; cf. the impossibility of separating the continuum 'continuously', where the characteristic property of continuity is that only very limited information about arguments is used.

- For (arithmetic) *formal systems with a specific interpretation*, incompleteness intervenes. So, if a closed formula $A \vee B$ is derivable and the systems is sound, either A holds or B; but, by incompleteness this does not ensure that either A is derivable or B. So, if a system has the disjunction property, its incompleteness with respect to disjunctions balances (as it were) its incompleteness with respect to the disjuncts.

Gödel probably noticed quite early the facts just discussed, but I am not sure. As it happened, I noticed them before I met him, and mentioned them to him soon afterwards. He took pleasure, as always when somebody else spotted a point that he liked himself (without, as usual, mentioning an independent discovery).

On several occasions Kleene has referred to a 'well know logician', evidently meaning Gödel, and his doubts the disjunction and existence properties. But he never elaborated just what was being doubted; neither the general distinction made above, nor the specific fact that Gödel had doubts about the existence property for HA (cf. p. 241).

b) Markov's rule and principle

In connection with (b) on p. 219, Gödel's result fails for systems with *function symbols*, which involve so to speak 'hidden' $\exists$-symbols.[11]

The most fruitful exception concerns $\forall\exists$-*theorems*,[12] to which the conservation result for the negative fragment has been extended, in various ways, over the

[11]Various forms of the axiom of choice are quite weak when added to systems of intuitionistic logic, but their negative translations are not (for example, in Spector [1962]).

[12]The results are also known under the proprietory names:

- *Markov's rule*: if $\forall x(A \vee \neg A)$ and $\neg\forall x \neg A$ are derivable then so is $\exists x A$
- *Markov's principle*: $[\forall x(A \vee \neg A) \wedge \neg\forall x \neg A] \to \exists x A$.

Their validity depends on the kind of predicate A considered; for example, neither holds for A with lawless parameters.

Incidentally, in accordance with the introduction to Section 2, there is a neat propositional analogue:

last 30 years. These theorems are, without exaggeration, typical of *algorithmic propositions*, prominent in intuitionistic rhetoric. Once again, the interpretation of the results is more demanding than the proofs, the latter having often been found independently by several people.

There were two stages in the interpretation:

- *adequacy-in-principle*
  Following the preoccupation with proof-theoretic strength, in (b) above, there was emphasis on the 'adequacy' of intuitionistic logic; in particular, for proving that a program is totally defined.[13]

  For classical and intuitionistic systems that 'correspond' in a natural and precise sense, the same programs are provably total. This establishes adequacy of intuitionistic logic as understood in the foundational tradition.

- *inadequacy-in-fact*
  More recently, and in line with the manifesto of Section 1, it was realized that adequacy in the *foundational* sense ensures *algorithmic* inadequacy.

  In fact, there are relatively simple proofs d of $\forall\exists$ theorems $\forall x \exists y R(x, y)$ (for example, of definition by transfinite recursion $< \epsilon_0$) that are hard to unwind. In other words, it is costly to extract or execute a program π_d such that $\forall x R(x, \pi_d x)$: this is algorithmic inadequacy.

  More formally, and more neatly, recent proofs of Markov's rule by Dragalin and Friedman show quite generally how easily *any* classical proof of a $\forall\exists$ theorem can be converted into an intuitionistic one; so intuitionistic logic is algorithmically no better than classical reasoning (for typical algorithmic problems); cf. Appendix 4 of Kreisel [1985b] for more details.

Above, in accordance with Section 1, the recent efficient transformations: $d \mapsto d_i$ (of a classical proof of $\forall x \exists y R$ or of an intuitionistic proof of $\forall x \neg \forall y \neg R$ into an intuitionistic proof of $\forall x \exists y R$) are used for a *critique* of foundational aims.

Within the foundational tradition, the advantage of the recent proofs of Markov's rule over earlier proofs would be seen in the use of more elementary metamathematical methods.

According to Section 1, this traditional view is itself distinctly problematic as long as there are no realistic doubts about the old methods, restrictions being as good candidates for justification as extensions.

1. if $(P \vee \neg P) \wedge (Q \vee \neg Q)$ and $\neg(\neg P \wedge \neg Q)$ are derivable then so is $P \vee Q$
2. $[(P \vee \neg P) \wedge (Q \vee \neg Q) \wedge \neg(\neg P \wedge \neg Q)] \to (P \vee Q)$

The principle 2 is not generally valid. Closure under the rule 1 follows by (an *ad hoc*) use of the disjunction property or, at the other extreme, by specializing some general fancy proof of the kind discussed in the next paragraph but one.

[13] Cf. Section 4 on effective rules for the perfect digital computer, and the distinction between (equational) programs with number e for which $\forall x \exists y T(e, x, y)$ can be proved classically or, respectively, intuitionistically (here Kleene's T-predicate is meant).

This remark brings us to a venerable general worry: *a seesaw in interpretations or shifts of emphasis* (*Akzentverschiebung* in psychoanalytic jargon). Just now, two stages in the interpretation of closure under Markov's rule were mentioned.

Where will it all end?

After all, the first interpretation as establishing adequacy (or stability of the notion of provably total recursive function) had a comforting look of finality about it.

The problem is genuine, and perfectly well recognized under such code words as 'dialectics'. But all these dialectical fireworks draw attention away from the *facts of scientific experience*. Time and again interpretations have settled down, just as expositions of various branches of science quite often reach a stable form; to be enriched, occasionally, by a more sophisticated vocabulary.

For example, already Goursat's *Cours d'Analyse* grouped the elementary parts of the subject in more or less the current order, except that today we give *names* to those groups: theorems valid in all Frechet spaces, topological spaces, metric spaces and so forth; realizing as it were the biblical idea of paradise (in *Genesis* 2, 19), where God brought Adam the objects He had created, and Adam gave them names; presumably, thus coding (his knowledge of) the principal properties of those objects.

Incidentally, Gödel himself had a horror of shifts of emphasis (which he would have called 'shifting one's ground', if the question had arisen), and saw in them a principal reason why philosophy made so little progress. Without exaggeration, it is more likely that attachment to a few (sterile) interpretations and to problems fundamental for them, has hampered the progress of philosophy than too many or too imaginative shifts of emphasis. (At least occasionally, the silent majority's practice of ignoring those interpretations is tantamount to attachment by - benevolent - neglect.)

General provability and formal derivability ([1933b])

Gödel's note [1933b] contains a *translation of intuitionistic propositional logic into one of the systems of (classical) modal logic*. The additional operator $\Box$ is variously interpreted as some kind of necessity or provability; cf. (c) below.

As it stands, the note does not go far. Gödel had simply focused on one item in the rhetoric of intuitionistic logic; in particular, on the alleged opposition between *truth* and *provability* or, in modern jargon, between truth and assertability conditions. He then tried out the first formalism at hand with the smell of that opposition. Later, the note was refined by others who showed that *the translation was faithful*. Gödel's own result was enough to establish simple metamathematical properties of intuitionistic logic (for example, the underivability of $p \vee \neg p$ by Heyting's rules).

But the note is a good peg on which to hand various observations of more permanent interest.

a) Soundness and incompleteness

Gödel himself used $\Box$ to improve his original formulation of the first incompleteness theorem. Instead of talk about the Liar, which sounds merely frivolous to most of us, Gödel here interprets his independent sentence as an instance of the most natural property in the world, *soundness* (also called *reflection principle*):

$$(\Box F) \Rightarrow F,$$

where $\Box$ is now formal derivability in *Principia Mathematica* (and related systems).

This form establishes *incompleteness for intuitionistic systems of arithmetic* too, in the sense that some valid sentence is not formally derivable; evidently, here it is not enough that some sentence be formally independent.

The formulation is also superior to the second incompleteness theorem (about consistency), in the following respects:

- First, in being applicable to a broader class of systems; for example, not requiring demonstrable completeness with respect to Σ^0_1 sentences.[14]
- Secondly, by referring directly to $(\Box F) \to F$ (for $F \in \Pi^0_1$); this is ensured by consistency (modulo Σ^0_1-completeness), and is the only reason for regarding consistency as sufficient for any kind of soundness.[15]

---

[14] At this point, completeness is not required for all Σ^0_1 sentences, but only those expressing formal derivability. As Visser has pointed out to me, and contrary to (my) first impression: *not all* Σ^0_1 *sentences are demonstrably equivalent to some* $\Box A$. In fact, for any R (not necessarily Σ^0_1):

$$\text{if } \Box R \to \Box\bot \text{ is derivable, then } (\Box A) \to R \text{ is not derivable.} \qquad (10.1)$$

If it were, by the properties of $\Box$:

$$\Box\Box A \to \Box R,$$

and by the assumption on R:

$$\Box\Box A \to \Box\bot.$$

Then

$$\Box\Box\bot \to \Box\bot$$

would follow from

$$\Box\Box\bot \to \Box\Box A,$$

which holds generally.

The letter 'R' is chosen because (the Σ^0_1 version of) Rosser's sentence satisfies the hypothesis of 10.1, and thus is not demonstrably (implied by, and in particular) equivalent to $\Box A$. But cf. (e) below.

[15] This is implicit in Gödel's [1972a], under the embarrassing heading 'The best and most

b) Formal derivability and general provability

In a different vein, Gödel's note remains of interest because it isolates a property (in the elementary formalism of modal propositional logic) that distinguishes between formal *derivability* and general *provability*; most elegantly, by use of Löb's theorem. Not only is some instance of $\Box p \to p$ underivable, but one has

$$\Box(\Box p \to p) \to \Box p. \tag{10.2}$$

This is an opposite extreme, as it were, of the property $\Box(\Box p \to p)$ of general provability since the latter, together with 10.2, implies $\Box p$.

c) Truth and general provability

At least so far, this is a distinction without much difference. Formally, all axioms of the modal logic considered by Gödel (in other words, all properties of general provability formulated there) remain valid if '$\Box$' is dropped altogether.

Admittedly, '$\Box$' cannot generally be introduced (for example, in the conclusion of $p \to p$) and preserve validity. But nothing is explicitly formulated about general provability that does not also hold for truth; in contrast, for example, to:

- 10.2 in (b) above for formal derivability $\Box$
- a language with propositional quantifiers, since evidently $\neg\forall p(p \to \Box p)$ holds, but not $\neg\forall p(p \to p)$.

In fact, it seems to be open whether anything can be said about general provability in the language considered that does not hold for truth. Of course this will not be (mis)interpreted as showing that the two notions have the same meaning! (cf. (a) on p. 219).

It was a genuine discovery of the 70's to recognize that formal derivability admits a neat theory at all. This was *philosophical* progress, correcting the simple-minded view that general provability 'ought' to be studied.

Incidentally, the kind of general provability meant here is not likely to be concerned with the outer limits of provability. The latter grow, and so it would be *prima facie* inappropriate to apply classical logic to statements containing $\Box$ (which is not to say that therefore intuitionistic logic is appropriate!).

---

general version of the unprovability of consistency in the same system' (cf. Section 1 of Chapter 9); 'embarrassing' because:

- it refers to *formal* systems, and so is obviously not most general (cf. Mostowski [1952])
- it gives no hint under which conditions this version (that is, $(\Box F) \to F$ for $F \in \Pi^0_1$) is equivalent to consistency.

So, far from being best, it is not even good. The best that could be said is that it is the version most directly relevant to Hilbert's program, where Π^0_1 sentences are privileged. So it should be noted that, for formal derivability $\Box$, $(\Box F) \to F$ is Π^0_1 if F is.

The selection of rewarding phenomena among those that present themselves to out untutored attention (here, of formal derivability within general provability) is a recurrent theme of this chapter.

d) A blindspot

This concerns the thoughtless (and by now largely forgotten) literature against *mixing object language and metalanguage*; as if the union of two sets were not a set. Gödel gave sensible examples in [1944] (about every sentence containing a relational word, and the like), but none is as memorable as the modal language considered above; particularly, when $\Box$ is interpreted as formal derivability.

Naturally, as with other unions of two sets (for example, of cabbages and kings), there is a genuine problem of finding non-trivial laws that hold for the union (in the particular language considered).

e) Some technical remarks

They concern the elegant theories of formal derivability.

It is an undoubtedly memorable fact that the three axioms expressing:

- the (distinctive) theorem of Löb
- closure under modus ponens
- completeness (at least) for those Σ_1^0 sentences that express formal derivability

should axiomatize all valid theorems of the language; and uniformly for a broad class of (formal and some other) systems, at that.

But, at least so far, this axiomatization has not helped to find new memorable theorems in the language itself comparable, for example, to earlier observations about the negation of consistency being conservative for Π_1^0 sentences (which can be done in the language for those expressing underivability; cf. note 14).

One would have hoped that, by completeness, a property established for some cunningly chosen system could then be generalized to all formal systems considered; as properties of the field of real numbers are generalized to all real closed fields.

The second technical remark is a reminder. Though cut-free systems have been recognized to be significant for current logic and, by Kreisel and Takeuti [1974], to have memorable properties in the language considered, little is known about formal theories for cut-free derivability.[16]

[16] A titbit: without recognizing the general significance of cut-free derivations, von Neumann is on record as having seen the principal difference between his and Gödel's proofs of the second theorem in these terms.

Infinitely many monadic propositional operators ([1932a])

Gödel's half-forgotten note [1932a] states that *Heyting's calculus for* $(\neg, \wedge, \vee, \to)$ *does not have an adequate finite truth table*. Now, infinite truth tables, such as Boolean algebras, are just as good or better (for example for decision procedures, the most prominent purpose of truth tables fifty years ago). Here there is no particular virtue in a finite truth table for the *whole* language; success requires an efficient way of first finding one for any given formula, and then evaluating it. So the result stated in [1932a] is obsolete.

Of more lasting interest is a step in the proof that provides a sequence A_n of pairwise (formally) inequivalent formulas with just one propositional variable p. In other words, *there are infinitely many monadic operators*: $p \mapsto A_n$. Or, more pedantically, they are different for all classes of propositions and interpretations of the operators $(\neg, \wedge, \vee, \to)$ for which the calculus is complete.

As a memorable corollary, there is a sharp contrast with the classical case where there are just four monadic operators: the two constants $\top$ and $\bot$, p and $\neg p$.

Viewed as above, Gödel's proof suggests immediately the operator $\circ$:

$$p \longmapsto \bigvee\{A_n : A_n \text{ not equivalent to } \top\};$$

cf. Goad [1978] on the wide spectrum of meanings for which $\circ$ is not equivalent to any operator built up (finitely) from $\neg, \wedge, \vee, \to$.

The remainder of this subsection attempts to give some perspective on new propositional operators; in line with the manifesto in Section 1, not only as a topic of logical research, but for examining ideologies. Naturally, we begin with what is known.

a) New propositional fragments

By Wojtylak [1984], the fragment $(\neg, \wedge, \vee, \to, \circ)$ has a respectable metamathematical theory. Wojtylak [1982] provides references to other work on *monadic operators* of intuitionistic logic (defined by infinitary conjunctions and disjunctions, or by propositional quantification). As far as mere legitimacy is concerned, $(\neg, \wedge, \vee, \to)$ is seen to be just one fragment of (the propositional part of) intuitionistic logic among many others.

The result of de Jongh [1980] on the *unbounded totality of binary operators* is naturally interpreted by a metaphor from set theory. While a fragment is a subject for research (to be compared to a set which can be grasped as a unity), its complement (that is, the totality of new operators) is not.

b) Experience in classical logic (for a proper perspective)

The *functional completeness* of its propositional part is a quite exceptional phenomenon in classical logic. Thus, already when applied to sets and relations, the question arises which properties distinguish the Boolean operations (say, intersection and complementation) among all operations on the collection of subsets of a given set (or some Cartesian power of it); cf. Craig's attempt in [1965]. The obvious parallel in intuitionistic logic comes up in the so-called topological interpretations, with operations on *open* sets.

The following points, concerning classical predicate logic, are in keeping with the familiar fact that *the propositional part of intuitionistic logic exhibits many of the formal complexities of classical predicate logic*. So the reference below to so-called abstract model theory is not dragged in out of the blue.[17]

The best known result here is *Lindstrom's maximality property* of the classical fragment $(\neg, \wedge, \forall)$, to be compared to attempts of sanctifying the fragment $(\neg, \wedge, \vee, \rightarrow)$ in intuitionistic logic. The comedy involved has been described often enough; perhaps, most recently in Kreisel [1985]. It need not be repeated here, except perhaps for this. Relevant extensions of experience such as considering *new quantifiers* or *fragments of* $\mathcal{L}_{\omega_1\omega}$ (and remembering those extensions!) have been more rewarding than the kind of brooding common in the so-called theory of meaning.[18]

Last but not least, there is a broad parallel suggested by the (general) view of intuitionistic logic $[\infty]$ as being concerned with a class of propositions beyond those of classical logic; following Aristotle (cf. *Metaphysics* Γ, 5, 1009a, 16–22 or Γ, 7, 1012a, 21–24). Incidentally, a compact formal expression of this view is found in old-fashioned systems of intuitionistic logic, with modus ponens as its only rule of inference, and literally a subset of the axioms for classical logic. This then evidently allows for more interpretations with larger ranges for the (propositional) variables.

c) Propositions and numbers: some parallels (for orientation)

Leaving aside pretentious drivel about the origin of 'the' concept of number (at least till somebody has as imaginative an idea, *mutatis mutandis*, as Darwin), one may think quite reasonably of various kinds of numbers that populated the intellectual life of the 18th or 19th century; including mildly embarrassing names like 'real' and 'imaginary', resurrected in Hilbert's terminology of real and ideal elements. The traditional perennials about existence, subsistence or what have you of those numbers simply draw attention away from the work that has been done about them. Nothing comparably imaginative has so far been done with propositions. So, to convey the parallel in question, it is best to begin with some reminders about numbers; specifically, about successful choices of particular kinds

---

[17] Incidentally, another fact of experience, namely how quickly abstract model theory rose and fell in the 70's, is in keeping with (a) above about it not being a subject at all.

[18] For examples of such sterile brooding readers may look at Goldfarb [1979] on the meaning of the quantifier in the 20's and, in case of intuitionistic logic, at Sundholm [1983] and Weinstein [1983].

of numbers.

Most familiar is the series of so-called *extensions* (of the number system) by means of closure properties required for solving larger classes of problems; for example, from natural to whole to rational to algebraic to real and complex numbers, and so forth. 'So-called' because, if numbers are thought of in connection with length (as they were in Euclid, not counting), something like the real numbers comes very early; to be *paraphrased* later in terms of suitable sets or sequences of rational numbers.

So to speak in the opposite direction, there was work on *limits to such 'extensions'*; for example, in terms of division algebras with the magic numbers (for the dimensions): 1, 2, 4, 8. Success depends on a proper selection of the properties (here, of $+$ and $\times$) to be preserved in the extension. Before bandying about words like 'fundamental', it is a salutary exercise to remember situations where it is relevant to think of commutative division algebras (over the reals) with a $\sqrt{-1}$, and where the geometric representation (x, y) with the familiar laws for $+$ and $\times$ provides effective knowledge.

Trivially, there are many more kinds of numbers than can possibly be used for effective knowledge. Selection requires thought on what we need to 'do' with them. Certainly, one thing we 'do' with them is to operate on them; in short, the choice of new operations (that is, functions) is an integral part of the extension. Briefly,

> we should not introduce new numbers without doing anything new with them.

For example, the passage from the algebraic to all real numbers would simply not be exploited well if all operations were still required to be algebraic (even in the weak sense of having algebraic values at algebraic arguments). This would exclude the exponential and trigonometric functions (since, for example, 0 is the only algebraic α for which $\sin\alpha$ is algebraic).

In the parallel meant above, the collection of logical operations $(\neg, \wedge, \vee, \rightarrow)$ corresponds to some familiar collection of, say, rational or algebraic functions (or literally to number-theoretic functions mod $2 : 1 - x, x \cdot y$, etc.).

For the view of intuitionistic logic here considered, as concerning larger classes of propositions (for example, about choice sequences), the parallel has an evident implication:

> If we cannot think of anything to do with new operators, the chances are that: either there is not much of interest to be done with the extended class of propositions, or we have not even begun to understand the possibilities.

Informed readers will remember here that till the 30's most logicians had not even begun to understand the *possibilities of intuitionistic logic beyond finitist mathematics*. For philosophy in the sense of Section 1 this fact is relevant, but also the obvious attraction of *sanctifying the familiar fragment* (parallel to the use of Lindstrom's theorem for classical predicate logic discussed above); cf. Schroeder-Heister [1984] on 'completeness' and 'strength' of the fragment. But the metamathematical properties involved in those notions are so hackneyed that they do not constitute any *test* of (the relevance of) the fragment; rather an expression of attachment; cf. p. 223.

d) Another view of intuitionistic logic: proof analysis by abstraction

Here one does not think of a literal generalization (in particular, of a larger domain of objects), but views an axiomatic analysis as identifying abstract properties that are relevant to given theorems (about a specific domain); with additional information as pay-off for eliminating some axioms.

For example, many elementary results about the rationals use only the field properties of $\mathcal{Q}$ (not even that $\mathcal{Q}$ is a number field). Then any Σ_1-theorem $\exists xA$ allows a sharpening to

$$\bigvee_{1\le i\le N} A[x/t_i],\text{ for } t_i \text{ depending rationally on the parameters of } A.$$

If the proof of $\exists xA$ uses only intuitionistic logic, a further sharpening is possible: a single t will do ($N = 1$); recall (a) on p. 220

But after nearly thirty years of experience with this kind of search for additional information, also by others using *sheaf-theoretic models* (a fancy way of talking about the continuous dependence of y on x in combinations like $\forall x\exists y$), I am sceptical. At least so far, one has fallen between two stools:

- On the one hand, when this sort of additional information is really needed, intuitionistic logic is *not refined enough* (recall p. 222 on its algorithmic inadequacy).

  And, to come from the sublime to the ridiculous, there is nothing in the ordinary mathematical tradition to stop one from recording such information if one has it.

- On the other hand, the ritual of intuitionistic logic prevents one from *testing* its ideology which, as emphasized at the outset, requires not only explicit definitions, but a constructive proof that they do their job.

So (by comparison, and for the time being) the view of intuitionistic logic as dealing with a larger class of propositions seems more rewarding; at least, for the following object lessons in Section 1.

10.3 Natural Languages (With Some Reminders on Natural History)

By Section 1, intuitionistic logic is a successful study of (the logical features of) a popular dialect of natural mathematical language.

Relevance in terms of survival value

Reference to natural mathematical language (for assessing methods of studying other linguistic phenomena) is at any rate plausible; provided of course it is remembered that such things as the language of texts on axiomatic set theory in the 50's are artifacts.

Perhaps the single most significant point here is *survival value* since, by experience, this has been a successful guide in biological studies (and natural languages are certainly a biological phenomenon). Judged by survival value, mathematical language (at least, of elementary mathematics) is more convincing than talk about cats doing something or other on mats.

Universal semantical schemes and small arsenals of meanings

As to the success of intuitionistic logic, it applies not only to the syntactic aspects adumbrated in Section 1, but also to *various meanings* (associated with them or, more precisely, appropriate) *in various situations*. As above, a couple of provisoes have to be remembered.

At the present stage, a realistic measure of success is to

> find *relatively few* meanings that are appropriate in *relatively many* situations;

in contrast to talk about a necessarily amorphous family of meanings.[19] Some of these meanings can be thought of as corresponding to other external 'parameters' besides the words used; such as the tone of voice, expression of face, gestures.

But readers should beware of the (most simpleminded, and) superficial 'alternative' to the small arsenals of meanings above, namely the introduction of an additional variable for *situations* (or, for that matter, tone of voice and the other external parameters mentioned earlier). This remains empty unless something substantial can be said about the situations that arise; cf. p. 225 on general provability, or the so-called abstract theory of constructions (which introduces a variable for proofs but nothing less banal about them than the relation between a

[19] All this does not exclude the possibility of much more sharply defined specifications in completely different terms; comparable to those in molecular biology of family likenesses (such as the Bourbons' nose, or Habsburgs' hare lip), so central to Schrodinger's *What is Life*.

proof and the assertion proved [∞]). This is as sterile as the business of situations twenty years later.

The weakness of those general schemes is underlined by the occasional particular problems that do benefit from attention to proofs (for example, q- and *fp*-realizability) or to situations (in which, for example, a counterfactual conditional is satisfied). This is so because the schemes do not help spot those problems, but may distract from the fact that the latter are the exception rather than the rule.[20]

General lessons from intuitionistic logic

It is an *illusion* (surely, not wilful deception) to present studies on natural languages as pioneer work, to which correspondingly lax standards are to be applied. This simply overlooks the fact that, realistically speaking, such subjects as intuitionistic logic are also studies of natural languages, and their *level of sophistication* provides more appropriate standards.

More specifically, experience in intuitionistic logic underlines the need for *selecting rewarding aspects* of linguistic phenomena. For example, such classes of propositions as considered in intuitionistic logic certainly turn up in natural languages, and many more besides. But (by (c) on p. 228) already those classes are suspect, as being too diffuse for rewarding theory.

This point is elaborated below in the broader context of natural history superseded by the scientific tradition, after the next brief digression for general perspective.

Natural and logical sense: a neglected distinction

The most obvious instance is that of *closure under the usual logical operations*. If the propositions p and q have logical sense, so has $p \vee q$; in classical logic in terms of truth values, in intuitionistic logic in terms of proof conditions. But, as a matter of simple experience, this is not so for natural sense; for example if

[20] *Digression about the skills needed to use (any) theory*. No phenomenon that presents itself *naturaliter* (in contrast to experiments, which are set up to exclude forces not treated in the theory at issue) comes with a label telling us *which* theory (if any) applies or, equivalently, which forces dominate it. So, clearly, some skill beyond knowledge of the theory itself is required; even in the case of planetary astronomy Tycho Brahe had to shift from the observed (also called 'apparent') motion to its 'correction' (for parallax), since only the latter lends itself to theoretical analysis.

It may be common to assume that linguistic phenomena (and others of the so-called human sciences) are very different in this respect, because we have conscious beings speak to us (and not dumb planets). But, if common, it is simply a common piece of scientific immaturity.

Viewed in these terms, the business of situations is a *step back* from the small arsenal of meanings above. The latter reduce the additional skills to a proper choice from that arsenal, while mere mention of 'situations' says nothing about their particular aspects that may be relevant.

$p =$ this glass is 5 cm high and $q =$ this glass is transparent.

Trivially, combinations of propositions that have logical but not natural sense can be *given* logical sense (even uniquely). The question is: at what price?

As in (c) on p. 228, experience in mathematics seems relevant; only now propositions are not compared to numbers, but to *sets of points*; as in the topological interpretation on p. 228. Logical sense concerns brutal existence; natural sense, for sets of *points*, involves geometry. It is a common place that not all sets of points are geometrically significant (cf. note 18 of Chapter 8).

Readers may try out other parallels; for example, between logical sense and *measurability* in the sense of Riemann or Lebesgue. Incidentally, geometric sense is then not altogether irrelevant to logical sense! If sets topologically equivalent to a disc are regarded as geometrically significant, more can be said about their *measure-theoretic* properties than about the class of all measurable sets.

Though the word 'natural sense' is not used in the literature I know, the idea is clearly implicit in a good deal of work on *partial* predicates and functions.[21]

Natural history

This is barely mentioned nowadays, except by historians of science. Yet it presents a style of thought (or, as one says, an ideal of understanding) which was once dominant, and still has great appeal. It relies on regularities in nature that strike our untutored attention, most often in the visual sphere. The world we see is determined by forms and colours; we recognize things in this way. For centuries, zoology, botany, but also mineralogy consisted of painstaking descriptions, and later classifications; always in these terms.

It is a fact that forms and colours are particularly unrewarding (or, at least, demanding) subjects of theoretical study. For example, the relation between (the chemical composition of) a thing and its colour involves quantum theory. Data that strike us less or not at all (for example, mass and its centre of gravity or electric charge, not to speak of atomic structure) are more amenable theoretically or, as one says, are physically more important. Obviously, one can use a metaphor like Plato's cave for almost anything; but it is not too farfetched to see it as a

[21] Gödel himself touches questions of sense towards the end of the introduction to [1940], when explaining the relation between abstract set theory and the cumulative theory of types: if an *atomic* formula (that is, a formula $a \in b$) has no type-theoretic sense, it is declared to be 'abstractly' false, and compound formulas are then evaluated in the usual way.

What he does not touch is where the convention goes wrong; this happens if p has no sense, t is the truth set, and so neither $\overline{p} \in t$ nor $\overline{\neg p} \in t$ has natural sense, but nevertheless the 'adequacy' condition of Tarski

$$(\overline{\neg p}) \in t \Leftrightarrow \neg(\overline{p} \in t)$$

is imposed. Evidently, if such simple and familiar points are overlooked in the manufacture of paradoxes, there is good reason to doubt Gödel's high expectations from a solution of the paradoxes [∞].

reminder of the need for looking at informative data, and not just those (shadows) that happen to dance before one's eyes.

As for appeal, it is just wonderful if, for a particular question, those things in front of us are enough; cf. also the simple-minded cult of the *black box*, and the more inspired view of the world advocated by René Thom.

It is not claimed here that there are no areas of experience where the tradition of natural history is effective. What is disturbing (at least, to me) is the scientific innocence of the linguistic fraternity. By and large, they simply do not have a clue about the relation between their style and a well known (if obsolete) tradition. A similar kind of innocence is behind the Faith in using mathematical methods.

It is not this literary form which distinguishes natural history from the natural sciences, but the *selection* of phenomena treated. After all, there are some pretty mathematical formulas in natural history, from d'Arcy Thompson to René Thom. On the other hand early chemistry, in its search for chemically pure substances (and eventually culminating in the atomic view of matter), used very little mathematics.

Last but not least, natural history has an up-hill fight; it competes with our (immense) ordinary knowledge of just those aspects of the phenomena that it considers; for example, in the case of *linguistics*, with the works of literate people who have a genuine feeling for language.

Concerning possible uses for *computer languages*,[22] p. 222 on algorithmic inadequacy provides an obvious warning. More generally, failure on two counts is to be expected:

- bad science, because human and digital data processing ('hidden' in black boxes) are different

[22] On a couple of occasions Gödel mentioned computer languages, presumably after the subject had come up in conversations with others. It has been reported (for example, by Zemanek) that Gödel more or less advocated predicate calculus as a programming language. He never suggested anything like that to me. But he did say (of course, expressing a mere feeling, without any basis in experience of the subject) that programming was *Sache der Geschicklichkeit* (in other words, a skill), and not likely to benefit from theory at all; let alone, logical theory. The word 'skill' jars, since practically everything needs some skill in his sense.

Actually, one can be more specific here, by reference to *Prolog* (short for: programming in logic). It is successful; not because it uses predicate logic, but because it does not use all of it. This is verified by studies of various attempts to add negation (to the Horn sentences used).

Be that as it may, Gödel certainly did not expect programming to benefit from theoretical studies of natural languages; or, more pedantically, from realistic theories. A bad theory (so to speak, how *der kleine Moritz* imagines natural languages to function) may well contain a bright idea that has some use for some program for some hardware for some computational problem. As somebody once said in a paper on the nervous system, with a far-fetched theory of r.e.m. dreams:

> if Nature does not use our idea, perhaps it can be used somewhere in *AI*.

- bad technology, because details of the hardware are not exploited efficiently.

While ten years ago such references to hardware were dismissed as red herrings, today it is (almost) universally recognized that a good use of many processors working in parallel requires new programs. Parodying p. 231 about 'situations', one could introduce a variable for 'hardware' instead of looking for a small arsenal of programming languages which is suitable for relatively many varieties of hardware (incorporating subroutines), and for relatively many computational problems.

10.4 Effective Rules ([1934a], [1936], [1972a])

For the present it is enough to consider rules that define functions whose arguments and values are *natural numbers* (or even only *numerals* $\overline{n}$ built from a constant 0 and a successor function s; or other *words* over a finite alphabet, such as formulas or derivations). The case of so-called higher types is reserved for the next section.

The topic of effective rules occupied Gödel throughout his life; with increased sophistication, at least in his formulations (except of course for the lapses in the 70's). Thus (to judge by footnote 18 of Church [1936]) in the mid 30's Gödel was simply ill at ease with loose talk about effectiveness, while thirty years later he was ready to make explicit distinctions (cf. Section 3 of Chapter 9). Today we can be more explicit still.

Effectiveness involves reference to the systems for which a rule is meant; or, perhaps more correctly, to our idea(lization)s of them; as always, preferably with a few kinds of such systems being adequate to many situations. A by now familiar, particularly elementary kind is (our idea of) a so-called real time digital computer. Gödel's own interests lay elsewhere.

They will be examined below under three headings:

1. equational rules
2. computation in formal systems
3. rules effective for the perfect mathematician (a particular subspecies is perfect in intuitionistic eyes, and called 'creative subject' in the literature).

It should be noted straightaway that intuitionistic logic enters in a quite trivial way into 1 and 2, via the difference between classical and intuitionistic proofs of the $\forall\exists$ theorems expressing the termination of formal computation procedures; cf. p. 221.

But 3 is absolutely pivotal for anything remotely like the original intuitionistic enterprise; not, as is sometimes thought, a marginal aberration of the aging Brouwer. As a corollary, any reservations about the business of the creative subject put *ipso facto* anything like the intended enterprise in question.

Equational rules ([1934a])

Define a, say, monadic function f by use of auxiliary functions $\vec{g} = (g_1, \ldots, g_p)$ and numerical variables $\vec{x} = (x_1, \ldots, x_q)$ (together with the constant 0, and the successor s) in the form of a finite system of equations $E(f, \vec{g}, \vec{x})$.

One familiar sense of E 'determining' a function f requires only that, for some auxiliary functions $\vec{g}$, f is the unique solution of the system E:

$$\exists! f \exists \vec{g}\, \forall \vec{x} E(f, \vec{g}, \vec{x}). \tag{10.3}$$

This is generally not enough to compute a value $f(\overline{n}) = \overline{m}$ from finitely many substitution instances of $E(f, \vec{g}, \vec{x})$; that is, from a conjunction

$$\textstyle\bigwedge_{0 \leq i \leq N} E(f, \vec{g}, \vec{\overline{x}}_i) \tag{10.4}$$

for a suitable N (depending on n), where $\overline{a}$ denotes the numeral with value a, and $\vec{\overline{x}}_i$ an appropriate sequence of numerals in place of $\vec{x}$.

For example, if E is

$$f(x) = 2f(s(x)),$$

only the constant function $f(x) = 0$ satisfies $\forall x E$, where $x \in \omega$ and $f : \omega \to \omega$. But each finite set of substitution instances

$$\bigwedge_{0 \leq x \leq N} E(f, x)$$

is satisfied by any f such that $f(x) = 2^{M-x}$ for $M \geq N$.

One thus considers only *effective* systems E for which, for every n, a value $f(\overline{n}) = \overline{m}$ can be derived from finitely many substitution instances (10.3 ensures that such a value in uniquely determined):

$$\forall n \exists m \exists N \exists \vec{x}_1 \cdots \exists \vec{x}_N \forall f\, \forall \vec{g}\, [10.4 \ \Rightarrow \ f(\overline{n}) = \overline{m}]. \tag{10.5}$$

This notion describes the class of functions computable from equations without reference to any computation rules:[23]

- Gödel [1934a] appealed to a more or less arbitrary calculus to derive $f(\overline{n}) = \overline{m}$ from 10.4.

- For the tradition of so-called informal rigour, it is more satisfactory to note that (for given $n, m, N, \vec{x}_1, \ldots, \vec{x}_N$)

  $$\forall f\, \forall \vec{g}\, [10.4 \ \Rightarrow \ f(\overline{n}) = \overline{m}] \tag{10.6}$$

  is relatively easily decidable, and to construct accordingly an equation calculus that is demonstrably *complete* for it; cf. Kreisel and Tait [1961], with refinements in Robinson [1968] and Statman [1977] (concerning, respectively, the use of monadic $\vec{g}$ and quantitative properties of the calculus).

[23] Instead of 'computable', *finitely determined* recommends itself; cf. 'validity' in logic in place of 'provability'.

Since the requirement 10.5 on effective E is Π_2^0 in 10.6[24] and 10.6 is decidable, the f defined by such E are recursive (pedantically, for some other usual definition of 'recursive'). The converse is read off from Kleene's normal form. Thus one has an *equivalent presentation of recursiveness*.

There are also refinements of this kind of equivalence, in the literature on (what has been called) Church's superthesis; cf. Kreisel [1971]. Gödel himself avoided such matters (from his point of view, wisely). Once one begins to look at particular sets of rules, one inevitably sees how little one knows of the totality of possible rules (or even of what one wants to know about them); no matter how nicely the particular sets considered behave.

Formal computability ([1936])

We consider now computability by use of so-called *entscheidungsdefinite* (or, more simply, invariant) expressions; originally, with respect to *Principia Mathematica* and related systems.

For (characteristic functions of) predicates, the expressions considered are formulas F with a single free variable such that, for each $n \in \omega$,

$$\text{either } F(\overline{n}) \text{ is derivable or } \neg F(\overline{n}) \text{ is derivable,}$$

with an obvious variant for functions (from ω to ω).

Two novelties, compared to the previous approach, should be noted:

- Before the 30's, formal rules of inference were thought of primarily as means for *checking* (rather than *generating*) derivations; let alone, computations.

  Understandably, since the procedure involved in this kind of formal computation is quite *unrealistic*: all derivations are thought of as laid out in ω-order, and the computation consists in looking for the first derivation whose end formula is $F(\overline{n})$ or $\neg F(\overline{n})$.

- Compared to the specific equation calculus above, the notion of formal computation has a glamorously *general look*; even if one considers only (consistent) finite (but otherwise arbitrary) extensions of some given formal system like *Principia*.

Gödel's afterthought in [1936] on *absoluteness* should be viewed in this light. The general idea is an *Aha-Erlebnis* for all of us;[25] and the property is most

[24] The string of quantifiers $\exists N \exists \vec{x}_1 \cdots \exists \vec{x}_N$ (with variable length $N+1$) is actually a single existential quantifier over the set of (codes of) finite sequences of possible values for $\vec{x}$.

[25] Indeed, so is the idea of *speed-up* by use of new 'abstract' axioms (in the logical sense of involving higher types). But Gödel's early formulation in [1936] is simply clumsy.

For one thing, the speed-up is illustrated most simply by any undecided formula $(\forall x \in \omega)[f(x) = 0]$. Computation of f according to its defining equation is slow. If, by use of new axioms, we know $(\forall x \in \omega)[f(x) = 0]$, we have the unsurpassably fast computation: $x \mapsto 0$.

impressive, by comparison with the embarrassing drivel about the 'evidence' for Church's Thesis allegedly provided by the equivalence between different definitions of effectiveness; without a thought of any safeguard against a systematic oversight.[26] (If each juror has a chance $1/2$ of judging correctly, why are verdicts wrong more often than once in 2^{12} cases?)

In contrast, absoluteness constitutes a striking *closure property* of the class of formally computable functions; actually, not only for finite extensions (of, say, *Principia*), but for all those enumerated by functions that are invariantly defined in a system already recognized as formal. This 'raw' attraction of absoluteness goes well with the use made of it later by Tarski to transfer recursive undecidability results to large classes of systems; cf. Tarski, Mostowski and Robinson [1953].

On the other hand, Gödel let his enthusiasm run away with him when he claimed (in [1936] and, especially, [1946]), that formal computability was unique among epistemologically interesting notions by being absolute (in the sense of being independent of the language considered). What else is the word *functional completeness* (as applied, for example, to classical propositional logic) about?[27] Of course, in conversation Gödel agreed that he had had a blind spot. But he is not alone in having forgotten the great impression (at least, on logically sensitive people) when we first learn such an easy and convincing answer to the question:

What is a propositional operator?

but cf. p. 189 of Chapter 8 on worries about this answer being so easy that it is liable to be singular.

The perfect mathematician ([1972a])

This is usually presented as an immensely subtle idea, and rules effective for that animal (but, tacitly, not for digital computers) are sought in the outermost reaches of Higher Thought.

In fact, practically none of the rules used every day (and thus stated in some natural language) is literally effective for any digital computer. The discovery that

---

The more elaborate formal exercises of [1936] fall between two stools; they are superfluous for the general point, and they do not help to discover realistic possibilities of speed-up. (Of course, the exercises are more 'weighty' than the aside above.)

[26]The drivel about 'evidence' for Church's Thesis obscures a genuine virtue of having many equivalent definitions or, more simply, descriptions of the same notion (whether or not they define the originally intended matter). When solving problems about the notion, use can be made of knowledge of the different concepts involved in those descriptions; cf. also p. 40 of Chapter 4. It is an object of research to *discover* which descriptions suit particular problems, even though it may well be that other descriptions tend to force themselves on us. Intensional logic, which is preoccupied with those other descriptions, is thus not an illusion, but often simply sterile.

[27]Cf. the subsection on p. 227 for the contrast in the case of intuitionistic logic.

many can be replaced by *different* rules so as still to define the same function (generally, without preserving the computation processes even approximately) was the sensation of formalization a hundred years ago, and is an essential ingredient of the computer business. Programmers are paid, sometimes handsomely, to find suitable replacements.

The idea that those rules in natural language including natural mathematical language are defective for brutal reasons, such as lack of precision or some other unreliability, is sheer dogma. Of course, they are not *formally* precise. But how adequate is this idea(lization) of precision for a realistic view of reliability?[28]

The problem is elsewhere, and readers of this chapter have been prepared for it:

What of interest can be said about the perfect mathematician?

The choice of concepts or 'language' in which this information is to be expressed is part of the problem.

Here is some background, necessarily less banal than the business of cats on mats on p. 231, including some of Gödel's own ideas. Reminiscences of conversation with him on and around the topic are at the end of Chapter 9.

a) The intuitionistic version

For the intuitionistic idea of the perfect mathematician, the so-called *creative subject*, the following type of non-mechanical rule has become standard since the end of the 60's: the map from

> formal derivations d of existential formulas $\exists x A$ (possibly with parameters) built up according to some intuitionistically interpreted (possibly formal) systems

to

> terms t such that t defines the object (or family of objects) x satisfying A, supplied by the proof $\overline{d}$ represented by d.

Certainly, no digital computer accepts this rule as it stands, since it requires an understanding of the maps $d \mapsto \overline{d}$ and $\overline{d} \mapsto x$, supplied by the interpretation of the system considered. Digital computers do not handle *this* kind of understanding or interpretation. So, already the question:

is the rule *equivalent to some computer program*: $d \mapsto t$?

goes beyond the domain of digital computing. Twenty-five years ago it was conjectured that some such rule might define a non-recursive function.

[28]Cf. p. 210 of Chapter 9 for a striking *petitio principii* in this connection.

There is an obvious formal parallel to the question above, involving the arsenal of functional and realizability interpretations of various systems of intuitionistic logic, but also transformations in the style of cut elimination. Each such operation O supplies a map $\mu_O : d \mapsto t$ though, of course, not generally a definition of the object supplied by $\overline{d}$ itself (for example, if t is simply the smallest numeral $\overline{n}$ for which $A[x/\overline{n}]$ is derivable in the system considered).

Here it was conjectured (twenty-five years ago) that most of the μ_O are (even) *extensionally different*, because the various interpretations are well known to satisfy laws not generally valid for intuitionistic logic; for example, the interpretation discussed in the next section satisfies Markov's rule, recursive realizability satisfies Church's Thesis, and so forth.

However, those conjectures were refuted during the 70's. The operations μ_O were shown to be *equivalent even up to conversion*, mainly by Mints. Furthermore, the map: $d \mapsto \overline{d}$ could be examined by use of so-called theories of abstract constructions (of little use for anything else), with the result that the μ_O were seen to be *equivalent to the rule stated above* in terms of the (intended) intuitionistic interpretation. This is the kind of safeguard against the possibility of a systematic error that is lacking in the so-called evidence for Church's Thesis considered on p. 238.

Perhaps the single most memorable corollary to all this *prima facie* satisfactory work (on the stability, as it were, of the idea of the perfect mathematician) is this. Once one looks closely at the map: $d \mapsto t$, one sees how marginal the labelg4map algorithmic aspects of proofs are; mathematically quite trivial changes in d lead to algorithmically wild changes in t, cf. Kreisel [1985b]. This fact is clearly embarrassing for several variants of the intuitionistic ideology. From *their* point of view it supports Gödel's worry (reported on p. 237) about leaving well enough alone; but not for philosophy in the sense of Section 1.

b) Gödel's own version

Evidently, (a) imposes what appear to be gratuitous restrictions on the perfect mathematician, by requiring perfection in intuitionistic eyes.

In [1972a] (cf. Section 3 of Chapter 9) Gödel goes to the opposite extreme (under the motto: *Wenn schon, denn schon*, i.e. you might as well be hanged for a sheep as for a lamb), and considers rules of the form:

> compute the characteristic function of $\vdash_2 A_n$ for a suitable sequence of formulas A_n (where $\vdash_2$ means second order validity, and $\vdash_2 A_n$ is to be decided by suitable axioms of infinity).

Gödel assumes familiarity with the subject, relying (in effect, though not in so many words) that St. Thomas' *adaequatio et rei et intellectu* would furnish the required axioms.

Gödel enlarges on that *adaequatio*, by pointing out that the intellect grows (when familiarizing itself with the material at issue). However, this does not add much; *not that it grows, but how* is the crux.

It is doubtful whether our knowledge of the possibilities of the mathematical imagination has reached the threshold for pursuing any idea(lization) of the perfect mathematician. Or, to put it in current jargon, there is nothing sufficiently specific that we know about biological data processing (nothing like Planck's discovery about black body radiation in another domain) to have confidence in such ideas.

The same applies, of course, to earlier *jeux d'esprit* in this area, now known as *autonomous progressions*.

10.5 Effective Rules of Finite Type

This matter was the principal topic of my conversation with Gödel, which is reflected in the style of the present section.

Gödel's own account in October 1955 of early background

In the first 20 minutes of our first meeting, in October 1955, he sketched some formal work he had done in the forties, and later incorporated in the so-called *Dialectica* interpretation (with a total shift of emphasis).

He was familiar with my own interest, also since the forties, in what I called *functional interpretations*. They rely on a kind of $\exists\forall$ normal form where (in contrast to Skolem's normal forms) the quantifiers need range only over recursive objects, albeit of higher type.

Gödel's interest in the forties, as described to me (but also in his notes for a lecture at Yale on the occasion of his honorary doctorate), was quite different: he wanted to fill the superficially principal gap left by his negative translation (cf. p. 221). In his own words in his notes in the *Nachlass*, he wanted to find out to what extent intuitionistic logic was really constructive. He dropped the project after he learnt of *recursive realizability*, that Kleene found soon afterwards.

Today the relations between the two schemes are summarized by the general facts about the existence property in (a) on p. 220, extended in the points below (which, for convenience, repeat some of the general material).

a) $\exists$-theorems

By the end of the 30's Gödel had doubts not only about the existence property of Heyting's formal arithmetic HA (cf. p. 221), but this:

Does a formal derivation d in HA of $\exists x A$ ensure some term t_d, defining a number x_d, such that x_d satisfies A (without $A[x/t_d]$ being necessarily derivable in HA)?

This was the main problem that the material in the 40's that later developed into the *Dialectica* interpretation was said to solve (in the relevant notes for the lecture at Yale).

As originally presented, neither of Gödel's and Kleene's schemes achieves quite what Gödel intended. For any derivation d of $\exists xA$ in the system HA of arithmetic considered, both obtain terms t_d^G and t_d^K (where Kleene's t_d^K is simply the number of a partial recursive function depending on the parameters of $\exists xA$) such that:

- $A[x/t_d^G]$ holds for the *Dialectica* interpretation
- $A[x/t_d^K]$ holds for the realizability interpretation
- neither holds necessarily for the interpretation intended by Brouwer and Heyting.

In particular, the original work left open whether (for appropriate translations of t_d^G and t_d^K into the language of arithmetic) $A[x/t_D^G]$ or $A[x/t_d^K]$ or both are formally derivable, which would of course ensure that they hold for the intended meaning.

Incidentally, the formulation above gives a concrete purpose to Gödel's warning in [1958] against confusing his interpretation and the orthodox meaning. The warning serves also, at least indirectly, as a correction of his blunder about *eine gering abweichende Interpretation* in [1933] (cf. p. 219).

b) A variant of realizability

Without emphasizing the issue in (a), Kleene soon found a variant (so-called *q-realizability*, with associated terms $t_d^{q_K}$) such that:

- $A[x/t_d^{q_K}]$ is not only q-realized, but formally derivable.

The translation of $t_d^{q_K}$ into arithmetic language depends essentially on the particular coding of partial recursive functions used (in q-realization).

Gödel's scheme has not been modified equally simply, least of all by him (who saw in such work only *Kleinarbeit*, even when done by others).

Since the work reported in (a) on p. 239 (on the stability of $\exists$-theorems, i.e. on the equivalence of the various operators μ_O), the whole matter is moot:

$$t^G, t_d^K, t_d^{q_K} \text{ and many more are } \textit{equal up to conversion}$$

(tacitly, by suitable rules, and for suitable numberings of the partial recursive functions in the case of realizabilities).

c) Provably total recursive functions

Gödels' scheme remains more convenient than Kleene's if not merely some realization of provable $\exists xA$, but an idea of the class of so- called provably (total) recursive functions is wanted. Gödel's schemata for primitive recursive functions provide an elegant description of the class of recursive functions provably total in Heyting's (or classical) arithmetic, and thus *a mathematically memorable* (not only accessible) *instance of a formally undecided* $\forall\exists$ *sentence* (expressing the computability of primitive recursive terms).

However, at least so far, the description of the class (of recursive functions provably total in arithmetic) extracted from Gentzen's analysis (in terms of α-recursion for $\alpha < \epsilon_0$) has been far more useful, for combinatorial and number-theoretic problems about rapidly growing bounding functions, than Gödel's scheme of higher types.

Less trivially, there seems to be a genuine obstacle to modifying Kleene's scheme for the purpose: the use of partial functions (as opposed to using *all* total recursive functions) is necessary for realizing the laws of intuitionistic logic, even of its propositional part. Where does one find suitable proper subclasses of the class of partial recursive functions, retaining their most highly advertized virtue: a *universal element* that enumerates the subclass from a few initial functions?[29]

d) Realizability for negative formulas

A most striking difference between Gödel's scheme and Kleene's realizability concerns negative formulas F^- (except for $\forall$ formulas, which are left uninterpreted by both):

- The version of realizability on p. 214 of Troelstra [1973] has, literally, *nothing* as the only possible realization for any such F^-, and so extracts no information (except realizability); neither from F^- alone, nor from a proof of F^-.

- In contrast, for elementary A and B, Gödel's scheme treats the (incidentally, very common) negative formulas

  $$\forall xA \to \forall y\neg\forall z\neg B \quad \text{and} \quad \forall y\neg\forall x\forall z\neg(A \to B)$$

  like

  $$\forall y\exists x\exists z(A \to B),$$

[29] All this applies to functions and functionals of lowest type. Now, Kleene's scheme S9 expresses auto- enumeration, but S1-S9 do not generate all partial recursive objects when applied to the principal classes of operations of higher type; for example, the countable functionals (beyond the lowest type).

Other, comparably elusive differences between the lowest and higher types will come up on p. 252, in connection with bar recursion.

> thus contributing to a principal concern of mathematics: of unwinding *prima facie* non-constructive proofs of $\forall\exists$ theorems.

Naturally, often the additional information supplied by Gödel's scheme is not needed; for example, for many crude provability results or, as mentioned repeatedly in this chapter, for $\exists$-theorems (of intuitionistic systems). Then realizability is more efficient than Gödel's scheme or the no-counterexample interpretation (let alone, proof-theoretic methods).

e) Interpreting terms of higher type

Gödel made a point of warning me that he had not given any thought to the objects meant by (his) terms of finite type. The only interpretation he had in mind was *formal*, as computation rules obtained when the equations are read from left to right. Gödel had the impression, in 1955, that ordinals $< \epsilon_0$ could be assigned to those terms so that each computation step reduced the ordinal. But this was done only much later by Howard (first in [1970], then more elegantly in [1980]).

For reference in Chapter 12, concerning a comedy of errors: at the time I did not listen to Gödel's warning, since I knew how I was going to understand his terms. The key words are: recursiveness and continuity, the two pillars of the constructive part (as understood in the mathematical tradition) of algebra and topology.

Gödel's scepticism in 1955 about logic

A few minutes after that first conversation on p. 241, we found ourselves waiting for the Institute bus that took us to the other end of Princeton where Gödel (and, at the time, also I) lived. He added a *caveat emptor* about the

> *Aussichtslosigkeit* (that is, hopelessness) of doing anything decisive in foundations by means of mathematical logic

generally, and by use of the ideas he had just talked about in particular. (One might trick intuitionists into believing that his scheme was constructive.)

Again I did not pay much attention, in accordance with my expectations of foundations already at that time. In 1954, at a congress, I had described foundational 'issues' (there, in connection with finitist proofs) by:

one man's meat is another man's poison.[30]

And soon I was going to describe (my) interests in such matters as a 'calculated risk'. My reservations differed from the more familiar variety, since I saw no logical

[30] From the French: *poisson* (fish), *poison* (poison).

defects such as intrinsic lack of precision (cf. the introduction to Benacerraf and Putnam [1964] for a different formulation of an equivalent view of my interests).

As for decisiveness, there are two complementary points:

- Unlike Gödel I never expected foundations to do *better than ordinary science* with questions like (his favorite):

  What is the world really like?

- What I have always expected (and continue to expect) from logic is a *correction of various naive conceptions or aims*, partly enshrined in the foundational literature.

  There need be nothing indecisive about such corrections; for example, to take a familiar parallel, about correcting the aim of astrology to predict human destiny from the position of the planets (rather than predict their orbits).

  In fact, no ordinary scientific result can be quite as decisive or final; even if it is right as it stands, there are usually better new questions.

  And when the refuted aims have popular appeal, as they often do, the refutations have a wide market; cf. the peroration of Chapter 6.

I believe, the main lesson I have learnt in the last thirty years (incidentally, *after* having tried the opposite scheme of pursuing pedantic distinctions, as in the appendices of Kreisel and Krivine [1966]) is this:

> Remarkably often, the defects are so elementary that a *bon mot* can be the appropriate literary form of a refutation.

In such cases (not in all! Cf. the end of Chapter 6) the ritual of (the literary forms of) mathematical logic simply distorts the 'epistemological situation'.

Principal progress during the years 1955–1957

All the results listed below are stated in (or are corollaries of) Troelstra's compendium [1973]. But a selection is needed to present [1958] as the gem it still seems (to me). By p. 254, this is not Gödel's own selection.

a) Hereditarily effective and continuous operations

Various classes of functions of finite types were described, that both fit the general idea of constructivity and satisfy the axioms in [1958] for primitive recursive functions. No 'reduction' is involved, since those functions are defined in arithmetic terms; for example, in the case of *HEO*, or of the countable functions, in the language of second order arithmetic. But the axioms needed to prove closure under primitive recursion are conservative over first order arithmetic.

This kind of work was enough for various formal *independence results*, and for describing the functions and functionals defined by Gödel's schemata in familiar, ordinal-theoretic terms.

b) Extension of first-order arithmetic to all finite orders

To avoid the impression of absurdity created by using functions of finite type to interpret mere arithmetic (parallel to proving consistency of ω-induction by ϵ_0-induction),[31] the emphasis was shifted from Heyting's arithmetic HA in [1958] to HA^ω, its by now familiar extension to finite types, with or without various forms of choice (cf. note 11).

The most memorable result, described as 'principal' from the start, is the *classical* equivalence

$$A \Leftrightarrow \exists s \forall t A_0 \tag{10.7}$$

where A is formulated in the fragment $(\neg, \wedge, \forall)$, $\exists s \forall t A_0$ is its interpretation according to Gödel's scheme, s and t range over the countable functionals, and s may be required to be *recursive*. The proof uses so little that it applies literally also to arithmetic A, with s and t ranging over HEO.

c) Inadequacies

Practically all crude questions about (b) were settled. For example, functions s and t of *bounded* type are not enough. Also, for A in the language of second order arithmetic, 10.7 need not hold when s and t range over HEO; nor when s is defined by Kleene's schemata S1-S9, and t ranges (as required originally by Kleene) over *arbitrary* functions.[32]

d) Sharpenings and omissions

As to *theorems* A of some formal system or other, two points were clearly recognized:

- By general theory, 10.7 can be sharpened; the range of s can be restricted to some r.e. subset of the range in 10.7 (cf. the subsection on p. 251).

[31] Cf. p. 272 of Chapter 12 for an example of that impression, and its cure, in the case of Spector.

[32] Berger has observed that, since the schemata define dense bases for each countable type, 10.7 can be sharpened: for countable t, an s is defined by S1-S9 applied to *countable* arguments. Actually, S9 can here be replaced by μ-recursion (in A_0).

- By inspection of mathematical practice, most analytic theorems can be proved in 'weak' systems that are conservative over arithmetic anyway, and interpretable by use of primitive recursive functions together with a few auxiliaries (such as the so-called fan theorem functional in (c) below).

Admittedly, no really striking new uses of the interpretation were found, comparable to those that the (formally much less appealing) no-counterexample interpretation, which uses only functionals of lowest type, had provided in the preceding decade.

Principal omission: by a fluke, *HRO* (which is the analogue to *HEO* when extensionality is dropped) was overlooked.

I reported on these matters in 1957, both at Cornell and at Amsterdam. There was evident, fairly general interest.

Gödel's last full-fledged paper ([1958])

In 1958 an opportunity presented itself to Gödel to give his own exposition, including second thought about the work he had started in the forties.

- In effect, but perhaps also by intention, there was remarkably little overlap with what I had said in my talks and with what others (for example, Kleene) had said about realizability.

- In particular, and in contrast to Gödel's original work, a main stress was on a primitive notion of effective rule of finite type without extensionality. Under Church's Thesis this reduces to *HRO*, the object that had been overlooked.

The opportunity referred to above was a *Festschrift* for Bernays' 70th birthday. Bernays had been associated with Hilbert's program on finitist foundations, and it might be added that Hilbert himself had already introduced schemata of which Gödel's are a special case; incidentally, without discussion, as if they were obviously finitist in the sense he meant.[33] It is hard to think of a better stage for

[33] Cf. the sketch for a collection of sets that satisfy CH, at the end of Hilbert [1926]. Hilbert's ground type consists of all constructive (not only the finite) ordinals.

Presumably, the subject of finitist rules (and possibly even of proofs) would become a little more rewarding, if the following distinction adumbrated in [1958] were pursued (cf. the progress in Section 4, by distinguishing between three kinds of systems for which rules are intended to be effective).

The finitist literature refers both to *finiteness* and to *visualization* (*Anschauung*):

- The idea of a hereditary finite operation (without restriction on proofs) is developed successfully in recursion theory.

- The idea of visualization derives from geometry, and is (without exaggeration) at the opposite extreme from finite (tacitly, discrete) mathematics; at least, as these things present themselves. (This may change when more is known about human data processing.)

Gödel's exposition.

We shall here concentrate on the principal novelty, both absolutely speaking, and to me personally when I saw [1958]: the *primitive notion of effective rule* mentioned above. Gödel had never breathed a word to me about his project of exploiting such a notion.

In retrospect, Gödel's step is seen to fit very well his philosophy in set theory, where he expected wonders from (working with and) brooding over the (primitive) notion: 'subset-of'. Similar wonders could then be expected in the area of constructive mathematics, from similar attention to the notion of effective rule. And the emphasis on higher types here fits his faith in higher types in set theory expressed in axioms of infinity (cf. Section 10 of Chapter 6).

As to the formal details of [1958], most of them are superseded by the later literature, which had to correct some oversights.

For the record, I still find the paper agreeable to read. When this came up in conversation, Gödel replied: No wonder (*kein Kunststück*), there are no proofs. But this alone would not make a gem.

Today, after more than 25 years, I regard [1958] as a most artistic package of a jumble of ideas, some of which will now be explained.

a) Asymmetry between rules and (the ranges of) their arguments

One feature that Gödel emphasized increasingly in conversations during the decade after [1958] appeared, was the possibility of exploiting the amorphous character (or, if preferred, our ignorance) of the *totality of all effective rules*. More fully, a rule is accepted only if it is understood to be well defined for all effective arguments (of appropriate type), even though little is (or can be) known about this possibly growing totality. This situation is only superficially paradoxical, to adapt the wording of footnote 5 of [1958] about propositional and other logical operators (for the class of propositions) meant by Brouwer and Heyting. There are two obvious illustrations from related areas.

First, ignorance-in-principle is an effective *source of knowledge* in the theory of lawless sequences:

- $\neg\forall x\neg[\alpha(x) = 0]$ is an immediate consequence of α being given by a finite initial segment.
- Also $\forall\alpha\forall\beta[\forall x(\alpha x = \beta x) \vee \neg\forall x(\alpha x = \beta x)]$ is seen this way; α and β are either given as identical objects, or it is impossible to prove $\forall x(\alpha x = \beta x)$.

Secondly, literal *models* of the asymmetry envisaged[34] are used in:

[34]The existence of these literal models in familiar terms evidently reduces one's expectations of miracles from the primitive notion; cf. note 21.

- That version of recursive analysis which requires its *recursive operations* to be defined on *arbitrary reals*, and not only (in the Russian tradition) on the recursive reals.
- Closer to home, the recursively countable functions of type $\sigma \to \tau$ are defined on arbitrary countable functions of type σ (cf. also (c) on p. 254 on the relevance of all this to bar recursion).

It will not have escaped the reader's notice that quantifier free systems like Gödel's T in [1958] have two obvious, obviously different interpretations: the variables may range over:

- all effective rules
- the rules that have been recognized-in-principle as effective (in other words, over the closure under various elementary operations of the class of those that have been literally recognized as effective).

A more delicate point concerning the constants will come up in (b) below.

b) Definitional and demonstrable equality between terms, possibly containing parameters

There is no mystery here; at least, for those familiar with the literature on *normal forms* (and equality up to renaming variables) in the λ-calculus: *definitional equality* is equality of normal forms of those $\lambda x.\, fx$ and $\lambda x.\, gx$ that happen to have normal forms. Though

$$\lambda x.\, gx = \lambda x.\, gx \ \Rightarrow\ \forall x(fx = gx),$$

the converse is not generally valid (except, of course, in specially concocted extensional models).

The transfer of those ideas to typed systems like T is evident, especially if one goes back to Gödel's original formal interpretation (in (e) on p. 244) in terms of computation rules. This *minimal definitional equality relation* for models of T was examined by Tait [1967]. Here every term has a normal form (in contrast to the λ-calculus), and the equality relation is recursive (but not provably recursive in formal arithmetic). Tait's proof uses the machinery of arithmetic (and more) in his definition of hereditary computability, and not inspection of any primitive notion of effective rule of finite type;[35] cf. Gödel's expectations from the primitive notion of effective rules [∞].

The literature seems to have neglected *non-minimal equality relations*, where the constants are interpreted by rules that permit other reduction or computation

[35] In *sharp contrast to sets* where, for example, Zermelo's axioms are verified on sight for all limit ordinals from a description of segments of the cumulative hierarchy.

steps besides those explicitly included among the (equational) axioms (when read as computation rules from left to right). This is the 'delicate point' at the end of (a).

In [1958] Gödel mentions definitional equality between terms of type different from 0, but does not get to grips with it. The context of [1958], a functional interpretation of arithmetic, is not very suitable: the interpretation does not contain such equations. Of course, the matter arises if HA^ω is extended by equations with a corresponding functional interpretation [∞].

In contrast, *axioms of extensionality* can be formulated already in the old context, but require rather different functionals for their interpretation; in a sense made precise by Howard [1973].

A further contrast is found in this area between the *axiom* and the *rule* of extensionality, the latter being satisfied by the primitive recursive functions and many other such classes.

c) The Fan Theorem functional

The last sentence of [1958] states, without comment, that the fan theorem is interpretable. This requires the so-called fan theorem functional; or, more simply, a modulus of continuity for all functionals Φ of lowest type, that is $(0 \to 0) \to 0$, applied to all functions bounded by f (of type $0 \to 0$).

The fan theorem is certainly evidently interpretable by a recursively countable functional, and equally evidently not interpretable in *HEO* or *HRO*. An argument is needed to show that it is not interpretable by an object generated from the countable functionals by Kleene's schemata S1-S9; cf. Gandy and Hyland [1977].

However, it is not at all evident that the fan theorem is interpreted by an effective rule of the kind considered in [1958]; equivalently: whether there is an effective modulus of uniform continuity for effective Φ and f. Under Church's Thesis it is not so interpretable, since then the effective rules are those of *HRO*.

In terms of (a) above, ignorance could be a source of knowledge. Specifically, if we know sufficiently little about the totality of effective functions (of type $0 \to 0$) then Φ can be recognized to be effective only if it is also recognized to be continuous (for the product topology). And then we also have a modulus of uniform continuity.[36]

A good deal more was said on related matters in conversations with Gödel. But it is better left for the digression in the next subsection.

[36] The schemata of Gödel's T are not explicitly required to be continuous (nor to be applied only to continuous arguments). But, for any primitive recursive Φ, there is also a primitive recursive M_Φ of type $(0 \to 0) \to 0$ that satisfies demonstrably the requirements of a uniform modulus of continuity

$$\forall g \forall h\{[\forall x \leq M_\Phi(f)][g(x) = h(x)] \wedge \forall x[g(x) \leq f(x) \wedge h(x) \leq f(x))] \Rightarrow \Phi(g) = \Phi(h).$$

What was not said, partly because I had not cottoned on to the parallel between the primitive notions of set and of effective rule, is more interesting. As matters stand, the parallel holds out little hope for effective rules. The last twenty years have shown that our knowledge of sets in segments of the cumulative hierarchy (for example, as expressed in current axioms of set theory) is simply more rewarding when applied to suitably defined structures, in particular models of set theory that satisfy additional conditions. We literally know more about the constructible sets: they demonstrably satisfy the current axioms, but also GCH while we do not know whether the cumulative hierarchy satisfies CH, nor even whether their second order version decides GCH; cf. p. 98 of Chapter 6. Exaggerating very little, the primitive notion of set serves to answer the question (if one insists on an answer):

from which stock of sets are the constructible sets defined?

Practically, most problems about the constructible sets are not very sensitive to the answer (but depend only on the stock satisfying certain closure conditions). But if one wants to commit oneself to some universe of objects, the primitive notion gives one the means at a small price.[37]

Similarly, in the constructive theory of functionals we have Brouwer's inductive definitions of the type $(0 \to 0) \to 0$; cf. the introduction and the appendix of Kreisel and Troelstra [1970].

From which stock of functions of that type are they selected?

Gödel's primitive notion advertized in [1958] is a good choice. In conclusion, there is no mystery about definitional equality ; but there also is not much prospect for any spectacular uses of that notion.

A sequel to [1958] by Spector

Bar recursion (or, more soberly, *recursion on well-founded trees*) became prominent when Spector [1962] used it to describe the provably recursive functions and functionals of lowest type for *formal* classical analysis. This sharpens not only 10.7 on p. 246, but also note 32 (since Spector's bar recursion is definable by S1-S9 on the countable functionals).

As will be recalled, Spector's version formally extends Brouwer's original bar recursion for decidable trees labelled by objects of a decidable species (for example, the natural numbers) to trees labelled by objects of higher type.

- At the time, the burning question was: Which objects?
- Starting with Gödel's [1958], the answer is: effective rules.[38]

[37] This responds to the worries at the end of Appendix 2A of Kreisel and Krivine [1966].

[38] As is obvious from my footnotes to Spector [1962], written some three years after [1958] had appeared, I had remained attached to the countable functionals (as labels).

Independently of all ideology of intuitionistic logic concerning the evidence of bar recursion (for example, in Gödel's banter reported on Chapter 12 [∞]), there is a question of 'raw' interest (broached already in note 29 about differences between functionals of lowest and higher types; specifically, between $(0 \to 0) \to 0$ and $(0 \to \sigma) \to 0$ for $\sigma \neq 0$). More pedantically, it concerns contexts in which rewarding differences occur, and concepts in terms of which they can be stated.

There is certainly no lack of candidates! For example:

- $0 \to 0$ is not isomorphic to 0 (unless the functions $0 \to 0$ considered are effectively enumerable), while $0 \to \sigma$ is isomorphic to σ.

- Decidability of the type 0 (in other words, of natural numbers) is not problematic, while that of σ is.

- Perhaps most pertinent to what Brouwer first called a proof of the bar theorem, is the possibility of an *inductive definition* of the species of well-founded trees labelled by objects of type 0 (and, admittedly, $0 \to 0$) at the end of the previous subsection; that species is demonstrably closed with respect to all operations that come to mind.

  The literal analogue is simply not true for higher types; at the very least, one would need some new idea about the notion of *freely chosen path* involved in well-foundedness.

This digression will now be concluded by odd formal facts that seem to me relevant here, as I mentioned to Gödel when the occasion arose. He explicitly rejected them; it might be added: fully in accordance with his expectations of wonders to be seen by looking at the primitive notion of effective rule (undistracted by the *Kleinarbeit* that goes into (a)-(c) below).[39]

[39]Provided non-ideological differences between functionals of lowest and higher types are formulated imaginatively, their examination need not be mere *Kleinarbeit*.

On the contrary, it could provide an exception to a general philosophical insight (and would thus be an insight, too). Specifically, at least usually, formal differences between instances of the same scheme (differing in logical or type complexity) *falsify the epistemological situation*.

For example, as proof theory has shown, instances of induction of different logical complexity have different proof-theoretic strengths, even though they all derive their evidence from one and the same principle. Without exaggeration, the whole matter of proof-theoretic strength is an artifact with respect to the evidence of proofs.

However, while generally valid, the insight leaves open the possibility of *discovering* situations where the formal differences are relevant, in which case the latter may fairly be said to have helped in the discovery.

In the particular cases of bar induction of lowest and all finite types, we start with the formal knowledge of their different proof-theoretic strengths; roughly, of Π^1_1-CA and Π^1_∞-CA. There is a chance of a reinterpretation in non-ideological terms [∞]; by reference to proofs (which are constructions, even when the proofs are non-constructive). In any case, *some* reinterpretation is needed to get away from the so-called consistency problem of classical analysis, which rests on highly dubious doubts.

a) Brouwer's fully analyzed proofs, and cut-free proofs

This item concerns the passage to higher types in the case of the bar theorem. Realistically speaking, Brouwer's idea of a *fully analysed* proof (without detour via complicated terms) of a Π^1_1 theorem (specifically, of $\forall\alpha\exists x R(\alpha, x)$ for decidable R) is not compelling when taken literally: if R is $\alpha[\alpha(0)] = x$ then $\forall\alpha\exists x R(\alpha, x)$ is evident without further analysis, but not fully analysed in his sense.

However, there is a perfectly good formal analogue, Gentzen's notion of (infinitary) *cut-free proof* (without detour via logically complicated formulas); at least one of its versions specializes to Brouwer's notions for the proofs considered. As already Aristotle knew, proofs using *modus ponens* need not be further 'analysed' (by cut elimination) to be convincing.

When passing to higher types α^σ (of type $0 \to \sigma$) the question is:

- What is now the appropriate notion of cut-free proof?
- Do at least the usual principles admit cut elimination?

If the α^σ are *defined* objects (say, neighbourhood functions α of countable α^σ, thus satisfying a suitable - analytic - condition C^σ) then it seems open whether a cut-free proof of

$$\forall\alpha[C^\sigma(\alpha) \to \exists x R(\alpha, x)]$$

will generally look at all like Brouwer's fully analysed proofs.

Here it is understood that the 'usual' principles are meant to include not only continuity, but also other mathematical axioms; especially those that concern the generally lawlike *data* for the choice sequences considered.

In summary, of course the formal analogues do not settle Brouwer's claims about arbitrary (convincing) proofs. But (in view of experience in set theory already cited) arbitrary proofs of $\forall\alpha\exists x R(\alpha^\sigma, x)$ may be less rewarding. Trivially, all this applies *mutatis mutandis* to the primitive notion of effective rule and defined models more or less inspired by it;[40] cf. the end of the previous subsection [∞].

b) Howard's neglected alternative to Spector's proof

This item concerns a surely noteworthy (if not often noted) aspect of Spector's proof, quite independent of agonizing re-appraisals of the principles used. The proof has obvious mathematical wit, and so there is surely something behind it. True enough; but certainly not the result stated.

For example, by Howard and Kreisel [1966], classical analysis has easy reformulations in terms of *bar induction*; specifically:

[40] Recently, several more models have appeared (for example, by Bezem [1985]) that seem rewarding, even if not necessarily for the properties generally emphasized (by logicians). After all, Gödel's incompleteness theorem is rewarding enough, though certainly not for what logicians consider to be the great 'mathematical' discovery of the fix-point 'lemma'.

- of type 0 for analysis without any choice
- of type 1 for analysis with the axiom of dependent choices.

Howard [1968] has given straightforward interpretations of those reformulations by use of *bar recursion* of Spector's type.

So, it is a problem to discover a context where the combinatorial wit in Spector's original proof is actually relevant.

c) Gödel's oversight concerning the role of higher types

This last item again involves a parallel with (Gödel's experience in) set theory,[41] but a more formal oversight. It is the role of higher types, with the first few steps illustrated already in [1958]: from primitive recursive functions of lowest type to all finite types (equivalently, in ordinal-theoretic terms, from ω to ϵ_0). There are two points:

- For the general context of intuitionistic logic, as already mentioned, (this kind of) proof theoretic strength goes with algorithmic inadequacy.
- But even judged only for such strength, the parallel is deceptive, because in set theory higher types derive their strength from closure under the power set construction; for example, without the latter, models of replacement (which pushes up the types) can be defined by use of comprehension.

  In the context of intuitionistic logic (that Gödel had in mind), one does not have any analogue for the power set.

Traditional philosophy ([1972])

'Not with a bang' describes Gödel's last attempt (in [1972]) to squeeze out results of cosmic significance from [1958]. Two droplets will convey the flavour.

a) Analytic axioms and proofs

In [1972] *proofs* represented in T are claimed to be *analytic* in the sense of Kant; in other words, they

> use only (properties of) concepts implicit in those used to state the theorem proved.

[41] For 'straight' limitations of higher types in intuitionistic logic (not relying on analogies with set theory), cf. the autonomous progression in Problem 3 of Kreisel [1968], and Friedman's models in Π^1_1-CA that serve to solve it.

Now, proofs of Π^0_1 theorems, expressing the insolubility of diophantine equations, demonstrably may contain (equational) axioms for primitive recursive functions of *unbounded* type. If the latter are implicit in $+$ and $\times$, what is not?

In [1972a] *axioms* (of infinity in set theory) are explicitly regarded as *analytic* in the sense of

'explicating' the concepts occurring in them;

as if not every property of a concept contributed to its 'explication'. In this case logical deductions from analytic axioms (for example, of the insolubility of a diophantine equation from axioms of infinity), are not generally analytic proofs (respecting purity of method).

Overlooking the distinction between analytic proofs and analytic axioms may have been a mere oversight. The *philosophical* sin (of omission) is that Gödel does not even begin to examine the relevance of Kant's ideal or, more generally, his question:

How are proofs possible?

cf. p. 64 of Chapter 6. Incidentally, as has often been observed, Kant's ethereal ideal corresponds to the venerable tradition of *purity of method* in mathematics (cf. p. 64 of Chapter 6; in agriculture such purity is required already in the Old Testament, *Deutoronomy* 22, 9-11).

The last paragraph is not merely irreverent. It suggests an examination of Kant's ideal by reference to the *whole body of mathematical experience*, where purity of method has been pursued since the Greeks, and its defects especially, the notorious loss of Beziehungsreichtum, have become apparent. Without exaggeration, the problem is to discover corners where the principle of purity of method is appropriate (for example, in connection with the extraction of algorithmic information from proofs of $\forall\exists$ theorems); cf. the marginal character of algorithmic aspects of proofs [∞].

b) Demonstrable and definitional equality

This item involves, I believe, only an oversight.

In [1972] Gödel suggests that some kind of reduction is achieved by the decidability of definitional equality. Though the words are vague, the meaning seems plain enough. He apparently forgot that the converse to

$$\lambda x.\, fx = \lambda x.\, gx \;\Rightarrow\; \forall x(fx = gx)$$

is not valid; cf. (b) on p. 249. But even if this particular impression is wrong, the following is surely generally right.

The wish to draw conclusions of cosmic significance is as sure a way to make mistakes as the kind of lack of interest that Chapter 12 sees behind the error in the last sentence of [1933a] about adding equality to the so-called Gödel case of $\forall^n\exists^2\forall^m$ formulas.

A final, sober view of [1958]

Even without considering all the work that refers to it, [1958] is a memorable reminder, using a minimum of scientific experience, of the *potential of higher types in constructive mathematics*. As its title stresses, this aspect had been neglected. Of course, the no-counterexample interpretation also used higher types (and a paper, Kreisel [1952], was devoted to the need for something of this sort). But the role is not nearly as memorable as in the scheme of [1958], where the *only* formal difference from ordinary primitive recursive arithmetic is the use of higher types.

In line with all this, computability of (even the purely numerical) terms is not derivable in formal arithmetic. Since the rules of T are memorable (not merely accessible), [1958] provides *a memorable formally independent sentence in the mathematics of computation*, comparable to soundness in metamathematics, or Gentzen's ϵ_0-induction in the area of infinite descent. And it is useful, if regarded as a *warning* concerning the algorithmic inefficiency of such schemes as that of [1958]; cf. p. 244 on Gödel's concern in 1955 about 'tricking' intuitionists. Admittedly, it is not as memorable as his second theorem in the branch of mathematics called 'metamathematics'.

Whatever its scientific value, the notion of effective rule with definitional equality is so close to the surface of our logical subconscious that, as a matter of logical hygiene, it is salutary to take a look at it. [1958], so it still seems to me, helps us do this in a remarkably painless way (compared, for example, to generalized recursion theory of the seventies; cf. Kreisel [1985a]).

Chapter 11

Constructive Aspects of Gödel's Main Results

Gödel gave *increasing* attention to intuitionistic (or, more precisely, to the even stricter finitistic) requirements on metamathematical methods; specifically, in connection with:

1. the completeness of predicate logic

2. the incompleteness of formal systems for arithmetic

3. the consistency of the axiom of choice and the generalized continuum hypothesis relative to the usual axioms of set theory.

Our concern here is only the extent to which those requirements are rewarding for 1–3; both in the long run, and for short term effects. The latter were always Gödel's main concern, at least, in conversations on titles of papers and on terminology (cf. Chapter 13). It is beyond the scope of this chapter to go into the delicate relations between the facts of Gödel's style here considered, and his later views or memories of them. Of course, those relations are fascinating, and may even be rewarding when treated in an inspired way. But only the most coarse-minded among us could be tempted to speculate on such matters in terms of ordinary knowledge above them (cf. the end of Chapter 4).

In summary:

- Gödel's discussion of 1 (especially in his dissertation) was remarkably penetrating and explicit, albeit a little clumsy by current standards.

- He was much less explicit about 2 at the start, but added pertinent improvements later (cf. Chapter 10 [∞]).

- He was almost perversely pointless with 3, with comic consequences (related on p. 104 of Chapter 6).

[0]Originally published in *Gödel remembered*, Weingartner and Schmettered eds., Bibliopolis, 1987, pp. 121–131, as Appendix I to 'Gödel's excursions into intuitionistic logic'.

11.1 Completeness

The question of *completeness* of (Frege's) rules for predicate logic is as old as the hills; it asks whether, for all logical formulas F,

$$(\models F) \;\Rightarrow\; (\vdash F) \tag{11.1}$$

where $\models$ means (classical) *validity* in arbitrary structures, and $\vdash$ means formal *derivability*.

Brouwer raised it in his dissertation, without however stopping to *paraphrase* 11.1 so that it becomes even a candidate for being settled by methods of intuitionistic logic. Hilbert's paraphrase is described on p. 81 of Chapter 6.

In the introduction to his dissertation Gödel considers

$$(\vdash F) \vee \neg(\models F), \tag{11.2}$$

which is classically equivalent to 11.1, and derives *decidability* of $\vdash$ from any proof of (tacitly, a paraphrase of) 11.2 in intuitionistic logic.[1] It was a brave attempt. But the intuitionistically invalid switch from 11.1 to 11.2 would exclude also intuitionistic completeness proofs for undecidable systems of intuitionistic logic (such as the negative fragment, cf. p. 218 of Chapter 10).

However, by Chapter 6 [∞], the idea above is enough to derive decidability of $\vdash$ for Hilbert's paraphrase of $\models F$.

An historical titbit: an objection not foreseen in the dissertation

According to Mostowski, in a conversation in Tarski's presence, the latter and his students had no confidence in Gödel's paper when they saw the relevant issue of the *Monatshefte* in Warsaw. Why? Gödel had not formally defined validity!

Anybody who is surprised by this, knows *ipso facto* that he simply has no feeling for the subject. I had the good luck, more than a quarter of a century later, to experience the similar reactions of clever people to completeness proofs for intuitionistic logic in the fifties, and incompleteness proofs in the sixties,

[1] Here the logical symbols $\vee$ and $\neg$ mean disjunction and negation of intuitionistic logic, and $\models F$ is 'soundly' paraphrased by, say, Val (F); in particular,

$$(\vdash F) \;\Rightarrow\; \mathrm{Val}\ (F).$$

The basic observation here is that, $\vdash$ being r.e., a proof of

$$(\vdash F) \vee \neg \mathrm{Val}\ (F)$$

in intuitionistic logic would provide, for each F, an object d_F which is a proof of either $\vdash F$ (i.e. a formal derivation of F, a decidable matter) or of $\neg\mathrm{Val}\ (F)$ and hence, by soundness, of $\neg \vdash F$.

also without formal definition of (here, intuitionistic) validity (key word: basis results).

Of course, when Tarski met Gödel in Vienna soon afterwards, confidence was established. In my case, personal explanations (sometimes repeated, *verbatim*, half a dozen times) also helped to establish confidence (of some people); cf. also Chapter 4.

Evolution of a perspective on the completeness theorem

On p. 61 of Chapter 6 we have discussed the relative glamour that the completeness theorem had when it appeared, and would have had 50 years earlier. Here we do not go into such a general perspective, but only into aspects to which intuitionistic logic is relevant; both for the wider meaning of 'constructive' in the sense of definable, and the stricter intuitionistic (or even finitist) sense concerned with provability.[2] A broader perspective, especially on paraphrases, is developed in Chapter 4.

First, Gödel's result that the most popular paraphrase of completeness around 1930 (Hilbert's) is false, has some anti-ideological consequences. The requirements of intuitionistic logic are sterile in connection with 11.1 because the conjectures, expressed by paraphrases that come to mind easily, are simply false. In the meantime there has been a shift of emphasis, in terms of (the absence of) *recursive models*[3] [∞].

Secondly, the particular rules set down by Frege are simply not adequate for the study of delicate aspects of logical proofs; consequently, claims about such aspects in terms of those particular rules are merely pretentious (cf. p. 92 of Chapter 6). So it is (philosophical!) progress to *discover* questions with answers that are less out of proportion with what we know of logical phenomena. The shift in the previous paragraph leads to such questions in terms of *recursion-theoretic complexity* of models and, more generally, of sets of valid theorems. This is a definability property, and thus a concern of intuitionistic logic, at least in the weak sense.

Thirdly, as a refinement, we have that distinctions between total and semivalutations and their recursion-theoretic complexities lead to analysis of different proofs of completeness.

Bibliographical remarks

Shoenfield's term 'characterization' instead of 'completeness' theorem is obviously intended to convey the second point, but has not caught on.

[2]The two senses were elaborated at the beginning of Chapter 10.

[3]Incidentally, the Chapter related to this subject was written without my having seen Gödel's dissertation [∞]. And, as usual, he never told me about his anticipation of the general idea 30 years earlier.

Hasenjäger [1953] used the third point to analyse convincing differences between Gödel's and Henkin's proofs of completeness.

Both points are elaborated in Kreisel, Mints and Simpson [1975] and, more recently, in Kreisel [1985].

11.2 Incompleteness

Naturally, also for intuitionistic logic the formulation of incompleteness with respect to sentences expressing the soundness of Π^0_1-theorems[4] is superior to Gödel's original (Hilbert-style) formulation in terms of consistency.

Also, the purely 'syntactic' reformulation of completeness for arithmetic favoured by Hilbert, namely:

$$\text{for all closed formula } A,\ (\vdash A) \vee (\vdash \neg A), \tag{11.3}$$

would not be expected in intuitionistic logic

In connection with Hilbert's program it was necessary to prove incompleteness by finitistically acceptable methods, and Gödel emphasized that he had done so. (This is almost, but not quite true; see below.) Inasmuch as he achieved this, his metamathematical methods are valid for intuitionistic logic, too.

Today it seems appropriate to go into some fine points he (and others) neglected at the time.

Weakness of negation in intuitionistic logic

This aspect is notorious. For example, a refutation of the Π^0_1 statement 11.3 above need not furnish a counterexample A_G such that:

$$\neg(\vdash A_G) \wedge \neg(\vdash \neg A_G).$$

On the other hand, since 11.3 is not required by completeness, it is not immediate that a counterexample implies incompleteness. But this follows from a little

Exercise 11.2.1 *Write* Comp (A) *(completeness for A) for:*

$$A \to\ (\vdash A),$$

Then a counterexample A_G implies:

$$\neg[\text{Comp }(A_G) \wedge \text{Comp }(\neg A_G)].$$

(Hint: Since

$$[\neg(\vdash A) \wedge \text{Comp }(A)] \to \neg A,$$

[4] Cf. p. 142 of Chapter 7, p. 199 of Chapter 9 and Chapter 10 [∞].

a counterexample together with

$$[\text{Comp } (A_G) \wedge \text{Comp } (\neg A_G)]$$

implies

$$[(\neg A_G) \wedge (\neg\neg A_G)],$$

which is absurd.)

The matter simplifies for the usual systems, which are demonstrably complete for Σ^0_1 (but not generally for $\neg\neg\Sigma^0_1$) formulas, if a variant A^*_G is used that expresses literally:

$$\neg A^*_G \text{ is derivable.}$$

Then A^*_G is Σ^0_1, and so $\vdash$ Comp (A^*_G). Since A^*_G is $\vdash \neg A^*_G$, Comp $(\neg A^*_G)$ reduces to $\neg\neg A^*_G$, which is false, but not formally refutable (if the system considered is consistent). Since $\neg A^*_G$ (like A_G) is formally equivalent to *consistency*, the usual formulation in terms of A_G can be recovered (see below).

Gödel's work was explicit enough not to need the exercise. In fact, $\neg$Comp (A_G) was proved from the consistency of the system considered. The proof of the implication is even finitist because, for any (proposed) derivation of Comp (A_G), not only the inconsistency of the whole system is concluded, but a specific derivation of an inconsistency.[5]

Independence of Gödel's sentence

This is more delicate, because of the assumption of ω-consistency. Is the proof of

$$(\omega\text{-consistency}) \;\Rightarrow\; \neg \vdash (\neg A_G)$$

valid for intuitionistic logic? It is, for example, by a straight application of the negative translation (cf. p. 218 of Chapter 10). But at the time Gödel overlooked the question.

When I met him in 1955, he brought up the matter. It had bothered him until he noticed a footnote in my paper on the no-counterexample interpretation that gave an explicit finitist interpretation of ω- consistency. Actually, all this is a bit superfluous: for establishing underivability of $\neg A_G$, the most immediate requirement is

$$\text{soundness of derivable negations of } \Pi^0_1\text{-formulas.}$$

At the time I talked of *1-consistency*; cf. Smorynsky's detailed obituary of this notion in his [1977].[6]

[5]This is a case of the familiar sharpening, for quantifier-free A and B, of $\forall x A \to \forall y B$ to $\forall y[(\forall x \leq f(y))A \to B]$ for suitable f.

[6]For the record, my silly terminology never came up in conversation with Gödel on such matters; cf. Chapter 13.

Improvements in formulating incompleteness

In the introduction to his paper on formal incompleteness, Gödel related his first theorem to (one version of) the program of Frege and Russell about the structure of mathematical concepts; specifically, in terms of formal systems (as opposed, for example, to second order axiomatization; cf. Section 3 of Chapter 6):

> One might ... conjecture that these axioms and rules of inference are sufficient to decide *any* mathematical questions that can at all be formally expressed in these systems. It will be shown ... that this is not the case ...

Near the end of the paper he warned against relating the second theorem to (again, one version of) Hilbert's program; specifically, concerning the possibilities of purity of method for finitistically formulated problems (as opposed, for example, to the business of a final solution for all foundational problems, which is convincingly refuted already by the first theorem):

> I wish to note expressly that ... [the second theorem does] not contradict Hilbert's formalistic viewpoint. For this viewpoint presupposes only the existence of a consistency proof in which nothing but finitary means of proofs is used, and it is conceivable that there exist finitary proofs that *cannot* be expressed in the formalism [considered here] ...

Contrary to a would-be sophisticated view, these matters are not of 'purely historical interest' (every bright teenager being interested in them). But they do not constitute a principal interest of the incompleteness theorems. This has been compared to Pythagoras' program expressed in the slogan:

(rational) numbers are the measure of all things,

and the irrationality of $\sqrt{2}$. The latter remains of interest, but not primarily because it refutes such an exaggerated (and, therefore, simpleminded) program.[7]

As matters stand today, the incompleteness theorems are not literally *fundamental discoveries* rewarding unlimited elaboration or analysis, but *samples*; to be compared to the irrationality of $\sqrt{2}$ (that is, $n^2 \neq 2m^2$) as a sample of diophantine problems.

1. The first theorem is a corollary to (recursive) *undecidability results about arbitrary* Π^0_1 *sentences*. This was later improved by a variety of incomparable results: on word problems in group theory, on diophantine equations in number theory, and in many other branches of mathematics (cf. Davis, Matyasevic and Robinson [1976], and Kreisel [1985a]).

[7] Incidentally, 'purely historical interest' is probably better suited to the more difficult works of Archimedes and the bulk of all the excellent mathematics in the 19th century; they are inaccessible to the outsider, and superseded by later results for the specialist. Cf. the end of Chapter 4 on historical matters.

2. The first theorem is also a special case of incompleteness of *not necessarily formal systems*; for example, of such systems extended by all true Π^0_1-sentences (cf. Mostowski [1952], and p. 75 of Chapter 6 on implications for the rate of growth of bounding functions in the case of Π^0_2 theorems that are formally independent of all true Π^0_1-sentences).

3. Apart from the improved formulations (quoted at the beginning of this section), the second theorem has not only been appropriately extended to non formal systems in 2, but also to *cut-free systems* that have been developed since Gödel's original paper (cf. Kreisel and Takeuti [1974]).

4. As stressed on p. 79 of Chapter 6, the second theorem serves as a cross check on proposed consistency proofs. This is more useful than it seems, just because consistency is so weak (too weak, for example, as a soundness property). Consequently, many metamathematical results (for example, various kinds of *normalization*), imply consistency formally. Thus the second theorem serves as a cross check on proposed proofs of such results, too.

5. The second theorem has been sharpened to *conservation results*. Thus, while the theorem only implies that the addition of (the false formula) $\neg\mathrm{Con}_S$ to a consistent system S is consistent, in fact no new Π^0_1 sentences can be proved in the extended system (cf. p. 79 of Chapter 6).[8]

6. The single most satisfactory way available today for highlighting the second theorem uses so-called (propositional) provability logic, as follows. The second theorem is the special case of *Löb's theorem*:

$$\Box(\Box p \to p) \to \Box p$$

when $p = \bot$, and Löb's theorem for a system S follows from that special case applied to $S \cup \{p\}$ (which is consistent if $\neg\Box p$, and would prove its own consistency if $\Box(\Box p \to p)$ held for formal derivability $\Box$ in S).

Now, as shown by Solovay, together with a couple of pedestrian properties of $\Box$, Löb's theorem axiomatizes completely its propositional theory. In this sense the second theorem is (as it were, demonstrably) central for the subject of formal derivability in the usual systems, which have those pedestrian properties.

Certainly, none of 1–6 is as exciting as (Gödel's) claims about the significance of the incompleteness theorems for the nature of mind and/or matter, or

[8] *Philosophical corollary*. The interest of this reformulation of the second theorem is independent of dubious (that is, ideological) doubts about the legitimacy of current principles S. On the contrary, for such S and the usually interpretation of the formalism, $\neg\mathrm{Con}_S$ is false and so the consistency of $S \cup \{\neg\mathrm{Con}_S\}$ is genuinely problematic.

as the even more remarkable claims (of, say, Hofstadter) in connection with digital intelligence; 'remarkable' because a theorem that states what (even perfect) computers *cannot* do, is supposed to provide evidence for the unlimited potential of AI, which relies on what real time computers *can* do.

Of course, the incompleteness theorems *do* tell us something of interest about (limitations of) human minds; in particular, how exceptionally gifted people (cf. p. 50 of Chapter 6) who talked about the topic (endlessly) could miss such simple proofs. More specifically (cf. the subsection on p. 62 of Chapter 6), how grand models for the structure of mathematics and the laws of thought could be proposed *without any check* on the mathematical properties of those models.

Incidentally, this is the way young Gödel saw matters himself. As he once told me, when he submitted the announcement [1930a] he was prepared (but not hoping) for a publication of those theorems by somebody else before his appeared. In other words, he did not think of them as far beneath the surface (cf. p. 55 of Chapter 6, with his brother's phrase of Gödel 'hiding' his light under a bushel; another way, as it were, of expressing Gödel's own view that he had to work very little for those results).

Gödel's use of the Chinese remainder theorem

Viewed within the context of his incompleteness paper, the attention paid to the language of rings ($+$ and $\times$) appears disproportionate; too much for the general problem, not enough for Hilbert's 10th problem (cf. the parallel between formal derivability and solvability of diophantine equations at the beginning of Section 4 of Chapter 6).

But viewed as part of Gödel's mathematical education, his use is most *satisfaisant pour l'esprit*. As emphasized by Taussky [1987], Gödel followed Furtwängler's lectures on class field theory, where the Chinese remainder theorem is used for the very same purpose as in Gödel's paper: to code sequences of elements by a single element.

Forty years later this part of Gödel's paper can be seen as a step towards Matyasevic's result (alluded to in 1 above) about diophantine equations that did solve Hilbert's 10th problem.

11.3 Relative Consistency

The introduction to Gödel's monograph [1940] on the consistency of the axiom of choice and the continuum hypothesis relative to the remaining axioms of set theory gives pride of place to the *strictly finitist* character of the proof. Several expositions dutifully repeated this emphasis. Objectively, this is surely *le côté le moins intéressant*.

Given the potential of [1940] and the delay in further work on it, the introduction was hardly effective. For Gödel himself, the stress on consistency had comic consequences (cf. p. 104 of Chapter 6 and Section 1 of Chapter 8).

Finitist proofs of relative consistency

Around 15 years after [1940] first appeared, Gödel himself felt ill at ease, and asked me one of his offhand[9] questions about it:

> If $S \subseteq S'$ and $\mathrm{Con}_S \to \mathrm{Con}_{S'}$ (in other words, if the consistency of S' relative to S) is proved in S itself, is there also a finitist consistency proof?

The trivial answer is: No. For example, if S_1 is consistent and $S = S_1 \cup \{\neg\mathrm{Con}_{S_1}\}$, S is consistent by Gödel's second theorem, and $\mathrm{Con}_S \to \mathrm{Con}_{S'}$ can be proved in S for arbitrary S' (since $\neg\mathrm{Con}_S$ is provable in S). But $\mathrm{Con}_{S'}$ need not even be true.

I did not agonize over the proof, and normally I should not have published it. But at the time I was preoccupied with establishing the notion of *conservation*, which I found better adapted to summarizing the interest of then current relative consistency proofs. And the temptation to ask such questions as Gödel's seemed an additional weakness of the notion of relative consistency. So I published [1958].

Almost 20 years later, [∞], I found a better wording of Gödel's question. But first some reminders:[10]

- Even *finitist relative consistency proofs do not assure conservation*: if R_S is a Rosser sentence, both

  $$\mathrm{Con}_S \to \mathrm{Con}_{S\cup\{R_S\}} \text{ and } \mathrm{Con}_S \to \mathrm{Con}_{S\cup\{\neg R_S\}}$$

  have quite elementary proofs)basically, because both R_S and $\neg R_S$ have a simple form; respectively, Π^0_1 and Σ^0_1). So the bare fact of relative consistency (and even of its elementary provability) gives no information about conservation.

  In contrast, the inner model constructions of Gödel or Cohen (preserving $\in$, ω etc.) do give useful conservation results.

- *Interpretations*[11] (at least, for one finitely axiomatized system in another) *immediately yield a finitist relative consistency proof*. For example, in

[9]Cf. p. 209 of Chapter 9.

[10]The only property of 'finitist' used below is that, say, primitive recursive arithmetic is finitist.

[11]In the terminology of Tarski, Mostowski and Robinson [1953].

> Gödel's case of GB and $GB \cup \{V = L\}$. This was, in fact, the point of the introduction of GB.
>
> The converse is false.[12]

Sharpening relative consistency quantitatively

There is a rewording of Gödel's question, with an additional quantitative condition on the relative 'lengths' of hypothetical proofs[13] of an inconsistency (say, $\perp$) in S and S':

$$\forall d'[\mathrm{Prov}_{S'}(d', \perp) \to (\exists d \leq f(d'))\mathrm{Prov}_S(d, \perp)], \tag{11.4}$$

where Prov_S and $\mathrm{Prov}_{S'}$ are the proof predicates of S and S'. 11.4 has the advantage of being Π^0_1 (for an elementary f), while the unrestricted assertion of relative consistency

$$\mathrm{Con}_S \to \mathrm{Con}_{S'}$$

is only Π^0_2. I then noticed that [∞]:

> if 11.4 (of course, not merely relative consistency) is proved in S for an elementary f, then (11.4 and hence) $\mathrm{Con}_S \to \mathrm{Con}_{S'}$ has an elementary proof.

This simply because 11.4 is Π^0_1, and a Π^0_1 theorem of S can be deduced from Con_S by elementary means.

At the time, 11.4 was intended as answer to the question which summarized what (I thought I or, with luck) we have learnt from Hilbert's second problem [∞]:

> What more do we know if we have an elementary proof of relative consistency?

Again, it turns out that a restriction on the function f in 11.4 is critical; less so the method of proof, a proof in S being enough (for an elementary proof).

---

[12] If not only relative consistency has a finitist proof, but $\mathrm{Con}_{S'}$ itself is proved in S, then (by the formalization of the completeness theorem) appropriate interpretations for S' can be defined in S.

[13] *Warning*. These results involving so-called *lengths of derivations* should be interpreted as giving significance to the parameters called 'length'; at least, in one of the usual senses of significance: the parameters are used to state consequences we want to know about.

However, the parameters do not provide any measure of complexity; for example, in the sense of intelligibility of the proofs 'represented' by the derivations. This is so because the 'representation' is far too crude to serve for any analysis of such delicate phenomena as intelligibility. Specifically, a (possibly quite short) *description* of a (possibly quite long) *formal derivation* is at least as convincing as the latter, and so the 'length' of the latter used above is an artifact in connection with intelligibility.

What more do we know from 11.4 if we restrict f even further?

Harvey Friedman has given a satisfying answer:

> if 11.4 is proved in S for an *elementary recursive* f (of course, not merely elementary in the sense of finitist, in Gödel's off hand question) then an interpretation for S' can be defined in S; with variants when the class of f is extended.[14]

Naturally, Friedman's result adds nothing to the monograph [1940] which, being an inner construction, produces directly a model with particularly useful conservation properties.

Philosophical assessment of relative consistency

Far from being fundamental (except in the sense of being familiar to us since our teens) relative consistency is, on the face of things, even less compelling than consistency:

- As cannot be repeated too often, consistency is justified by the observation that it is sufficient for the truth of Π^0_1 sentences. (Otherwise, one thinks of consistent liars.)
- In the case of finitist consistency proofs of S' relative to a dubious S, one thinks of a passage in *Mr. Midshipmen Easy*, where a virtuous wet nurse was to be hired. A girl applied, who turned out to be unmarried. Did she not have a child? Yes, but only a very little one.

  S' is so little more dubious than S!

Friedman's theorem need not be presented as pursuing mindlessly the ideology of finitist relative consistency proofs. It is also a *contribution to philosophy*, establishing some significance (that is, some consequence) of elementary relative consistency proofs for a sensible purpose; specifically, with a proper meaning of *elementary*, and for the purpose of defining interpretations for S' in S. (His interpretations *may* be rewarding!) Here, finitist would simply be too crude.

As for Gödel's increased attention to requirements of intuitionistic logic on metamathematical methods: it was not misplaced in the previous section, but simply fell between two stools in the present one. The emphasis on finitist proofs of relative consistency is a *philosophical error*: the requirement is either too strong (since only the fact, not the method of proof is relevant to uses of conservation) or too weak (since for the quantitative version 11.4 of relative consistency, subdivisions within finitist mathematics are critical).

[14] Roughly speaking, S is replaced by the ramified hierarchy over S of level α, where α is the ordinal usually associated with S; for example, ω^ω when S is primitive recursive arithmetic.

Chapter 12

Last Sentences of Gödel's Publications

In this (but also throughout previous) chapter(s), problematic points in Gödel's work are given prominence. This is unusual, but I see it as a corollary of a specific and of a quite general fact. By and large, Gödel's expositions have been so effective that his unproblematic contributions have found their way into texts, and (this is the general point) usually in improved form.

At least statistically, the tradition of going back to the sources, so often appropriate in literature and the arts, cannot be expected to be equally effective in the sciences, including mathematical logic, where 'progress' has a pretty clear meaning (cf. the early part of Section 2 of Chapter 10). In view of differences between subject matter, not to speak of authors, only the short sighted would expect to rely equally on 'digging' painstakingly into sources for all aspects of all intellectual activities; even in connection with their history (cf. the end of Chapter 4), let alone, their exposition.

On a more specific (even personal) level, what is conventionally regarded as a sin (and, quite objectively, can indeed cause trouble for many people concerned) often does not trouble me at all. For example, as I see the long list of defects, not only in this chapter, I am particularly impressed by their quality. Even today they continue to suggest worthwhile reflections, and much more so than many a tame perfectly sound contemporary publication. (Of course, Gödel's defects are not recommended to the rest of us: they separate the men from the boys.)

Besides, at least to me, those defects related to his carelessness present a most welcome relief from that would-be philosophical constipated 'precision' (out of all proportion to background knowledge) that had always repelled me in some of Gödel's popular writings; for example, in [1944] and [1947].[1] Specifically, it is

[0]Originally published in *Gödel remembered*, Weingartner and Schmettered eds., Bibliopolis, 1987, pp. 157–161, as Note 5 to 'Gödel's excursions into intuitionistic logic'.

[1]The oversight in the latter mentioned on p. 104 of Chapter 6, about judging CH by its (non existent) arithmetic fruits, is not exactly in a last sentence. It was certainly not offhand,

a relief to think of that style as a pose to impress philosophers, even if in fact it only attracted philosophical cripples; rather than as coming from the heart or, in one of his favourite phrases, employed *mit Lust and Liebe*.

To judge by past experience, I am obviously disturbed by seeing a picture of somebody I knew well that conflicts with my own memories.[2] In Gödel's case, such things as [1958] (which, by Section 5 of Chapter 10, I continue to regard as a gem) are, of course, central to my picture of him; but such lists as the one below belong to it, too. Together they do a little, at least for me, towards balancing the singularly mediocre recent Gödeliana.

A final word on 'straight' (in other words, formal) errors such as the first one below. Sure, a more cautious person than Gödel would either not have been tempted to put into print an ill-considered answer to a (as it has turned out) unexpectedly interesting question, or would have employed one of the standard academic conventions for avoiding offhand (formal) mistakes; by making equally ill-considered conjectures or, most simply, by asking a question.[3]

But, objectively speaking, not without loss! Just think of the joy that our fellow brethren, who perhaps do not have many other joys in life, derive from having discovered an 'actual' error by Gödel. Again, making such errors cannot be recommended to the rest of us: their discovery would give less joy.

Summary

Remarkably many of the last sentences of Gödel's publications are defective. Here they are used mainly as pegs on which to hang sundry observations. But some will serve also as memorable object lessons; for example, about reading too much into the printed word, or about the way a whole story can be lost when things get into print. It is relatively rare that impressions of such things can be checked against fuller details, and some readers may wish to do so. The digressions at the end of this chapter are directed to such readers too.

---

but the result of an ingrained blind spot, as explained there.

[2]For example, when I first saw the selection by Wittgenstein's literary heirs among his remarks on the foundations of mathematics I described it as 'a surprisingly insignificant product of a sparkling mind' [∞]. More than 20 years later I had the good luck to find a documentary correction [∞].

[3]In this connection it should be remembered that thoughtful mathematicians are sensitive to the abuse of the word conjecture (for off-hand questions); for example, A. Weil (cf. p. 454 in Volume III of his *Collected Works*). The degree of thoughtlessness involved in this abuse (in particular, with respect to evidence for such 'conjectures') is as staggering as anything (I know) in the philosophical literature, albeit in less pretentious language. Thus 'evidence' is bandied about, without a moment's hesitation over the warnings (in elementary texts on statistics) about minimal precautions needed before such talk is profitable at all.

Gödel's prefix class ([1933a])

The best known example concerns [1933a], with a decision procedure for (the validity of) $\forall^n\exists^2\forall^m$ formulas of predicate calculus. Goldfarb [1984] has shown that, contrary to the last sentence of [1933a], the addition of equality makes this prefix class undecidable (and thus somewhat exceptional).

When doubts were raised in the 60's, I had no view about the truth of the matter, but suggested the following recipe for making a mistake: note that the full equality axioms follow from the atomic cases, and thus from universal axioms, and absorb them in one of the blocks $\forall^n$ or $\forall^m$ above (in other words, forget that a universal premise becomes existential in prenex normal form).

I once even began to speak to Gödel about this, but got sidetracked by (my own) speculations on the circumstances that favour such mistakes.

The Finiteness Theorem ([1930])

The finiteness (nowadays called compactness) theorem at the end of [1930][4] is formally correct, but defective in being unnecessarily restricted to countable sets of formulas. It came up in conversations with Gödel on two occasions.

I was struck by the fact that soon after [1930], in the note [1932] on propositional logic, Gödel stated its result explicitly for arbitrary sets. Also, I knew that he often took the opportunity of improving earlier formulations in later notes, even if they were only tenuously related.[5] As he himself described the matter, he had first stated the propositional result for countable sets of formulas too, but found, on rereading the fortunately short note attentively, that the proof nowhere used countability. So he reworded the theorem, but was not interested enough in the generalization to look for parallels.

Here it is to be remarked that (at the time) the general formulation might well have clouded the issue, with worries about writing down uncountable sets of formulas (even though, realistically speaking, arbitrary countable - not necessarily r.e. - sets are not very different in this respect). So the generalization might have limited the market for [1930].

By contrast Malcev, who used the finiteness theorem in algebra, was not liable to trouble *his* public. The finiteness theorem for arbitrary sets was needed to get his algebraic conclusions in the (unrestricted) form usual in algebra.

Before I met Gödel I had simply assumed (cf. Kreisel [1956]) that the finiteness theorem was consciously intended to answer the question:

> Given that the formal rules of Frege are complete for validity, what about consequence?

[4]Incidentally, the finiteness theorem was an afterthought that did not appear in Gödel's dissertation.

[5]For example, he put his undecidable sentence into the form $\Box p \to p$ in [1933b], and stated the absoluteness of formal computability in a note on speed up ([1936]).

the latter being defined for all sets of formulas (that one cares to consider). It certainly was not quite clear what a formal correlative might be. The finiteness theorem avoids any agonizing on this score. Gödel agreed of course, but did not remember thinking in these terms at the time.

The Constructibility Axiom ([1938])

At the end of [1938], $V = L$ (there called 'A') is described as 'a completion' of current axioms of set theory. Evidently this was out of tune with Gödel's song (and dance), starting a few years later, about the search for axioms valid for the full cumulative hierarchy (or, at least, for far-out segments). If one has a particular notion in mind, one speaks neither of 'a' nor of 'the' completion.

What he did have in mind was coded in the terminology 'L' for: lawlike. At the time he toyed with the idea that L contained all legitimate definitions of sets. And he clearly changed his mind before he gave his lecture [1946] at the bicentennial celebration at Princeton (though he was always reluctant to recognize, let alone, to enjoy any change of mind).

Incidentally, the worry in [1938] about the 'vagueness' of the notion of *arbitrary set* is removed in [1947] by a reminder about the clarity of the notion *subset of*.

The Fan Theorem ([1972])

In one of the later versions of the English translation of [1958], Gödel attributes the last sentence (stating that the fan theorem is interpretable) to me. The details are more entertaining than the fact (reported, with background, in Section 5 of Chapter 10).

At the time, I ignored his warning that he had treated his terms of higher type purely formally because I had already decided on an interpretation, for which the fan theorem was obviously valid; and I told him so a few days later. He remembered the remark when he came to write [1958], but forgot to check whether it applied to his interpretation too. We spoke briefly about this. I pointed out in print, probably on several occasions, that the last sentence was not evident (without, however, mentioning the background), and there the matter rested.

During his final illness he brought it up, apparently worrying how to draw attention to my 'contribution'. The way he saw (or, at least, put) it was this. He regarded Spector's posthumous [1962] as an important contribution (as he had already said in a *postscriptum*). In his view, given Spector's background (in particular, all he had learnt from Kleene's lectures about ordinary bar recursion and its relation to the fan theorem, and of course the idea of passing to all finite types in [1958]) that last sentence was enough to trigger Spector's result.

As we will see immediately, Gödel's view about Spector's education was wrong. But on this view, the next order of business was to find an appropriate wording for my 'contribution'. He first proposed to add 'for a slightly different interpretation'. The sequel was predictable. I asked if this was supposed to be a translation of *eine gering abweichende Interpretation*; meaning of course his blunder discussed on p. 219 of Chapter 10. The allusion had to be explained to him, and he was not amused. Obviously, it was not the person I had known for 15 years. I never asked him what wording he chose in the end.

Here is a sketch of some things I know about Spector's background, before he embarked on [1962]. When we first met at Cornell in 1957, [∞], he told me he was sick of Turing degrees, but also told me of his difficulties with the no-counterexample-interpretation, which he had once presented in Kleene's seminar. At that time he concluded (without being contradicted) that it must be nonsense, since it reduced arithmetic to Π^1_1 statements. Well, Gödel's finite types (namely, Π^∞_1) sounded worse still. I reminded him that it's not what you do, but the way that you do it, that counts; in particular, not the mere mention of the *language* of higher types, but the particular *properties* (or axioms) used, and we looked together at some striking examples. We kept in touch, and he visited me in 1959 at Los Altos near Stanford.

In the meantime, it had occurred to me that I had derived the no-counterexample-interpretation from Ackermann's version of Hilbert's ϵ-substitution method, involving in an essential way an order of priority. One had thus a relation to a principal element of Spector's logical background, Friedberg's priority method.

Also, the popular article on Hilbert's program [∞] had appeared, with a reference to some *lurking lemma* in Ackermann's work. The phrase took Spector's fancy. In fact, he had studied the matter before the visit, and thought that he saw a lurking lemma (without putting it into words; in fact, till his death he said that he had employed that lemma as a principal trick in his functional interpretation of the negative translation of the axiom of choice; cf. [1962]).

In short, his work grew out of a great deal of familiarity with ideas and results surrounding functional interpretations, helped perhaps by a few hints stemming from my own experience in this area; not least, the central place I had given to continuous functionals in my presentation at Cornell.

Ambiguities in Gödel's conversations and writings

The present digression, on Gödel's (incidentally, to me very congenial) literary taste will now be introduced (and concluded) by reference to Spector's paper. But it is a fluke that this kind of transition is possible, since the matter is general.

At the beginning of [1962], Spector quotes Gödel and Bernays as saying that bar recursion of higher type is just as evident as Brouwer's bar recursion (of lowest type). Spector himself felt encouraged, and (cf. footnote 2 of [1962]) so was I.[6]

I remember Gödel's glee when he pointed out that it could also be interpreted as follows: Brouwer's bar recursion is no more evident than Spector's generalization (cf. p. 60 of Chapter 6, with a reference to de Gaulle). I did not bother to ask Gödel how he had intended the remark originally, in line with the view of intelligent literary critics about the 'life' of a literary product being (best regarded as) independent of its author's thoughts.

Actually, Gödel's conversations were full of such ambiguities and, once sensitized to this, one finds them also in his writings. For example, the parenthetical qualification in *inhaltliche (intuitionistische) Überlegungen* of his lecture [1931a] at Königsberg will mean to the intuitionistically indoctrinated reader 'and, hence, intuitionistic', but to the seasoned logician 'here, for once, intuitionistic'. Granted those intellectual reflexes attributed to young Gödel throughout previous chapters, it is in the cards that this splendid ambiguity just came to him without any brooding.

Personal remarks

The first concerns ambiguities in my references to Gödel's intervention in choosing the title for Spector's paper.[7] Here is the full story. In accordance with the views we had come to share about the consistency business, Spector's simple title was: *Provably recursive functionals of analysis*. Gödel did not find this exciting, and proposed the addition: *a consistency proof of analysis*. If at the time I had known his Konigsberg lecture (in which he scoffed at Hilbert's claims that consistency was a sufficient condition for soundness), I should have quoted it back to him. But I didn't. Of course, I appreciated his flair for attracting attention, but my views about the sham of the consistency business have remained uncompromising. So, to water down his addition, I proposed the further qualification: *by an extension of principles formulated in current intuitionistic mathematics*, to which Gödel agreed (albeit reluctantly).

The second remark is more speculative. Presumably, all of us who have a liking for (hearing or making up) ambiguities view them as a spontaneous game of hid-and-seek as it were; as in *cache-cache*, and Talleyrand's or Fouché's *La parole a été donnée a l'homme pour cacher sa pensée*. (In contrast, for de Gaulle, cited above, it was not only a game.) But sometimes it seems to me there is a darker side to it, especially for those of us who have to do with foundational fundamentalists, notorious for their (cult of) literal-mindedness. The reaction to them is an almost inhuman coldness, viewing them as a different species, although one knows that many of them are worthy people. The game can be a

[6]See p. 251 of Chapter 10 for my later reservations.

[7]In line 8 of footnote 1 of Spector [1962] the printer omitted the words 'by adding:' after the semi colon (not colon!), and I did not correct it. This puzzled at least one reader of p. 78 of Chapter 6.

relief, engendering an illusion of *complicité* with our own species somewhere out there.

Chapter 13

On Some Conversations With Gödel

These notes contain reminiscences of Gödel, starting in the mid fifties. They may balance the picture of him provided by recent Gödeliana, badly weighed down by material from the seventies, when he had become (in the words of a secretary) less 'formidable'. Readers familiar with this stuff will notice that some of its silliest items are debunked by immediate corollaries to observations in this and previous chapters (cf. also the end of Chapter 4). Charity forbids my giving chapter and verse; anyway, in most cases I have forgotten them.

13.1 On the Proper Order of Priority in Logical Research

In the 70's Gödel spoke and wrote rather freely of this matter, and not particularly convincingly.

He spoke to me about it mainly after I had demonstrated my own interest spontaneously. Specifically, my statement in support of his election to the Foreign Membership of the Royal Society (on p. 44 of Chapter 5) pointed out how Gödel's principal results were related to simple philosophical distinctions that others had ignored. Gödel's comments on that statement (on p. 46 of Chapter 5) express the kind of pleasure he felt when others had, by themselves, come to views similar to his own.

The following reminiscences give some idea of the way Gödel liked to muse about such matters in the mid sixties.

[0]Originally published in *Gödel remembered*, Weingartner and Schmettered eds., Bibliopolis, 1987, pp. 161–169, as Notes 6 and 7 to 'Gödel's excursions into intuitionistic logic'.

Formal results by inspection of informal notions

Gödel was of course much impressed by Cohen's results. And not only because, as people often say, Cohen 'beat' him (after all, Gödel *is* on record as having been interested in the problem itself; in fact, as having considered the continuum problem fundamental; cf. [1947]).

At least as he put it to me, there was more to it. He had thought that with a problem as fundamental as (he regarded) CH, the proper strategy was to reflect on the answer for the (or an) intended meaning, and then to translate it into formal terms.

This had been his way for the constructible sets; in more traditional terms: the transfinite extension of the ramified hierarchy, with simple types being replaced by cumulative types in the style of Zermelo. Already as a student Gödel had felt sure that Skolem's argument for defining elementary submodels of any infinite cardinality would establish the axiom of reducibility (cf. p. 105 of Chapter 6). The formal paraphernalia for converting this into a relative consistency proof was heavy only as long as he wanted to avoid the use of replacement, his first contact with axioms of infinity (cf. p. 178 of Chapter 8).

Since Gödel had come to believe that CH was false for the full cumulative hierarchy, the proper strategy would be to reflect on the latter, and to convert this reflection into a relative consistency proof for $\neg CH$. No *general method* of constructing models would be needed. Cohen had provided such a method (cf. p. 109 of Chapter 6).[1]

Contrast between philosophical 'positions' and logical practice

Gödel's early exercises on intuitionistic logic (described in Section 2 of Chapter 10) are at the opposite extreme to the order of priority above, that he had come to advocate later. The style of [1958], and especially of the notes that he added for the English translation [1972], serves as a foil.

Gödel was quite aware that also his own attempts in the forties to prove the independence of the axiom of choice did not employ the strategy above, but a reinterpretation of the logical particles, in clumsy syntactic terms to boot. It is fair to say that the idea behind it is very well expressed by means of Boolean-valued models.

The see-saw continued into the fifties, with his encouragement of work on intermediate r.e. Turing degrees for new ideas on CH (cf. p. 60 of Chapter 6).[2]

[1] Gödel knew what a mathematical method was! *He* never used (at least, not when I used to see him) this word for the fixed point lemma, nor for the enumeration of formal objects (that is, for the idea of Gödel numberings), nor even for the formal definition of the arithmetic operations derived by that enumeration from those on the formal objects.

[2] This suggests at least a couple of thoughts.

In the opposite direction, as it were, Cohen professed in [1971] to be a *formalist*, after he had used *models* in his independence proofs successfully, and had given a feeble relative consistency proof at the end of [1964].

Gödel's way is more congenial to me; he bandied *platonism* about after troubles with the *syntactic* methods alluded to, and with interpreting his primitive recursive terms of finite type purely *formally*.

Effects of 'ad hoc' solutions for fruitful problems

This item is touching, but seems to have a moral too. On one occasion Gödel mused about not having published those syntactic reinterpretations; people might have misunderstood him to mean that those were the right interpretations. I never misinterpreted his remark to concern his conscious reason at the time, but rather abstract possibilities; in particular, situations where attention is drawn away from a potentially fruitful problem by an *ad hoc* or otherwise unsatisfactory (albeit correct) solution. (A formal error often draws attention *to* the problem; cf. p. 268 of Chapter 12.)

Individuals of a certain temperament do, in fact, worry that a better solution may not be so widely acclaimed as a first solution of problem. Other temperaments derive confidence from knowing the answer, or like the idea of doing better than a well known author (especially if the latter has taken the trouble to treat the problem in print), and so forth. With *current* mass activity, Gödel's simple view of triggering chains of events may well apply to *somebody*, though not in the cases to which he applied it specifically.

For example, by p. 272 of Chapter 12, not to Spector. Nor to the (published) monograph [1940], with an introduction preoccupied with legalistic precautions: it is quite remarkable how little work was done on the constructible sets in the 40's and 50's though, as Jensen has shown, there was a lot more to discover about them.

Personal remark. Partly because of the musings above, I began to record my expectations of various projects, especially in periods of consolidation; by p. 140 of [1971], with the explicit purpose of checking them against later experience. As so often (in accordance with T.S. Eliot's memorable phrase), the temperamental preferences above have pertinent objective correlatives.

On the positive side, there are certainly objective relations between (perfect set) forcing and, say, some (of Spector's) methods in the theory of Turing degrees.

On the negative side, and perhaps more interestingly, Turing degrees of r.e. sets require at least a certain measure of symmetry between (the complexities of) the sets and the mappings considered; in contrast to CH for the full cumulative hierarchy; cf. p. 122 of Chapter 6 ('certain measure' because, as some of us sensitive souls have complained, r.e. sets and their complements are severely asymmetric with respect to proof of membership, while Turing reducibility is not).

13.2 On Titles, Terminology and Other Expository Devices

Our ordinary view of human nature (as described at the end of Chapter 4) requires a kind of causal interpretation of the anecdotes below; involving motives (rather than 'mere' reasons). I do not often find the view compelling; certainly not here. And I have nothing to contribute to such interpretations.

Gödel often spoke of expository tricks to be found in his publications, and I never bothered to ask if he had had them consciously in mind at the time or meant their general relevance (to repeat: not cause and effect, just as one does not ask for cause and effect when relating equilateral and equiangular triangles).

The anecdotes below will be used to enrich the usual view of Gödel's work, and also of its relation to later work in logic. The section proceeds by easy stages from the sublime to the ridiculous, with a digression at the end.

The meaning of a theorem is its proof

This slogan was current in Gödel's student days, and it fits well the boring statements of theorems in the constructivist literature (and, of course, its interpretation of the logical particles as, literally, operating on proofs).

Gödel had a strong dislike of it. At least in conversation with me he insisted that only results be mentioned, since their pattern might be obscured by the proofs (not only mine, but also by nice ones done by others).

What about information contained in proofs, but not stated in the theorem? For Gödel the first order of business was to state elegant and memorable theorems. Afterwards people can look at the proofs for additional information of interest to them.

His practice followed this principle, very much in contrast to Herbrand and Gentzen who (before and after the completeness and incompleteness theorems) used all purpose expressions (*Verlegenheitsausdrücke*) like *thèoréme fondamental* and *Hauptsatz*, without any explicit indication of what made those theorems so 'basic'.

Some 10-20 years after them, I attempted to find concepts more adequate for expressing at least some of the additional information provided by their kind of metamathematics; for example, *functional interpretations* rather than consistency, *conservation* rather than relative consistency, and many more besides. I do not think that Gödel felt comfortable with those concepts, though by now they are familiar enough to be considered memorable.

A blind spot. During the period of our frequent conversations I had not yet realized clearly enough (for a rewarding discussion) that contemporary mathematics has its own response to the slogan above:

> Find the concepts needed to state theorems that express the meaning of (in the sense of what is called, crudely, essential in) a proof.

Generalizations (tacitly, in terms of skillfully chosen concepts) often do just that; cf. Kreisel [1985a]. This is philosophical progress (for the kind of philosophy meant in Section 1 of Chapter 10); not least, for the face-saving powers of the low-keyed mathematical style. Nobody feels ashamed about having to search for some (or even a particular kind of) generalization. But many feel ill at ease when they do not know what is essential about a proof, or what it means.

My titles and terminology

The following specific conversations on titles and terminology touching topics of 'constructivity' seem pretty typical.

Soon after we first met, Gödel made fun of the title of [1950]. The content appealed to him; a twist on his incompleteness theorem that would have been perfectly accessible in 1930, with an undecided sentence U in Δ^0_2, but described in [1950] in terms close to what would today be called: of degree $\leq \mathbf{0}'$. The proof leaves open the parameters (systems and/or codings considered) determining whether U is true or false (cf. Manewitz and Stavi [1980] for partial results). As noted later, the twist yields what is still the most 'logical' model-theoretic proof of Gödel's *second* theorem (cf. Smorynski [1977a]).

Gödel found it odd that one could be clever enough to find the results, but not a sensible title. Obviously, one does not spoil such a remark in conversation by boring analyses of the circumstances. But here it seems worth adding that (I thought) I had not achieved one of my aims in [1950]; specifically, of a consistent formula without any recursive model (in fact, even without any recursive valuation for its *atomic* properties). What I had actually proved was that GB including the axiom of infinity (or, more pedantically, its Skolem normal form) had no such model. And at the time I did not know that GB had any model at all! I did not know the cumulative hierarchy; [∞]. Nor did I have the experience needed to find the words appropriate in such circumstances.

Gödel complained, equal pertinently, about *basis* in 'basis theorem'. The notion is popular enough, at least, after Kleene's exposition [1955], but probably not the way I have always looked at it. It was to extract the sensible side of Ockham's razor without the absurdity of supposing that things do not exist when they are not needed (to handle the phenomena and problems that have so far struck us).

After he saw the properties of *absolutely free* (choice) sequences in Kreisel [1965], Gödel proposed the term *lawless*. (Neither of us knew at the time Brouwer's earlier analysis.) In choosing between the two terms one faces the familiar issue between freedom and license (lawlessness). In the late 50's I certainly felt that even the thought of restraint, e.g. the possibility of a diet, was a restriction of freedom, and so I naturally used 'absolutely free'. With the years the lure of licence has diminished, and I used 'lawless' in [1968]. May the healthier and livelier (called 'hippier' in the 60's) readers not be mislead by it!

Several points are to be added:

- At the time, I did not connect the proposed terminology with his code meaning for 'L' (for 'lawlike'); the pair 'lawlike' and 'lawless' is not only catchy, but easily translated: *gesetzmässig* and *gesetzlos*. (I take it he saw no need to use a code like 'L' here, because I was not going to be reticent about the meaning I intended anyway.)

  Incidentally, today I prefer the terminology *open data*, because it expresses explicitly the relevant enrichment.

- Gödel was touchingly pleased by the innocent successes of so-called informal rigour applied to lawless sequences; in other words, of the strategy he advocated in connection with sets (cf. Section 1). For example, the axioms for open data, or the decidability of extensional equality for lawless sequences, are recognized by inspection. Certainly, so far, the notion of lawless sequence has been quite comparable to the primitive notion of set as a *source of axioms* (though not as a scientific tool); and much more rewarding in this respect than the primitive notion of effective rule (cf. 249 of Chapter 10 on the contrast between the latter and that of set).

- At least as I see things now, the principal philosophical interest of lawless sequences derives from the object lesson they provide for the topic of natural languages (elaborated in Section 3 of Chapter 10). To repeat, a precise and elegant development is perfectly possible, but paraphrases are just as effective.

- At least so far, nothing much has come of Gödel's great expectations for lawless sequences of higher type objects.

Digression on 'manipulating the reader'

The next two subsections will be extreme examples of the kind of musing (on 'manipulating' the reader) that Gödel enjoyed very much. But, for a proper perspective, it seems necessary to get a couple of generalities out of the way. Thinking about such manipulation (from advertizing to formal education) involves not only highly publicized 'normative' elements about desired effects, but costly empirical elements in assessing actual effects; 'costly' because either many people are involved or, as in the case of education, only long term observations are of much use.

So, on the negative side, it's hard to establish any conclusion (and therefore often difficult to refute silly opinions).

On the positive side, as already mentioned in Section 1 about current mass activity, Gödel's short term view on getting immediate attention by at least a few able people is perhaps more effective than it used to be; the activity will lead to a kind of natural selection, not only of concepts (cf. the story of determinacy on p. 116 of Chapter 6). With relatively few exceptions, silly (ideologically inspired) shibboleths have been quietly dropped in the last 15 years. For example:

- One no longer speaks of 'deriving consequences from the axiom '$V = L$' when in fact one is proving properties of L (which generally have those consequences as trivial corollaries, but not conversely).
- Slowly one is beginning to talk of the rate of growth of bounding functions for Π^0_2 theorems, and not only of formal independence (when in fact one has proved independence from all true Π^0_1 theorems; cf. p. 75 of Chapter 6).

The progress achieved in this way becomes spectacular if contrasted with one of the exceptions; especially with Nerode and Harrington [1984], which revives remarkably many thoughtless first impressions that were corrected by logical research in this century.[3]

Here are the two promised titbits, both about the incompleteness paper.

A role of formal detail in [1931]

When I once mentioned to Gödel that the introduction to his incompleteness paper was fully convincing he agreed, but thought that the masses of formulas in later parts had served to avoid futile discussions about any relations between his work and Finsler's [1926]. Nobody looking at both would even dream of worrying about such relations. Several things should be added.

First, a concise formulation of those relations needs, as so often in such cases, considerable familiarity with the subject (cf. note 14 on p. 71 of Chapter 6 about a quite closely related business with Zermelo). So it is not sensible to have outsiders worry their heads over relations to Finsler (even if academic etiquette requires some kind of reference; tacitly, for insiders).

Secondly, on the negative side as it were, the price paid for introducing all those formulas was high; it is (at least, to some extent) the superstition that Gödel's proof is subtle.

Last but not least, as I realize now, the general presentations of the incompleteness theorems in Hilbert and Bernays [1939] (which, according to Bernays, incorporates several suggestions made by Gödel during their translatlantic crossing in 1935) are not much longer than Gödel's original introduction; of course, the

[3]The drivel about the fundamental character of relative consistency is put in its place in Section 3 of Chapter 11.

verification of the general conditions by a specific system may take some effort. This is a, philosophically, more satisfactory balance.[4]

Digression on details. Gödel once spoke about a very long typescript of Chomsky's that circulated in the sixties and consisted almost wholly of critical observations. Though formulated in quite different terms, Gödel's reservation was that the subject simply had not reached the *threshold* where this kind of detail was rewarding. Without exaggeration, much the same applies to the bulk of natural history, which of course is particularly proud of its painstaking detail.

A potential use for a Part II of [1931]

As is well known, the title of the incompleteness paper suggests a sequel, but none has appeared. I never asked Gödel about the general circumstances, which might allow one to judge to what extent it would be sensible to speak of 'causes'. Anyhow, he volunteered a view on the matter.

If there had been massive and systematic misunderstanding of the paper, Part II could have been used to give a full statement and proof of the second theorem, so to speak as its principal purpose; and some of the (actual, not merely imagined!) misunderstandings would have been corrected incidentally.

Viewed this way, my leaving Gödel's work on intuitionistic logic out of Chapter 6 has turned out to serve a similar purpose. I have used Chapter 10 (not so much for correcting misunderstandings, but) for reiterating certain points of the former.[5]

Logical work in styles different from Gödel's own

Some ideas of Gödel about logicians with or without his flair for flashy formulations:

1. He often called *Gentzen* a better logician than himself. Obviously, Gentzen was not more interesting; but his results were not in the air (nor on the surface: it took a long time to see convincing implications).

2. Gödel had such a high regard for *Kleene*'s contributions, that his wife complained about the Institute for not making Kleene a professor there. In particular, always appreciative of a twist to his incompleteness theorem, Gödel talked with relish of the formulation in terms of disjoint r.e. sets that are not recursively separable.

[4] For other abstract formulations cf. note 17 on p. 74 of Chapter 6, 7.2.5 on p. 140 of Chapter 7, and 9.1.1 on p. 199 of Chapter 9.

[5] This purpose certainly did not occur to me when I started on Chapter 6, nor even on Chapter 10, which sets out in the opening paragraph my (conscious) reasons for neglecting intuitionistic logic in the former.

3. By 1957 he had so much confidence in *Scott* that he said he expected him to prove soon the formal independence of CH. And, in fact, Scott contributed fairly soon to Gödel's favourite proof (by use of Boolean-valued models).

4. When he told me how much he liked *Takeuti*'s contributions to proof theory, I asked for a summary. Knowing my logical tastes, Gödel said instead that Takeuti had so much talent for seeing through complicated combinatorial situations that he did not need the kind of abstract view I wanted (and that he, Gödel, could not formulate adequately such a view). With due regard for conversational licence, I'd say much that same after my experience in preparing the joint paper Kreisel and Takeuti [1974].

Perhaps 1–4 balance a little the embarrassing exaggerations in Gödel's interview with *Time Magazine* after *Friedberg*'s solution of Post's problem: as (bad) luck would have it, Gödel went out of his way to say that this kind of achievement happened once in a generation; more or less when Mucnik's paper was in press.

13.3 Tricks of the Memory

The subject has fascinated many; most recently, the writer Marguerite Duras (*L'amant*), who made a point of recording her memories as they presented themselves to her, though it would have been easy to correct many quite obvious errors (for example, chronological discrepancies).

From a solemn point of view like Freud's, one would look for specific causes of each slip; a kind of hybris to suppose that they can be found even approximately without knowledge of memory structures (cf. the end of Chapter 4, on the inner life of planets). Of course, as always, there are exceptionally favourable circumstances [∞].

Here we treat those tricks quite differently, along the lines of this chapter.

Gödel [1938] and Hilbert [1926]

A particular trick of memory involves Gödel's recollection (admittedly, in his less 'formidable' period) that his work on CH had nothing to do with Hilbert's sketch in [1926] (cf. note 33 of Chapter 10). The facts are plain enough:

- In notes for a lecture at Brown University (1938) in the *Nachlass*, Gödel explicitly says that he had recently discovered a presentation of the work that was closely related to Hilbert's sketch.

- At the end of his review [1940] the genuinely cautious, not merely calculating Bernays says the same thing; and, after all, anybody competent can verify some objective relations.

Granted the similarities between Hilbert's sketch and Gödel's proof, the differences are surely much more impressive and consequential:

- As in the ε-substitution method, which was a main topic of his paper anyway, Hilbert thought of the ordinal-theoretic functions involved as 'essentially' finitistic.

- For Gödel's (successful) work it is essential to think of them, in modern terminology, as α-recursive (for constructible cardinals α).[6]

Thus for Gödel's primary concern (cf. the beginning of Chapter 11) of making his work accessible to a wide audience (and so necessarily without much background), it is certainly most appropriate to draw attention *away* from the similarities to Hilbert's sketch, since a *correct* appreciation of them demands substantial background knowledge. All this looks bad from a more solemn point of view.

First, we have a conflict with academic traditions about acknowledging priorities. But since human beings (and situations in which we find ourselves) differ, the best we can hope for from traditions, including laws, is that they are appropriate in many cases. Besides, academic traditions are not primarily concerned with a wide market.

Secondly, Gödel's later views of the facts occur, if I remember correctly, in letters (meant for posterity to boot), not in a casual conversation.

Instead of (or after) enjoying the thrill of indignation, let us take a look at the solemn view.

Documentation versus impressionistic anecdotes

In political history, letters and secret memoranda or tape recordings have proved to be useful, often more so than relying on public pronouncements; at least after extensive sifting, with due attention to the temperaments of the authors. Certainly, *one* element that contributes to reliability is that the documents are produced by many different people, and affect even more. Recriminations after the event automatically generate new material. *This* element is not present to the same degree in scholarly doings, and further restricted by academic etiquette.

A principal assumption of the history of ideas is that, in contrast to politicians, scholars do not try to manipulate their public. But even when this is true (and Section 2 serves as a *caveat*), the scholars, and above all the pioneers, often simply

[6] In 1936, and later in the 70's (cf. [1972]), Gödel himself felt that he had provided a first step towards the kind of further 'collapse' to finite structures needed for Hilbert's program. But, before Girard's (admittedly, blatantly transfinite) iterations of limit processes applied to dilators, there was hardly a hint for achieving that kind of collapse. (The collapse of cardinals into recursive ordinals in Bachmann's hierarchies was not enough, at least not for me, to inspire confidence.)

do not know how to say what they know (cf. p. 278). So, realistically speaking, as a reflection of their thoughts, words of scholars (whether written or spoken) may be less reliable than those of politicians.

Without exaggeration, when the history of ideas apes the kind of documentation familiar from ordinary history it is liable to become a parody. All this is, of course, in keeping with the suspicion of the history of ideas among the silent majority of historians.

Here it should be added that many professional historians are also unsympathetic to the use of history for political rhetoric (which, for the record, I like; just as much as for entertainment, despite Nietzsche's diatribes in *Unzeitgemässe* . . .).

In mathematics such rhetoric tries to convey a view (for example, on the relative promise of different methods) by presenting a suitable selection of historical titbits; including (of course) hopelessly false starts, according to that view.[7]

The historians' antipathy goes well with the fact that pseudo-history can be quite as effective rhetorically as real history, provided it catches our imagination.

The following points about letters and publications are quite down to earth:

1. Bertrand Russell complains in the introduction to (the second edition of) his book on Leibniz that the latter never wrote a *magnum opus* of universal interest. In Russell's view, because Leibniz wrote too many letters to princesses; presumably (if there is to be a conflict at all) relying too much on references to specific interests of those ladies.

2. Depending on the temperament of authors, formal publications (especially on intimate subjects like broad views) are not always as different from letters as 1 suggests. Speaking from personal experience, while writing I often catch myself addressing particular readers; sometimes dead ones, sometimes present company. Even if I am a bit exceptional in this respect, my aberrations probably merely magnify a widespread phenomenon.

3. The same surely applies to the following anecdote about an, admittedly, exceptionally simple-minded person. By a fluke I had recently the rare luck of learning the interpretations by that person of some letters I had written to a friend. Phrases that just happened to have caught the ear of the friend in conversation, were given portentious (and always inappropriate) meaning. Others that were of genuine concern to the person, were completely overlooked.

[7]Naturally, there is plenty of relevant literature on the matter, particularly in recent years. For example, (some) survey lectures and reviews in the *Bulletin of the American Mathematical Society*; or those histories where the author likes to identify himself with the hero (and so has a lopsided historical view), but happens to have something significant to say about the mathematical content of the hero's thought.

Of course, this does not mean that it is totally unrewarding to read letters not addressed to one; but it underlines the unusual skills needed to guess the intentions. Again, this does not mean that an interest in intentions is illegitimate or, perhaps, logically 'senseless'. It just isn't rewarding if the chance of anything beyond a very rough general idea is slight.

Part III

CONCLUSION

Chapter 14

Contemporary Logic

The bulk of contemporary logicians are engaged in the internal development of the (mathematical) disciplines currently thought of as the proper province of logic. These disciplines constitute a (mathematically) heterogeneous mixture. What is common is that they developed from the well-known range of foundational schemes (originally for arithmetic, then for the whole of mathematics) associated with Frege and Russell, Cantor and Zermelo, Poincaré, Hilbert, and Brouwer. As with any other schemes, mathematical tools are needed to formulate them precisely enough in order to make a detailed analysis rewarding (with a proper proportion between the degree of accuracy of the claims to that of the data).

The schemes were intended as *rival* views on the nature of mathematics and of mathematical reasoning, usually associated with traditional (rival) views on the nature of the world and our knowledge of it. But (and this is one of the most decisive contributions of contemporary logic for the foundational tradition) the mathematical laws actually derived from the would-be rival views, admit also an interpretation as *variants* of the original scheme (by Frege and Russell); specifically: as simplifications, refinements, and extensions.[1] This development follows the pattern of most would-be dramatic philosophical rivals in the past; from nature vs. nurture in phenomena of perception, to selection vs. modification (of objects occurring naturally) in the anthropology of tools: with extended experience, the drama disappears. As a corollary, the currently widespread interest of logicians in the different logical disciplines, associated with those 'rival' views, is not necessarily related to lack of moral fibre.

On the contrary, familiarity with the broad canvas of all the logical disciplines, and at the same time with the mainstream of contemporary mathematics (together with such manifestos as Bourbaki [1948] concerning the matter) reveals (what may properly be called) a *logical view* of mathematics, which stresses aspects cultivated in contemporary logic, as opposed to those emphasized in con-

[0]This chapter is based on lecture notes for the seminar 'Contemporary logic', Winter 1983–84, Stanford University, and it is published here for the first time.

[1]Cf. Section 2 below [∞].

temporary mathematics.

Reminder. Certain elements originally introduced for the sake of foundational schemes have by now been totally absorbed into contemporary mathematics (and culture generally). For example: the language of sets, the notation for ordinary logical operations with their simple vocabulary and grammar, the use of models for independence results (without reference to necessarily dubious specifications of rules of proof), the notion of formal rule or computer program (for the so-called perfect computer).

Scope of the chapter

We will consider, roughly speaking, a period of 50 years, starting from the mid thirties: by then enough notions and results had been established to assure the visibility of logical disciplines as branches of pure mathematics (and not only as tools for foundational schemes).[2] *Samples*:

- relating inaccessible segments of the cumulative hierarchy to (what are nowadays called second-order) familiar axioms of set theory (Zermelo)
- non-categoricity of first-order logic (Skolem and Löwenheim)
- the finiteness theorem, together with recursive enumerability of logical validity (Gödel)
- completeness of familiar axioms for the whole first-order theory of substantial mathematical notions (real closed fields) by use of quantifier elimination (Herbrand and Tarski)
- existence of a recursively enumerable nonrecursive set, and hence incompleteness of formal (i.e. recursively enumerable) systems of arithmetic (Gödel)
- improved generation of logical valid formulae 'without detours' (Herbrand and Gentzen)
- formally elegant rules for intuitionistic logic (Heyting), with unsuspected relations to ordinary logic (Gödel and Gentzen) and certain systems of modal logical (Gödel).

The style of publications during the period here considered changed at around half time, when the effect of the summer school at Cornell in 1957 (whose Proceedings were published by the *Institute for Defense Analysis*) began to be felt. Afterwards a much larger proportion of specialists in one branch of logic had

[2] This is in keeping with the broad tradition of science. For example, some notions and results of the theory of functions of a complex variable were originally introduced for developing the idea of (two dimensional motion of) a perfect fluid, and have been separated from the latter (especially when that idea proved to be a poor idealization).

relatively adequate knowledge of progress in the other branches, or at least the successful ones.

Evidently, for details and bibliographical information readers are referred to specialized accounts of those different branches. But there are enough features which are common to many of those branches, and different from the mainstream of mathematics (both w.r.t. methods and to applications) to require a broad exposition of contemporary logic as a whole. This is the purpose of the present chapter, which is organized as follows.

In Section 1 some of the mathematically most striking *features* will be indicated (informally) by means of samples; *both* w.r.t. neglected parallels in branches of mathematics outside logic (so to speak debunking the often exaggerated mathematical originality involved), *and* w.r.t. the imagination displayed in finding some relevant uses of logical discoveries originally introduced for dramatic, and thus dubious, (foundational) idealizations or issues. Readers are invited to check for themselves that, though generally the same methods of proof are used, the theorems so proved that are relevant to the old and new issues are generally different (reversing the advice: if you want to know the meaning of a theorem, look at the proof).

In Section 2 samples are given of *contributions* made by foundational schemes, correcting the widespread impression that the principal weakness of those schemes was lack of precision; with the implied hope of (near) miracles if we only knew more about the notions and problems regarded as basic in those schemes. In particular, experience in contemporary logic corrects the significance (for the broad concerns in the philosophy of mathematics) assigned by some of the schemes to such matters as completeness and incompleteness, not to speak of the importance of formal semantics for those matters, or of the hoary business of the paradoxes. Also in Section 2, results from contemporary logic are used to introduce the idea of a (common) *logical view*, of which the so-called rival schemes are variants.

In Section 3 the *assumptions* of the logical view, expressed in such logical ideals as a (logically) universal language, or the (logical) possibility of formal rigor, are examined by reference to some perennial questions about improving our mathematical capacities, including the (mathematical) language currently regarded as natural, but radically changed within a couple of generations. The examination consists in contrasting logical requirements on answers to those questions with solutions implicit in the mathematical literature. Since the latter does not have the concepts needed for to formulating those questions, contemporary logic helps indirectly to make explicit the philosophical relevance of the mathematical solutions. Here, in contrast to Section 2, mathematical logic is not intended as a tool of the philosophy of mathematics which solves philosophical issues by means of (neat) metatheorems; rather, it presents certain phenomena of mathematics and (both good and poor) mathematical reasoning in a so to speak chemically pure, memorable form.

Finally, in Section 4, Computer Science and Artificial Intelligence are considered, as the prime examples of the passage *from foundations to technology*; that is, an unrealistic idea about (biological) *processes* of reasoning turns out to be realizable by existing (electronic or photonic) technology, where the realization achieves some of the *results* of those processes equally well or better (for example, more cheaply). Again, as in the remainder above, some elements of contemporary logic (for example, of recursion theory with its use of the same code for arguments and programs) have been absorbed into the subject. But beyond this point contemporary logic provides above all logical hygiene, in *proving* that most (logically) ideal problems do not lend themselves to an algorithmic treatment at all.

14.1 Mathematical Features

Despite slogans about the 'unity of mathematics' there is considerable diversity not only within mathematics, but within branches of mathematics. *Samples* (apart from bright ideas, both for ingenious constructions and imaginative new problems which - so to speak, by definition - do not fit into any general scheme):

- There are relations, sometimes called *applications*, of one part of the branch to others, and of the branch to other branches of mathematics; in particular, those recognized as central (the so-called mainstream).

- But also there is the matter of a *new interpretation*, the need for a shift of emphasis, when the current interpretation has reached the point of diminishing returns. This side of research is particularly acute when the current emphasis (for example, in the choice of problems) derives from - the mathematical formulation of - a false conception, such as a false physical theory or logical view. In this respect mathematical logic is exceptional in the 20th century, since contemporary mathematics has concentrated on the axiomatic analysis or generalization of earlier mathematics that has already proved to be successful.

The samples above are meant to indicate classifications of mathematical results (incidentally, cutting across familiar logical classifications). They are not a blueprint for progress because, quite trivially, there are sensible and silly shifts of emphasis.

1.1) Mid thirties to the end of the fifties

As might be expected the diversity in question is easier to see in the years of consolidation during the first half of contemporary logic (before Cornell) and, with the help of hindsight, easier to judge reliably.

a) Reversing an earlier trend: positive aspects of negative theorems

In model theory and recursion theory (the theory of formal rules), *non-categoricity* and *incompleteness* were reinterpreted. The question was:

> What more do we know if a property can be defined by first order formulas than, say, just set-theoretically? Or if an arithmetic theorem can be proved by some formal rules than if it is merely true?

The *finiteness theorem* was applied directly (Robinson, Tarski), but also (more ingeniously, and even earlier) certain properties defined by higher order formulae could be shown to be definable by *infinite sets of first order formulae* (Malcev). Thus a specific mathematical 'reduction' was combined with logical generalities about first order formulae.

More importantly, the *reasoning by analogy* made available by set-theoretic language (tacitly, over a structure S), that is, transfer of set-theoretic properties P to all structures isomorphic to S was extended (by Lös) for properties P' formulated in elementary logic over S: P' transfers to all ultrapower of S too.

For any system $\mathcal{F}$ of formal rules for arithmetic there is a recursive function $\mu_{\mathcal{F}}$ which dominates the *rate of growth of bounds* (of y in terms of x) for any provable Π^0_2 formula $\forall x \exists y A$, with extensions to functionals $\mu^+_{\mathcal{F}}$ for provable Π^1_1 formulae $\forall f \exists y A$ with a function parameter f. For effective use, mathematical analysis was needed to find (provable) equivalents in Π^0_2 or Π^1_1 form to assertions presented in different (syntactic) form. Formal independence (not from $\mathcal{F}$ itself, but) from $\mathcal{F}$ together with all true Π^0_1 (or Σ^1_1) formulae gives lower bounds in terms of $\mu_{\mathcal{F}}$ (or $\mu^+_{\mathcal{F}}$).

Though more special, the shift in emphasis in connection with the method of *quantifier elimination* is quite comparable. Originally introduced for decidability results, it was soon noticed (by Tarski) that their consequence for definability was more rewarding, and not affected by the obvious algorithmic efficiency, a kind of parallel to Ockham's razor of eliminating claims that are unrealistic (even if not literally false).

b) Can't one do better?

In proof theory with proofs of results selected by ideological preoccupations, *consistency* or *relative consistency* proofs (of familiar axioms with familiar models) were reinterpreted.

Consistency concerns only Π^0_1 formulae, in the sense that a Σ^0_1 theorem of a consistent system need not be true (e.g. $S \cup \{\exists x A(x)\}$ is consistent if $\exists x A(x)$ is formally undecidable in S). But consistency *proofs* for $\mathcal{F}$ also generally supply the *bounding functions* $\mu_{\mathcal{F}}$ and $\mu^+_{\mathcal{F}}$ discussed in (a).

For relative consistency proofs, two new interpretations were discovered at an early stage. The first is expressed by so-called *conservation results*; for example, formal set theory including the axiom of choice has a familiar model, and so a proof of relative consistency to set theory without that axiom is a teratological exercise, but it is not obvious from the familiar model whether new arithmetic theorems can be formally derived by additional use of that axiom. The second uses of a particular method for proving relative consistency (spotted by Tarski) to transfer *recursive undecidability results* (to theories which have a so-called inner model or interpretation in the given theory).

c) Relations of logical notions to more central parts of mathematics

Evidently, inasmuch as the work of (a) and (b) concerned specific branches of the mainstream (as it did), it comes under the present heading. But people looked also for broader relations, without much success.

Thus Tarski and his students aimed at algebraic closure conditions in models of formulae F equivalent to the syntactic (prenex) form of F. Conversely, starting with specific structures especially in algebra (for example, ideals) Robinson introduced notions which are meaningful for all first order formulae, and specialize to the familiar notions when applied to the specific structure.

The much more difficult matter, of spotting phenomena (for example, open problems) to which such formal generalizations are actually relevant, made little progress. Recursive undecidability of *word problems* (for various kinds of semigroups and groups) were not only formally stronger than those for the corresponding full first order theory, quoted at the end of (b), but also mathematically more substantial. Evidently, this goes with the improvement in the opposite direction, for (positive) decidability and definability results provided by quantifier elimination, from purely existential to arbitrary formulae.

Remark. The next section is not sharply separated from the present one, since many metatheorems involve usually only a simple device from mainstream mathematics (for example, König's lemma for defining models of consistent formulae). So a relation between, say, recursive and such models induces a relation between recursion and compactness.

d) Relations between different parts of logic

They were pursued quite successfully. The most familiar line was to apply proof theory, that is, particularly transparent *generations of the set of logically valid formulae* (for example, by means of cut-free rules) to results concerning only provability and hence, via completeness, model-theoretically meaningful; cf. interpolation (Craig).

More delicate *relations between recursion-theoretic and model-theoretic notions* (then called 'absolute', nowadays more often 'invariant') were stressed throughout the 30's (by Gödel). In the same area models were classified by recursion theoretic measures, of prefix complexity or degree; for example, lower bounds for all models of particular (consistent) formulae, and upper bounds for some model of arbitrary ones; cf. also the *Remark* above.

A more special, but also more imaginative relation was discovered by Robinson in connection with *quantifier elimination*, mentioned already at the end of (a): his model-theoretic interpretation made possible a quite new way of establishing such elimination for familiar structures (by allowing the exploitation of specific knowledge about them to establish the relevant model-theoretic property). *Reminder*: quantifier elimination allows one to infer, e.g. in the case of $\mathcal{R}$, that if any real satisfies $P(x)$ then there is an algebraic one that does too.

e) Internal development of logic

This is not the place to go over such permanently useful contributions as Kleene's *disjoint r.e. sets which are not recursively separable* (extracted from a minor improvement by Rosser of the incompleteness theorem), or the introduction by Lorenzen and Novikov of *infinite proof figures* into proof theory (calibrated by ordinals in articles of Schütte and others). But two contributions, so to speak, at opposite ends of the spectrum, are relevant to a contemporary prospective.

In set theory, the *constructible hierarchy* L (of Gödel) transferred - in effect if not by intention - one of the most useful lessons of algebra to logic, learned for example from the algebraization(s) of the continuum. For $\mathcal{R}$ the familiar axioms (of Dedekind) are immediate; but soon many questions about $\mathcal{R}$ become unmanageable. On the other hand for first order questions about $\mathcal{R}$ very little of $\mathcal{R}$ is used. Experience shows that it is rewarding to consider real closures[3] of arbitrary ordered fields K, obtained from K generally by transfinite iterations of adjunctions of elements. The most obvious properties of $\mathcal{R}$ may not since verification or even modification, but the later development is smoother: a case of *reculer pour mieux sauter*. The general algebraic technology of adding (skillfully chosen) transcendental elements is an obvious (conscious or unconscious) source of ideas for enriching (Gödel's original) work on L. Be that as it may, already his own work (which, after all, treated infinitary set-theoretic operations in [1940], and not only finitary algebraic operations) constitutes a very significant advance in (manipulating functions on the ordinals) over anything that had been done before.[4]

Remark (for reference on p. 301). Most of the work on the hyperarithmetic hierarchy (by Kleene) is most simply reinterpreted by reference to the fine structure of L up to the first nonrecursive ordinal, though Kleene did not see that aspect. In technical terms the difference is this: the hyperarithmetic hierarchy is a version of the *analytic* ramified hierarchy, while L is (a version of) the cu-

[3]Completeness of the theory is not relevant; cf. the pythagorean closure.

[4]The closure conditions on the ordinals involved are of course satisfied by countable ordinals, even though the latter are harder to describe than the cardinals which satisfy the relevant conditions.

mulative *set-theoretic* hierarchy. Because of the stability of the notion 'the least nonrecursive ordinal' (with 'elementary recursive' or Σ_1^1 in place of 'recursive'), the work does not use specifically recursion-theoretic properties.

In contrast to the above, work on the *semilattice of Turing degrees* (either of arbitrary or r.e. subsets of ω) transferred the language, but not the spirit of algebra to logic. The interest of the work goes beyond its new (priority) arguments, probably the mathematically most original contribution of logic during the years of consolidation.

For a broad perspective, consider the following *parallel* between the programs of Pythagoras and Hilbert (number is the measure of all things): the irrationality of $\sqrt{2}$ refuted the former, the existence of a non-recursive r.e. set (i.e. incompleteness) the latter. Here it should be recalled that Hilbert's broad program:

> you lose *nothing* by formalization; in particular, completeness of Peano's or Dedekind's axioms w.r.t. derivability is not lost by passing to the corresponding first order schemata

required that every arithmetically or analytically definable set was recursive! though Hilbert himself never realized this implication. This is in fact the most natural sense in which incompleteness refutes Hilbert's (broad) program. Be that as it may, Pythagoreans and Hilbertians were faced with (so far unexpected) irrationals and non-recursive sets. What is one to do with them? Which of them should one study? What questions should one ask about them? Euclid's choice in Book X of the *Elements*, *la croix des mathématicians* (cf. Knorr [1983]), made one *choix* (roughly speaking, pythagorean dependence): precise, coherent, intelligible (as Knorr shows), and utterly sterile. Turing and Post, in the case of non-recursive sets, made another, patently modelled on the notion of rational dependence (of arbitrary irrationals), but ignoring the lessons of number theory which show that, in general, algebraic dependence on the one hand, and (above all) measures of irrationality and of transcendence can be more rewarding. Nothing so far done in the theory of Turing degrees has even a remotely similar flavor to the discovery (by Thue) how measures of irrationality of algebraic irrationals have implications for bounds on the size of solutions of diophantine equations.

Speculation. It would be hard to imagine Archimedes trying to give a 'systematic' exposition in the style of Euclid's Book X, at a time when so few specific irrationals were known. After all, a systematic style (preferably loose, as in the bulk of Euclid's *Elements*) may magnify the advantages of a guiding good idea. But in other cases it runs the risk of introducing systematic errors (especially if one relies on the rigid system as a substitute for a convincing idea, as in Euclid's Book X); errors in the choice of notions and problems, if not in (the formal validity of) the deductions. By and large, history of mathematics is at its best when reminding us of (forgotten) schemes that failed. After all, we remember the successes in outline, and the details involved are rarely repeatable, hence generally misleading.

f) An imagined experiment

What would mature mathematicians have thought about (a)-(e) at the end of the 50's (if they had known about the material)? In objective terms: How does this material fit in with mathematical experience?[5]

As to (a), there was widespread appreciation of the fact that 'negative' logical theorems, advertized for showing basic 'errors' in our ordinary conceptions, had perfectly sober interpretations too. But just how far could one go in this direction? In the 20's, the Polish school had already pursued generalizations with a logical flavor; of analysis in descriptive set theory, of number theory in ordinal and cardinal arithmetic. The German school had gone in the opposite direction of algebraization. By the 50's, it was quite clear that not all is gold that glitters. *Principal demand*: More substantial uses.

As to (b), one can always do better: but at what price (marginal utility)? In particular, the new interpretations of relative consistency proofs derive their interest from (logical) *assumptions* about the relevance of the two (logical) categories involved: derivability from (necessarily incomplete, and thus possibly unduly arbitrary) formal systems, and the division between recursive and non recursive (in particular, in connection with the ordinarily relevant sense of 'decidable'). In the light of mathematical experience accumulated by the mid fifties, questions of 'principle' about (logically idealized) derivability or decidability had become suspect, as *distracting from more subtle information* provided by (actual) proofs and solutions. The logical notions are not 'first' steps, because they go so to speak in the wrong direction. More concretely, in the particular case of the axiom of choice the conservation result distracts attention from what is *gained by use of choice*. It is not simplification by shortening of proofs (which is purely linear), but the elimination of specific choice functions which contribute nothing to the result stated. Thus the conservation result draws attention away from the *open* problem where those choice functions may be relevant (for example, in algorithmic matters; cf. Section 4 below).

Reminder. In this century mathematics has developed its own *proof theory* under the heading of axiomatic analysis in terms of basic *structures mères*, which is not merely different from, but in *conflict* with the categories used in logical proof theory (or decidability theory). Also, it finds itself on the defensive because it lacks anything comparable to the neat metatheorems of logic; cf. the (grammatically well formed) grunts like 'This is trivial' or 'This is ugly'.

As to (c), one has the same kind of reservation as with (b): not *whether*, but *which* relations are established (between logical and more central disciplines).

(d) and (e) are, formally, an internal matter, and so do not invite a view from

[5] This is not to be confused with the reputation of logic among mathematicians, a mixture of hearsay, the charisma of individual logicians, interest of the general public, and so forth. Besides, some mathematicians have broader intellectual interests than pure mathematics, and so the non-mathematical interest of logic, elaborated in Section 3 below, intervenes too.

outside. But experience with relations between other branches of mathematics is certainly salutary.

About (d): around the turn of the century two kinds of relations were established between algebra and geometry (not forgetting Cartesian coordinates!):

1. (Hilbert's) equivalences between algebraic and geometric facts modulo skew fields and projective geometry (reminder: commutativity and Pappus' theorem).

2. (Poincaré's) algebraic topology.

Though topology was then a much more recent subject than was logic in the 50's, let alone today, so far nothing like 2 turned up under (d).

About (e): parallels with mainstream mathematics were described there. As is emphasized by specialists on Turing degrees, the most obvious parallel in other mathematics is (finite) *combinatorics*. Though its practitioners seem fascinated by (or easily addicted to) the subject, and quick to present any new idea as some kind of revelation (revolution), most outsiders seem unconvinced. After all, there is no guarantee that the subject lends itself to a uniform treatment; an obvious alternative is to embed certain parts of the subject in one or more branches of (mainstream) mathematics, other parts in others (cf. number theory, where certain problems have been recognized to have algebraic, others to have analytic character). In combinatorics, the recently discovered relations between van den Waerden's theorem about partitions P of ω and both ergodic theory and topological dynamics are good examples (leading to conjectures about special classes of partitions P defined, respectively, in ergodic and topological terms). Similarly, in the case of degree theory, the 'unity' conjured up by such logical devices as the first order theory of (r.e.) degrees is unconvincing. For example, the existence of *some* incomparable degrees generalizes to all partial orderings of subsets ω in which each element has at most countably many predecessors (at least, if the sets, functions and orderings involved are taken from a stock of sets for which CH fails). The existence of incomparable r.e. degrees seems more delicate but the choice of recursiveness is dubious, and literally irrelevant, at least in the trivial sense that the result generalizes to relative recursiveness in an *arbitrary* $A \subseteq \omega$.

1.2) Busy years since Cornell

Some of the (conscious or unconscious) reservations of mature mathematicians at the end of the 50's, listed in (f), were shown to be not quite convincing by later work in logic.

For example, in the case of model theory the relevance of the theory of *p-adic fields* was shown strikingly by Ax, Kochen, and Ershov in the sixties, and by

MacIntyre and Denef more recently (in applications to conjectures of, respectively, Artin and Serre).

Both model-theoretic and proof-theoretic *independence proofs* (in appropriate formal systems $\mathcal{F}$ together with all true Π^0_1 statements) established unsuspected lower bounds $\mu_{\mathcal{F}}$ for theorems in corners of combinatorics and number theory concerned with variants of Ramsey's theorem and the representation of integers with different bases.[6] The theorems involved, though of exceptionally slight intrinsic interest, were discovered to exhibit a phenomena previously found only in metamathematics and related branches: dramatic differences between the mathematical and computational difficulty of theorems.

The model-theoretic independence proofs use of course *non-standard models* (of $\mathcal{F}$) specially tailored to the particular statements considered. They belong to that part of non-standard arithmetic or non- standard analysis which parallels the familiar strategy in mathematics of embedding a given structure (tacitly, together with a given class of problems) in a richer structure with more 'elbow room' (e.g. extending the factorial to the Γ function, and the like). In short, detailed knowledge of the particular nonstandard model is exploited to see facts concerning its standard elements. There is another part of the subject where no special properties of non-standard models are used, and thus conservation results, as in (b) above, apply. Occasionally, they have a similar function to that of the axiom of choice mentioned in (f): in avoiding details that are irrelevant to the problem considered.[7]

Warning. For the mature mathematician the fact that logical ideas have been used effectively is *more convincing* than any claims about the (logical) need for those ideas, for the simple reason that the analysis of 'need' in logical terms is dubious; in effect, it is in conflict with the categories of mathematical proof theory mentioned in (f) *à propos* of (b).[8]

[6] *Reminders*. All the variants of Ramsey's Theorem follow trivially by compactness from its infinite version. The finiteness of Goodstein's sequences follows trivially by use of ordinals $< \epsilon_0$ or $< \omega^\omega$, corresponding to the variants considered.

[7] *Warning*. This view revises the common view (for example, in Robinson and Roquette [1975]) concerning close relations between non-standard parts of models of arithmetic and function fields. The particular use of the former in Robinson and Roquette avoids details of specific function fields (which would be needed for quantitatively more informative theorems). It is beyond the scope of this chapter to go over the large and very uneven literature on nonstandard mathematics, even though it is not so large and not so uneven as the literature on fuzzy sets.

[8] Cf. Cantor's proof of the existence of transcendental numbers by cardinality arguments, 10 years after Liouville had shown that $\sum 10^{-n!}$ was transcendental. Evidently, there was no 'need' for higher cardinals, except for stating a proper generalization: the result applies not only to the algebraic numbers, but to every countable subset of the reals.

Aside. The use of uncountable iterations of the power set operation in Martin's proof of Borel determinacy would be a contribution to knowledge even if it were not logically needed; tacitly, w.r.t. the remaining axioms of set theory.

a) New reversals of old trends

In particular, the trend of concentrating on those aspects of recursion theory which are related to *decidability* (by the perfect computer), and the trend of concentrating on *cardinality* (in keeping with the logical ideal of universality: you may not know which - say, geometric - properties of the continuum are relevant, but you can always ask what its cardinal is). *Samples*:

- Higman's discovery of the significance of r.e. word problems for questions concerning the embedding of a finitely generated group in a finitely presented group.

- The shift in emphasis from categoricity in power to algebraically more relevant notions such as stability and transcendence degree (used in Morley's proof of Lös' conjecture).

Higman's work uses only elementary properties of r.e. sets well known in the forties, while Morley's use of (countably) transfinite constructions differs sharply from the earlier level of model theory.

Less well known is the shift away from interpreting so-called *normalization* (or *cut elimination*) procedures of formal derivations in terms of equality of proofs represented by those derivations (cf. Section 2 below on the operational semantics of Lorenzen and Prawitz involved in that interpretation). *Sample*:

- Lambek's discovery of the significance of normalization, incidentally for somewhat teratological propositional systems for so-called relevant implication (and conjunction), for the word problem of free closed categories and their coherence properties (needed, for example, in the study of loop spaces).

Though this significance was spotted by mathematicians, logicians provided slick and correct proofs.

Despite the considerable literature on *topoi* and *sheaf models* in relation to intuitionistic logic, the latter will be treated for its relevance to Sections 2 and 3 below. For one thing the shift from the original preoccupations (for example, of Brouwer) with continuity (in connection with choice sequences) does not seem - at least, to me - comparably imaginative to the samples above. In the same vague vein: the place of set-theoretic operations involved in the classical logical operations among operations on sets seems to be - again, at least, within my experience - much more central than that of the corresponding sheaf- theoretic operations in the subject of sheaves (reminder: function completeness of truth functional propositional logic).

b) Squeezing out the last drop from ideas and proofs in logic

Obviously, in the busy years under consideration this kind of thing has become a principal occupation, but sometimes with satisfactory results. *Samples*:

- If *relative consistency* statements are enriched quantitatively (so as to refer to the relative lengths of hypothetical proofs of inconsistencies, and not only to provability) then they imply strong reductions and inner models (Kreisel, Friedman).

- A different kind of development, but also attractive, is the elegant axiomatization of a number of apparently isolated results about *formal provability* (in familiar classical logic with *modus ponens*) such as Löb's theorem. Without exaggeration: this work is needed to establish their significance (in the sense that the whole subject can be set out in the language of modal logic, with those earlier results as only axioms). Of course, this leaves open a judicious choice among the consequences.

c) Relations of logic to more central parts of mathematics

They still constitute a relatively small part of logical activity, some of which has been mentioned in (a). Nor has logic (unlike number theory) raised problems which required (or, at least, inspired) significant developments of broad interest.

However, logicians think here of certain corners of mathematics such as universal algebra (including uncountable abelian groups), general or set-theoretic topology, descriptive set theory, and the theory of infinite games. This seems absurd to most contemporary mathematicians. But while (contrary to Popper or Lakatos) there is no evidence of basic errors in judging the *validity* of proofs, as knowledge accumulates the *center of gravity* (interest) is liable to shift. For example, Borel sets may become more central. Given this distinction, there are now massive relations between such corners and logic, especially the set-theoretic part.

Many of these relations are stated in terms of *independence results*. However, the same work admits a different interpretation; for example, as raising the question whether, in the corner of mathematics concerned, families of sets restricted to L or satisfying Martin's axiom etc. lead to a *more rewarding development* of that area (than the simpler full cumulative hierarchy). This too is a shift, if not a reversal, of interest, familiar from ordinary mathematical experience.[9]

Experience shows that, in general, no *one* family can be expected to be privileged. For example, even in such a small area as free bases for uncountable abelian groups: Whitehead groups are sensitive to the choice (Shelah), but the group of bounded sequence of integers with pointwise addition is not (Specker and Nöbelung). It remains to be seen whether a *relatively small* number of families is adequate for *relatively many* situations that concern us. This matter will be

[9] *Sample*. Hilbert's 13th problem about building up functions of n variables (for $n > 2$) from binary functions. He spoke of continuous functions, but the problem would seem more rewarding for algebraic functions (at least, in the light of his examples).

pursued in Section 3, especially w.r.t. the highly exaggerated doubts about the subjective elements in the *selection* of our concerns.

d) Relations between different parts of logic

They have kept logicians almost as busy as (b), and there can be no question of summarizing this material here. The *Remark* on p. 294 leads naturally to the next sample of some delicate (and hence neglected) aspects of such relations.

So-called *metarecursion theory* on the recursive ordinals (extended to *α-recursion theory* on admissible segments α of the ordinals) established the possibility of imitating, on suitable infinite segments, the most delicate priority arguments of ordinary recursion theory on finite segments (cf. Gödel's transfer of the lesson from algebra to infinitary set-theoretic operations in (e) above). This contribution of α-recursion theory during the sixties fits well the following change in climate between 1960 and 1970. At the beginning set-theorists hardly ever cared to 'know' (i.e. acknowledge) any ordinals which were not cardinals, and so questions concerning the *fine structure* of L appeared marginal. By the end of the 60's there was relatively general confidence in such questions. Granted this indirect contribution, the project of generalizing recursion theory exceeded the wildest expectations for it. It is then a quite separate matter to assess the reverse project of using knowledge of that fine structure to elaborate metarecursion theory, or even the theory of ordinary Turing degrees (Simpson, Shore). After all, though ladders (which are reusable) should not generally be thrown away (cf. *Tractatus*, 6.54), a child that has at least found a comfortable home had better not be put back in the slums.

e) Internal developments of logic

They were, implicitly, used throughout (a)-(d); not only *forcing* and *Boolean-valued models* in set theory (Cohen, Scott, Solovay), but also the use of *indiscernibles* and *stability* in model theory (Morley, Shelah), and various extensions of the *priority method* in recursion theory (Sacks, Lachlan).

Some internal developments made heavy use of methods already perfected in other branches of mathematics. This applies obviously to the *sheaf model* already mentioned, the construction of families of partial functions with application (satisfying the axioms of the λ-calculus and much more besides) by use of *injective limits*, the construction of families of partial and total functions of all finite types and satisfying $X^{Y\times Z} \simeq (X^Z)^Y$ by use of limit spaces and general facts about *cartesian closed categories*, and, more recently, (Girard's) studies of the category of orderings with order preserving mappings and a special class of functors, so-called *dilators* (that preserve direct limits and pullbacks, and are therefore) determined on well-orderings by their behavior on finite orderings.

1.3) The place of contemporary logic in pure mathematics

First and foremost, some of its most *elementary notions* and their properties have become an integral part of mathematics; this includes sets and isomorphism, logical language, the use of models, the notion of formal rules (operating on each other). Further, it has contributed a few very general, and hence necessarily simple, *observations*; for example, on arbitrary first order axiom systems, concerning definability and transfer properties. As a matter of actual experience such results have *continued* to be useful for solving significant problems (in combination with methods specific to the area involved); thus providing a kind of exception which proves the rule that what is true in general is liable to be trivial in each particular case.[10] Finally, though by no means a particularly new branch of mathematics, up to now the more elaborate parts of logic have not been strikingly rewarding (compared to the efforts made by selected people). In short, *contemporary logic is a useful, but relatively minor adjunct to mathematics*, so to speak, *as a tool for extending* (mathematical) *reasoning*.

This is in sharp contrast to the place of (contemporary logic among attempts to analyze mathematical reasoning; for example, the phenomena involved in proofs and rules (for deciding classes of problems). Whether right or wrong or even totally misguided, the logical answers in the form of neat metatheorems are *unique* among all such attempts (at least, in their aesthetic appeal) partly because of their independence of detailed (mathematical) experience. So much so that (for most of us) alternative attempts, in the mathematical or philosophical literature, are *hors de combat* unless the superficially plausible logical answers have been shown to be basically inadequate. As with any other theoretical program, the mathematical properties of the proposed theory are needed for such an investigation. (Newtonian mechanics has to be developed before one knows its verdict on the perihelion of Mercury!)

We turn now to this use of contemporary logic.

14.2 Trade with Logical Foundations

Section 1 above concerned the disengagement of (mathematical) logical disciplines from their use for logical foundations. The present section is complementary to it, and is organized as follows.

We first take the different branches of foundations as objects of study, without deciding in advance on whether they are variants or rivals, or on the degree to which any of them contribute to the broader concerns of the philosophy of mathematics which they were intended to advance. 2.1 reports on the (mathematical)

[10] In this respect contemporary logic is comparable to its (younger) variant and, occasionally, rival for attention: *category theory* or, more precisely, the different branches of the latter besides 'pure' category theory.

notions and results of contemporary logic related to (or actually derived from) foundational ideas, and 2.2 reports on the foundational interpretation of mathematical developments in logic, usually for correcting first impressions concerning those foundational schemes.[11]

We then use this material to go back, in 2.3, to the questions of: variants versus rivals, or of the relevance of the foundational schemes to their intended aims.

2.1) From foundations to (mathematical) logic

Again it is convenient to separate the years of consolidation from the busy years. By and large it was not the fashion during the years of consolidation to set out relations of mathematical notions to foundational ideas, in keeping with the mathematical style of the time (but ignoring the fact that most of contemporary mathematics deals with well tested notions that lend themselves to a mathematically pure treatment). But some exceptions will be mentioned, too.

a) Before Cornell

The set-theoretic branch of foundations, especially the ideas of Poincaré and Russell on a *ramified hierarchy*, is obviously related to all branches of logic. *Samples*:

- Gödel's *constructible hierarchy* (with the remarkable discovery that the notorious axiom of reducibility is actually valid for L) is a so to speak crude formulation of the ramified hierarchy; 'crude' inasmuch as it ignored the question (clearly of primary concern to the pioneers[12]) how far the hierarchy is to be iterated.

- Kleene's *hyperarithmetic hierarchy* is most naturally regarded as the recursive segment of L (for the hierarchy restricted to sets of natural numbers). The concern mentioned, at least when interpreted loosely (as the pioneers - God knows! - would have done), is respected here because of the extraordinary stability of the order types of well orderings definable in the recursive segment (Spector).[13]

[11]Since foundational schemes claimed to be conceptions of (the nature of) mathematics, this is perfectly parallel to progress on any idea (i.e. conception or theory) we may have of the physical world, where mathematical theories are developed to formulate that conception and mathematical results are then interpreted, for example, to reject the conception. There would have been no problem about the perihelion of Mercury if Newtonian celestial mechanics had not been developed to a very considerable degree!

[12]But certainly no more than a modern atomic theory ignores the primary concern of atomists that atoms be indivisible! It is not lack of moral fibre that led physicists to drop this concern.

[13]The concern becomes primary in so-called *autonomous progressions* (for predicative proof theory), enough to examine the *relevance* of that primary concern. To spell it out: of course,

- Gödel's Princeton lecture [1946] (*to* mathematicians but without mathematical formulae) derives the notion of *ordinal definability* from a variant. The lecture is also the first place which draws attention to the fact that, contrary to first impressions, not the power set operation, but the number of its iterations (axioms of infinity) is problematic for full set theory.

Without much exaggeration, the 'primary concern' above is *more relevant* for the development of ordinary set theory.

The *finitist* branch of foundations (which had been so exceptionally fruitful in the early part of the thirties, in connection with hereditarily finite *operations*) also spawned an autonomous progression for *finitist proofs*. The remarks just above about relevance apply here too. A perhaps more sensible mathematical notion, applicable to weak quantifier free systems, comes from the finitist branch, namely that of *direct proof* in number theory (in the sense of Gauss: no new concepts should be introduced in the course of a proof). Tarski's *cylindric algebras* express another, curiously literal, finitist view: the stock of symbols of a theory, including its variables, should be finite!

The intuitionistic branch, obviously because of its alleged obscurity, gave rise to several mathematical notions, so to speak depending on the particular part of the intuitionistic literature which happened to have caught the attention of a particular author. *Samples*:

- From the literature of choice sequences, Tarski's *topological interpretation*.
- From (Heyting's) explanations in terms of proofs, via Gödel's modal exposition, Tarski's *calculus of systems*.
- From the general preoccupation with constructivity or effectiveness in the intuitionistic literature, combined with Church's thesis, Kleene's *recursive realizability*.[14]
- The new half-forgotten part of recursion theory about the distinction between *simple and hypersimple sets* (in other words, different presentations of *finite* sets) also comes from the intuitionistic literature.
- The by no means forgotten use of *higher types*, as in Gödel's Dialectica interpretation, is imposed by the property of intuitionistic *implication* which is not reducible. In contrast, in classical arithmetic, *negations of prenex*

that concern is part and parcel of a coherently predicativist scheme; but relevance involves an additional examination *whether* (or which variant of) coherence is appropriate.

[14]The latter gives *no* interpretation to $(\neg, \rightarrow, \wedge, \forall)$ consistent with Kleene's own report that he did not understand the intuitionistic literature, but relied on a remark of Hilbert about *existential* operators (including $\vee$) as providing partial information. This applies to the intuitionistic existential operator, not (contrary to Hilbert's impression) to the classical operator which need not provide *any* realization.

formulas are a constructive reduction class, and so lead to functionals of lowest type.

b) After Cornell

Gödel's idea of paying attention to *large cardinals*, even without any specific suggestion (except about what one should *not* do), was taken up in set theory itself; but without any guidance at all except a weak 'analogy'. Specifically, a 'large' cardinal should stand in the same relation to ω as ω stands to finite cardinals (≥ 1), say to 2 (the only finite strongly inaccessible cardinal). The same idea also influenced model theory and recursion theory, by putting a premium on questions which are sensitive to the stock of sets considered (for example, cones of Turing degrees).

Ideas about *predicative* (in the sense of invariant) *definability* dominated for a while the part of model theory which uses infinitely long formulae (and so-called generalized recursion theory needed to provide the syntax for such languages).[15] Starting with *inductive* (semi-invariant) *definability*, a parallel exposition was given, usually with the (somewhat modest) purpose of establishing *equivalence* (!); in other words, different descriptions of the same objects, but without any indication (so far) of areas where one or the other description is superior.

The intuitionistic branch has provided very little, except such additions as *Kripke's schema*. As will be seen in 2.2, here most of the trade went in the opposite direction (of mathematical results improving foundational ideas). Given the wealth of notions derived in the years of consolidation *from* intuitionistic ideas (sampled above), it seems proper that after Cornell people were kept busy by these notions.

Finitist foundational ideas (in any sense that is even remotely reminiscent of Hilbert's) have not given rise to mathematical ideas in logic. But, for example, Girard thinks of the functors (in the category of orderings and order-preserving maps) which he introduced and called *dilators* as being in the finitist tradition: the requirement that they preserve direct limits (and pull backs) ensures that their values on well-orderings are completely determined by their behavior on finite orderings, since every well-ordering is a direct limit of finite orders. Of course, those dilators belong no more to a coherently finitist tradition than the whole of L belongs to a coherently predicative mathematics. But, again, they are relevant to the question whether or not a coherently finitist presentation serves the broad concerns which it aims to advance.[16]

[15]Similar ideas, though less explicit, are involved in the tradition of descriptive set theory; for example, Borel sets. Without exaggeration: the only idea that had been overlooked (though it can easily be made plausible to anyone steeped in descriptive set theory of the 20's), is the parallel between *finite* and Borel (in place of *recursive* and Borel).

[16]Interpreted in this way, the finitist tradition has been quite prominent in many parts of mathematics; thus it would be a *mis*representation of the matter to discuss it in the narrow context of mathematical logic. As a corollary: it must be expected that a part of mathematical

2.2) From mathematical logic to (logical) foundations

Some return from mathematical logic to improving foundational ideas is skipped here. For example, the *analysis of completeness proofs for topological interpretations* by isolating axioms for choice sequences needed for the proof: these axioms could then be seen to apply to a particular notion of choice sequence that had been mentioned in the literature (by Brouwer), without having been correctly judged. With these defects, the material will again be split into parts before and after Cornell.

a) Before Cornell

We first have to *correct first impressions* about (the defects of) foundational ideas. Without exaggeration: the single greatest obstacle to progress in foundations (which includes of course seeing the genuine defects of schemes) is the - persistent - belief in *imaginary* defects, and a smug feeling of triumph at having overcome the latter. One such error, incidentally reappearing in almost all the schemes, is this: mathematics just couldn't be done that way! For example, by: restricting oneself to intuitionistic logic, paying attention to the form of set- theoretic axioms used, using the language of types in place of sets (let alone, ramified types), and so forth. As a (neglected) corollary, one thinks that existing mathematical knowledge is already sufficient to reject the schemes involved. This is not the case, and this draws attention away from the problem of discovering *genuine* defects.[17] During the years of consolidation a very simple device was stressed and perfected for this purpose; namely, to set up *subsystems* (in a familiar language) which have straightforward models or interpretations in the (allegedly) formidable schemes. *Samples*:

- Gödel's early observation on the conservative character of the *fragment* $\neg$, $\wedge$, $\forall$ of classical over intuitionistic arithmetic made the point clearly enough, especially after Π^0_2 sentences were included. The important point which was stressed by the pioneers, but (consciously or unconsciously) obscured by peddlers of 'elegant' intuitionistic expositions, was this: not only are these conservation results derivable in (say) primitive recursive arithmetic, but the transformations are patently clear.[18]

logic that pursues primarily this tradition has been 'preempted' by developments in other branches.

[17]In the natural sciences we have to set up special experiments to decide between theories, because as nature as it presents itself generally does not.

[18]Gödel's own paper on the subject may have contributed to the confusion because he conjectured that such results would not extend to the theory of species or, generally, 'stronger' systems. Gentzen's observation, noted independently by several people in the forties, was overlooked when he pointed out that ordinary analysis did not need much more than weak subsystems of formal arithmetic. Goodstein's expositions were admittedly clumsy. And so on.

- Gödel's work on the *constructible universe* L is another perfect illustration; only the stilted formulation in terms of relative consistency obscured this point.
- The axiom or the rule of Σ_1^1-*choice* in connection with predicative mathematics is an example of the particular prominence of this kind of logical hygiene towards the end of the fifties.

Foundational bearing of this logical hygiene: even if there are good reasons for rejecting the foundational schemes involved, it cannot be *philosophically* satisfactory to do so for a bad reason.[19] General but neglected conclusions to be drawn from such work will be listed in 2.3.

We turn now to samples of results in logic with genuine foundational bearing. The first example shows how notions regarded as fundamental (here: validity, the sacred cow of formalistic semantics) are reassessed by inspecting proofs of results about them. *Gödel's completeness proof* did not appeal to any formal definition of validity, but used (implicitly) the following properties of validity:

- valid formulas F (with the relations R occurring in them interpreted as certain relations R_F over the integers) are true
- formulas obtained from valid ones by the rules of inference (considered) are valid.

This is definitely more adequate then picking on a particular definition of validity.[20] Beginning in the fifties, completeness proofs for intuitionistic propositional logic (later extended to minimal logic) were set out on the same pattern, except that R_F involved an additional parameter α, and:

- valid formulas F (with R interpreted as R_F) are true for every α

is used instead.

The second example shows how foundational refinements can be extracted from mathematical results. The *completeness proof of Tarski and Eilenberg* for Heyting's propositional logic w.r.t. the topological interpretation provides an R_F in the style just described above, as long as the additional parameter α satisfies certain simple formal laws. Inspection of these laws led to the recognition that

[19]Cf. T. S. Eliot's definition of (intellectual) treason: doing the right thing for the wrong reason. (It is not a lawyer's definition.)

[20]*Sample*. If validity is defined in the (finitely axiomatized) system GB of classes (as truth in all set theoretic structures) then one cannot prove in general that a valid formula F with a single binary relation R is true in GB (with R interpreted as $\in$); or, equivalently, that a formula false in GB is (not valid, i.e.) false in some set theoretic structure. For example, if F is the negation of (the conjunction of) the axiom(s) of GB then the above statement reduces to the assertion of consistency of GB, which cannot be proved in GB itself by Gödel's second incompleteness theorem.

certain simple choice sequences, nowadays called 'lawless' satisfy them.[21] For reference below: those objects were later seen to satisfy also additional laws which are not needed here, for example extensionality.

The last example is Lorenzen's attempt at introducing an *operational semantics for elementary logic* based on Gentzen's discovery of cut elimination. Gentzen made a side remark about 'introduction' rules for a logical operator defining its meaning by its use, a phrase made popular by the Vienna circle without an even remotely plausible sense. Lorenzen made it at least plausible that some coherent (though not necessarily relevant) sense could be given to the phrase; cf. (b) below on writings of Prawitz in a similar vein.

b) After Cornell

A good deal of the work involved dotting the i's and crossing the t's of results established (convincingly) already before Cornell. This applies particularly to the material at the beginning of (a) above (weak subsystems, conservative extension of classical over intuitionistic logic, conservative extension of theories in higher type languages, and so forth). After absorbing this material readers should have no difficulty in recognizing the pattern, and so no explicit reference will be made to these later refinements. We can thus concentrate on other directions.

So to speak in principle, Gödel's point on the possible bearing of *axioms of infinity* was established up to the hilt: the simplest proof of the formal consistency of ZF is by appeal to its smallest model in the cumulative hierarchy. The issue was to get an idea of the *range of relevance* of his point. *Samples*:

- Scott's refutation of $V = L$ from the existence of a *measurable* (tacitly, uncountable) *cardinal* is certainly memorable, despite the ad hoc character of the hypothesis.

- Martin's very natural use of the segment of the cumulative hierarchy up to ω_1 in the proof of Borel Determinacy is perhaps even more convincing, despite the fact that no *new* large cardinal axiom was involved: Martin used *previously unemployed cardinals*.

- Solovay's very natural use of *inaccessible cardinals* to define a family of sets in which all sets of reals have a measure (tacitly, in the family) is striking enough.[22] And Shelah's verification of that need (tacitly, within

[21] *Parallel*. The integral $\int_b^a f(x)dx$ of f was originally thought of, purely mathematically, as the inverse of differentiation. Riemann gave another description in terms of upper and lower sums, provided those sums for f (with $f(x) \geq 0$ in $a \leq x \leq b$) converge. The only properties used are: additivity w.r.t. intervals, monotonicity w.r.t. f, and the convention that $\int_0^1 1dx = 1$. These properties can be recognized as being satisfied by the measure of the *area* under the curve $a \leq x \leq b$ and $0 \leq y \leq f(x)$.

[22] Cf. Cantor's use in note 8 of uncountable cardinals to solve trivial questions in transcendence theory trivially.

> the ordinary set theoretic tradition, in particular, without determinacy!) is reassuring: one has not overlooked some really simple alternative within the tradition.

All this has raised Gödel's (foundationally witty) *aperçu* to a genuinely more substantial level.

In much the same vein, the quite obvious possibility of having families of *partial or total functions that apply to themselves*[23] (or, so to speak less solipsistically, to each other) was realized in a quite substantial manner by Scott (models of the λ-calculus) and Friedman, thus raising at least potentially the level of foundational debates about some 'universal need' for type distinctions to a perfectly acceptable threshold. It is a quite separate matter whether these foundationally highly relevant results are at all rewarding in mathematics or Computer Science (where, for example, the requirement of extensionality is usually quite extraneous). The results are also well suited to reassess, in the sense of (a) above, the basic logical *assumption* that the existence of *some* model is a fundamental issue, in the sense that, once you have it, the 'rest will look after itself'; cf. the case of term models. All this provides a good example how successes of (contemporary) mathematical logic can be used to *refute* the logical ideals which inspired the problems that were solved successfully.

On the topic of partial and total functions (so to speak at the opposite extreme) another logical assumption, namely, that one or the other of these objects must be 'fundamental', is neatly put in its place. For an algebraic theory with a universal enumeration, you need partial functions. For computational efficiency you must not insist on partial functions.[24]

In proof theory perhaps the single foundationally most striking observations were these:

- a semantic interpretation tailored to *cut-free rules*, including the use of not necessarily well-founded proof figures
- the discovery of mathematically *trivial changes in proofs* which involve computationally dramatically different results (when the proofs are unwound to yield algorithms).[25]

---

[23]Obvious examples of such partial and total functions are obtained from the theory of partial recursive functions (with a little care to ensure extensionality), and from any semigroup (if the element a is interpreted as the function $x \mapsto ax$).

[24]*Reminder*. To define multiplication for partial functions $m \cdot 0 = 0$ must be replaced by something like $0 \cdot 0 = 0$ and $(m+1) \cdot 0 = m \cdot 0$ since, for undefined m, $m \cdot 0$ is undefined but 0 is not; cf. Statman.

[25]There are many open problems here, e.g. in connection with the finite variants of Ramsey's theorem by Paris and Harrington, and by Ramsey himself. Both versions can be proved by compactness (from the infinite version), and a mathematically simple modification of the second proof leads essentially to the standard proof of the finite version (by recursion).

In intuitionistic logic there is clearly the *potential* for a genuine foundational advance, but it has not been realized. Specifically:

- There is a great deal of work (combining old and new methods) which establishes such formal properties as *conservation* and *non-conservation* of masses of systems (incidentally, copied somewhat haphazardly from the classical literature). It would be a foundational advance if descriptions of the systems in specifically foundational terms allowed one to read off those results.[26] In fact, logicians are just kept busy calculating.

- There has indeed been some clarification; for example, on the place of the usual *intuitionistic connectives* $\neg$, $\wedge$, $\vee$, $\rightarrow$ among all intuitionistic propositional operator (definable by infinitary operators on second order logic), or on one of the intended interpretations of those operators in terms of operations and proofs (of logic free propositions). The conclusions are parallel to those concerning predicativity or finitism: all this (simple) work shows how to implement the primary concerns of the pioneers, but it leaves wide open the question to which extent the pioneers knew what was good for them. It is simply a *superstition* to think that *everything that is coherent is good for one*.

Probably unconsciously, most people in the subject continue to think of intuitionistic logic as *deep and obscure*, with many hidden treasures if only its concepts were 'clarified'. This overlooks the fact that intuitionistic language is very close to one of the dialects of (natural) mathematical language in current use (and even more widely used before Cantor). And the reaction overlooks the possibility that this particular dialect is just not very effective for its most highly advertized aim under the slogan of 'constructivity'.

Finally, work has been done on the general (foundational) ideal of giving a meaning (technically: a semantics or, more precisely, a semantics in a particular tradition) to so to speak *formal tricks* (more precisely, what appear to be tricks for the particular tradition[27]). *Samples*:

- Henkin's straightforward semantics for the usual (many sorted) first-order version of *higher-order logic*, leading to the terminology of 'nonstandard' models. The later development by Robinson highlights the two separate issues of internal development (for example, considering not only formal

[26]Cf. *imagined experiments* of studying formal systems of set theory without knowing their interpretations in terms of the cumulative hierarchy. For such a study it could be as difficult as Gentzen's consistency proof for arithmetic to see that replacement cannot be derived from the other axioms (while knowing these interpretations it is sufficient to note that replacement does not hold at level $\omega + \omega$).

[27]For example, though the exclusion of *tertium non datur* is evident for the intuitionistic meaning, it seems a trick for standard semantics; specifically, a trick to ensure recursive realizability.

systems, but all true formulae) and discovering areas where the new notions are actually relevant. The difficulty of the latter is illustrated by the (so far) unrewarding areas of invariant subspaces of operations and diophantine problems.

- Boolean valued semantics for *forcing constructions* introduced by Scott, Solovay and Vopenka.[28] This was later used for a notion of nonstandard model different from the ones above, by Scott and Takeuti. The latter discovered the relevance to von Neumann and AW^* algebras.

- The formal discovery of *cut-free proofs* has been given both a model-theoretic interpretation (including the case of proof trees that are not well-founded) and an operational interpretation (by Lorenzen and Prawitz). Prawitz has also given an interpretation to (uniqueness of different sequences of) cut elimination steps.

- *Kripke's models* are, in fact, trees or partial orderings of familiar kinds of models.

Readers should be warned [∞] of the interplay (e.g. in the second sample above) from mathematics to foundations and back to mathematics in general terms, and of the fact that the important thing is not *that* a semantics is used, but *which* (equivalence classes of formulae).

2.3) A notion of logical foundations in the light of contemporary logic

Provided 'logic' is interpreted broadly, but quite naturally, there is a good deal of *historical continuity* (both positively and negatively), with the paradoxes barely constituting a singularity.[29] On the positive side: Leibniz is quoted in connection

[28] This semantics can be compared to an earlier intuitionistic semantics (Kreisel [1965]) in terms of lawless sequences (which, however, has been less rewarding). Incidentally, it is not *necessarily* an accident that cleverer people have worked a Boolean valued models than on lawless sequences; just as it is not necessarily an accident that, during the last 3 centuries, cleverer people have specialized in astronomy then astrology.

[29] Contrary to many contemporary accounts, the paradoxes which dominate so much of the logical literature (and of introductions to even respectable texts in logic) are a mere ripple in that broad logical tradition. On the contrary, the paradoxes constituted a so to speak *irresistible temptation* to present any (bright) idea in the logical tradition as being relevant to the paradoxes; cf. the temptation to present any discovery on cell growth, however striking on broad biological reasons, as (possibly) contributing to a cure for cancer or aging or whatever else is regarded as a 'problem'. Questions about the broad nature of mathematics existed before the paradoxes, but the latter produced the impression that *validity* of (familiar) principles was a central and rewarding issue. Anyway, the paradoxes were a *godsend*, because they provided the *pretex* for sounding logic itself for validity! What they showed was that one might need logic for discussing *logical mistakes*!

with both recursion theory (recursive decision methods) and nonstandard analysis; Frege in connection with elementary logic, still a central part of the subject; Hilbert's *Foundations of Geometry* were a hit before the end of the last century; finally, Russell wrote his *Principles* (on: What is mathematics) before he discovered his paradox. On the negative side: mathematicians had (generalized) doubts about logic and sets before that paradox (for example, did not wish to touch Cantor's higher cardinals), and Kronecker had specific doubts (which he could have expressed more effectively if he had known more logic).

The obvious hallmark of logic (already in Aristotle's business of Being) is *generality*, both in the case of propositions (Being) and of processes (methods or knowledge). The idea is that one proceeds *from the general to the particular*: one could not recognize a particular proof if one did not have the (or, at least, a) general concept of proof to start with. There is the additional idea that we need (or even, perhaps, that there are) only very few such general ideas, and that we have become aware of them by the end of the 19th century (at the latest, since to some the *Wisdom of the Ancients* is sufficient).

From this point of view, Frege's question: What is the number 1? is perfectly natural. On the one hand, natural numbers have a very general domain of application. On the other hand, if they too (and Kant's space and time) were needed (as additional general categories over and above the logical ones), where would it all end? What would one be talking about when using such expressions as '(the theory of) knowledge' or 'on what there is'?

Airy-fairy as all this sound, at least superficially, it is not so different from (the ideas behind) a unified fundamental theory in physics. Less superficially, one is impressed by the difference that the unified theory in physics talks about 'deep' phenomena in the literal sense of being far removed from everyday experience (which certainly does not have a uniform look). The attempt - going back to Aristotle - of presenting logical concepts as deep in the same sense (namely, by regarding abstract notions as far removed from sensory experience) conflicts with the observation - incidentally, also due to Aristotle - that sense experience already involves abstractions; in current jargon, the information processing in the sense organs corresponds to (processes which we call) abstractions.

From this same point of view, a 'reduction' of mathematical knowledge to logical generalities would explain its superficially most striking property (alias 'problem') of being certain, and not refutable by experience; for if the knowledge applies to all Being, it should not be sensitive to particular experience.[30]

[30] Admittedly, this point of view is strange if one thinks of concrete cases, of the precautions needed to fit the mathematical notions employed to the phenomena considered. But for the moment we stick to the 'spirit' of the view.

a) Internal Refinements

Granted the desirability (or even possibility) of such a reduction to very general concepts, the question arises: Can't we do better? Specifically, when applied to a particular branch of mathematics the 'reduction' should (so to speak automatically) specialize further: to concepts that fit the particular branch. This point of view of *purity of method* was stressed by Hilbert in the *Foundations of Geometry*. The whole program of *finitist reduction* is a special case of this scheme: finitist proofs for finitistically meaningful (and valid) statements.[31]

In the tradition of Aristotle (*Metaphysica*, Γ, 4, 1012a), the logical notions considered by Frege were well defined. The question arises whether, contrary to Aristotle's and Frege's impressions (alias convictions), logical operations could be coherently extended to propositions about *incompletely defined terms*. The bulk of Brouwer's publications concerns one positive answer to this question, in the case of choice sequences.[32]

Thus the scheme of Hilbert presents itself as a *refinement*, and that of Brouwer as an *extension*, of the scheme of Frege and Russell or Cantor and Zermelo.

The ramified hierarchy of Poincaré and Russell is also a refinement in the *style* of Hilbert, but with the line being drawn differently (instead of finitist: *reducible to the ordinals*). Thus we have completeness not only for Σ_1^0 sentences, but for Σ_3^1 sentences or, more prettily, Σ_1 over ordinals (in place of Σ_1 over finite ordinals). Closer to Poincaré: for Δ_2^1 sentences, the meaning is not changed by introducing new elements into the universe (cf. Gödel's 'absolute' or 'invariant' definability).

b) Comparison with the popular views

Hilbert's and Brouwer's own interpretation of their schemes presented them as *rivals* to the set-theoretic scheme.

Hilbert's scheme is a pointless rival, because it simply does not apply to perfectly sensible parts of mathematics. If regarded as a refinement, it leads to the *question*:

> What more can we do with a branch of mathematics that *does* fit the scheme?

[31] To get off the ground the program requires formalization of proofs, since otherwise the program could not even be stated in finitist terms.

[32] *Remark* concerning (what Bourbaki calls) the 'profound intelligibility' of mathematics (tacitly, in their sense, but here even extended to that of choice sequences). The logical operations were defined by Brouwer in terms of (idealized) proofs, so to speak 'idealized' from the familiar psychological phenomena of convincing mental constructions, the kind of thing that Frege explicitly excluded from his science.

Another interpretation of the resulting logical laws, in terms of some kind of constructivity, is formally possible but less convincing: too many constructions remain hidden in those logical operations. The specialization to *recursive* realizations is tied to the fact that only *formal* (r.e.) systems are considered.

Philosophically even more rewarding is the question:

> Are such branches distinguished by a special kind of certainty, as Hilbert claimed for this scheme?

If not, we have the corollary: the *hypothesis* that the scheme is relevant to questions of certainty is thereby refuted.

Brouwer's scheme, as presented by him, is not a rival at all because it talks about different objects (with a different interpretation of the logical operations). To make it worthy of attention Brouwer had to argue against the intrinsic incoherence of the so-called classical scheme, and this makes his own scheme superior *by default*. Be that as it may, this was not successful. If viewed as an extension, the question arises: how does Brouwer's (also coherent!) scheme compare with a *paraphrase*[33] in set-theoretic terms? More specifically: for which problems may it be relevant? Here the scheme must be regarded as a *pars pro toto*: not confined to (subjectively, slanted) choice sequences, but adapted to random sequences or propositions (as in the quantum theory) about terms which are prevented from being completely definite by the *uncertainty relation* of Heisenberg. Put differently: the choice between schemes is not made on brutal grounds of *validity*. The choice presents therefore a nontrivial particular instance of vague talk in the philosophy of science about 'simplicity'.

As far as *actual* knowledge is concerned, Poincaré's and Russell's ramified variants of Cantor and Zermelo, as developed by Gödel, seem certainly viable in the corner of mathematics consisting of descriptive set theory, and the like (cf. Jensen's many results in L). Put simply: our knowledge of V is more rewarding when applied to L (and not to V itself).[34]

c) Looking back

The *notion* of logical foundation (or, if preferred, of logical aspects) of mathematics adumbrated in 2.3 simply raised the question:

For which problems are these logical aspects relevant?

It does not *claim* that they are relevant for all of mathematics; let alone, for the particular issues of certainty or validity prominent in fundamental discussions (at least, after the paradoxes).

Admittedly, *if* logical aspects were significant for those 'grand' questions, it would be an error (from every point of view!) to ignore this, and play about

[33] The matter of paraphrase is standard. For example, continuous operations were *intended* to deal with imprecise data.

[34] References to actual knowledge are usually dubbed 'pragmatic'; this means, ordinarily, 'shortsighted'. It may be that our actual knowledge of V and L is simply too restricted to serve for any sound evaluation of what is rewarding, of which (of course, valid) results are of permanent interest. But on the other hand it is not obvious that this is the case, either.

with modest refinements in group theory or number theory. Conversely, if logical aspects are not relevant to those matters then the (logical) problems and conjectures derived from misguided conceptions about such matters are liable[35] to be sterile; not only false, but (when false) *trivially* refutable. *Sample*:

- Hilbert's 'grand' conjecture: it is consistent to assume that every problem P can be solved, i.e. either P or $\neg P$ can be proved. Actually, it is consistent to assume that everything can be proved!

Interested readers should check that many of the results in Section 1 and 2 are naturally interpreted as exhibiting the relevance of logical aspects to the areas of mathematics involved. Obviously, able people are not fettered by any particular scheme or notion; so there are bright ideas in the logical literature which do *not* fit this interpretation.

14.3 Logical View and Mathematical Practice

On the principle:

What do they know of logic who only logic know?

this section will briefly describe schemes (used, but *not* mentioned!) in the mathematical literature which are almost at the extreme opposite to the requirements that constitute the *logical view*.

We have presented the logical view in 2.3 as a *disturbing* element for mathematical practice.[36] Perhaps paradoxically, more disturbing than the (exaggerated) claims and demands of traditional logical foundations, which after all are primarily concerned with *analyzing* and *justifying* ordinary mathematical experience and methods, and to *supplement* them with answers to questions such as: What is the number 1?, but *without interfering* with day-to-day operations (except possibly when presenting the elementary 'foundational' first steps). Hilbert,

[35] *Liable* and not *inevitable*, just as in the case of gifted theologicians like St. Thomas who made significant observations about (the nature of) thought when holding forth on the Divine Mind.

[36] I downplayed the danger of the 'invasion' of logic into mathematics, on the ground that mathematics ignored logic anyway [∞]. It is true that they ignored the details, but not the general scheme (given the name 'logical view' above). There is no contradiction here; that is, it is possible to pay attention to the general view without knowing about the details, just because (the latter fit quite well into the general view, and) *this general view is simply part of Western intellectual tradition* (cf. Plato's *Republic*: when people lose their reason, they concentrate on particulars). The mathematicians have come to rely on extracting (or abstracting) general ideas from particulars. In other words, Plato overlooked the possibility of concentrating on particulars up to a certain point (of diminishing returns), and then doing something else; just as before the period of digging into particulars, general considerations will have drawn attention to (so to speak) the particular particulars as rewarding objects of study.

for example, was quite explicit: once the possibility of eliminating infinitist procedures was established, one would go on as before, perhaps more so.[37] All this may be compared to the view in the 19th century of fundamental physics which would leave then-current physics intact but give theoretical reasons for the value of, say, certain empirical constants. (Segregation between macroscopic and microscopic physics.)

If, however, is regarded merely as a notion, as a particular style of mathematics (with a preference for objects and problems about them selected in terms of logical categories) then the logical view acquires a down-to-earth function: a kind of *strategy for selection* of notions and problems. *Samples*:

- *recursive decidability* for all elementary formulae about a given structure
- classification of theorems according to *logical equivalence* (modulo some particular 'basic' system).

Mildly vague as the logical view may be (at least, in the exposition of 2.3), it is immeasurably more precise than the (practically comparable) strategies used in mathematics itself. They are formulated as *helpful hints*, usually under the guise of pedagogic recommendations on 'exposition' (without any reference to empirical evidence, though pedagogy is obviously an empirical affair).

This is not all. The logical view *and* (though this is rarely emphasized) those helpful hints have a direct bearing on broad concerns about *improving*[38] *natural (mathematical) language*, the most visible and outward sign of our understanding of mathematics. But here there is a difference: the logical view presents itself as a candidate in this area, which the 'helpful hints' do not.[39]

The purpose of this section is to give three samples of broad concerns, which we name in everyday language: we compare the logical ideal for an answer with the forms of answer that have developed in the course of mathematical experience.

3.1) Vagueness: virtues and defects

Though, by convention, written formal expositions tend to be relatively free of (what are, in the ordinary sense of the word) vague considerations, the bulk of natural mathematical discourse is full of vague expressions. They have an obvious virtue (to use Russell's phrase from the introduction to the *History of Western Philosophy*) in 'the face of uncertainty': *flexibility*. They also have an obvious

[37] The, literally, radical reform implied in this proposal of ignoring (the meaning of) anything other than Π^0_1 sentences was hardly noticed by anybody, including himself; much less than Brouwer's 'reform'.

[38] To repeat what cannot be repeated too often: improvement is not confined to correcting previous *errors*; improvements can consist in extending our interest in *new* (and rewarding) *directions*.

[39] Not even mathematicians recognize the bearing of those hints on questions formulated in the pretentious language of philosophy or Artificial Intelligence.

defect (to which purveyers of 'theories' about inexact notions or fuzzy predicts seem to be blind): *detailed analysis* is liable to be unrewarding, as being out of all proportion to the accuracy of the 'data'.[40]

a) Remedies

The logical view has (implicitly, but obviously) two remedies. First, *generality*: if what we have to say is true for everything, it will be true also for those things that we happen to be uncertain about.

Secondly, a (so to speak) second line of defense: by reflection, we shall find a *fundamental* (or at worst a few fundamental) notion(s) which are behind the vague notions in common use; these fundamental notions are assumed to be among the familiar abstract notions already part of the *Wisdom of the Ancients* (and neglected - not by Newton but - by science since Newton).

Historical Remarks:

- Frege's idea that (proper) notions must be meaningful for every conceivable object, fits well with the particular virtue of generality mentioned above.

- Cantor's idea that this must fail, so to speak, because of human frailty (the Absolute Infinite being the domain of the Almighty) goes well with his proofs of modesty ($n + \omega = \omega$, while $\omega + n > \omega$ if n stands modestly behind ω).

- Russell's idea of *having to look* for a domain of significance is a vestige of his empiricism, that does not go well with giving a central place to logic.

b) Reservations

First, about the *assumption* that generality is a panacea. What about the possibility that what is true in general is liable to be trivial in each particular case? Of course, this does not mean anarchy or, say, the naturalist's preoccupation with particulars. But it *does* mean that one has to look for an *appropriate level of generality*, comparable to Russell's idea above.

Secondly, even if something approximating the idea of a fundamental notion is behind vague notions, why should it be *exactly* one of the familiar notions prominent in the Wisdom of the Ancients?

So to speak taking over the basic assumption of the logical 'oppressor' (with its preoccupation with precision and validity), the reservations often take the form of

doubts about the *possibility* of a precise analysis

[40] One other feature is that, if knowledge of the world around us happens to grow at a particular period, the common meaning of 'vague' expressions tend to change so to speak before our eyes; not only in mathematics, but in daily life.

of those familiar notions, shifting the emphasis to those relating innocent doubts from the much more basic

doubts about the *adequacy* of those notions

(tacitly, for their intended purpose).

c) Helpful hints: combining flexibility and precision

Here the idea is that, as far as those familiar notions are concerned, in each particular case

only relatively few properties of the notions are relevant

and, among the particular cases of interest,[41]

relatively few sets of such properties cover the bulk of the cases.

Evidently, in any particular case, judgement[42] is required to see which of those (few) sets is relevant. *Sample*:

- From the present point of view, the study of *models of incomplete axiom systems* (for arithmetic, set theory, etc.) is not at all interpreted as having to do with lack of precision in the notions which are the so-called intended models (though, of course, occasionally there may have been serious lack of precision[43]). On the contrary, a small arsenal of such models may be much more effective for mathematical progress than preoccupation with the familiar model.[44]

3.2) 'Paraphrase' versus 'analysis' of familiar notions

Contrary to the parrot-like repetition of the would-be truism that, for example, there can be no 'proof' of the identification between effectively computable and recursive functions, this kind of question belongs to a *successful* tradition going back to antiquity. *Samples*:

[41]The matter of interest is rejected as 'subjective'. Be that as it may, it is perfectly comparable to our selection for attention of phenomena in the external world *among* all those things in heaven and on earth that Hamlet proposed to poor Horatio for study.

[42]Again by comparison with natural science, the choice of theory which is relevant to a particular phenomena (that is, the discovery of the dominant factors determining the particular phenomena, and thereby the relevant theory) is also left to judgement based on experience, and not mechanized in a theory.

[43]For example, the long preoccupation with alleged imprecision of impredicativity, and total disregard for the *length* of the cumulative hierarchy; or Wittgenstein's worry about the meaning of successor (e.g. whether $100 + 1 = 102$) compared with the hackenayed worry about $\omega!$.

[44]*Reminder*. The importance of non-Euclidean geometry does not lie in rejecting any 'privileged' place for Euclidean geometry; after all, most such models *continue* to be defined in Euclidean space (possibly of higher dimension).

- in geometry, definitions of *length* or surface *area*
- in rational(!) mechanics, of *perfect rigid body* or *perfect liquid*
- not to mention the general scheme of analytic geometry, where *geometric notions* are expressed in logical terms.

a) Refinement

The logical view puts a premium of the following kind of (logical) refinement. What could previously only be seen, can now be said; specifically, in terms of *representation theorems* derived from explicit answers. *Samples*:

- Deriving the Riemann integral from evident properties of the (familiar) notion of *area* from the axioms of monotonicity and additivity.
- Deriving the axioms of *order* (instead of using 30 pages as Russell did in *Principles*).

Traditionally, one gave *one* convincing set of axioms, instead of relying on the equivalence of many different descriptions (as is familiar from recursion theory); on the principle that one good reason is better than 20 poor ones.

To distinguish this kind of activity from formal deductions, one speaks of *informal rigour*.

b) Reservations

This kind of informal analysis, long dominant in mathematics, has been downplayed since the thirties. The most obvious external sign: disappearance of familiar words like 'dimension', and introduction of *neologisms*.[45]

The most obvious reservation applies when one has no confidence in the familiar notion; tacitly, in its (familiar) meaning or any other context in sight. Probably, this applies to choice sequences.

The matter of (a) admits a convincing reinterpretation, without any hypothesis on the 'fundamental' character: if descriptions in terms taken from different areas of knowledge specify the same object, one has a ready made arsenal of methods to study that object, namely all those areas. And if the object has been recognized to be relevant to one significant area, we are in business. This view of the matter is unaffected by the obvious weakness of (a) in overlooking the possibility of a *systematic* oversight which spoils all of the (systematically equivalent) descriptions. In other words, the weakness consists in a *lack of safeguards* against such systematic oversights.[46]

---

[45] Possibly, a temporary point of diminishing returns; to be taken up later, in combination with another interpretation coming, say, from gauge theories in physics.

[46] However, when (as described above) one has no confidence in a notion, it is hardly clear that it is worth the trouble to find such safeguards.

Current mathematical practice is usually described by means of (grammatically well-formed) grunts about: 'judging by success'; or 'on aesthetic reasons'; or even (in Hilbert) about 'judging by the fruits', in a reference to the Bible (*New Testament*) where not only good but also bad fruits are mentioned! Here the whole attempt at an informally rigorous analysis which *justifies* a paraphrase is rejected (with grunts). Superficially, this seems quite silly: it can surely do no harm to have such an analysis (for example, if one uses the paraphrase anyway); as it can surely do no harm to have additional money or power (assuming that, like all mathematicians, we are full of good will). But the rejection is consistent with a more subtle consideration, reminiscent of (perfectly sound) concerns in politics or economics: regarding some piece of analysis as a *worthy end in itself* draws attention away from delicate relevant *open* questions, e.g. what to do next when the money has lost its value by devaluation. From this point of view, *logical preoccupations are the opium of the scientist* wishing to make scientific progress. As a corollary, current mathematics insists that the paraphrase has been used *successfully*. For, while (in the abstract) reference to 'success' is weak because (again, in the abstract) it is easy to imagine circumstances where success is difficult to judge, there is a tacit requirement that the paraphrase has been used not only successfully, but so that success is *easy to judge*.

Digression. A moment's though shows that the present considerations (3.2) apply very generally to other parts of language beside the (natural) language of mathematics. It would therefore not only be inefficient, but unnecessarily clumsy to present the points above merely in the special context of mathematics. So readers are invited to look for parallels in current studies of natural language: the *possibility* of finding both grammatical rules and (an intended) meaning for dialects (as intuitionistic mathematics uses a dialect of mathematical language), which draws attention away from the fact that a paraphrase can be better.

3.3) 'Enrichments by descriptions' versus 'equivalence classes of descriptions'

(As in the digression above) this topic is a special case of a very general preoccupation, going back to the Greeks, on the priority of 'the' *extensional* (object) or 'the' *intensional* (specification). The whole thing is a mixture of experience and so-called normative considerations; not so much 'normative' in the sense of what one ought to do to make progress, but in the sense of what the world 'ought' to be like (to be called rational, as in 'rational mechanics'). In particular, one argues that the specification should be prior because that is how we come to know about objects; and, in fact, that is how we identify objects. The argument is not particularly persuasive if one actually looks at experience: in so-called preattentive (or, better, inattentive) perception we remember the object or some abstract feature, and not the details nor even the particular aspect or side of the object which was perceptible on the occasion in question.

The *assumption* that an exposition should proceed from a list of axioms requires us to single out a particular description or definition though, as knowledge develops, that particular choice may become arbitrary.[47]

a) Relevance

As mentioned repeatedly, not only validity, but *relevance* and *significance* are of central interest. Now, there is a certain subjective flavor attached to the latter virtues which creates a malaise (as if subjective elements were necessarily arbitrary, and hence unsuitable for theoretical study). To avoid this issue, mathematicians have adopted the device of setting out generalizations and consequences (corollaries) to *exhibit*, respectively, relevance and significance, *without saying* that this is the purpose of the exercise; a device to avoid (often sterile) arguments whether the purpose is actually achieved.[48]

Given any particular theorem, not everything that is known about the structure of the objects concerned will be relevant to the theorem (key word: axiomatization). But also (as in the paradigm of the passage from functions, in the sense of rules, to their graphs) not everything about the genesis or 'history' of the object will be relevant. So it seems plausible that, as knowledge grows, different aspects will be discovered to be relevant. *Sample*:

- In the case of functions the *graphs* were relevant (apart from numerical analysis, where the notation chosen was equally important). But in more delicate problems at least *types* (that is, bounds on the range of the functions) had the advantage that problems (e.g. functional equations) could be solved by more elementary means. And *if* these additional data are available (and effective) it is poor science to ignore them.

b) Equivalence classes

It is a feature of 'fundamental' theories that fundamental equivalence relations are *discovered*; for example, for chemical phenomena, the chemical structure of a substance (within limits independent of, say, its temperature).

As a matter of *general* policy, the passage to a particular equivalence class (and thus the choice of a particular equivalence relation) does not seem to have worked well. *Samples*:

[47] When definitions are called mere 'conventions', it is to be distinguished whether the *choice of object* defined is an arbitrary matter, or whether (given the object) the *choice among possible definitions* is arbitrary (to some extent).

[48] This tactic is parallel to the use of neologisms after the 30's for naming concepts like homotopy, in order to avoid the arguments that arose in the 20's about claims implicit in using familiar words like dimension or area (of a surface).

- In the case of natural numbers, if one thinks *only* of *counting* one-by-one then the successor structure on notations is indeed primary; but not if one starts calculating, let alone if one remembers their use for measurement (e.g. in geometry).
- Trivially, pushed to its extreme, the equivalence relation used will be mere equality, and then we finish up with *cardinal arithmetic* (the logician's paradise, as mentioned already).

c) Enrichments

The opposite passage, to increasing enrichments by (or, for that matter, occasional suppression of) new data, is a neglected instance of an *empirical element* in mathematics, a possibility of *learning from experience*. Phrases like:

the tricks of mathematics are eternal

suggest that it is disturbing not to have this possibility! This is consistent with Lakatos, Popper and Putnam wanting the truths to be in need of correction. The slight variant:

the (eternal) tricks of mathematics are eternally interesting

is less disturbing. Experience *has* a place in correcting assessments of interest or relevance of a piece of knowledge for the broad body of knowledge, as already discussed in (a).

14.4 From Foundations to Technology

The prime examples we consider are Computer Science and Artificial Intelligence. Broadly speaking, the two subjects are quite different. Computer Science tries to do things well that people do badly or, realistically speaking, simply could not do at all. Artificial Intelligence - for example, realized by robots - tries to do what (so to speak, abstractly) people do well or even immeasurably better than robots, but under certain circumstances cannot do for one of diverse reasons; for example: boredom, or because they cannot survive the surroundings (on Mars, inside a nuclear reactor), or because they are uneconomical. So it should not be surprising if all that is common to both subjects turns out to be pretty trivial in each.

a) Computer Science

The logical view has, by and large, been unproductive for Computer Science (in the sense of: science of algorithms), except at a very early stage when it showed the value of hardware that would realize Boolean operations, recursion, and (of course, above all) the possibility of using the same material as codes for arguments and operations. *Samples*:

- Large *lower bounds* for practically all problems about arbitrary formulas of elementary logic. This is a *conflict* because, on the logical view, such problems have the logically ideal form.

- The choice of language, in the sense of *explicit definitions*, turns out to be of the essence. Again, for the logical view explicit definitions are negligible.

Contemporary logic has been used to *establish* these results (even though experienced computer scientists may have been convinced of it before).

b) Artificial Intelligence

Artificial Intelligence is a typical example of the step from foundations to technology. The situation goes back to one of the main features of the logical view, particularly clear in Frege. Granted that we know little about human information processing, it is simply unpromising to speculate (like Kant) what we *use* in reasoning, in particular in solving problems. This is in any case irrelevant to the question he raised about what (kind of solution) is *possible*. Obviously, if we are sensible we do not confine ourselves to what is so to speak logically needed, but we use the resources we have. Now, granted further that we do not have the technology to construct cheaply systems which use the full range of resources available in higher organisms, then it becomes interesting to go back to Kant's question (not his particular claims!) about looking for *minimum* possibilities. These may be realizable technically.

In this respect the logical view concerning a technical task (to achieve a given output for a given input) appears perfectly appropriate; especially if we really have no idea how to do it; for example, if we found no model for it in a phenomena in nature that we understand well.

Of course, a good deal of Artificial Intelligence (which is by no mean homogeneous) proceeds by starting with an algorithm together with a robot that can do unexpected things; then one looks for applications in the literal sense of commercial applications.[49] This direction is only slowly being counted as part of Artificial Intelligence, because it does not involve the detour via a familiar psychological solution of some (partly intellectual) task. In other words, it does not appear to be an analysis of an intellectual process, but so to speak a *paraphrase*.[50]

[49]So here too we have an interplay between methods and aims adjusting one to the other; cf. Section 2.

[50]Evidently, contemporary logic is also not homogeneous, and exhibits a similar interplay.

14.5 Conclusions

Contemporary logicians think of their subject as a collection of mathematical disciplines, not so much related by the methods used to prove the main result, as by the subject matter; more precisely, the subject matter of most general (and therefore most elementary) logical notions of relations including predicates and (various kinds of) general processes. The generality involved is elementary because it is the kind illustrated by what is common to two objects when they are seen from far away; and not what may be common when their atomic or subatomic structures are meant. Naturally, most of the logical disciplines have developed on the pattern of other mathematical disciplines: the general *notions* remain of interest or have even been absorbed in the body of current thought, but the general *theory* of the (general) notion has long reached the point of diminishing returns, to be compared for example, to the subject of *all* groups (i.e. the general notion of groups).

Section 1 of the present chapter documented the heterogenous character of contemporary mathematical logic, by reference to a few samples (which interested readers can easily supplement from the literature, provided they look for the so to speak qualitative difference in level of mathematics sampled in Section 1).

But the main stress was on the relevance of contemporary logic to a very much broader topic and a neglected distinction within it. Specifically:

1. the broad idea of a *logical view* or, more simply, of *logical aspects* of (mathematical) thought, with the particular emphasis on generality mentioned earlier

2. the additional idea of giving a central place to those aspects for so-called *foundational* aims; both in the sense of a systematic exposition beginning with (that is, founded on) logical notions, and in the sense of being secured (made 'certain') in that manner.[51]

[51]The following two common oversights exaggerate the role of logical aspects on general (or so to speak *a priori*) grounds, and of logical foundations on empirical grounds.

The general *must* come first. For example, before recognizing a particular argument as a proof one must know what a proof is, i.e. proceed from the known to the proved. But even granted this piece of wisdom, it is of little relevance when the principal issue is *selection*: from which particular knowledge, by which procedures. *Reminder*. Tarski's definition of truth is unsatisfactory *only* if one expects too much, that is, expects what was generally claimed for the importance of the question: What is Truth? The solution has corrected a false evaluation of the question.

The problems of logical foundations are said to have had heuristic value (at least, for some people) because they introduced clever ideas *à propos* of those problems; cf. medieval theologians who had interesting ideas on thought and action, and formulated them in the context of speculations on the mind or designs of God. But it is a quite delicate matter to assess what these same people would have done outside the constraints (of foundations or theology), and how many people talented for the broad area (of logic or philosophy) were put off by the constraints.

This broad idea 1 is as old as Western culture; cf. Plato's *Republic* about men having lost their reason when they become concerned with particulars. Now, in fact (at least when interpreted literally) practically the whole of experimental science and mathematics did the latter (though there may occasionally have been a general interest 'behind' those specific preoccupations). Without exaggeration: nothing remotely like that logical view had got a foothold in mathematics at all before the work of Cantor or abstract sets, and of Boole and Frege on logic. There was no mathematical discipline of those logical aspects,[52] and they were not prior to the bulk of mathematics.

Here it is to be noted that work of the pioneers on the subject was not primarily concerned with either of the foundational aims in 2. The work was certainly not viewed by the silent majority of the contemporaries as contributing to foundations. On the contrary, it was suspect; partly because, as already mentioned, one had had no experience with successful mathematics of such broad generalities.

Perhaps, the most recurrent theme of early work on logical aspects was the refutation of Kant's claim for some kind of need for specific (reasoning) abilities, which incidentally are pretty obviously used constantly, such as visualization of one sort or another. The logical pioneers interpreted this 'need' as logical, that is, for obtaining in principle as it were mathematical results (tacitly, after being reformulated logically).[53] Unfortunately for the development of the then-young logical view, the pioneers chose to stress that particular theme, and not the *obviously* striking discovery that there was anything (in fact, a good deal) to be done with superficially so unpromisingly general subject matter.

The other distraction came from the business of the paradoxes. They led to (or at least provided a temptation for) converting age-old elements of the logical view (such as the issues of purity of method, or incompletely defined predicates) into panaceas for the paradoxes (specifically, finitistic purity and incomplete or potential totalities). The outcome followed the principle: if one worries about nonexistent dangers, one is liable to miss the real ones.

Nevertheless work associated with these divisions permits also a reinterpretation: as a contribution to intellectual tools extending our capacities for (mathematical) reasoning, but not necessarily as particularly suited for foundational aims in the literal sense. Specifically, the logical notions proposed as *sole* foundations are regarded as (more or less new) elements in the arsenal of mathematical notions, whose range of *relevance* (not merely of definition) is to be determined by research.

Contemporary logic has achieved two things. As mentioned:

---

[52]One may indeed look at elementary number theory as something very general about finite sets: but this was not the way it was looked at (and neither the theory of amicable numbers nor of ternary quadratic forms fits into it).

[53]The recognition that a proposition stated in terms prominent in Kant was correctly expressed by a logical reformulation was outside the logical domain.

1. it has established that something can be done, that is, (general) logical notions have some rewarding theory

2. by shifting emphasis, so to speak in the course of nature, without external pressure it has itself established the limits of relevance, the point (or range) of diminishing returns.[54]

For reasons explained in Section 4, the positive side is particularly prominent in Artificial Intelligence, the negative side in Computer Science (two, in other respects, easily comparable subjects).

[54]This is not paradoxical. Generally any theory has to be developed before its weaknesses are apparent (except for really brutal errors like Frege's axioms, or Galileo's first formula: $v = \alpha s$ for freeling falling bodies).

Bibliography

Ackermann, W.

[1924] Begründung des 'tertium non datur' mittels der Hilbertschen Theorie der Widerspruchsfreiheit, *Math. Ann.* 93 (1924) 1–36.

Ax, J., and Kochen, S.

[1965] Diophantine problems on local fields, *Am. J. Math.* 87 (1965) 605–630.

Bachmann, F.

[1975] Frege als konstruktiver Logizist, in *Frege und die moderne Grundlagenforschung*, Thiel ed., Meseinheim am Glan, 1975, pp. 160–168.

Barwise, J.

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

Barwise, J., and Perry, J., eds.

[1983] *Situations and attitudes*, M.I.T. Press, 1983.

Bell, J.S.

[1977] *Boolean-valued models and independence proofs in set theory*, Clarendon Press, 1977.

Benacerraf, P. and Putnam, H., eds.

[1964] *Philosophy of mathematics*, Prentice Hall, 1964.

Bernard, J.F.

[1973] *Talleyrand*, New York, 1973.

Bernays, P.

[1935] Sur le platonisme dans les mathématiques, *Enseign. Math.* 34 (1935) 52–69.

[1940] Review of Gödel [1939], *J. Symb. Log.* 5 (1940) 116–117.

[1961] Die hohen Unendlichkeiten und die Axiomatik der Mengenlehre, in *Infinitistic methods*, Pergamon Press, 1961, pp. 11-20.

Berry, M.V.

[199?] Some quantum-to-classical asymptotics,

Bezem, M.

[1985] Strongly majorizable functionals of finite type. A model for bar recursion containing discontinuous functionals, *J. Symb. Log.* 50 (1985) 652–660.

Bishop, E.

[1967] *Foundations of constructive analysis*, McGraw Hill, 1967.

Boole, G.

[1854] *An investigation of the laws of thought, on which are founded the mathematical theories of logic and probability*, Walton and Maberley, 1854.

Bourbaki, N.

[1948] L'architecture des mathematiques, in *Les grands courants de la pensee mathematique*, Le Lionnais ed., Blanchard, 1948, pp. 35–47, transl. in *Am. Math. Monthly* 57 (1950) 221–232.

Brouwer, L.E.J.

[1905] *Leven, Kunst, en Mystiek*, Delft, 1905, partly transl. in [1975], pp. 1–10.

[1907] *Over de grondslagen der wiskunde*, Dissertation, Amsterdam, 1907, transl. in [1975], pp. 15–101.

[1908] De onbetrouwbaarheit der logische principes, *Tijdschr. v. wijsb.* 2, 1908, transl. in [1975], pp. 107–111.

[1908a] see Volume II, p. 150.

[1913] Intuitionism and formalism, *Bull. Am. Math. Soc.* 20 (1913) 81–96, also in [1975], pp. 123–138.

[1924] Intuitionistische Zerlegung matematischer Grundbegriffe, *Jber. Deutsch. Math. Verein.* 33 (1924) 251–256, also in [1975], pp. 275–280.

[1927] Über Definitionsbereiche von Funktionen, *Math. Ann.* 97 (1927) 60–75, transl. in van Heijenoort [1967], pp. 446–463.

[1947] Address to Prof. G. Mannoury, *Synthese* 6 (1947) 190–194, also in [1975], pp. 472–476.

[1948] Essentieel negatieve eigenschappen, *Ind. Math.* 10 (1948) 322–323, transl. in [1975], pp. 478–479.

[1948a] Consciousness, philosophy and mathematics, *Proc. Int. Congr. Math.* 10 (1948) 1235–1249, also in [1975], pp. 480–494.

[1975] *L.E.J. Brouwer Collected Works, Volume I*, Heyting ed., North Holland, 1975.

[1981] *Cambridge lectures on intuitionism*, van Dalen ed., Cambridge University Press, 1981.

Brouwer, L.E.J., and de Loor, B.

[1924] Intuitionistischer Beweis des Fundamentalsatzes der Algebra, *Proc. Ned. Acad. Wetensch.* 27 (1924) 186–188.

Browder, F.E., ed.

[1976] , Proc. Symp. Pure Appl. Math. 28, 1976.

Burgess, J.P.

[1981] The completeness of intuitionistic propositional calculus for its intended interpretation, *Notre Dame J. Form. Log.* 22 (1981) 17–28.

Cantor, G.

[1879] *Nachr. Ges. Wiss. Göttingen* (1879) 127–135.

[1885] Review of Frege [1884], *Deuts. Literat.* 6 (1885) 728–729, reprinted in Cantor [1932], pp. 440–441.

[1889] Letter to Dedekind, reprinted in Cantor [1932], pp. 443–447.

[1932] *Gesammelte Abhandlungen*, Zermelo ed., Springer, 1932.

Chang, C.C. and Keisler, H.J.

[1973] *Model Theory*, North Holland, 1973.

Chevalley, C.

[1936] Demonstration d'une hypothese de M. Artin, *Abh. Math. Sem. Univ. Hamburg* 11 (1936) 73–78.

Church, A.

[1953] Non-normal truth tables for the propositional calculus, *Boll. Soc. Mat. Mex.* 10 (1953) 41–52.

Cohen, P.J.

[1963] The independence of the continuum hypothesis, *Proc. Nat. Acad. Sci.* 50 (1963) 1143–1148.

[1964] The independence of the continuum hypothesis II, *Proc. Nat. Acad. Sci.* 51 (1964) 105–110.

[1971] Comments on the foundations of set theory, *Proc. Symp. Pure Appl. Math.* 13 (1971) 9–15.

Conway, J.H.

[1976] *On numbers and games*, Academic Press, 1976.

Craig, W.

[1965] Boolean notions extended to higher dimensions, in *The theory of models*, Addison et al. eds., North Holland, 1965, pp. 55–69.

Crawshay-Williams, R.

[1970] *Russell remembered*, London, 1970.

Crick, F.

[1981] *Life itself*, Simon and Schuster, 1981.

Crick, F., Griffith, , and Orgel,

[1957] Codes without commas, *Proc. Nat. Acad. Sci.* 43 (1957) 416–422.

Crick, F., and Orgel,

[1973] Directed panspermia, *Icarus* 19 (1973) 341–346.

Davis, M., Matyasevic, Y. and Robinson, J.

[1976] Hilbert's tenth problem. Diophantine equations: positive aspects of a negative solution, *Proc. Symp. Pure Appl. Math.* 28 (1976) 323–378.

Dedekind, R.

[1872] *Stetigkeit und irrationale Zahlen*, Braunschweig, 18872.

[1888] *Was sind und was sollen die Zahlen*, Braunschweig, 1888.

Denef, J.

[1984] The rationality of Poincaré series associated to the p-adic points of a variety, *Invent. Math.* 77 (1984) 1–23.

Descartes, R.

[1637] *Discours de la méthode*, Leyden, 1637.

Dirac, P.M.

[1978] *Directions in physics*, Wiley, 1978.

Dreben, B., Andrews, P. and Anderaa, S.

[1963] False lemmas in Herbrand, *Bull. Am. Math. Soc.* 69 (1963) 699–706.

Dries, L. van der and Wilkie, A.J.

[1984] Gromov's theorem on groups of polynomial growth and elementary logic, *J. Alg.* 89 (1984) 349–374.

Einstein, A.

[1944] Remarks on Bertrand Russell's theory of knowledge, in *The philosophy of Bertrand Russell*, Schilpp ed., Evanston, 1944, pp. 279–291.

Eklof, P.

[1976] Whitehead's problem is undecidable, *Am. Math. Month.* 83 (1976) 775–788.

Ellison, W.J.

[1975] *Les nombres premiers*, Hermann, 1975.

Ershov, Y.L.

[1965] On the elementary theories of local fields, *Alg. Log.* 4 (1965) 5–30.

Feferman, S. and Spector, C.

[1962] Incompleteness along paths in progressions of theories, *J. Symb. Log.* 27 (1962) 383–390.

Finsler, P.

[1926] Formale Beweise und die Entscheidbarkeit, *Math. Zeit.* 25 (1926) 676–682, transl. in van Heijenoort [1967], pp. 438–445.

Frege, G.

[1879] *Begriffschrift*, Halle, 1879, transl. in van Heijenoort [1967], pp. 1–82.

[1884] *Grundlagen der Arithmetik*, Breslau, 1884.

[1893] *Grundgesetze der Arithmetik*, Jena, 1893.

Friedman, H.

[1977] *Ann. Math.* 105 (1977) 1–28.

Gabbay, D.

[1976] A note on Kreisel's notion of validity in Post systems, *Studia Logica* 35 (1976) 285–295.

[1981] *Semantical investigations in Heyting's intuitionistic logic*, Reidel, 1981.

Galvin, F. and Prikry, K.

[1976] Infinitary Jonsson algebras and partition relations, *Alg. Univ.* 6 (1976) 485–494.

Gandy, R.O. and Hyland, M.

[1977] Computable and recursively countable functions of higher type, in *Logic Colloquium '76*, Gandy et al. eds., North Holland, 1977, pp. 407–438.

Gentzen, G.

[1935] Untersuchunngen über das logische Schliessen, *Math. Zeit.* 39 (1935) 176–210, 405–431, transl. in [1969], pp. 68–131.

[1935a] First version of Sections IV and V of [1936], in [1969], pp. 201–213.

[1936] Die Widerspruchsfreiheit der reinen Zahlentheorie, *Math. Ann.* 112 (1936) 493–565, transl. in [1969], pp. 132–213.

[1969] *The collected papers of Gerhard Gentzen*, Szabo ed., North Holland, 1969. **Geroch, and Hartle,**

[19?] Volume II p. 248

Getty, P.J.

[1963] *My life and fortunes*, New York, 1963.

Girard, J.Y.

[1987] *Proof theory and logical complexity*, Bibliopolis, 1987.

Goad, C.A.

[1978] Monadic infinitary propositional logic: a special operator, *Rep. Math. Log.* 10 (1978) 43–50.

Gödel, K.

[1929] Dissertation,

[1930] Die Vollständigkeit der Axiome des logischen Funkionenkalküls, *Monash. Math. Phys.* 37 (1930) 349–360, transl. in [1986], pp. .

[1930a] Einige metamathematische Resultate über Entscheidungsdefinitheit und Widerspruchsfreiheit, *Anz. Akad. Wiss. Wien* 67 (1930) 214–215, transl. in [1986], pp. .

[1931] Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I, *Monash. Math. Phys.* 38 (1931) 173–198, transl. in [1986], pp. 145–195.

[1931a] Diskussion zur Grundlegung der Mathematik, *Erkenn.* 2 (1931) 147–151, transl. in [1986], pp. .

[1932] Eine Eigenschaft der Realisierungen des Aussagenkalküls, *Ergebn. Math. Kolloq.* 3 (1932) 20–21, transl. in [1986], pp. .

[1932a] Zum intuitionistischen Aussagenkalküls, *Anz. Akad. Wiss. Wien* 69 (1932) 65–66, transl. in [1986], pp. .

[1933] Zur intuitionistischen Arithmetik und Zahlentheorie, *Ergebn. Math. Kolloq.* 4 (1933) 34–38, transl. in [1986], pp. 287–295.

[1933a] Zum Entscheidungsproblem des logischen Funkionenkalküls, *Monash. Math. Phys.* 40 (1933) 433–443, transl. in [1986], pp. .

[1933b]Eine Interpretation des intuitionistischen Aussagenkalk"uls, *Ergebn. Math. Kolloq.* 4 (1933) 39–40, transl. in [1986], pp. .

[1934] Review of Skolem [1933a], (1934) , transl. in [1986], pp. .

[1934a]On undecidable propositions of formal mathematical systems, in *The undecidable*, Davis ed., Raven Press, 1965, pp. 41–74.

[1936] Über die Länge von Beweisen, *Ergebn. Math. Kolloq.* 7 (1936) 23–24, transl. in [1986], pp. .

[1938] The consistency of the axiom of choice and of the generalized continuum-hypothesis, *Proc. Nat. Acad. Sci.* 24 (1938) 556–557.

[1939] Consistency-proof for the generalized continuum-hypothesis, *Proc. Nat. Acad. Sci.* 25 (1939) 220–224.

[1940] *The consistency of the axiom of choice and of the generalized continuum hypothesis*, Princeton University Press, 1940, 2nd edition 1951.

[1944] Russell's mathematical logic, in *The philosophy of Bertrand Russell*, Schilpp ed., Evanston, 1944, pp. 125–153.

[1946] Remarks before the Princeton bicentennial conference on problems in mathematics, in *The undecidable*, Davis ed., Raven Press, 1965, pp. 84–88.

[1947] What is Cantor's continuum problem, *Am. Math. Mon.* 54 (1947) 515–525.

[1949] An example of a new type of cosmological solutions of Einstein's field equations of gravitation, *Rev. Mod. Phys.* 21 (1949) 447–450.

[1949a]A remark about the relationship between relativity theory and idealistic philosophy, in *Albert Einstein, philosopher-scientist*, Schilpp ed., Evanston, 1949, pp. 555–562.

[1950] Rotating universes in general relativity theory, *Proc. Int. Congr. Math.* (1950) 175–181.

[1958] Über eine bisher noch nicht benütze Erweiterung des finiten standpunktes, *Dial.* 12 (1958) 280–287, transl. in [1986], pp. .

[1964] What is Cantor's continuum problem, revised and enlarged edition, in *Philosophy of mathematics*, Benacerraf et al. eds., Prentice-Hall, 1964, pp. 258–287.

[1972] second edition of [1958]

[1972a]remarks on the undecidability results

[1974] Preface to the second edition of Robinson [1966].

[1986] *Collected works*, volume I, Oxford University Press, 1986.

[1989] *Collected works*, volume II, Oxford University Press, 1989.

Goldfarb, W.D.

[1979] Logic in the twenties: the nature of quantifier, *J. Symb. Log.* 44 (1979) 351–368.

[1984] The unsolvability of Gödel class with identity, *J. Symb. Log.* 49 (1984) 1237–1252.

Grattan-Guinness, I.

[1979] In memoriam: Kurt Gödel, *Hist. Math.* 6 (1979) 294–304.

Hajnal, A.

[1956] On a consistency theorem connected with the generalized continuum problem, *Zeit. Math. Log. Grund. Math.* 2 (1956) 131–136.

Hasenjäger, G.

[1953] Eine Bemerkung zu Henkins Beweis für die Vollständigkeit des Prädikantenkalküls, *J. Symb. Log.* 18 (1953) 42–48.

Hasse, H.

[1964] *Vorlesungen über Zahlentheorie*, 2nd edition, Springer, 1964.

Hawking, S.W. and Ellis, G.F.R.

[1973] *The large scale structure of space-time*, Cambridge University Press, 1973.

Heckmann, O. and Schücking, E.

[1962] Relativistic cosmology, in *Gravitation: an introduction to current research*, Witten ed., Wiley and Sons, 1962, pp. 428–469.

Heijenoort, van, J.
[1967] *From Frege to Gödel*, Harvard University Press, 1967.

Herbrand, J.
[1930] *Recherches sur la théorie de la démonstration*, Thesis, Paris, 1930, transl. in van Heijenoort [1967], pp. 525–581.

Heyting, A.
[1930] Die formalen Regeln der intuitionistischen Logik, *Akad. Wiss. Phys. Math.* (1930) 42–56.
[1956] *Intuitionism. An introduction*, North Holland, 1956.

Henkin, L.
[1952] A problem concerning provability, *J. Symb. Log.* 17 (1952) 160.

Hermann, I.
[1949] Denkpsychologische Betrachtungen im Gebiete der mathematischen Mengenlehre, *Schweiz. Z. Psych.* 8 (1949) 189–231.

Higman, G.
[1961] Subgroups of finitely presented groups, *Proc. Roy. Soc.* 262 (1961) 455–474.

Hilbert, D.
[1899] *Grundlagen der Geometrie*, Leipzig, 1899.
[1904] Über die Grundlagen der Logik und der Arithmetik, *Verhandl. III Intern. Math. Kongr.* (1904) 174–185, transl. in van Heijenoort [1967], pp. 129–138.
[1926] Über das Unendliche, *Math. Ann.* 95 (1926) 161–190, transl. in van Heijenoort [1967], pp. 367–392.
[1930] Probleme der Grundlegung der Mathematik, *Math. Ann.* 102 (1930) 1–9.
[1931] Grundlegung der elementaren Zahlentheorie, *Math. Ann.* 104 (1931) 485–494.

Hilbert, D. and Ackermann, W.
[1928] *Grundzüge der theoretischen Logik*, Springer, 1928.

Hilbert, D., and Bernays, P.
[1934] *Grundlagen der Mathematik*, Berlin, 1934.
[1934] *Grundlagen der Mathematik*, volume II, Berlin, 1939.

Hodge, W.V.D. and Pedoe, D.
[1947] *Methods of algebraic geometry*, Volume I, Cambridge, 19437.

Holmberg, M.A., ed.
[1951] *Les prix Nobel en 1950*, Stockholm, 1951.

Howard, W.A.
[1968] Functional interpretation of bar induction by bar recursion, *Comp. Math.* 20 (1968) 107–124.
[1970] Assignment of ordinals to terms for primitive recursive functionals of finite type, in *Intuitionism and proof theory*, Kino et al. eds., North Holland, 1970, pp. 443–458.
[1973] Hereditarily majorizable functionals of finite type, in Troelstra [1973], pp. 454–461.
[1980] Ordinal analysis of terms of finite types, *J. Symb. Log.* 45 (1980) 493–504.

Howard, W.A. and Kreisel, G.
[1966] Transfinite induction and bar induction of types 0 and 1, and the role of continuity in intuitionistic analysis, *J. Symb. Log.* 31 (1966) 325–358.

Hübner, , and Wuchterl,
[198]

Hume, D.
[1777] *An enquiry concerning the principles of morals*, London, 1777.

Janik, A., and Toulmin, S.

[1973] *Wittgenstein's Vienna*, Simon and Schuster, 1973.

Jensen, R.

[1972] The fine structure of L, *Ann. Math. Log.* 4 (1972) 229–308. **Jeroslow, R.G.**

[1973] Redundancies in the Hilbert-Bernays derivability conditions for Gödel's second incompleteness theorem, *J. Symb. Log.* 38 (1973) 359–367.

Jong, de, D.H.J.

[1980] A class of intuitionistic connectives, in *Kleene symposium*, Barwise ed., North Holland, 1980, pp. 103–111.

Kanamori, A. and Magidor, M.

[1978] The evolution of large cardinal axioms in set theory, *Springer Lect. Not. Math.* 669 (1978) 99–275.

Ketonen, J., and Solovay, R.

[1981] Rapidly growing Ramsey functions, *Ann. Math.* 113 (1981) 267–314.

Kleene, S.C.

[1952] Recursive functions and intuitionistic mathematics, *Proc. Int. Congr. Math.* (1952) 679–685.

[1955] Hierarchies of number-theoretical predicates, *Bull. Am. Math. Soc.* 61 (1955) 193–213.

[1976] The work of Kurt Gödel, *J. Symb. Log.* 41 (1976) 761–778.

[1978] An addendum to [1976], *J. Symb. Log.* 43 (1978) 613.

Knorr, W.

[1978] Archimedes and the spirals: the heuristic background, *Hist. Math.* 5 (1978) 43–75.

[1983] 'La croix des mathématiciens': the Euclideans theory of irrational lines, *Bull. Am. Math. Soc.* 9 (1983) 41–69.

Komar, A.

[1964] Undecidability of macroscopical distinguishable states in quantum field theory, *Phys. Rev.* 133 (1964) 542–544.

Kreisel, G.

[1950] Note on arithmetical models for consistent formulae of the predicate calculus, *Fund. Math.* 37 (1950) 265–285.

[1951] On the interpretation of non-finitist proofs I, *J. Symb. Log.* 16 (1951) 241–267.

[1952] On the interpretation of non-finitist proofs II, *J. Symb. Log.* 17 (1952) 43–58.

[1952a] On the concept of completeness and interpretation of formal systems, *Fund. Math.* 39 (1952) 103–127.

[1952b] Some elementary inequalities, *Ind. Math.* 14 (1952) 334–338.

[1953] Note on arithmetical models for consistent formulae of the predicate calculus II, *Proc. Int. Congr. Phil.* 14 (1953) 39–49.

[1956] Some uses of metamathematics, *Brit. J. Phil. Sci.* 7 (1956) 161–173.

[1958] Relative consistency proofs (abstract), *J. Symb. Log.* 23 (1958) 109–110.

[1959] Analysis of the Cantor-Bendixson theorem by means of the analytic hierarchy, *Bull. Acad. Pol. Sci.* 7 (1959) 621–626.

[1960] see Volume II, p. 208.

[1965] Mathematical logic, in *Lectures on modern mathematics*, Saaty ed., Wiley, 1965, vol. 3, pp. 95–195.

[1967] Informal rigour and completeness proofs, in *Problems in the philosophy of mathematics*, Lakatos ed., North Holland, 1967, pp. 138–186.

[1968] Functions, ordinals, species, in *Logic, Methodology and Philosophy of Science III*, van Rootselaar and Staal eds., North Holland, 1968, pp. 145–159.

[1970] Hilbert's Programme and the search for automatic proof procedures, *Springer Lect. Not. Math.* 125 (1970) 128–146.

[1971] Some reasons for generalizing recursion theory, in *Logic Colloquium '69*, Gandy et al. eds., North Holland, 1971, pp. 139–198.

[1971a] Review of Kreisel [1971], *Zentr. f. Math.* 199 (1971) 300–301.

[1972] Which number-theoretic problems can be solved in recursive progressions on Π^1_1-paths through $\mathcal{O}$?, *J. Symb. Log.* 37 (1972) 311-334.

[1973] Perspectives in the philosophy of pure mathematics, *Log. Meth. Phil. Sci.* 4 (1973) 255–277.

[1974] A notion of mechanistic theory, *Synthese* 29 (1974) 11–26.

[1975] Some uses of proof theory for finding computer programs, in *Colloque International de Logique, Clermont-Ferrand*, Guillaume ed., 1975, pp. 123–134.

[1977] On the kind of data needed for a theory of proofs, in *Logic Colloquium '76*, Gandy et al. eds., North Holland, 1977, pp. 111–128.

[1980] Grundlagen der Mathematik, in *Handbuch wissenschaftstheoretischer Begriffe*, Vandenhoeck and Ruprecht, 1980, pp. 393–400.

[1982] Review of Pour El and Richards [1979] and [1981], *J. Symb. Log.* 47 (1982) 900–902.

[1985] Review of *Strukturtypen der Logik* by Stegmüller and Varga, *Grazer Phil. Stud.* 24 (1985) 185–195.

[1985a] Review of *Fundamentals of generalized recursion theory* by Fitting, *Bull. Am. Math. Soc.* 13 (1985) 182–197.

[1985b] Proof theory and the synthesis of programs: potential and limitations, *Springer Lect. Not. Comp. Sci.* 203 (1985) 136–150.

[1986] Philosophie: eine Ergänzung der Wissenschaft ?, *Int. Wittgen. Symp.* 9 (1986) 51–56.

Kreisel, G., and Krivine, J.L.

[1966] *Elements of mathematical logic*, North Holland, 1966, 2nd edition 1971.

Kreisel, G., and MacIntyre, A.

[1982] Constructive logic versus algebraization I, in *The Brouwer centenary symposium*, Troelstra et el. eds., North Holland, 1982, pp. 217–260.

Kreisel, G., Mints, G.E., Simpson, S.G.

[1975] The use of abstract language in elementary metamathematics: some pedagogical examples, *Springer Lect. Notes Math.* 453 (1975) 38–131.

Kreisel, G. and Tait, W.W.

[1961] Finite definability of number-theoretic functions and parametric completeness of equational calculi, *Zeit. Math. Log. Grund. Math.* 7 (1961) 28–38.

Kreisel, G., and Takeuti, G.

[1974] Formally self-referential propositions for cut free classical analysis and related systems, *Diss. Math.* 118 (1974) 1–50.

Kreisel, G. and Troelstra, A.S.

[1971] Formal systems for some branches of intuitionistic analysis, *Ann. Math. Log.* 1 (1970) 229–387 and 3 (1971) 437–439.

Kripke, S.

[1982] *Wittgenstein on rules and private language*, Harvard University Press, 1982.

Lakatos, I., ed.

[1967] *Problems in the philosophy of mathematics*, North Holland, 1967.

Lévy, A.

[1957] Independence conditionelle de $V = L$ et d'axiomes qui se rattachent au systeme de M. Gödel, *Comp. Rend. Acad. Sci.* 245 (1957)1582–1583.

Löb, M.H.

[1955] Solution of a problem of Leon Henkin, *J. Symb. Log.* 20 (1955) 115–118.

Lomonosov, V.I.

[1973] Invariant subspaces of the family of operators that commute with a completely continuous operator, *Funk. Anal. Prilož.* 7 (1973) 55-56. **Lopez-Escobar, E.G.K.**

[1976] On a very restricted ω-rule, *Fund. Math.* 90 (1976) 156–172.

[1982] Further applications of ultraconservative ω-rules, *Arch. Math. Log. Grund.* 22 (1982) 89–102.

Lorenz, K.

[1968] Dialogspiele als semantische grundlage von Logik kalkülen, *Arch. Math. Log. Grund.* 11 (1968) 32–55, 73–100.

Lorenzen, P.

[1951] Über endlich Mengen, *Math. Ann.* 123 (1951) 331–338.

MacIntyre, A.

[1971] On ω_1-categorical theories of fields, *Fund. Math.* 71 (1971) 1–25.

Mahlo, P.

[1912] Zur Theorie und Anwendung der ρ_0-Zahlen, *Ber. Verh. Sachs. Akad. Wiss.* 64 (1912) 190–200.

Malcev, A.

[1936] Untersuchungen aus dem Gebiete der mathematischen Logik, *Mat. Sbor.* 1 (1936) 323–336.

[1941] On a general method for obtaining local theorems in group theory, *Ivanov Gos. Ped. Inst.* 1 (1946) 3–9.

Maistre, J. de

[1821] *Les soirées de Saint-Pétersbourg ou entretiens sur le gouvernement temporel de la providence: suivis d'un Traité sur les sacrifices*, Paris, 1921.

Manewitz, L. and Stavi, J.

[1980] Δ_2^0 operators and alternating sentences in arithmetic, *J. Symb. Log.* 45 (1980) 144–154.

Martin, D.A.

[1975] Borel determinacy, *Ann. Math.* 102 (1975) 363–371.

[1976] , *Proc. Symp. Pure Appl. Math.* 28 (1976)

[1985] new proof of Borel determinacy, *Proc. Symp. Pure Appl. Math.* 42 (1985)

Matyasevic, Y.

[1970] Enumerable sets are diophantine, *Dokl. Acad. Nauk* 191 (1970) 279–282, transl. *Sov. Math. Dokl.* 11 (1970) 354–357.

McCarthy, M.T.

[1952] *The groves of Academe*, New York, 1952.

Milnor, J.

[1958] Some consequences of a theorem of Bott, *Ann. Math.* 68 (1958) 444–449.

Montague, R. and Vaught, R.L.

[1959] Natural models of set theories, *Fund. Math.* 47 (1959) 219–242.

Moore, G.

[1980] Beyond first-order logic, *Hist. Phil. Log.* 1 (1980) 95–137.

Mostowski, A.

[1952] *Sentences undecidable in formalized arithmetic*, North Holland, 1952.

Motz, L. and Motz, R.O.

[1990] Sensible mathematics, *Nature* 345 (1990) 300.

Musil, R.

[1930] *Der Mann ohne Eigenschaften*, Berlin, 1930.

Nedo, M., and Ranchetti, M.

[1983] *Ludwig Wittgenstein: sein Leben in Bildern und Texten*, Surkamp, 1983.

Nerode, A. and Harrington, L.

[1984] The work of Harvey Friedman, *Not. Am. Math. Soc.* 31 (1984) 563–566.

Neumann, von, J.

[1927] Zur Hilbertschen Beweistheorie, *Math. Zeit.* 27 (1927) 1–46.

[1928] Die Axiomatisierung der Mengenlehre, *Math. Zeit.* 27 (1928) 339–422.

Newman, M.H.A.

[1969] *Luitzen Egbertus Jan Brouwer*, part II, *Biogr. Mem. Fell. Royal Soc.* 15 (1969) 46–53.

Nkrumah, K.

[1970] *Consciencism; philosophy and ideology for decolonization*, Montly Review Press, 1970. **Nobeling, G.**

[1968] Verallgemeinerung eines Satzes von Herrn E. Specker, *Inv. Math.* 6 (1968) 41–55.

Parikh, R.J.

[1979] Review of Kreisel [1977], *Math. Rev.* 58 (1979) 3203–3204.

Paris, J.B., and Harrington, L.

[1977] A mathematical incompleteness in PA, in Barwise [1977], pp. 1133–1142.

Pour El, M.B. and Richards, I.

[1979] A computable ordinary differential equation which possesses no computable solutions, *Ann. Math. Log.* 17 (1979) 61–90.

[1981] The wave equation with computable initial data such that its unique solution is not computable, *Adv. Math.* 39 (1981) 215–239.

[1983] Non-computability in analysis and physics: a complete determination of the class of non-computable linear operators, *Adv. Math.* 48 (1983) 44–74.

Prawitz, D.

[1974] On the idea of a general proof theory, *Synth.* 27 (1974) 63–77.

Rabin, M.

[1977] Decidable theories, in Barwise [1977], pp. 595–629.

Rados, G.

[1906] Zur ersten Verteilung des Bolyai-Preises, *Math. Ann.* 62 (1906) 156–176.

Ramsey, F.P.

[1928] On a problem of formal logic, *Proc. Lond. Math. Soc.* 30 (1928) 338–384.

Robinson, A.

[1955] *Théorie Métamathématique des idéaux*, Gauthier-Villar, 1955.

[1966] Non-standard analysis, North Holland, 1966, 2nd edition 1974.

Robinson, A., and Roquette, P.

[1975] On the finiteness theorem of Siegel and Mahler concerning diophantine equations, *J. Number Th.* 7 (1975) 121–176.

Robinson, J.

[1968] Recursive functions of one variable, *Proc. Am. Math. Soc.* 19 (1968) 815–820.

Rokeach, M.

[1964] *The three Christs of Ypsilanti; a psychological study*, New York, 1964.

Rosser, B.J.

[1936] Extensions of some theorems of Gödel and Church, *J. Symb. Log.* 1 (1936) 87–91.

Russell, B.

[1900] *A critical exposition of the philosophy of Leibniz*, Cambridge, 1900.

[1903] *The principles of mathematics*, London, 1903.

[1905] On denoting, *Mind* 14 (1905) 479–493.

[1906] On some difficulties in the theory of transfinite numbers and order types, *Proc. Lond. Math. Soc.* 4 (1906) 29–53.

[1912] *The problems of philosophy*, London, 1912.

[1919] *Introduction to mathematical philosophy*, London, 1919.

[1944] My mental development, in *The philosophy of Bertrand Russell*, Schilpp ed., Evanston, 1944, pp. 1–20.

[1945] *A history of western philosophy*, New York, 1945.

[1951] What desires are politically important?, in *Les prix Nobel en 1950*, Holmberg ed., Stockholm, 1951, pp. 259–270.

[1959] *My philosophical development*, New York, 1959.

[1967] *Autobiography 1872–1914*, vol. I, Atlantic Monthly Press, 1967.

[1968] *Autobiography 1914–1944*, vol. II, Atlantic Monthly Press, 1967.

[1969] *Autobiography 1944–1968*, vol. III, Atlantic Monthly Press, 1967.

Schroeder-Heister, P.

[1984] A natural extension of natural deduction, *J. Symb. Log.* 49 (1984) 1284–1300.

Schütte, K., and Van der Waerden, B.L.

[1953] Das Problem der dreizehn Kugeln, *Math. Ann.* 125 (1953) 325–334.

Scott, D.S.

[1962] More on the axiom of extensionality, in *Essays on the foundations of mathematics*, Bar Hillel et al. eds., Magnes Press, 1962, pp. 115–176.

Senovilla, J.M.M.

[1990] A new class of inhomogeneous cosmological perfect fluid solutions without Big Bang singularities, *Phys. Rev. Letters* 64 (1990) 2219.

Shoenfield, J.

[1957] Open sentences in partial systems of arithmetic, *J. Symb. Log.* 22 (1957) 112.

[1959] On the independence of the axiom of constructibility, *Am. J. Math.* 81 (1959) 537–540.

[1962] The problem of predicativity, in *Essays on the foundations of mathematics*, Bar Hillel et al. eds., North Holland, 1962, pp. 132–139.

Siegel, .

[1929]

Simpson, S.G.

[1984] Which set existence axioms are needed to prove the Cauchy-Peano theorem for ordinary differential equations?, *J. Symb. Log.* 49 (1984) 783–802.

Skolem, T.

[1922] Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre, *Proc. Congr. Scand. Math.* 5 (1922) 217–232, transl. in van Heijenoort [1967], 290–301.

[1933] Review of Gödel [1933], *Forts. Math.* 59 (1933) 865–866.

[1933a] Über die Unmöglichkeit einer vollständingen Charakterisierung der Zahlenreihe mittels eines endlichen Axiomensystems, *Norsk Mat. For. Skrif.* 10 (1933) 73–82.

Smorynski, C.

[1977] ω-consistency and reflection, *Proc. 1975 Log. Coll. Clermont-Ferrand*, 1977, pp. 167–181.

[1977a] The incompleteness theorems, in Barwise [1977], pp. 821–865.

Solzhenitsyn, A.I.

[1968] *The first circle*, London, 1968.

[1968a]*The cancer ward*, London, 1968.

Specker, E.P.

[1950] Additiven Gruppen von Folgen ganzer Zahlen, *Port. Math.* 9 (1950) 131–140.

Spector, C.

[1962] Provably recursive functionals of analysis: a consistency proof of analysis by an extension of principles formulated in current intuitionistic mathematics, *Proc. Symp. Pure Appl. Math.* 5 (1962) 1–27.

Statman, R.

[1974] *Structural complexity of derivations*, Ph.D. Thesis, Stanford University, 1974.

[1977] Herbrand's theorem and Gentzen's notion of a direct proof, in Barwise [1977], pp. 897–912.

Sturm, C.

[1835] Mémoire sur la résolution des équations numériques, *Mem. Acad. Roy. Sci.* 6, 1835.

Sundholm, G.

[1983] Constructions, proofs and the meaning of the logical constants, *J. Phyl. Log.* 12 (1983) 151–172.

Tait, W.W.

[1967] Intensional interpretation of functionals of finite type, *J. Symb. Log.* 32 (1967) 198–212.

Tarski, A. Mostowski, A. and Robinson, R.M.

[1953] *Undecidable theories*, North Holland, 1953.

Taub, A.H.

[1951] Empty space-times, *Ann. Math.* 53 (1951) 472–490.

Taussky, O.

[1987] in *Gödel remembered*, Weingartner and Schmettered eds., Bibliopolis, 1987, pp.

Troelstra, A.S.

[1973] *Metamathematical investigations of intuitionistic arithmetic and analysis*, Springer Lecture Notes in Mathematics 344, 1973.

[1977] *Choice sequences, a chapter of intuitionistic mathematics*, Clarendon Press, 1977.

[1981] The interplay between logic and matehmatics: intuitionism, in *Modern logic - A survey*, Agazzi ed., Reidel, 1981, pp. 197–221.

[1983] Analysing choice sequences, *J. Phil. Log.* 12 (1983) 197–260.

Troelstra, A.S., and van Dalen, D.

[1988] *Constructivism in mathematics*, 2 volumes, North Holland, 1988.

Turing, A.M.

[1936] On computable numbers with an application to the Entscheidungsproblem, *Proc. Lond. Math. Soc.* 42 (1936) 230–265.

[1939] Systems of logic based on ordinals, *Proc. Lond. Math. Soc.* 45 (1939) 161–228. **Van den Hoeven, G.F., and Moerdijk, I.**

[1984] Constructing choice sequences from lawless sequences of neighbourhood functions, in *Logic Colloquium '83*, Müller et al. eds., Springer, 1984, pp. 207–234.

Van der Dries, L.

[1982] Some applications of a model-theoretic fact to (semialgebraic) geometry, *Ind. Math.* 44 (1982) 397–441.

Vaught, R.L.

[1974] Model theory before 1945, *Proc. Symp. Pure Math.* 25 (1974) 153–172.

Wang, H.

[1974] *From mathematics to philosophy*, Routledge and Kegan, 1974.

Weil, A.

[1974] *Basic number theory*, Springer, 1974.

Weinberg, S.

[1976] The forces of nature, *Bull. Am. Acad. Sci.* 29 (1976) 13–29.

Weinstein, S.

[1983] The intended interpretation of intuitionistic logic, *J. Phil. Log.* 12 (1983) 261–270.

Weyl, H.

[1918] *Das Kontinuum*, Leipzig, 1918.

[1946] Review of *The philosophy of Bertrand Russell*, *Am. Math. Monthly* 53 (1946) 208–214.

Whithead, A.N., and Russell, B.

[1910] *Principia Mathematica*, vol. I, Cambridge, 1910.

Wigner, E.

[1960] The unreasonable effectiveness of mathematics in the natural sciences, *Comm. Pure Appl. Math.* 13 (1960) 1–14.

[1982] , *Nobel Conf.* 17 (1982)

Wittgenstein, L.

[1921] Logisch-philosophische Abhandklung, *Ann. Naturphil.* 14 (1921) 185–262.

[195] *Philosophical Investigations*, 195 .

[1956] *Bemerkungen über die Grundlagen der Mathematik*, Oxford, 1956.

[1967] *Zettel*, Oxford, 1967.

[1980] *Vemischte Bemerkungen*, Chicago, 1980.

Wojtylak, P.

[1982] Collapse of a class of infinite disjunctions in intuitionistic propositional logic, *Rep. Math. Log.* 16 (1982) 37–49.

[1984] A recursive theory for the $\{\neg, \wedge, \vee, \rightarrow, \circ\}$ fragment of intuitionistic logic, *Rep. Math. Log.* 18 (1984) 3–35.

Zermelo, E.

[1896] Über einen Satz der Dynamik und die mechanische Warmetheorie, *Ann. Phys.* 57 (1896) 485–494.

[1904] Beweis, dass jede Menge wohlgeordnet werden kann, *Math. Ann.* 59 (1904) 514–516, transl. in van Heijenoort [1967], pp. 139–141.

[1908] Untersuchungen über die Grundlagen der Mengenlehre I, *Math. Ann.* 65 (1908) 261–281, transl. in van Heijenoort [1967], pp. 199–215.

[1912] Über eine Anwendung der Mengenlehre auf die Theorie des Schachspiels, *Proc. Int. Congr. Math.* 5 (1912) 501–504.

[1930] Über Grenzzahlen und Mengenbereiche, *Fund. Math.* 16 (1930) 29–47.

[1932] Über Stufen der Quantifikation und die Logik des Unendlichen, *Jber. Dt. Mat. Verein.* 41 (1932) 85–88.

[1935] Grundlagen einer allgemeinen Theorie der mathematischen Satzsysteme, *Fund. Math.* 25 (1935) 136–146.

I.S.B.N. 978-65-900390-1-9