KREISELIANA
Essays
About and Around
Georg Kreisel

Edited by

Piergiorgio Odifreddi

Contents

Preface

by Piergiorgio Odifreddi

Like everyone, only a lot more than most, Georg Kreisel is a man of multiple and conflicting personalities: an extremely influential logician and philosopher of mathematics, a convoluted and cryptic essayist and reviewer, an immensely prolific epistolizer, a fascinating and tireless conversationalist, a misanthropic friend of billionaires and actresses, an enthusiastic intellectual philanthropist invariably disappointed by his disciples and often disappointing to them ...

It thus comes as no surprise that the present celebration is based on conflicting wishes of the celebrants and the celebrated. Kreisel has known about the work in progress for quite a while, but has mentioned it to me only incidentally, in these (for him, typical) terms:

> I know that some people are discreet about this sort of thing, but I am not. In particular, I want you to know that *I know* that an old friend of mine has told you of my dislike for this sort of thing, and that he *was right*.
>
> As you know I take little interest in (your) motives, conscious or unconscious; for example, whether you think that – contrary to overwhelming evidence – I'd be somehow pleased to see that I am not forgotten, or whether you are an inveterate busybody, or or or ...
>
> When I described myself – as they say, provocatively – as the spokesman for the silent majority, many people including you regarded this as *totally* absurd. But almost as many people *assume I must be like the majority*, in particular, like

> those who want to be remembered (at all, or on their 70th Birthdays). (26.VI.93)

Granted, Kreisel had spotted an inconsistency in people's beliefs. But, as I told him in my reply, being remembered depends more on one's deeds than on one's wishes, and certainly, with his behavior and writing, he has left both unforgettable marks and unforgivable scars.

As for my *conscious* motives, I have to confess that I started organizing this collection out of exasperated affection: after years of having heard from him about his own insights, as well as my (or other people's) misunderstandings, I longed to see different perspectives. The occasion of his 70th Birthday (on 15.IX.93) provided an opportunity as good as any.

As one might have expected, all these essays (including descriptions of Kreisel that range from an excellent cook, through a zen monk, all the way to Lord Krishna) have only deepened the mystery.

Certainly an ability to generate opposite perceptions and feelings is one of Kreisel's outstanding characteristics. For example, while logicians are convinced he was the inspiration for the manipulative character of Julius King in Iris Murdoch's novel *A fairly honourable defeat* (Penguin Books, 1972), and some of them personally recall that it was Kreisel himself who told them the story, Ms. Murdoch claimed to be shocked and surprised when I let her know about it, and insisted:

> Of course the character was not 'based on' Kreisel: the scheming mind and activity of J.K. is *utterly different* from G.K.'s particular honesty, sincerity and goodness. (2.V.92)
>
> Kreisel's witty and lovable mannerisms and his nobility may have been an inspiration for similar aspects of the personality of Julius King; but the unpleasant aspects of this character are *obviously not* related to Kreisel. (22.IX.93)

Often the same person will go from one extreme to the other, at different times. To give an example close to home: after reading the strong words that Kreisel used to write me in his letters, my wife could not understand how I (or anybody) could stand such a spiteful man; only to find him attractive and fascinating once she finally met him in person, and he was free to cast his spell.

Such an experience, however, appears to have been rare: if not in the switch of perception, certainly in its direction. This accounts for some notable and noticeable absentees in this collection among Kreisel's former students, coauthors, colleagues and friends, including people who have dedicated books to him, or have otherwise acknowledged a high intellectual debt.

These absentees may have chosen the easiest way out of their obligation to the controversial man: paraphrasing Wittgenstein, who was Kreisel's friend and confidant in the last ten years of his life, they passed over in silence what they could not speak about. Or, perhaps, they have only been more responsive to Kreisel's own desires in this matter.

More bravely, the contributors to this volume have often dared crossing the boundaries of dry scientific exposition, to roam through the vaster and more exciting land of human understanding. Obviously, they could only take the first steps toward their complex goal, and one day somebody will have to provide a more coherent view of the man Kreisel. In the meantime, this collection of first-hand accounts (notwithstanding, or because of, the often unorthodox and somewhat atypical character of some of them) provides a flavour of his complex personality and achievements.

One thing is certain: this volume is a sincere effort to communicate to a wider circle Kreisel's unique personal and intellectual influence, and an encouragement to get past the roadblocks that he puts in our way to the hidden gems of his subtle and interesting thought. We all hope he feels pleased by at least part of the tribute, but if he only feels systematically misunderstood ... I'm willing to bet that *he*, unlike most members of the silent majority, would still get from this a kind of indirect pleasure.

Kreisel's Vita

Georg Kreisel was born in Graz (Austria), on September 15, 1923. He received his mathematical education at Trinity College, Cambridge (England), where he earned a B.A. in 1944, an M.A. in 1947, and a Sc.D. in 1962. He worked as an Experimental Officer for the British Admiralty from 1943 to 1946, and held academic positions at various places, including Cambridge (1946–48), Reading (1949–54, 1957–58, 1959–60), Princeton (1955–57, 1963–64), Paris (1960–62), Los Angeles (1968), and Stanford (1958–59, 1962–85), where he became Professor Emeritus in 1985. In 1966 he was elected Fellow of the Royal Society of London.

REMINISCENCES

Kreisel, Lambda Calculus, a Windmill and a Castle

by Henk Barendregt

This paper gives some idea of the role that Kreisel played at the start of my scientific career. The facts are taken mainly from the period starting spring 1971 (when I worked on my Ph.D. thesis) until summer 1972 (when I ended a stay at Stanford as a postdoc).

1 The Setting

At Utrecht University in the late sixties I studied logic under reader Dirk van Dalen. In those days readers in the Netherlands did not have the *ius promovendi*, i.e. were not allowed to be official Ph.D. supervisors. Kreisel had accepted to spend a spring semester at Utrecht and van Dalen asked whether he was willing to be my Ph.D. supervisor. For this I had to send a description of my work to Kreisel and in a few days the answer was there. The 'master' found my work "congenial" and it could be the basis for discussions leading towards a thesis. Kreisel was supposed to be difficult, so I had heard. Actually this inspired me, while I prepared the manuscript before his coming to Utrecht in 1971.

When Kreisel arrived in the Netherlands, I had borrowed my fathers car (a Renault 16), because mine (a Renault 4) was total-loss due to an accident, and I picked him up from the airport. Kreisel was pleased by this reception 'in style' and for a moment he forgot to be difficult. When we stopped for gas, however, he started to ask sharp questions that were meant to frighten me. Since I did not have a high opinion about myself (I did not have a low one either), it did not matter to me that I could not answer all of his questions. The fact that I was unconditioned by his interrogation mellowed him down and in high spirits we arrived at Utrecht University.

It was well-known that Kreisel always requests a quiet place to sleep. After some effort Mrs. Dook van Dalen had found a candidate place in Bilthoven, a fancy village near Utrecht. But one never knows if it is really adequate. Therefore I showed Kreisel a picture of my home – a windmill in the countryside along the Waal river, the main branch of the Rhine – and offered him to stay there in case he wanted.

After the obligatory inspection of the house in Bilthoven, Kreisel decided to stay in the windmill. This brought Mrs. van Dalen in an awkward situation. The owner of the house had been asked to provide a quiet room for Kreisel. Being sensitive to who was to come into her house, the landlady had emptied her own bedroom on the garden side for the use of the famous professor. From the cool reaction of Mrs. van Dalen about Kreisel's decision to stay in the windmill one could deduce the fury of the old lady when her kind offer was declined.

The windmill was situated 45 minutes by car from the University, a relatively long distance in the Netherlands. Arrived there, Kreisel told me that the next day he was invited by a baroness for tea at her castle and asked me to join him. "Of course one of the reasons to ask you is that you can help to drive me there," was his frank admission. By a remarkable coincidence the castle was only 2 kilometer upstream from the windmill. I knew the place, but had no idea who lived there.

Against my expectations the baroness was a young woman, good-looking with two small children. After being introduced to her I felt a bit uneasy about the angle in which she shook hands. Was I supposed to give a handkiss?[1] The baroness was British and lived with her husband in London. Often during weekends the baroness and her children would come to their Dutch castle, while the lord remained in London. In the 15th century ancestors of the baron had been victors of a small battle in the Netherlands. Because of this historical fact his lineage was proclaimed to be 'Baron van Ophemert en Zennewijnen' since that time.

During tea Kreisel invited the lady to come to his first lecture in Utrecht, the following week. She accepted. "It is better for you to leave

[1]The next day I borrowed a book on etiquette from a girl-friend in Utrecht. My question was not answered, though.

the lecture-hall after twenty minutes, because then I will start to be rather technical," was Kreisel's advice.

On the morning of the day of the first lecture I dropped Kreisel at the castle. He and the baroness were going to have lunch in Utrecht and would come together to the University. At 2:00 p.m. the logicians in the Netherlands were eagerly waiting. At 2:15 Kreisel was still not there. I gave them some details. Nobody yet at 2:30. Finally at 2:40 a taxi entered the parking place of the department, directly followed by a car with a British license plate. Kreisel came out of the British car and paid the taxi and the baroness parked her car. This scene could be followed by all Dutch logicians from the lecture-hall on the sixth floor of the building. Troelstra remarked that Kreisel was still wearing the same trousers as in Stanford.[2] When at last the lecture started at 2:50, it went as follows. "Logic is the science of deduction. How we can derive from a statement A another statement B. ... " This went on for twenty minutes. When the baroness left – as foreseen – she had to pass a door next to the blackboard in front of the audience. Shyly she opened the door. Kreisel, however, was fully at ease. With an elegant bow he paid his respects to the lady. At 3:10 the real lecture started. It was on Rosser sentences, brilliant and full of firework. I could not follow it.[3]

A few days later Kreisel and I were invited by the baroness for dinner at her castle. Impressive stairs went down to a cellar with burning torches on the walls. Some of the dishes were exotic. During the dessert Kreisel showed a side of his, that was new to me. One of the children started to whine about the pudding. "Ma, I do not want the raisins in it; ma, I want you to take them out. . . ." It went on and on. Then Kreisel spoke very slowly to the little one, a boy about six years old: "Listen. What do you think is worse: your mother does not want to take the raisins out; or, your mother is not able to take them out." The little boy had to think for a while. "It is worse, if she does not want to take them out," was the answer. "Well," Kreisel continued, "it is simply the case that she is not able to take them out!" The boy could do nothing else but eat his pudding.

Kreisel and I often visited the castle. Since then the baron also came more frequently to the Netherlands. I was impressed by Kreisel's explanations to the lord and his wife on the research I was doing in

[2] Troelstra had been a visitor at Stanford a few years before.

[3] Later I mentioned this to Kreisel, telling him at the same time that in private conversations I could understand perfectly well his technical remarks. Also I said that I could not understand many of his writings. "Oh, but Gödel can, and Bernays can," was his reaction.

combinatory logic: "When we reason, we make steps, deductive steps. These can be smaller or larger steps. Barendregt studies the smallest possible such steps—the so-called atomic steps—and the way they can be combined to form larger ones." The high-born couple understood this, and at the same time I learned to be more flexible when expressing myself.

Also in other places of the world Kreisel introduced me to his 'upper-class' friends. The definition of this predicate varied from country to country. Sometimes they were aristocratic refugees from pre-war Europe, sometimes highly successful leaders of industry; but also there were high party members of some – then – powerful political party.

2 The Work

In the early 1970s the λ-calculus was considered to be a fringe area of logic. This in spite of the breakthrough by Dana Scott, who constructed in November 1969 the first set-theoretic model D_∞ of the λ-calculus. Kreisel, who makes a point of being interested in neglected areas of research, showed some genuine interest in the subject. This was very encouraging to me. I will discuss two main themes of my thesis (the ω-rule and (un)solvability) and their later developments.

The ω-rule

For my thesis, Barendregt [1971, 1971a], I wanted to construct a recursion theoretic model of the type-free $\lambda\mathsf{K}$-calculus.[4] Even today I have never been able to complete the construction. But the attempts proved fruitful. One of the candidate models was extensional and a hard one, i.e. generated by the closed terms. This implied the consistency of the following ω-rule.

$$FZ = GZ, \text{ for all closed terms } Z \quad \Rightarrow \quad F = G. \tag{1}$$

Since the model was resisting, I wanted to settle the consistency of this rule by other means. This was done using a proof-theoretic ordinal analysis. Then I wanted to know whether the proof was relevant. Perhaps rule (1) was simply derivable; in that case the consistency is trivial. It turned out that rule (1) was derivable (for $\beta\eta$-reduction), except possibly for some pathological terms F, G, the so-called *universal generators*. This almost proved the validity of the ω-rule. In Plotkin

[4] The construction was related to Kreisel's HRO (hereditarily recursive operations), which forms a model of the typed λ-calculus, see Troelstra [1973].

[1974], however, some impressive universal generators were constructed for which rule (1) in fact does not hold.

This ω-incompleteness had some repercussions on the notion of model of the λ-calculus. There are λ-models and λ-algebras. The λ-models are well-behaved and include Scott's set-theoretic models and the open term models. The λ-algebras are less well behaved and include closed term models. Nevertheless, the latter models have interesting properties, notably as pre-complete numerations in the sense of Ershov [1973, 1975, 1977]; see also Visser [1980]. In Barendregt [1977] and [1981] the notions of λ-model and λ-algebra are described in a correct but rather ad hoc syntactical manner. A nice description in first order logic of the notion of λ-model was given independently in Scott [1980] and Meyer [1982]. Koymans [1982] completed the story of finding the description of λ-calculus models. Based on work of Scott he gave a description of λ-algebras in terms of cartesian closed categories (CCC's)[5], in which λ-models form a special case.[6] The details of all this can be found in Barendregt [1984].

Finally, based on the work of Koymans [1982], in Curien [1986] so-called categorical combinators are developed for the use of implementations of functional programming languages. A successful application of this is CAM (categorical abstract machine) and the compiler CAML, used for implementing the proof-checker/developer COQ (based on the calculus of constructions, Coquand and Huet [1988]).

It also should be mentioned that in the proof of the partial validity of rule (1) the hypothesis ($FZ = GZ$, for all closed terms Z) was used in an extremely weak form: only one special closed term Z was used. Kreisel insisted that I should try to make use of more arguments in order to prove the full rule (1). I did not succeed and by Plotkins construction we know why. Nevertheless Kreisel turned out to be right that one can make use of more arguments. In Plotkin [1974a] it is proved that rule (1) is valid, provided that only one of F, G is not a universal generator, see Barendregt [1984], §17.3, for the details. In this proof the hypothesis of (1) is used for two different Z. In this line the best result is due to Nakajima [1975] and Wadsworth [1976]. They showed, using infinitely many Z, that the (axiom corresponding to the) rule (1) is valid in Scott's models D_∞; see also Barendregt [1984], §19.2.

[5] A λ-algebra is a *reflexive* object in a CCC, i.e. an object D such that for some arrows $F : D \to [D \to D], G : [D \to D] \to D$ one has $F \circ G = id_{[D \to D]}$.

[6] A λ-model is a reflexive object D *having enough points*, i.e.

$$\forall f, g : D \to D[[\forall z : 1 \to D \; f \circ z = g \circ z] \Rightarrow f = g],$$

where 1 is the terminal object. Notice the relation with rule (1).

So history proved that the interest of Kreisel in a neglected area of logic was fully justified.

Unsolvability

Another topic discussed with Kreisel for my thesis was the interpretation of the term $\Omega = (\lambda x.xx)(\lambda x.xx)$ in the recursion theoretic model. (This Ω is a $\lambda\mathsf{I}$-term and for these the candidate model is correct.) In the notation of Rogers [1967] this interpretation is

$$[\![\Omega]\!] = \varphi_e(e),$$

where e is such that $\varphi_e(x) = \varphi_x(x)$ for all $x \in \mathbb{N}$. The question was whether $[\![\Omega]\!]$ is defined or not.[7] It turned out that the answer depends on the choice of coding used to construct the universal function $\varphi_x(y)$. If a natural condition concerning lengths of computation is assumed for φ, then $[\![\Omega]\!]$ will be undefined; on the other hand for some 'non-standard' choices of φ this assumption is not valid and $[\![\Omega]\!]$ can be an arbitrary natural number[8]. See Barendregt [1975] for details.

Based on this result it follows that the interpretation of all λ-terms without normal form in the recursion theoretic structure with natural φ is undefined. As a consequence it seems to follow that terms without a normal form can be equated consistently.[9] This, however, is not the case. For example the equation between terms without a normal form

$$\lambda x.x\Omega\mathsf{true} = \lambda x.x\Omega\mathsf{false}$$

immediately gives $\mathsf{true} = \mathsf{false}$.[10] Analyzing the situation one sees that the two terms are *solvable*: $(\lambda x.x\Omega\mathsf{true})\vec{P}$ has a normal form for some $\vec{P}$ (and similarly for the other one). This notion turned out to be fruitful. All unsolvable terms can be identified consistently.[11] The

[7]This is related to a problem of Henkin, asking in arithmetic the provability of a statement H such that $\vdash H \leftrightarrow \Box H$, where $\Box H$ denotes formalized provability of H. As shown by Löb [1955] this is always the case.

[8]This non-standard φ is nevertheless an acceptable enumeration of the partial recursive functions of one argument in the sense of Rogers [1967].

[9]This presupposes that the recursion theoretic interpretation is a model for the $\lambda\mathsf{K}$-calculus.

[10]One needs an application to the combinator $\mathsf{K}_* \equiv \lambda xy.y$, which is a $\lambda\mathsf{K}$-term. In fact the recursion theoretic model does work for the $\lambda\mathsf{I}$-calculus. Hence we have the consistency of the $\lambda\mathsf{I}$-calculus $+\{M = N \mid M, N \text{ have no normal form}\}$.

[11]In order to prove this one needs the so-called 'genericity lemma': if M is unsolvable and N a normal form, then $FM = N \;\Rightarrow\; Fx = N$. In the $\lambda\mathsf{I}$-calculus the unsolvable terms are exactly the terms without normal forms, see Barendregt [1973]. This gives an alternative proof of the consistency of the theory mentioned in footnote 10.

consistency of the identification of unsolvable terms was proved later in Hyland [1976] and in Wadsworth [1976] by semantic methods (because $[\![M]\!]^{D_\infty} = \bot$, for M unsolvable). Moreover in these two papers it is proved that for closed terms M, N one has

$$\begin{array}{rcl} D_\infty \models M = N & \iff & \forall F[FM \text{ is solvable} \iff FN \text{ is solvable}] \\ & \iff & M = N \text{ belongs to the unique maximally} \\ & & \text{consistent extension of the } \lambda\text{-calculus} \\ & & \text{equating the unsolvables.} \end{array}$$

In this way solvability is a natural organizing principle for semantics of the λ-calculus. Along similar lines in Abramsky and Ong [1993] an alternative semantics is introduced that reflects features of implementations of lazy functional programming languages (one does not reduce 'under' a λ).

3 GK: Person and Influence

The person

Who is Georg Kreisel? Although the answer should be given by professional biographers, let me make some personal remarks.

Surely Kreisel is one of the most remarkable and enigmatic figures among logicians (and non-logicians). His behavior is non-conventional. Take for example his daily sleeping time: Kreisel goes to bed at 9:00 p.m. Sometimes I suspect that for him this is a convenient way to avoid social obligations and to get some extra work done.[12] In any case it is a fact, that Kreisel receives daily more than a dozen letters and/or articles. All this correspondence is usually answered by return mail.[13] This requires a discipline and concentration that I have experienced otherwise only among zen monks.

Another characteristic is that Kreisel seems to like to create a certain distance between himself and persons that he does meet. In many cases this is indeed the case. At the same time this creation of distance is applied to his own emotions as well. This quality I have experienced otherwise only in theravadin monks, albeit that the latter use a somewhat different method for doing so. Kreisel's method to successfully keep a proper distance from his emotions is by a logical analysis. If

[12] This hypothesis, however, I have not been able to verify.

[13] In these answers Kreisel often asks questions about a paper that are difficult for the author to answer.

this is skillfully done, then it is possible to disintegrate one's emotions into smithereens. And the resulting parts and pieces are harmless.[14]

Let me give some examples. Being in a certain country, Kreisel was asked by the late professor X – locally a well-known logician – the following. "Perhaps it is a stupid question, but can you tell me why this and this." Kreisel's reaction: "Ah, this question is not stupid at all, not stupid at all. The matter is so and so." The questioner continued eagerly: "Hence one also has that and that?" At this point Kreisel remarked in a flash: "Now *that* is a stupid question!"

This is what I would call creating a distance. And in this case Kreisel did so justifiedly: one does not want to be close to someone that either asks stupid questions while he knows it or wants to show off how smart he is. Kreisel gave this professor X both what he had hoped to hear and what he had feared to hear.

An example of Kreisel's way to decompose his own emotions is more difficult to give. Let me make it clear that his mastery of argumentation is of a high standard.[15] Given this skill, he can accomplish a lot.[16] A good example of an analysis of emotions into nothingness is the way Kreisel made the son of the baroness eat his pudding. But remember, the boy was only six. Given the sophistication of Kreisel's arguments one can imagine what is needed for a convincing transcending analysis of his own emotions. It is beyond my capacities to reproduce any of the cases in which I have witnessed this remarkable auto-analysis of Kreisel.[17]

Influence

One of the main pieces of advice that Kreisel gave me, was to use *reflection*. He claimed that logicians are often too busy with technical details in order to take some distance from their work. This way they lose a chance of obtaining better results.

An example of this is the following. In Barendregt [1973] I proved[18] that in the $\lambda\mathsf{I}$-calculus a term M has a normal form iff M is solvable. As

---

[14]The theravadin monks also analyze emotions into components, not by ratio but by *insight* based on *mindfulness*. As opposed to Kreisel's analysis based on logic, this requires less energy but a longer practice.

[15]Kreisel obtained some of his education from Jesuits.

[16]Due to an academic disagreement a colleague of Kreisel almost challenged him to court. It would have been interesting to compare the arguing skills of professional lawyers with that of Kreisel.

[17]In order to get an approximate idea of this phenomenon, one should read *Der Mann ohne Eigenschaften* by R. Musil (Rowolt Verlag, 1952).

[18]Using induction over a Σ^0_4-predicate.

title of the paper I chose: *'A characterization of terms in the λI-calculus having a normal form'*. Kreisel thought this was a particularly bad title. With a little more thought I could have given a much more significant and memorable one. Indeed, from the well-known result of Böhm [1967] and my own result that unsolvable terms can be consistently equated it follows that the λI-calculus plus η-conversion plus the equation of terms without a normal form constitute a Hilbert-Post complete theory.[19] So the title should have been something like: *'A natural Hilbert-Post complete extension of the λI-calculus'*.

Kreisel often made me aware of my academic career. "Take for example Heyting", he once said, "with results not more technical than yours he became well-known." Also Kreisel emphasized that one should publish one's results as soon as possible. With the present 'publish or perish' insanity,[20] this may sound obvious, but in 1971 it was not.

Each time I would see Kreisel again, after a few days, a few months or a few years, he would ask me "What is new?", cutting through our natural tendency to remain with our attachments. In fact, both scientifically and personally he always emphasizes *change*.

Claiming that all the advice that Kreisel had given me would fall under the heading 'Influence' is too much to be said. I wish it were true. But enough of his remarks are there in my memory, that occasionally things go along lines of his advice. Enough to have considerably profited from them.

Coda

Let me end with describing two sides of Kreisel that are apparently contradictory.

When I did drive with Kreisel in his car, I noticed that he would arrange things in a certain way, in order to have a better view. Remembering this a few years later when Kreisel came to Europe, I arranged things the same way in my car. His reaction: "Thank you very much. My cleaning girls never do this. Does one really need a Ph.D. in logic in order to do so?" This event, trivial as it may be, shows him as a pleasant companion.

One of the things that made him most angry, was my use of the sentence: "People will believe your opinion on this, because of your authority." At the time his strong reaction was not clear to me. This seems

[19] I.e. a maximally consistent theory.

[20] Presently one publishes too much in proceedings; this is notably the case in computer science. Good ideas deserve to be published in journals.

to show that he is a 'difficult' person and *not* a pleasant companion.[21]

The explanation of the apparent contradiction is easy. His negative feelings are caused by a general dislike of his. Kreisel abhors insincerity.[22] Since by and large most things in this world are done with insincere intentions, Kreisel is often 'difficult'. On the other hand Kreisel behaves well with the upper class, if they are sincerely upper class, and with the middle and lower classes, if they are sincerely so.[23]

The author Iris Murdoch wrote a novel in which a figure 'Julius' is said to be inspired by Kreisel. She wrote about this person: 'He is one of the people that opens your mail in your absence. But he tells you later that he did'.[24] If one has a taste for his style of sincerity and irony, then Kreisel is a very stimulating person to have around, both scientifically and personally. My friends that I had introduced to Kreisel all were charmed by him.[25]

Reflecting on what I just wrote, I search for a counterexample. Is it really true that his being difficult is always caused by insincerity of others? It almost sounds too good to be true for an ordinary mortal. In the case of Kreisel I could find no counterexample. So at least it is consistent.

References

Abramsky, S., and Ong, C.-H.

[1993] Full abstraction in the lazy lambda calculus, *Information and Computation*, 105 (1993) 159–267.

Barendregt, H.P.

[1971] *Some extensional term models for combinatory logics and λ-calculi*, Ph.D. thesis, University of Utrecht.

[1971a] *On the interpretation of terms in the lambda calculus without a normal form*, Ph.D. thesis, Part II, University of Utrecht.

[21] Once he said: "Because they do this, I will be difficult. And I can be difficult *as a woman*."

[22] Indeed, his anger was justified: one should not believe in someone because of authority, but because of that person's arguments.

[23] Kreisel gave genuine attention and protection to a secretary, that had problems with a boyfriend, and to a terminally ill student.

[24] *A fairly honorable defeat*, Penguin Books, 1972. Another of her books (*An accidental man*, Penguin Books, 1973) has as simple dedication 'To Kreisel'.

[25] E.g. (in this order) my father, the musician Nol Prager, my mother, the computer scientist Corrado Böhm, the conceptual artist Hans Koetsier. Buffee Nelson, then a three year old girl that later became my adopted daughter, once picked up the phone when Kreisel called. Before she passed the call to, me she happened to listen attentively. Later she asked: "Who was that?" When I had told her, she said: "I like Kreisel!"

[1973] A characterization of terms in the $\lambda\mathsf{I}$-calculus having a normal form, *J. Symbolic Logic*, 38 (1973) 441–445.

[1975] Normed uniformly reflexive structures, in Böhm [1975], pp. 272–286.

[1977] The type-free lambda calculus, in *Handbook of Mathematical Logic*, ed. J. Barwise, Studies in Logic, North-Holland, pp. 1092–1132.

[1981] *The Lambda Calculus, its Syntax and Semantics*, Studies in Logic 103, North-Holland, Amsterdam.

[1984] *The Lambda Calculus, its Syntax and Semantics*, revised edition, Studies in Logic 103, North-Holland, Amsterdam.

Böhm, C.

[1968] Alcune proprietà delle forme $\beta\eta$-normali nel $\lambda\mathsf{K}$-calcolo, Pubblicazioni dell'Istituto per le Applicazioni del Calcolo, Via del Policlinico 127, Roma, no. 696.

Böhm, C. (ed.)

[1975] *λ-calculus and Computer Science Theory*, Lecture Notes in Computer Science 37, Springer, Berlin.

Coquand, Th., and Huet, G.

[1988] The calculus of constructions, *Information and Computation*, 76 (1988) 95–120.

Curien, P.-L.

[1986] *Categorical combinators, sequential algorithms and functional programming*, Research Notes in Theoretical Computer Science, Pitman/Wiley, London.

Ershov, Y.

[1973] Theorie der Numerierungen I, *Zeitschrift Math. Logik u. Grundlagen d. Math.*, 19 (1973) 289–388.

[1975] Theorie der Numerierungen II, *Zeitschrift Math. Logik u. Grundlagen d. Math.*, 21 (1975) 473-584.

[1976] Theorie der Numerierungen III, *Zeitschrift Math. Logik u. Grundlagen d. Math.*, 23 (1977) 289–371.

Hyland, J.M.E.

[1976] A syntactic characterization of the equality of some models of the λ-calculus, *J. London Math. Society* (2), 12 (1976) 361–370.

Koymans, C.P.J.

[1982] Models of the lambda calculus, *Information and Control*, 52 (1982) 306-332.

Löb, M.H.

[1955] A solution of a problem of Henkin, *J. Symbolic Logic*, 20 (1955) 115–118.

Meyer, A.

[1982] What is a model of the lambda calculus?, *Information and Control*, 52 (1982) 87–122.

Nakajima, R.

[1975] Infinite normal forms for the λ-calculus, in Böhm [1975], pp. 62–82.

Plotkin, G.

[1974] The λ-calculus is ω-incomplete, *J. Symbolic Logic*, 39 (1974) 313–317.

[1974a] Personal communication.

Rogers, H.

[1967] *Theory of recursive functions and effective computability*, McGraw Hill, New York.

Scott, D.S.

[1980] Lambda calculus: some models, some philosophy, in *The Kleene Symposium*, eds. Barwise et al., Studies in Logic 101, North-Holland, Amsterdam, pp. 223–266.

Troelstra, A.S.

[1973] *Metamathematical Investigations of Intuitionistic Arithmetic and Analysis*, Lecture Notes in Mathematics 344, Springer, Berlin.

Visser, A.

[1980] Numerations, λ-calculus and arithmetic, in *To H.B. Curry: Essays on Combinatory Logic, Lambda-Calculus and Formalism*, eds. J.R. Hindley and J.P. Seldin, Academic Press, pp. 259–284.

Wadsworth, C.P.

[1976] The relation between computational and denotational properties for Scott's D_∞ models of the lambda calculus, *SIAM J. Comput.*, 5 (1976) 488–521.

The Right Things for the Right Reasons

by Jon Barwise

There have been many logicians who have had an impact on my career, but in one important regard, no one's influence has been more profound than Kreisel's. It is one I hope to acknowledge and describe here.

Writing about Kreisel is fun because he is such an interesting person. But it is also a bit dangerous. One is bound to give offense to someone or other, since Kreisel is a controversial figure. Many consider Kreisel one of the most influential figures in twentieth-century logic. Others are less generous. "Where," they ask, "are his important theorems or his memorable proofs?"

Kreisel has had his theorems and his share of graduate students. He also has many important if difficult philosophical papers, and he has probably written more letters to his colleagues than any other logician alive. But adding these together does not explain the awe in which he is held by some, or account for the criticisms of others. What is the truth of the matter? Besides acknowledging my special debt here, I hope to also convey my sense as to why Kreisel has played such an influential role in logic.

Stanford, 1963-67

I started graduate study in mathematics and logic at Stanford in the fall of 1963 and met Kreisel that winter. It was a thrilling time in logic in general, and at Stanford in particular. Model theory and recursion theory were in full flower, and the fresh techniques and results of Stanford's Paul Cohen promised a blossoming of set theory as well.

My own predilections were for the absolute, not the relative. As a result, I found the independence results in set theory dissatisfying. Partly for this reason, but also, I confess, partly to avoid the frantic competition starting in set theory (Kenneth Kunen was also student at Stanford), I chose to turn away from the independence results of set theory toward more absolute topics in model theory and recursion theory.

In those days there was a feeling among logicians that we had a good understanding of the classical case in both fields. First-order model theory was already a rich field, with many beautiful results. Recursion theory was even more fully developed. There was a natural interest in increasing our understanding by trying to generalize the work in each field: to stronger languages in the case of model theory, to domains other than the natural numbers in the case of recursion theory. In this way, it was felt, we could come to understand what was going on in the classical case more fully and develop tools that would apply in cases where the classical frameworks were too limiting.

Kreisel's writings and lectures on these topics gave off both heat and light. The heat took the form of scorn (expressed in scathing reviews) for some of the work done in these areas. On the side of light, Kreisel stressed that the two enterprises were inseparable. For example, there are many equivalent notions of "recursive" or "computable" on the natural numbers. Some of these are more semantic in nature, being based on notions of definability. Others are more computational. But the proofs of equivalence of the semantic with the more computational approaches depend on the Completeness and Compactness Theorems for first-order logic. In domains other than the natural numbers, this no longer works, in general, so the possibility of divergent approaches arises. Which one is "right"? In turn, the Completeness Theorem can be looked at abstractly as saying that the set of valid sentences is semi-recursive. While it is less obvious, the Compactness Theorem also has a recursion-theoretic statement. It could be that generalizations of logic will work out better with the right notions from generalized recursion theory.[1]

Sometime in 1965 I proved a result in this area – I have no idea what it was now. Excitedly, I took it in turn to Stanford's logicians, Sol Feferman, Dana Scott, and Kreisel. Feferman expressed genuine

[1]Kreisel had more substantive things than this to say about the matter, but I think the paragraph above gives a fair picture of my simplistic understanding of his views when I began my own work.

interest and listened to my proof. When I took the result to Scott he thought for a moment, then told me why it was true, giving me a better proof than the one I had prepared. Kreisel, on hearing the result, told me in no uncertain terms why it was the wrong thing to have proved – true or not.

This experience had a significant impact on my graduate work. For one thing, I asked Feferman to be my advisor. It was a good choice. Feferman was supportive, something I very much needed, but very firm about getting details right, which I also needed. He also urged me to take Kreisel's writings seriously, not something I was initially inclined to do, finding them almost impossible to understand.[2] For another, though, I started thinking not just about what was true, but also about what was worth proving.

Still, the results in my dissertation felt at the time more a matter of luck than of direction, planning, or forethought. In a logic seminar run jointly by Feferman and Kreisel, I happened to give presentations on two recent works. One was a paper based on Ken Lopez-Escobar's thesis, done under the direction of Dana Scott. This paper presented a cut-free sequent calculus for the infinitary logic $\mathcal{L}_{\omega_1,\omega}$, a richly expressive language that allows countably long conjunctions and disjunctions, but only finite strings of quantifiers. The other was the appendix to Richard Platek's just-completed dissertation. This dissertation, done under Kreisel's direction but never published, was on generalized recursion theory. The appendix, on admissible sets and recursively regular ordinals, had never even been proofread, as far as I could tell, but it was a terrific piece of work. I can still remember the excitement I felt in working through it. Given Kreisel's insistence that the projects of generalizing model theory and generalizing recursion theory were at heart the same project, and Feferman's urgings to take Kreisel's views seriously, it was only natural to try putting Lopez-Escobar's and Platek's work together, by seeing what would happen if you looked at admissible versions of $\mathcal{L}_{\omega_1,\omega}$.

Admissible sets are the well-founded, transitive models of a "weak" set theory, one I dubbed *KP* after Kripke (who had independently and somewhat earlier developed a similar theory) and Platek.[3] I quickly realized that the wrong approach was to try to prove Lopez-Escobar's

[2] I remember trying to read one of Kreisel's papers about informal, rigorous proofs, and having no idea what on earth he was getting at. Now, many years later, I find it an obvious and important distinction. Indeed, John Etchemendy and I make it a theme of our undergraduate logic text.

[3] While Kripke and Platek had similar theories, they were not the same, and it was Platek's version which I needed for my own work, and which I took as KP.

results within *KP*. It took me a little longer to see what did work: if one reformulated Lopez-Escobar's notion of proof for $\mathcal{L}_{\omega_1,\omega}$ carefully, without making gratuitous use of the axiom of choice, one could show that provability was absolute for admissible sets. This was the main lemma needed to prove what later became known as the Barwise Completeness and Compactness Theorems for countable, admissible fragments of $\mathcal{L}_{\omega_1,\omega}$.

Kreisel was particularly pleased and encouraging about these results. They were not difficult, but they supported the general line of thought he had been encouraging. In fact, when I understood things better, I came to see that my results were clearly foreshadowed in his work on ω-logic (see, e.g., footnote (b) on pages 119–120 of Kreisel [1960]). I suppose I should have understood it at the time, and Kreisel would have been quite justified in asking for more credit. Similarly, the assumption that it was a complete accident that I presented those two works, looked at from my current perspective, seems naive. It is just the sort of "coincidence" a good dissertation advisor like Feferman would engineer.

U.C.L.A., 1967-68

With Ph.D. in hand, I took part in the U.C.L.A. Logic Year, thanks to an N.S.F. Postdoctoral Fellowship. What a year that was. It was crucial for me, in many ways. I learned a great deal from the many logicians who were part of this year. I made lifetime friendships. And my own work got some useful recognition, especially when Harvey Friedman and Ronald Jensen used it to solve Gerald Sacks's long-standing open problem: is every countable admissible ordinal of the form ω_1^X for some set X of natural numbers X. Using my compactness result, they were able to show that the answer was affirmative.

Kreisel took part in the Logic Year for one quarter. Until he bought a car, I sometimes provided him with transportation. As a result, I got a glimpse of a different side of Kreisel: in the famous Polo Lounge, at home with Richard Montague, in the Hollywood star's home he rented for his stay, and at a cocktail party given in honor of Steve Kleene, where Kreisel flirted shamelessly with my office-mate's wife. Kreisel the social creature was quite different from Kreisel the logician. In these settings he was at his most disarming, full of charm and good-natured (usually) wit.

During this period, Kreisel gave a series of logic lectures. One day, speaking of a certain result on constructive order types, he remarked how it had proven useful, even though he considered the motivation for trying to prove it in the first place quite misplaced. He concluded by saying that it reminded him of a line from *The Cocktail Party*: This "is the greatest treason, to do the right deed for the wrong reason."

Now it so happens that this line of T.S. Eliot's is one that had haunted me even then, as it continues to do today. Feeling a little surer of myself than I would have a year earlier, and pleased at catching Kreisel in a mistake, I pointed out that it was not from *The Cocktail Party*, but from another Eliot play, *Murder in the Cathedral.* Without missing a beat, he said, "Oh, yes, that's right. I *heard* it at a cocktail party."

By mail, the 70's

During the 70's, I saw Kreisel occasionally at meetings in England or at Oberwolfach, but I was teaching at Yale and then at the University of Wisconsin, not places he frequented on his travels between Stanford and Europe. Consequently, most of my interactions with him during this period were by mail.

What an astonishing and exhausting correspondent Kreisel was, and maybe still is. It seemed to take longer for me to write and post a letter than it did to receive a lengthy response from Kreisel from half a world away. Given the number of people he was writing to, I wondered how it was physically possible. But more impressive than the quantity was the content of these letters. I never knew what to expect next. It was a hopeless challenge to match him idea for idea, or even just page for page. Still, he answered and only rarely expressed disappointment at some second-rate thought.

In 1973–74, I visited Oxford at the invitation of Dana Scott and worked on *Admissible Sets and Structures.* Part III of this book was directly inspired by Kreisel's ideas on definability as a form of recursion theory. My earlier results on countable, admissible fragments of $\mathcal{L}_{\omega_1,\omega}$ had as a consequence that various approaches to generalizing recursion theory coincide on countable, admissible sets. In the uncountable case, though, the approaches diverge, giving rise to competing notions. In his first paper, inspired by Kreisel's ideas, Kenneth Kunen had shown that in the uncountable case, the definability approach was more promising for applications to model theory. In my Part III, I followed up on Kunen's work, obtaining a number of results that gave further evidence for the Kreisel-Kunen theory.

Oxford proved a congenial setting for working on my book, in spite of the coal strike which kept us huddled by heaters and in bed early that winter. Robin Gandy and Dana Scott, the leading logicians at

Oxford at the time, were interested mainly in the material in the first two parts of my book. So I welcomed Kreisel's interest in this Part III. This interest kept our correspondence going and was particularly welcome.

One day at the math tea, Kreisel's name came up – not an unusual event in logic circles. Someone said that Julius King, the charismatic but treacherous protagonist of Iris Murdoch's 1970 novel *A Fairly Honourable Defeat*, was inspired by Kreisel. I rushed off to Blackwell's after tea to buy the book, discovering in the process that another of Murdoch's novels, *An Accidental Man*, published in 1971, was dedicated to Kreisel. This discovery made the notion that Murdoch's Julius King might have been modeled on Kreisel intriguing. If she had written such an unflattering portrait of him, why had she turned around and dedicated a book to him a year later? Interesting possibilities suggested themselves. I never found out which of these, if any, were actual. The straightforward thing would have been to ask Kreisel, but Kreisel does not encourage straightforward interactions, so I took another approach. In my next letter, I told Kreisel that I had come across the dedication in *An Accidental Man*, and wondered what he thought of the book. Had he read it? Would he recommend it? He responded (well before I had managed to finish reading of Julius King's tricks), saying that while *An Accidental Man* was a workmanlike novel (or words to that effect), he personally much preferred a different Murdoch book, *A Fairly Honourable Defeat.*

Stanford again and again, 1979-83

In the late 1970's, my continuing interest in generalizing first-order logic led me to the semantics of natural language. This has led, after some twists and turns, to the work I am doing on information theory, constraints, and heterogeneous inference. My idea then was that the roots of all the concepts we use in logic and mathematics, the concepts we had been trying to formalize in generalized recursion theory and model theory, are ultimately grounded in mechanisms that can be found in ordinary language, which in turn needs an analysis that sees meaning as part of the natural order. As mathematicians, we were only making some of these mechanisms more rigorous. As logicians, we tend to make them more formal. But to really *understand* them, I felt, logic should return to its roots and make sense of how meaning and inference function in the languages people actually use. (Today, I would modify this by also stressing other forms of representation people use in real life to solve problems and to communicate information.)

This shift of my own interests toward less traditional parts of logic happily coincided with an invitation to join the Stanford Philosophy Department in 1979. Kreisel was still at Stanford, but he was (and no doubt remains) highly skeptical that this would be a fruitful line of logical research. As a result, we found ourselves going different ways, with little work of mutual interest. But my office was just around the corner from his and we remained on friendly terms.

That spring, Kreisel gave me the first draft of a paper he had just written, on Gödel, if I remember correctly. It is not an exaggeration to say that I was astounded by this paper. As a graduate student, I had not been privy to any of his first drafts. Kreisel's published papers are as famous as his letters, but for entirely different reasons. They are notorious for the difficulty they present the reader. But this was wonderful: clear and absolutely compelling, unlike anything of his I had ever read. I told him so, including a few minor suggestions for form's sake.

Imagine my dismay when he gave me a copy of the final draft. The paper had become "typical Kreisel," full of long asides, footnotes, convoluted qualifications, and modifiers – all the things which make his articles so impenetrable. My dismay was due not only to the final paper itself, or even to the thought that my own quibbles had played a small role in the change, as to something else. What I found most upsetting was the thought that his other papers, which had cost me such pains to understand, may have had such a life history.

When I saw him next, Kreisel asked me how I liked the final version. I told him that the first version had been far easier to read. To my surprise, he was not at all upset. Indeed, he seemed almost pleased. It was almost as if I had caught him out in a mischievous prank.

I left Stanford in 1981, returning to Wisconsin, but then returned to Stanford in 1983 to direct the new Center for the Study of Language and Information. Kreisel's decision to take early retirement made my rejoining the Stanford Philosophy Department possible, since it freed up a tenure slot. My final dealings with him were just that, a "deal." I purchased his old car for $600. When I got the car home and opened the trunk I discovered he had left behind a dusty tome called *A Short Introduction to the History of Human Stupidity.* I suppose it says a lot about my relation with him that only now, ten years later, does it occur to me that Kreisel might have simply forgotten or misplaced the book.

Silence, 1983-present

I have neither seen nor heard from Kreisel since 1983. Initially, I heard rumors that he was not particularly happy in Austria. This worried me since I felt implicated in his decision to retire early. But I knew that he had no particular interest in the work I was engaged in. Later, I heard he had moved to Oxford and was his old self. Still, I have never written to him or heard from him.

It might seem from this that Kreisel's influence on me and my work ended with the end of our mutual interests in logical matters over a decade ago. But in fact, this is not the case. If anything, his influence has assumed an increased importance. But while the vignettes presented above may give a sense of my relationship to Kreisel, and my work's relation to his, they only hint at Kreisel's real impact on me. Even less do the provide an explanation for his impact on logic. For these purposes, I must resort to metaphor.

To me, Kreisel crackled with life like an oak fire. The fire that warms the room, cooks our food, or helps us read a book also sends out dangerous sparks and attracts unwary moths to their end. Just so, Kreisel could be illuminating, attractive, helpful, but also potentially dangerous. His pointed remarks were like sparks hitting whom or what they would, illuminating things you have never seen before, whether they be about mathematics, logic, people around you, or yourself. And sometimes they left scars. Like fire, Kreisel could have a devastating effect on those who got too near. I saw it happen more than once. It is Kreisel's personal fiery quality, something that goes beyond the record by which we usually judge our colleagues, that accounts for his influence on logicians, and so on logic.

If some got burned, for me he had a particularly liberating effect, one that has gone unacknowledged. There is a mystique in mathematical circles that identifies personal worth with intelligence, and intelligence with the difficulty of the proofs one has given. Kreisel's brilliance, coupled with his fierce scorn for thoughtless pyrotechnicalities (his word perhaps), helped me get beyond this crippling mystique. It gave me and gives me the heart to pursue what matters to me, at my own pace, without worrying too much about whether the theorems it leads to are profound, the proofs easy or hard.

This freedom was Kreisel's gift to me. If I had not been exposed to him, I might well have proved more theorems, but even fewer would have been the "right" theorems. They would certainly not have been proven for the right reasons. Of course, given the direction of my current work, Kreisel may well feel I squandered the gift. If so, I suspect my belated recognition of it will only make him smile.

References

Barwise, J.

[1975] *Admissible Sets and Structures*, New York, Springer, 1975.

[1967] *Infinitary Logic and Admissible Sets*, Ph.D. Diss., Stanford U, 1967.

Barwise, J., and Etchemendy, J.

[1992] *The Language of First-Order Logic*, 2nd ed., Stanford, CSLI, 1992.

Eliot, T. S.

[1958] *The Complete Poems and Plays, 1909-1950*, Boston, Harcourt, 1958.

Friedman, H., and Jensen, R.

[1968] A Note on Countable Admissible Ordinals, in *The Syntax and Semantics of Infinitary Logic*, Ed. J. Barwise, Springer Lecture Notes in Mathematics, New York, Springer, 1968, pp. 77–79.

Kreisel, G.

[1960] Set-Theoretic Problems Suggested by the Notion of Potential Infinity, in *Infinitistic Methods*, Oxford, Pergammon Press, 1960, pp. 103–140.

[1965] Model-Theoretic Invariants: Applications to Recursive and Hyperarithmetic Operations, in *The Theory of Models*, Ed. J. Addison et al., Amsterdam, North Holland, 1965, pp. 190–205.

Kunen, K.

[1968] Implicit Definability in Infinitary Languages, *Journal of Symbolic Logic*, 33 (1968) 446–451.

Lopez-Escobar, K.

[1965] An Interpolation Theorem for Denumerably Long Sentences, *Fundamenta Mathematica*, 57 (1965) 253–300.

Murdoch, I.

[1970] *A Fairly Honourable Defeat*, New York, Viking, 1970.

[1971] *An Accidental Man*, New York, Viking, 1971.

Pitkin, W.

[1932] *A Short Introduction to the History of Human Stupidity*, New York, Simon, 1932.

Platek, R.

[1966] *Foundations of Recursion Theory*, Ph.D. Diss. and Suppl., Stanford U, 1966.

Georg Kreisel: a Few Personal Recollections

by Francis Crick

I first met Georg Kreisel during the middle of the second World War. At that time I was working as a civilian scientist for the British Admiralty in a converted country house called West Leigh near the town of Havant on the south coast of England, not far from the large naval base at Portsmouth. My job, as a member of the Department of Torpedoes and Mines, was to devise new mine circuits for non-contact minesthe type laid by aircraft in shallow enemy watersthat could be got into production as quickly as possible.

There was a small cafeteria at West Leigh. There, one evening when I was working late, I found myself at dinner with a physical chemist and a young man I had not noticed before. The physical chemist was holding forth and gradually I sensed that both the young man and I did not think much of what he was saying. This was the beginning of my long friendship with Kreisel – I call him 'Kreisel' and not 'Georg' because in our family he is always so referred to, and indeed we still address each other in our correspondence in this formal way. I notice that I even pronounce his name in imitation of the way Kreisel does.

It was only later that I learnt something of his family background. At that time I knew that he had taken a war-time degree in mathematics while at Trinity College, Cambridge and had then been immediately recruited to do his war service in the Admiralty. He must have been about 19 or 20 at the time. Why he was sent to West Leigh I don't know. The man at the Admiralty who was then responsible for such appointments (Brundret) had been trained as a mathematician. The head of West Leigh – Edward Collingwood, later Sir Edward – was also a mathematician and a Trinity man. Even in those days Kreisel was sufficiently unusual that Brundret may have thought that Kreisel had better work under someone who was used to brilliant but unconventional young men.

I never did discover exactly what Kreisel was doing at West Leigh. There was a war on and I had plenty of urgent problems of my own to deal with. I believe that one of his first efforts was to apply the methods of Wittgenstein (whom he had known at Trinity) to the problem of mining the Baltic. I would dearly like to read this paper but Collingwood was believed to have locked it away in his safe (could it have been a threat to security?) and it has almost certainly not survived.

After a time Collingwood must have decided that there were not enough suitable problems for Kreisel's talents at West Leigh and he was transferred to the Department of Miscellaneous Weapons, housed in Fanum House near Leicester Square in the center of London. Here he was on tap to give mathematical and scientific advice to the naval officers and others working to develope radically new equipment. "It's funny," Kreisel said to me, "but they don't seem to have any feeling for which problems are soluble. They come to me, saying they have an easy problem and apologise for bothering me with it, but it turns out to be quite intractable. Or they may say that some other problem is too hopelessly complicated but it may only need the application of the conservation of momentum to solve it easily."

Kreisel was involved with calculating the effects of waves on the floating (Mulberry) harbors being designed for the Normandy landing. He even supervised some hydrodynamic experiments done, as I recall, in a large tank at Imperial College. The motion of the water was made apparent by having fish eggs in the water. (Not surprisingly, after a time they went bad.)

His stay at Fanum House produced a major influence on Kreisel. Also in the Department was another eccentric, an ex-naval officer (he had earlier been divorced out of the navy) who had become a painter and was employed to advise about camouflage. Whoever was in charge of allocating offices in Fanum House must have come in one morning and had a bright idea: why not put both these unusual characters together in one room?

The camouflage expert was an enthusiast. His two major enthusiasms, apart from painting, were food and sex. He tended to talk, at length, about both subjects in somewhat the similar terms. This opened new worlds to Kreisel. I suspect this experience was at the root of Kreisel's subsequent interest in cooking.

In spite of my being at Havant, while Kreisel was now in London, Kreisel went out of his way to keep in touch with me while I, on my part, was only too pleased to continue seeing him. Kreisel seldom spoke about his family background, so I was surprised when on one occasion he introduced me briefly to his brother. It was only later that I learnt that Kreisel had come to England as a refugee and had been sponsored by Stanley Baldwin. The only anecdote I recall about his boyhood in Austria concerned his holding their housemaid out of a window by her ankles – he must have been quite muscular at that time.

After the war Kreisel returned to Trinity as a research student and switched his interest to mathematical logic. He renewed his acquaintance with Ludwig Wittgenstein who also had rooms in Trinity. Wittgenstein did not dine at High Table. He said he had dined there once but "the conversation was so *awful*" that he never dined there again. He had chosen rooms in Neville's Court as far from the center of Trinity as possible and usually had meals at the cafe above the Regal Cinema. On one occasion Kreisel and I accompanied him there. Wittgenstein and Kreisel used to go for a walk together once a week. Kreisel must have softened his customary manner with The Master, since neither of them was easy to get on with. I once asked Kreisel if they ever quarrelled. "There was once a certain coolness" he admitted. That they remained on speaking terms says a lot for Kreisel's tact.

In 1947 I myself left the Admiralty and went to Cambridge, so I saw rather more of Kreisel. I don't recall that then, or since, we ever talked much about what he himself was doing, mainly because the foundations of mathematics have always been rather beyond me. Kreisel, on the other hand, took pains to follow what interested me. He also encouraged me, while still at the Admiralty, to take the plunge into scientific research, rather than stay on at the Admiralty or go into industrial work. He helped me once with a mathematic problem – the Fourier transform of a coiled-coil – on which I was stuck. Kreisel was not the only person I knew with considerable intellectual powers, but he was the only one I could talk to at length about many subjects. When I met him I was a rather sloppy thinker with a taste for wit and paradoxes in the style of Oscar Wilde. Kreisel would tactfully but sternly rebuke me for any careless thinking so that under his influence my ideas became more logical and better organised.

Sometime during this period I took Kreisel home with me to visit my mother, then living at Weston Favell, a village close to Northampton. My parents were not in the least academic – my father had been a boot and shoe manufacturer – but they were used to my having unusual friends. During the weekend Kreisel developed 'flu. His pallid, unshaven face above the bedclothes made it plain that he could not be moved for a few days. There was nothing for it but to leave him there while I returned to work. My mother was visibly apprehensive at having to take care of this alien creature. Fortunately all went well. When I returned the next weekend they had established an easy relationship. Kreisel dried the dishes while my mother washed them. My mother was impressed by the way Kreisel, without saying a word, handed her back any plate that had not been perfectly washed. Fortunately they had found a common topic of interest – the activities of the human digestive system – that they could discuss together at some length.

Kreisel had hoped to get a Trinity Research Fellowship which would have supported him, with no duties, for several years but the electors though otherwise. There is a story (I cannot vouch for it) that one of the referees had dropped the manuscript of Kreisel's thesis and was unable to put it together again in the correct order because the pages were unnumbered. Kreisel had to find a job. For a brief period he taught the boys at an English boarding school. He told me they referred to him as "The Pwoblem Child" because he would frequently say in class, "What is the pwoblem?". He applied for a job at Reading University. I recall talking to him after he returned from the interview. "How did it go?" I asked. After a pause he replied, "By the end they were thinking before they asked a question." In spite of this they elected him and he moved to a suburban house in Reading.

At Reading a surprising development took place. Kreisel took up cooking. A mathematician, and especially one working on mathematical foundations, cannot easily apply himself to his research for a steady eight hours a day. I believe Kreisel's routine at Cambridge was to work for a few hours in the mornings and then again for a period between tea and dinner. This would leave anyone with time on their hands, unless they got into administration of one sort or another, something Kreisel has always avoided. At Cambridge he claimed that, for a time, he went to watch wild-west films, sitting in the front row with all the small boys. (Wittgenstein is also supposed to have done this.) At Reading he used to shop in the afternoon and then, after a spell of work, he would cook himself a three or four course dinner. He soon became a really excellent cook and, not surprisingly (since he never took vigorous exercise) he put on about 40 pounds weight.

After that we would employ Kreisel as a guest cook (though my wife is also a good cook) whenever he paid a visit to Cambridge. Kreisel's method of cooking had several unusual features. He usually preferred to cook stripped to the waist, displaying a complex pattern of black hair that covered his upper body. He would only cook one dish at a time. This meant that there would be a prolonged pause after each course while Kreisel cooked the next one. Not surprisingly the meal usually lasted for three or even four hours. It is remarkable how much one can eat without discomfort if there are gaps between the courses to allow time for digestion. In his later years Kreisel has almost abandoned cooking, finding that fat disturbs his digestive system. He restricts what he eats and as a result has become slim again.

Kreisel, with his European background and his easy command of several languages, has always enjoyed travel in Europe. We would occasionally meet him abroad but I only spent a holiday in Europe with him on one occasion. This was in Switzerland in the summer of 1955. We stayed at a hotel not far from Interlaken and made excursions to nearby places. In one of these we went up the mountain railway to the Jungfrau. Kreisel at that time was carrying round with him, for medical reasons, an annular rubber cushion. Standing on the glacier in a dark overcoat, with his umbrella in one hand and the cushion on the other, he made a striking contrast both to the white of the snow and the other visitors on the glacier.

One day I went into my office at the Cavendish (the Cambridge University Physics Laboratory) and was surprised to see a letter for me with a Spanish postmark. Inside was a picture of Kreisel, beaming away between two rather good-looking young men. I do not read Spanish but I could understand enough to tell me that the letter, although addressed to me by name, was really for Kreisel. When I saw him next I tackled him about it. He was quite unabashed. “When I travel,” he said, “I often use your name.” Being good friends I was only amused by this little habit of his.

Kreisel had several general rules about travel. The main one, which to this day is still known in our family as “Kreisel’s Rule” was: never start traveling before 10 a.m. I believe this to be a good rule and try to follow it whenever I can although all too often the airline schedules make this difficult, if not impossible. His other rules I have found less useful, as they arise from his difficulty in sleeping. As is well known he always travels with black-out material so that he can curtain the windows of his temporary bedroom to exclude all traces of light. He advised me always to stay at expensive hotels, with deep carpets in the corridors which reduced the sound of the footsteps outside the bedroom.

Part of Kreisel’s troubles in sleeping are probably due to his auditory system. He told me that he was quite unable to recognise any tune, even “God Save the Queen” which in those days was often played in theatres and cinemas. Many years later, at our house in Cambridge, we were having lunch with friends of ours, one a neuroscientist and the other an acoustic engineer. When Kreisel described his symptoms we wondered if perhaps his ears emitted sound. This may sound far-fetched but it has since been shown that this can happen. We never did carry out the tests but I suspect that Kreisel does suffer from a form of tinnitus, as it is called.

One other occasion when we traveled together was our trip to Hollywood. I think Kreisel was then at Stanford. Shortly after we arrived we visited a friend of his, an actress, then "resting" between jobs. "My God," she said, "You look just like two income tax inspectors." This was because we had taken no thought about our clothes and were wearing the rather scruffy apparel which is almost *de rigueur* among young scientists and mathematicians. The next day we tried to rearrange our clothing so that we could take on a little of the local colouring but with our limited wardrobe – it never occurred to us to buy some new clothes – we still failed to pass as natives. In spite of this we were invited to one Hollywood party and then a small social gathering in a house in Beverly Hills. It was a pleasant afternoon (as it often is in Southern California) and some of the guests started to play croquet on the lawn. Croquet, although superficially a quiet, relaxed sort of game, produces rather violent emotions in many people, probably because a player is allowed to hit not only his own ball, but, indirectly, the balls of other players too. One stands helpless while one's well-situated ball is vigorously despatched to some quite unfavourable part of the lawn.

Faced with this, Kreisel adopted a quite novel strategy. He simply hit his ball at random. One might have thought that his opponents would welcome this behaviour, since it would increase their chances of winning, but this was not the case. They were furious, probably because Kreisel was disturbing what one might call "the logic of the game". Kreisel seemed quite unmoved by the tension he was creating and continued to play at random.

At one period Kreisel became interested in European society life. As he explained to me, such people are easily bored by the usual social round, and so welcome anyone who is both unusual and entertaining. Kreisel's very logical and somewhat ponderous way of talking – lightened from time to time with flashes of wit and malice – was a novelty, especially when he applied it to an unusually frank discussion of interpersonal relationships. He would occasionally go on holiday with a group of them to fashionable places but he was not completely accepted as one of them, in spite of his dressing somewhat better than when he was a young man. He told me that, at one hotel, he overhead the following conversation between two society ladies:

> "My dear, have you seen that extraordinary looking man?"
>
> "You mean the one who looks as if he might be a mathematician?"
>
> "Just so!"
>
> "My dear, he *is* a mathematician."

In recent years Kreisel seems to prefer a more academic environment, such as All Soul's, Oxford, though he still sees a few old friends from the outside world.

Kreisel is an excellent letter writer. I have kept all his letters to me. They are unlikely to be published in our life-times, because of their frankness, but they should make amusing reading for posterity. His earlier letters – often written on borrowed note paper – are quite short but very pithy and amusing. More recently they have become rather longer and more convoluted, but this is partly because he has had to explain to me some rather elementary points related to mathematical foundations – why he thinks Roger Penrose's arguments about the brain and computability are wrong, for example. A surprising feature to me is the number of slips he makes. He seems even more careless than I am in writing down algebraical expressions. But from what Teddy ("Sir Edward") Bullard, the geophysicist, told me, Kreisel made many slips in a joint mathematical paper they wrote together when Kreisel was a young man, so I suspect it is a habit of mind and not just a matter of age.

Kreisel is unique among my close friends in many ways but one deserves special mention. He has few worldly possessions. He does not own a house or a car – though at one time he used to drive a car. He has few clothes and, surprisingly, very few books and papers. He can easily transport most of his personal effects on a plane. Since he took early retirement from the Philosophy Department at Stanford he has spent most of his time in one place or another in Europe. Sometimes he finds a temporary niche at Oxford; sometimes in the visiting quarters of a monastery. Between seasons he may spend a few weeks at some hotel or other. He once told me that his ideal place to live in would be a quiet luxury hotel, fully staffed, but with no other guests!

I have urged him to buy a small house somewhere, if only to have a *pied-à-terre*, but to no avail. Perhaps some intelligent and long-suffering lady will turn up to take care of him in his old age. One would like to think of him mellowing gracefully in the Italian sunshine, having delightful (but low-fat) meals and talking at length to occasional guests (suitably lodged in a separate guest house) but I fear he will be unable to abandon his wandering way of life. I have known him now for about fifty years. Over that time I have been immensely influenced by his powerful intellect. We have many interesting and enjoyable times together. If I had never met him my life would have been very different. It is with deep gratitude and affection that I salute him on his 70th birthday.

Kreisel's Effectiveness

by John N. Crossley

> "I regard mathematical logic as *applied mathematics*, absolutely parallel to theoretical physics except that the correctness of a definition or an axiom is not judged by 'observation' but by introspection (as, indeed, physicists judge the plausibility of a theory)."

Kreisel wrote this to me in a letter from Stanford dated December 8, 1970. Letter-writing has long been Kreisel's forte, not mine. At the end of this paper I shall return to analyze, a little, the above quotation but I now present a case study of development in one small area of logic over a period of thirty years in which Kreisel has had a significant and rather varied input. The technical details I shall try to keep to a minimum, though the references should supply them all.

Cardinal and ordinal numbers, even more than set theory in general, seem to captivate students; they did me. However, I developed a strong interest in recursion theory and the question my adviser, Ken Gravett, posed, was to provide a constructive theory of cardinal and ordinal numbers. At that time (1960) communications were not as good as now and Oxford was an outpost as far as logic was concerned, just becoming noticeable because of Higman's work [1961] on the word problem for groups and not yet having a mathematical logic course.

The problem seemed intractable: after all, the interesting parts of cardinals and ordinals involve uncountable sets which are anathema to recursion theory. John Shepherdson (in the Summer of 1961) told me that Dekker and Myhill [1960] had already solved the problem. I got hold of the book and from it was led, eventually, to Tarski's two books:

Cardinal Algebras [1949] and his later work with C.C. Chang and B. Jonsson [1965], *Ordinal Algebras*.

It was clear that Dekker and Myhill had dealt with cardinal numbers: ordinals apparently had not been tackled. (In fact I subsequently discovered that Alfred B. Manaster and I had started work on the problem about the same time.) But there was the problem of how to give the basic definition of two orderings having the same *constructive* order type.

Kreisel was, at that time (1962) much in evidence at Oxford. I was attending and enjoying Michael Dummett's lectures on what my fellow listeners and I then regarded as the rather peculiar mathematics of Brouwer and his followers (see Dummett [1977]). At least it seemed peculiar in the 1960s. Michael appeared to this young logic student as Kreisel's main focus and it was many years before I learnt of Kreisel's wider range of contacts (see Watson [1968], p. 146 which records "the correct equations fell out, partly thanks to the help of Kreisel").

Michael felt that I should talk to Kreisel and arranged for me to meet Kreisel in his (Michael's) rooms at All Souls – the Oxford College with no students and a forbidding sign which to this day says that visitors are only allowed in the college between 2 and 5 p.m.

I went up the stairs and into Michael's room. I did very little of the talking. The next hour was essentially a monologue.

After that I went away and wrote my thesis.

While the previous sentence is true, there are two remarks that should be made. The first is that I had understood only a small fragment of what Kreisel had said. I do not think there is anything unusual in that. The second is that the reason I was able to write my thesis was that I now had a definition: the basic definition of what it means for two ordered sets to be constructively isomorphic.

Again there are two remarks to make. The first is that the definition I came up with was only one among several possibilities. The other is a remark made to me only last year by John Shepherdson: That mathematics is remarkably robust in that it seems to work well despite errors by its practitioners and variations in definitions.

In the case of a constructive version of the ordinals, the issues can be spelt out easily.

Dekker and Myhill had shown that the constructive theory of cardinals for recursive and recursively enumerable (r.e.) sets is boring: any two r.e. sets of the same cardinality are recursively equivalent. The proof is simple. If A, B are such sets where

$$A = \{a(1), a(2), \ldots, a(n), \ldots\} \text{ and } B = \{b(1), b(2), \ldots, b(n), \ldots\}$$

for some recursive functions a, b then map $a(n)$ to $b(n)$.

So the interesting sets are those which are not r.e.

For ordinals we also have to consider the order.

However, given a non-r.e. set A, two questions arise: for distinct elements a, b, first: are $a, b \in A$?; and secondly: is $a \leq b$ or $b \leq a$ in the ordering? Of course one then asks about constructivity. Any difficulties there may be are caused by elements which are *outside* A. For, in the case of cardinals, Dekker and Myhill require A to be contained in a r.e. set A' and inthe ordinal case one has to consider the ordered set $(A, \leq)$ and ask what happens to $\leq$ on A'.

I can now give the answer to the question of how to deal with the elements outside A, though I could not in the sixties. That is that it makes very little difference to the theory – provided that the elements in A' but not in A "do not interfere". This was first pointed out to us by Sol Feferman in a letter in 1963. Here by "us" I mean Parikh and myself, for we wrote an abstract (Crossley and Parikh [1963]) which opened up a route to constructing ordinals in a natural way.

So now I must discuss "natural well-orderings". These days I feel it is a bit like discussing natural scales in music: to the classically educated musician the ordinary tonic sol-fa suffices; to the twelve-tone composer Schönberg's full use of both black and white notes is "natural", provided certain proprieties are observed (Haimo [1990]) and to the Chinese the five tone octave is "natural". In the sixties, and, I believe, before that time, ordinal numbers such as ω (the order type of the natural numbers), $\omega^2, \omega^3, \ldots$ even $\epsilon_0 = \lim_{n\to\infty} \omega^n$ had "natural" representations. What was special? One obvious answer was that representations were easily *constructed*, and I use the word 'constructed' on purpose because, using the representations, one can easily give constructions which are readily seen to be recursive.

The study of constructive order types (*viz.* equivalence classes of well-orderings under recursive order-preserving functions, see Crossley [1965]) showed that there were a number of constructive functional constructions which yielded (equivalence classes of) these natural representations.

I pursued this vigorously (see e.g. Aczel and Crossley [1967], and for the latest view on the subject see Crossley and Kister [1986/87]); so did Feferman. He, however, seemed much more influenced by Kreisel. I suppose I was searching for some beautiful intrinsic property (Hardy [1904] and Veblen [1908] had also failed in this endeavour); Feferman [1968] gave concrete rules for "nice" systems of functions which embraced Veblen's and went further. This was a solution of the problem of natural well-orderings in the applied tradition of my opening quotation.

I was not satisfied, but I was by that time involved in developing further the ordinal analogue of Dekker and Myhill's work. Nerode [1961] had produced a far-reaching extension of Dekker and Myhill's theory and I began to work on the corresponding theory for COTs (constructive order types). (This ultimately appeared in Crossley and Nerode [1974].)

Kreisel did not approve of my approach to providing a constructive version of the ordinals. His review (Kreisel [1968]) of my thesis (Crossley [1965]) said (in my interpretation): what is the point of this development? I now believe he wanted to know what would make a constructive theory of order types work. And although he claims (in his [1968]) that Brouwer already had "a highly developed constructive theory of ordinals" and refers the reader to Heyting [1961], I believe that anyone trying to understand, and then make use of, Brouwer's work, would have a very difficult task indeed. It simply does not build a grand structure which can compete in elegance with, say, the work of Sierpinski [1965].

Kreisel had advocated the use of axiomatics – what are the axioms for a 'decent' theory of ordinals, how invariant are they when you change the class of functions from classical to, e.g., recursive ones? I did not even know how to approach this. One person who made moves in this sort of direction was Steve Simpson with his programme of 'Reverse Mathematics' who, using powerful independence result techniques, began to analyze exactly what axioms, or even what restrictions of specific types of axiom (e.g. replacement), were required to prove certain key theorems. Such endeavours were certainly beyond me then; they would still be difficult now for reasons which will shortly emerge.

Another move took place, initiated by Dana Scott. Subsequently Dana told me that he felt that the work on COTs had little to do with recursion theory, or at least that the set theory and the recursion theory could be separated. David Charles McCarty in his thesis (see [1986]) achieved that. He developed a part of intuitionistic set theory and then created a realizability model (in the style of Kleene [1945]).

Kleene's idea of realizability was that if one was developing a constructive (intuitionistic) theory, then one should be able to provide a model all of whose functions (and predicates) were (partial) recursive. His initial work showed that every theorem provable in intuitionistic (Heyting) arithmetic was realizable (see Kleene [1952]).

There seemed, at that time, to be an intrinsic restriction of realizability to number theory but Scott and Solovay's approach to building Boolean models (see Bell [1985]) opened up a way to extend the treatment to set theory.

McCarty was able to show that basic results in set theory gave, *via* realizability, the basic results in Dekker and Myhill's work. In particular it follows from McCarty's approach that all the axioms, and therefore also all the theorems, of the set theory are true in the realizability interpretation. Moreover, if we have a theorem of the form $\forall f \exists g \phi(f, g)$, where f, g range over functions, then from the recursive realizability interpretation we have that, if f is a (partial)recursive function, then we can *construct* a (partial) recursivefunction g, such that $\phi(f, g)$ holds. So the axiomatic theory yields direct constructions in the constructive context. For example, if one proves, for ordinals α, β, γ that $\alpha + \beta = \alpha + \gamma$ implies $\beta = \gamma$, then this yields the result that if there is a *recursive* isomorphism from $A + B$ onto $A + C$ where α is the "constructive cardinal" (recursive equivalence type in Dekker and Myhill's terminology) of A, etc., then from *that* f one can construct a *recursive* isomorphism from B onto C.

So at last we had a nice axiomatic theory as Kreisel had desired back in the late sixties.

When I wrote to Kreisel to apprise him of this he wrote back (on 11.v.1984):

> "I should have been MUCH MORE pleased if you (or somebody else) had found an axiomatic treatment for recursive order types 20 years ago."

I do not think that would have been possible. The computations in set theory books (such as Sierpinski's [1965]) are very simple in principle, though they can get very complicated by piling on layers of computation. Secondly, Nerode's work in extending combinatorial functions (definition below) to RETs (recursive equivalence types – the constructive analogues of cardinals) was very complicated and even though Nerode claims that his negative results (Nerode [1961]) correspond to independence results in (intuitionistic) set theory, no-one has yet written this out. I think we are, at last, in a position to clarify these issues but a planned paper (Crossley [1993]) is yet to be written. I think I can now describe the situation, but first the definition of combinatorial function.

Any function from the natural numbers to natural numbers can be written in the form $f(n) = \sum_{i \leq n} c_i \begin{pmatrix} n \\ i \end{pmatrix}$ where $\begin{pmatrix} n \\ i \end{pmatrix}$ is the combinatorial coefficient $\dfrac{n!}{(n-i)!\, i!}$. Such a functionis said to be *combinatorial* if every $c_i \geq 0$ and only finitely many $c_i \neq 0$.

So now the idea is as follows:

Take intuitionistic set theory – which basically means the classical axioms interpreted intuitionistically – though it also seems one has to include explicit constructions, represented by terms, such as $\langle x, y \rangle$ for ordered pair and $\{x : \phi(x)\}$ for a set given by a replacement axiom. Secondly, give a realizability interpretation.

Here McCarty's, and indeed Kleene's, versions of realizability appear inadequate. The difficulty seems to be with negation, because under either interpretation there are classically false formulae which are realizable (though, as pointed out above, all theorems are indeed realizable as one hopes). The next task is then to extract, *via* realizability, the appropriate constructive functions to give constructive analogues of classical results and the final task is explicitly to show that where this is impossible we have independent results for intuitionistic set theory (or at least second order logic).

These appear non-trivial tasks and although a start has been made much more remains to be done. For example, it turns out on deeper analysis that it is not negation which is the major problem, though it is still a problem, so much as the definition of realizability for closed, atomic formulae. That these apparently innocent formulae – as simple as $2 + 2 = 4$ – should be so troublesome was first discovered in writing Crossley and Remmel [1994].

In my [1988] I crowed that McCarty's work had shown the approach of Crossley [1965] was correct. I now believe that was premature but I think it was on the right track *if* what you want is to find out how ordinal numbers work in a constructive context.

It is not clear that that is what Kreisel wanted. I have always liked grand, elegant theories – such as Galois theory. So let me now return to the opening quotation and see how it looks analyzed in the light of the above developments.

Kreisel's characterization of mathematical logic as "applied mathematics" in 1970 has now become a generally accepted maxim. When I was a student many people were still looking for a "foundation" for mathematics. Kreisel simply asks how things work. Moreover, he is prepared to accept any, or many, reasonable explanation(s).

This did not sit well with the impression he gave in Oxford in the sixties where constructivity appeared paramount, but, of course, if you really know how things work you can carry out the appropriate constructions.

Next, the parallel with theoretical physics is amusing and provocative. Theoretical physics writ small just backs up mundane experiments but writ large, as most recently in Stephen Hawking's book [1988], goes far beyond anything which one could remotely call "practical" – and I have always felt that "applied mathematics" ought to (rather than does) have that connotation too. And then, of course, comes Kreisel's parenthetical qualification: "as, indeed, physicists judge the plausibility of a theory", claiming, with justification, let me say, this is how physicists behave too. So mathematical logic is concerned with the "real world". Perhaps.

Yet throughout the sentence Kreisel provokes one into thinking, and tantalizes us all with the hope of making real progress in understanding the world. I think that is what appealed to me.

Of course such "real progress" is a myth – in the true sense of the word – and very powerful. What Kreisel has managed to do is to disseminate that myth– not least by his countless letters. (Most of those I received in the sixties I have since, regrettably, destroyed.) He has not been too precisein his directions, though he has often asked absolutely specific and technical questions, yet even these have led to more, interesting, problems and he has always led one deeper and deeper into areas of interest, dangling new layers as he dangles dependent clauses (and parenthetical reservations) in sub-section after sub-section of his writing.

I used to try to find the point, the centre, but that was a mistake. As in the case of the Holy Grail, it is not the goal that is important, but the quest. For Kreisel's pushing along the path I am very grateful.

Bibliography

Aczel, P.H.G., and Crossley, J.N.

[1967] Constructive Order Types, III, *Archiv für mathematische Logik und Grundlagenforschung*, 9 (1967) 112–116.

Bell, J.L.

[1985] Boolean-Valued Models and Independence Proofs, in *Set Theory*, 2nd ed., Oxford.

Crossley, J.N.

[1965] Constructive Order Types, I, in *Formal Systems and Recursive Functions*, ed. J.N. Crossley & M.A.E. Dummett, Amsterdam, 1965, pp. 189–264.

[1988] Fifty years of computability, *Southeast Asian Bulletin of Mathematics*, 11 (1988) 81–99.

[1993] Using second order logic (in preparation). Abstract submitted to *AMS*, Washington, D.C. meeting April 1993.

Crossley, J.N., and Kister, J.E.

[1986/87] Natural well-orderings, *Archiv für mathematische Logik und Grundlagenforschung*, 26 (1986/87) 57–76.

Crossley, J.N., and Nerode, A.

[1974] *Combinatorial Functors*, Springer, Berlin, 1974.

Crossley, J.N., and Parikh R.J.

[1963] On isomorphisms of recursive well-orderings (abstract). *Journal of Symbolic Logic*, 28 (1963) 308. (See correction in *J.S.L.*, 31 (1966) 525–538.)

Crossley, J.N., and Remmel, J.B.

[1994] Proofs, Programs and Run-times, *Methods of Logic in Computer Science*, 1 (1994) 183–215.

Dekker, J.C.E., and Myhill, J.

[1960] *Recursive Equivalence Types*, University of California publications in mathematics, 1960, n.s. 3, pp. 67–214.

Dummett, M.A.E.

[1977] *Elements of Intuitionism*, Oxford, 1977.

Feferman, S.

[1968] Systems of Predicative Analysis, II. Representations of ordinals. *Journal of Symbolic Logic*, 33 (1968) 193–220.

Haimo, E.

[1990] *Schönberg's Serial Odyssey: the evolution of his twelve-tone method, 1914-1928*, Oxford, 1990.

Hardy, G.H.

[1904] A theorem concerning the infinite cardinal numbers, *Quarterly Journal of Mathematics*, 35 (1904) 87–94.

Hawking, S.

[1988] *A Brief History of Time*, Toronto, 1988.

Heyting, A.

[1961] Infinitistic methods from a finitist point of view, in *Infinitistic Methods*, Warsaw, 1961, pp. 185–192.

Higman, G.

[1961] Subgroups of finitely presented groups, *Proceedings of the Royal Society*, Series A, 262 (1961) 455–475.

Kleene, S.C.

[1945] On the interpretation of Intuitionistic Number Theory, *Journal of Symbolic Logic*, 10 (1945) 109–124.

[1952] *Introduction to Metamathematics*, Amsterdam, 1952.

Kreisel, G.
[1968] Review of Crossley [1965], *Zentralblatt für Mathematik und ihrer Grenzgebiete*, 145 (1965) 10–12.

McCarty, [D.] C.
[1986] Realizability and recursive set theory, *Annals of Pure and Applied Logic*, 32 (1986) 153–183.

Nerode, A.
[1961] Extensions to Isols, *Annals of Mathematics*, 73 (1961) 362–403.

Sierpinski, W.
[1965] *Cardinal and Ordinal Numbers* (2nd rev. ed.), Warsaw, 1965.

Tarski, A.
[1949] *Cardinal Algebras*, New York, 1949.

Tarski, A., Jonsson, B., and Chang, C.C.
[1965] *Ordinal Algebras*, Amsterdam, 1965.

Veblen, O.
[1908] Continuous increasing functions of finite and transfinite ordinals, *Transactions of the American Mathematical Society*, 9 (1908) 280–292.

Watson, J.D.
[1968] *The Double Helix: A Personal Account of the Discovery of the Structure of DNA*, London, 1968.

Kreisel on the Telephone: an Appreciation

by Anita Burdman Feferman

> "I learned how to assess people, a rather unpleasant pastime since it inevitably leads to disillusionment."
>
> Dimitri Shostakovich

I am blessed – some might say cursed – with a memory for detail, and so I remember my first encounter with Georg Kreisel. It was on the telephone, in the spring of 1956.

"I am terribly sorry to disturb you," the caller said, "but I wonder if I might speak with Professor Feferman?" "Professor Feferman" was, at that time, one of Alfred Tarski's graduate students, struggling to complete his PhD thesis, at the University of California in Berkeley. I was – and still am – his wife, soon to be labeled "*Frau Professor Doktor* Feferman." We were wet behind the ears, which is to say unsophisticated.

The caller's elaborate apology, his unusual hybrid accent, and the status he accorded my young husband made me jump to the conclusion that it was my father I was talking to and that he was trying to be funny. I decided to be funny, too. Speaking slowly and aiming for the same haughty tone as the jester, exaggerating the sound of each “t”, I said, “No, he isn’t at home at the moment. Would you care to leave a message?”

Repeating how sorry he was to intrude, the man declined to state his business and would not even leave his name. “No, no,” he said, “it is not important. It is a completely trivial matter.”

“Trivial?” Oh dear! That’s *not* my father, I realized, much embarrassed. For my father, no phone call to me–or indeed anyone else–was ever trivial. Moreover the word signaled the caller’s true identity. Only weeks before, Sol and I had heard from Dana Scott that a certain Georg Kreisel was to be visiting Berkeley and that “trivial” was one of his trademark adjectives, his *leitmotiv*. In fact, as we ourselves were to discover, he made such a *tsimmes* about it that for years the word evoked the man.

A day or two after the now not-so-mysterious phone call, Kreisel found Feferman at the university. They talked and in no time a “romance” was kindled, the dazzling older scholar making an enormous impression upon the young man. Eager to expand his logical horizons, the budding Ph.D. hung on this new mentor’s every non-trivial word. Likewise, the worldly Kreisel smiled and nodded approvingly at Feferman’s insights and quick appreciations. Almost instantaneously, they were in sync.

In September 1956, Feferman moved across the San Francisco Bay from Berkeley to Stanford University where he began his teaching and research career. A few years later, Kreisel, too, was invited to join the faculty at Stanford. There they became the closest of colleagues, seeing each other daily, exchanging ideas, trivial and important, sharing research grants, discussing every aspect of their professional interests and some aspects of their personal lives.

Then Kreisel’s telephone calls flew thick and fast, coming, as often as not, at dinner time: hundreds and hundreds – maybe thousands – of calls over the years. Often they came in clusters: an afterthought of an afterthought, a continuation of the business or gossip of the day. There was no longer any need to ask who was calling. I could tell by the way the phone rang. And although this constant, almost obsessive telephoning was annoying, it was also amusing both at the moment–for no-one was wittier, slyer or subtler than Kreisel at his best–and afterwards, because his provocative assertions always provided fodder for post-mortem discussion.

If I answered the telephone, Kreisel would begin, “Is that Frau Professor Doktor Feferman?” before launching into his “so sorry to disturb you, *Gnädige Frau*" routine which he knew was guaranteed to make me laugh. Sometimes it was indeed me he wished to speak to about practical matters such as where to shop, which dentist or doctor to consult, which auto-mechanic; or, after a bad night's sleep, where he could find a quiet, I mean a *deathly* quiet place to live. Conversations with Kreisel were never short; he would start by asking questions but then he told *you*. These calls became the model for our children's first play conversations on the telephone: "No-oh, I can't do dat now," the three-year old would say into the toy receiver, "I have to talk to Kreiso."

In his days at Stanford, there was always a whiff of scandal surrounding Kreisel which, it seemed to me, he gave off deliberately. His conversation dripped with innuendo whether he was talking about mathematicians and philosophers or dropping names of famous friends and acquaintances in other arenas. He kept company with a changing cast of interesting and beautiful women, and he liked to discuss his preference for young, attractive, delicate individuals, male or female. He made a great to-do of his meetings and friendship with people such as Iris Murdoch, whom he always called Miss Murdoch, emphasizing the "Miss", James Jones, the writer, and his circle of friends in Paris, Peter Duchin, the pianist, and his wife, Cheray, the Duchess of something, and the French Baroness of something else. And then there was his tale of the night he spent with Brigitte Bardot. Was all this true? we wondered. There was evidence that at least some of it was. In any case, while *he* told of his escapades with the rich and famous, his colleagues, young and old, traded stories about *him*. People would actually go around asking each other: have you heard any new Kreisel stories?

In Kreisel's presence, the atmosphere was always charged. Eyes lit, hair standing on end, a not quite full smile, he looked wired, ready to amuse or be amused, ready for mischief. This was a man who hated being bored and did things – odd and unpredictable things – to circumvent the humdrum. Like a carnival ride, it was heady, thrilling, a little out of control, but you came back for more.

He had an odd slant on most matters or, I should say, a slant very different from mine. He was conservative, elitist, called his or anyone else's household help "servants", and frequently made stereotypically disparaging remarks having to do with racial or ethnic origins–remarks, such as, "Isn't that exactly what you would expect from a Brooklyn Jew?" If I said, "How dare you? You're Jewish yourself," that didn't faze him. His disdainful answer would be, "Well of course I am, that's obvious, so it only confirms what I say." His attitude was that these were facts not opinions. He seemed to have none of the restraints or inhibitions most of us have about insulting people face to face or behind their backs, privately or in public. Women he didn't like were "old cows", slow thinkers were "sclerotic"; those he considered hopeless cases were "positively paralytic", and those who repeated themselves were "senile". But he did appreciate quick wit. He would pay someone a high compliment by saying "Ve-y [very] good," in response to a quip or clever argument. When he was in the mood, he could be gracious, and it was sometimes an unmitigated pleasure to sit next to him at a dinner party. He had beautiful hands and elegant table manners, and whether his behavior was polite or appalling, he rarely commanded less than rapt attention.

Professionally, it was exciting and flattering to Feferman to have his work followed so closely, with such interest by the eminent and notoriously critical professor. Even though it was a bit unnerving to entertain the thought that Kreisel's stiletto wit which he saw pointed at others might be turned toward him, Feferman sought Kreisel's critique on almost every single thing he wrote and probably even wrote in anticipation of his approval. In many ways he took up Kreisel's well-developed points of view about logic and the foundations of mathematics and, most particularly, about what kind of work was worth doing and which approaches were worth pursuing.

Although Kreisel did not like large groups, until he and Feferman had their falling out, he was part of our social circle. If he felt inclined to attend a big party, he always came early and was the first to leave. Mostly, we had small dinners together, sometimes at his place (he was a good cook with strong opinions about ingredients and procedures), sometimes at ours, and in a restaurants where he always made a fuss about the wine. At least half of the time, an opened bottle was tasted, rejected and had to be replaced.

Another way we entertained ourselves was by going to the movies. Kreisel loved to laugh almost as much as he liked mocking friends and colleagues. *The General*, a pre-"talkies" Buster Keaton film, simply undid him. We sat in the balcony of the old Varsity theatre in Palo Alto, munching popcorn, and he laughed, a high hee-hee-hee, like a teen-ager gone hysterical and completely out of control. I was amazed.

One summer, aided by Dana Scott, we planned a big party. In our garden, we had a Chinese smoke oven in which large pieces of meat could be cooked. Wanting something exotic that would measure up to Kreisel's haute-cuisine standards, we decided to barbecue a suckling pig. Kreisel was also supposed to be a co-host for the affair but he was noticeably absent during the early preparations. About an hour before the party was to begin, he arrived with a young woman whom he introduced as the London "deb of the year", a pale, frail lovely in a white cotton sundress – strapless as I recall.

Kreisel didn't approve of the way I was slathering the pig with marinade. He took over and treated the animal with a more delicate touch. A lengthy discussion ensued, about how long we should cook what I had come to think of as the unfortunate little creature. Kreisel's notion was that it should be served "rare"; pink, not grey. I, on the other hand, had been instilled with a fear of trichinosis and thought it should be well-done. Sol was on my side but less vociferous. This difference was never resolved. Meanwhile guests began to arrive and the pig was placed in the oven.

The party was a sensation – an event talked about for years afterward. Since it took a long time to get the meat cooked to *any* degree, guests drank and the normal cautions dissolved. A visitor from Harvard leaned just a bit too far back in his chair and crashed into the full length window behind him. Fortunately, in spite of much broken glass, he suffered little injury and the party raged on. Other entertainment included a lovely blond guest (not the English visitor) standing on her head in response to a $10 bet with Kreisel that she couldn't or wouldn't do it right then and there. She would and did. Modestly folding her skirt between her knees, easily inverting herself, she happily collected her winnings.

When the pig was pronounced done by Kreisel, he carved and Sol served. It was pink and juicy. Sol remembers thinking that after all his racing around there wasn't going to be any left for him even to taste. I didn't care either way. I was ready to become a vegetarian.

Later that week, we went to San Francisco for dinner, to Jack's, a classic San Francisco restaurant and one of Kreisel's favorites. We ate in one of the small private rooms upstairs, where, in the good old days, men about town brought their mistresses to dine, away from the prying eyes of the community. The waiter, experienced in these matters, played his role to the hilt. Although I would not have traded places with her, I realized I was jealous of the English deb.

The Kreisel-Feferman honeymoon endured for fifteen years before the relationship began to sour and it became unavoidably evident, even to Feferman who didn't want to see, that Kreisel was actively knocking and disparaging the very work he had inspired his younger colleague to create. Feferman, who had long outgrown the “disciple” role, was angry and hurt. Perhaps he should have been less stunned than he was by this turn of events since he had already observed Kreisel’s scorn for former allies, but naively, because of the closeness and the length of their association, he had had the impression he was exempt.

The divorce was by no means unique; it fit Kreisel’s pattern of behavior towards many promising young scholars: an initial burst of enthusiasm and later, ennui and disillusionment. In fairness, this was also Kreisel’s attitude towards himself. He often turned against his own past work with equal disdain. He had no nostalgia for or attachment to old ideas or old friends. Both, it seems, could easily be dismissed. At first Feferman wondered how he would do without the intensely provocative interchange that he had counted upon as part of the process of inspiration. Seeing Kreisel, dissecting work with him, writing to him was a habit not easily broken, but within two or three years he viewed the rupture as a positive good in terms of his work.

For years, contact between Feferman and Kreisel was at the lowest level of civility. They communicated in writing when it was absolutely necessary, but otherwise avoided one another. Certainly the telephone calls stopped – except for one extraordinary circumstance.

In 1979, on the streets of San Francisco, an armed robber attacked my husband and shot him at point-blank range. A bullet went into his chest, creased his liver, grazed a lung and lodged in a rib. Miraculously he recovered in a few months with only minor after-effects, but in the first weeks no one knew for sure that this would be so. Kreisel telephoned as soon as he heard the news. This time there was no exaggerated politeness. His words were, “Please don’t hang up on me. I have to know how your husband is.”

Kreisel was also among the first to visit Sol in the hospital, but once Sol was well, their icy relations were re-established. This is not as strange as it might seem at first thought. There was no way these men were going to be friends again but nevertheless it is clear that old attachments did mean something to Kreisel when life itself was threatened. For all his cynicism, he is very attached to life.

Coda

My relationship with Kreisel was entirely social and auxiliary to my husband’s. When their friendship ended, mine went the same way. Here too, however, there is an exception. In the mid 1980s, while I was working on the biography *Politics, Logic, and Love: The Life of Jean van Heijenoort*, I asked my then still living eponymous hero (or anti-hero, as some have viewed him) to name the people who influenced him most. He answered, "Trotsky first, of course,"[1] and after a moment's thought he added "and next, Kreisel. Yes, Kreisel. I have very warm feelings for him."

I was only a little surprised at the juxaposition of Trotsky and Kreisel. If ideologically and professionally these men were worlds apart, temperamentally they had much in common. Both brilliant and tireless correspondents, both persuasive by the force of their personalities, both scornful of anything less than the truth as *they* perceived it. In any case, since van Heijenoort had named him as mentor and friend, I felt I would be committing a serious error of omission if I did not solicit Kreisel's impressions and recollections of my subject. A few years after van Heijenoort's death, I wrote to Kreisel, not without trepidation, taking great pains to craft questions that I hoped would elicit serious answers. To my relief and satisfaction, he did indeed respond with observations and suggestions that were valuable precisely because his way of understanding an individual and the world is unlike mine. Over the course of a year, Kreisel's beautifully detailed and evocative letters gave ample evidence that he reciprocated van Heijenoort's warm feelings. I was deeply touched by a kindness I had not previously noticed and this led me to a renewed and rather different appreciation of his *sui generis* genius.

[1] Jean van Heijenoort was Leon Trotsky's secretary, bodyguard, and translator for seven years and for six years after Trotsky's death he was still totally committed to Marxist politics. Only when he was thirty-three did he turn to mathematics, logic, and philosophy as a profession.

Thoughts on the Occasion of Georg Kreisel's 70th Birthday

by Verena Huber-Dyson

Es klingt so prächtig, wenn der Dichter
der Sonne bald, dem Kaiser, sich vergleicht,
doch verbirgt er die traurigen Gesichter,
wenn er in düstern Nächten schleicht.

(Goethe, Weimar am 7. November 1815)[1]

Wenn der Menschengeist sich beschwert oder herabgedrckt fühlt, ... dann wendet er sich wohl gern dem Gebiet der Mathematik zu, in welchem ein deutliches und genaues Erfassen von Gegenständlichkeiten sich findet, und Gewinnung von Einsicht durch angemessene Begriffe in so befriedigender Weise erreicht wird. Hier fühlt der menschliche Geist sich heimisch ...

(Bernays, 1955)[2]

---

[1]It sounds so splendid, when the poet / the sun's, the Kaiser's pose assumes, / yet he is hiding twisted features, / when through troubled nights he stumbles to the dawn.

[2]When feeling burdened or downcast, ... the human mind will gladly turn to the

Kreisel is completing the seventieth year of his fruitful and tortured life. Halfway through this stretch I shared two years of his thoughts, joys and preoccupations to the fullest extent to which any human being can share another's life. They were years of deep foundational turmoil and professional realization for Georg Kreisel. For me it was a time for reflection and realignment that ultimately returned me to a self of deepened awareness.

I am glad of this occasion to express my personal appreciation for the liberating effect of these two years. After a number of weeks engrossed in reminiscences, I have reached the decision to restrict my remarks to a few observations on the interplay between Kreisel's life style and his work, hoping they may shed some light on the process of "foundational" thinking. With this choice I want to underscore my opinion that technical, methodological or philosophical writing is no place for biographical side shows. Anecdotal asides may amuse and illuminate an historical or social context, if conceptually connected with the main text. But the exposition of frailties, physical, mental or emotional, of a protagonist is justified, even in a biography, only if a good case can be made for its significance to the story being told. And that requires a much deeper commitment than any logician so far has been willing, or able to devote to a study of a complex, puzzling or even controversial personality.[3] Moreover, this is my contribution to a "*Festschrift*" not to a character study. But, in response to some editorial comments, I shall add at the end a very few personal data for the benefit of people reluctant to read between the lines or too impatient to gather from the ensuing narrative what they think they want to know about me.

The Surroundings

Why then have I placed two highly personal quotations at the beginning of this essay? Because Goethe's words sum up the inevitable predicament of the creative mind while Bernays, with his gentle, sure touch, puts his finger on the reason why many of us have chosen a life devoted to Mathematics rather than some other craft. But, while the

realms of Mathematics, where a lucid and precise grasp of objectivities is obtained and insight is gained so pleasantly through appropriate concept formation. Here the human spirit feels at home.

[3]Compare Hao Wang's strained attempt at "sticking to (embarrassing) facts" [1987] and Kreisel's innuendo's in an otherwise illuminating analysis [1980] with Rudy Rucker's wonderfully spontaneous account of his meeting with Gödel in [1982]. Rudy is informative on several levels and he really brings Gödel to life! Incidentally, Rudy is an intellectual grand child of Myhill's.

mathematician is fortunate to find refuge from human nightmares in the realm of Bernay's solace, if he has a philosopher's turn of mind, he will soon be thrown to worse terrors. Grappling with the Foundations of Mathematics not only requires high powers of abstraction but also a relentless commitment to introspection, an activity that easily spreads to all aspects of life and may well breed ulcers, insomnia and other diffuse ailments, if not counterbalanced by a vigorous out going life style. Some may have so natural a need for expansion that they will be propelled out of the net of philosophical naval gazing by claustrophobia. Mountaineering logicians spring to mind. Others will be saved by the ups and downs of teaching, an outlet not available to scholars privileged enough to be relieved of all mundane duties by a position such as a permanent membership at the Institute for Advanced Studies.

Einstein, Gödel, Kreisel, Tarski, Wittgenstein, and so many other pioneers in the Foundations of Mathematics and the Sciences all came to the English speaking world as expatriates from continental Europe. An adjustment to English ways (any place in the Commonwealth) was probably not so difficult to achieve with a bit of native sophistication, a good dose of ambition and an air of self assurance. But in the States tougher survival skills seemed to be needed. I am not so sure that any of them ever really got assimilated. But I am sure Einstein's strength of character and of conviction were a great asset, as was Tarski's peasant resilience and ebullience. Tarski had the further advantage of having to teach. To find yourself facing a crowded class room of American students is like being thrown into the sea and told to swim. Tarski built an entire school out of this experience, much to his own benefit and that of the world of Mathematics. At the same time he remained true to his Polish heritage. Scholars in an exclusive research institution miss out on this opportunity of growth by interaction with the real life of young America, unless they seek it in their own ways, be it by gathering a following of young disciples, turning to politics or seeking indirect connection through the mass media ...

Apparently none of these options were compatible with Gödel's temperament, so he became the cautious and famous recluse, a role encouraged and respected at the Institute, but probably not exactly conducive to anyone's vitality. It must be pointed out though, that in central Europe around the turn of the century this was the typical life style of the scholar, enhanced and embellished by a bit of genteel social life, and buttressed by a loyal wife of good breeding and, hopefully, some means. Gödel did have a devoted wife and one could only wish that she were given credit for offering him the warm and reassuring environment that he so obviously needed. But their domesticity was not embedded in a familiar nurturing social environment. That must have been particularly hard on Mrs. Gödel. There is a passage in Hao Wang's book [1987], p. 116, where he notes that during the academic year 49/50 Gödel was in unusually good health and of fine spirits. Well, that was when Kurt Reidemeister was joined by his wife for a second year at the Institute and the two couples saw quite a bit of each other at the Gödel's home.[4] They had been old acquaintances in the old world. In fact, Reidemeister had taken part in the "Diskussion zur Grundlegung der Mathematik" on Sunday September 7, 1930, in Königsberg (cf. Gödel [1986], p. 197).

Kreisel

When I first met Kreisel, in Princeton at the end of summer 1955, he fitted my Swiss preconception of a middle aged, upper middle class Austrian gentleman to the dot. And that again was my impression whenever I ran across him in later years. He does not age, it is very difficult to visualize Georg Kreisel as a young man, much harder than to see him as a first grader. Maybe he lost buoyancy very early on, and that may well have been inevitable for a bright Jewish kid growing up in Graz during the thirties, of all Austrian places in Graz. His parents had been foresighted and affluent enough to send their two boys to England shortly before the "*Anschluss*" which, according to Hao Wang, means "political and economic union".[5] There his native competitiveness, developed early on in school by the need to have something special to show for himself, had come in handy at Trinity College in Cambridge and was enhanced by the Tripos spirit. Apparently Freeman Dyson, Georg Kreisel, James Lighthill and John Myhill were the major contenders for the Senior Wranglership.[6] I suspect it was the general British public school atmosphere as well as his cosmopolitan experimentations with

[4] Mrs. Gödel used to drive the R.'s home to the housing project after these domestic evenings, protesting what a pleasure it was driving her car without having to worry for the life of a genius. Pinze R. would, quite innocently, talk about this at the cocktail parties of the season, much to Kurt R.'s chagrin.

[5] Cf. Wang [1987] p. 100. Ever since I asked Hao where he got that translation from, I am puzzling over yet another manifestation of Kreisel's macabre sense of humor. Was he pulling Hao's leg or his own?

[6] A distinction of strange magical powers that Math students would compete for. I do hope someone knowledgeable of Trinity during the forties will write about these friendships and contests. Sad to say, John Myhill,a couple of years younger, is no longer alive. He had found a totally unique and expansive way of coping with adjustment to America. He was indomitable, brilliant and original, a very sweet, if impossible, person. But, he came from a completely different, much more down to Earth, not to say rough, background than Kreisel and Dyson.

a vast variety of London life styles that helped expand the scope of central European class snobbisms in which Kreisel had been brought up. Having settled down to a permanent membership at the Institute for Advanced Studies in fall 53, Freeman Dyson encouraged Gödel to invite Kreisel to the Institute and in summer 55 Georg Kreisel arrived as a very sophisticated bachelor in a country where both sophistication (of manners) and bachelorhood are rare, if not suspect, phenomena.

As I met him then and knew him later, Kreisel was not a fighter, he thought of himself as physically frail. In the good old middle European tradition, the physical delicacy included "nerves". He also was and still is brilliant, articulate, perceptive and witty. His provocative and abrasive sense of humor is best taken as a test of the recipient's self esteem and self reliance. I do not know whether Kreisel has ever done any irreparable damage to anyone, I hope not. At times I have wondered whether his tendency to pick on a protagonist's physical and other frailties, often totally out of context, might be a form of displaced self irony, a common stance in defense of vulnerability. After all, he was a double exile.

Although he did not come with the explicit intention of staying, and the possibility of return to Europe was open to him, Kreisel did not have strong ties to his position at Reading, a rather puzzling domicile for a person of his peculiar qualities. The community of logicians in the States was involved in research very close to Kreisel's own concerns. Everybody was eager to listen to him, and he was quick to discover some of the most talented budding scholars and scientists. Kreisel was not a "born teacher" and he did not aspire to become an educator. He lacked both the patience and the exhibitionism, at least at the time I knew him, and I would be very much surprised if he ever sought anonymous audiences, which classes of students invariably are, at least at the beginning of a course to all of us, and apparently forever to some. He would restrict his communications to select circles of peers and promising Ph.D. students of colleagues. And indeed he had a green thumb with brilliant young scholars who were technically and philosophically sophisticated enough to comprehend Kreisel's foundational ruminations, while endowed with the diligence, the agility and the initiative it took to turn his ideas into viable research projects. Fruitful collaboration would take the form of discussions in personal confrontation, interminable telephone conversations and daily correspondence. I do not recall any of the usual academic bureaucracy involving grant proposals, reports and accounts surrounding and encumbering these pursuits. (Anyone familiar with academic bureaucracy in this country will suspect that there must have been some one sacrificing time and dirtying hands with this kind of business on our European scholar's behalf). Work was all just a matter of thinking hard and aloud in a tète à tète until the fogs would start lifting. Wonderful goings on to watch for a young algebraist. 56/57 was the year of Dana Scott and of Hilary Putnam at Princeton.

Summer 57 Kreisel took me along to the logic meeting at Cornell, my first encounter with the world of logicians, bewildering, exciting and the beginnings of my interest in intuitionism. I recall sneaking away to lie under a tree with Heyting's *Introduction* [1956] at every opportunity. Kreisel never thought much of sunlight and fresh air during those years. But he would bring the A. Robinson's home for long discussions with Abraham of the constructive content (or version) of proofs in field theory. Altogether the concept of constructivity was in the air. And it was the summer that Kreisel discovered Clifford Spector. What a promising encounter that was, if only Clifford had lived.

Incidentally, during my studies at the University of Zürich I had been trying in vain to find out about Logic and the Foundations. Finsler was reading set theory. But, after the first few lectures, he would start drawing snares in the left hand corner of the blackboard, expounding the vicious circle principle from which he then was unable to extricate himself for the rest of the entire term. He was not well during the forties and he seemed such a very nice gentle frail person. Attending his lectures was heart breaking. It was only much later, when reading his paper [1926] , that I came to appreciate the lucidity of his anticipation of Gödel's ideas, Hao Wang's rash interpretation of one of Gödel's illegible scribbles notwithstanding [1987] p. 17. We did have a parched Aristotelian, Karl Dürr (sic), in the philosophy Department, not the kind of teacher I was looking for. Speiser, my "doctor father" was regularly running a very elegant inter-disciplinary Neo-Platonic seminar which left me equally frustrated. Thirty years later I learned that he had been a pupil of Hilbert's! At the institution next door, the ETH, Bernays was lecturing. But, believe it or not, I only became aware of his existence after meeting Kreisel. There was a great variety of prejudice afoot in Switzerland during the war, we were extremely isolated and naive. This is how I had remained untouched by any philosophical "school".

European Interlude

After the Cornell meeting we went to Europe where Kreisel participated in the Constructivity Colloquium in Amsterdam (Heyting [1959]), about which, unfortunately and significantly, I have practically no recollection. In particular, even then I still did not hear, let alone meet Bernays. Browsing through the proceedings now I realize what a land mark of a meeting it must have been! A few weeks that we spent in Zürich in fall 57 are equally vague in my memory. Staying at the Waldhaus Dolder and going for constitutional walks in the woods of the Zürichberg, where I used to roam about during my student years, gave me an eery feeling as if I had come back out of some disconnected future. In retrospect it seems that altogether Kreisel was much more at ease with me in public in America than "back home" in Europe. (Incidentally here is an amusing association: before their tendencies had become "in" many of my gay friends would visit Europe as a place where they could let their hair down, while back home in the States they would feel uneasy and self-conscious!)

For the winter of 57/58 we settled in Reading, where Kreisel resumed his duties and I crawled into a corner with lots of blankets and Kleene's *Introduction* [1952]. By then I had been initiated by Heyting. To a "foundationally" naive but deeply inquisitive mind, whose adventures into the realm of structural Mathematics had come to a shivering halt in the shadow of a budding genius, the encounter with Intuitionism brought a revelation. Here, just a few leaps beyond the ivory cage of Princeton domesticity, was a clear and compelling vista of the process of mathematical reasoning, directly understood as a dynamic, living phenomenon rather than packaged into a disembodied calculus of "values" backed up by a static potpourri of set theoretic contraptions! Questions of soundness would not haunt me here as they did in classical logic and completeness – relative to what? – seemed a spurious worry. To quote Heyting [1930]:

> Intuitionistic Mathematics is a mental process, and every language, the formalistic one included, is an aid to communication only. It is in principle impossible to construct a system of formulas equivalent to intuitionistic Mathematics, since the possibilities of thinking cannot be reduced to a finite number of rules constructed in advance.

Heyting's formal rules make sense and reflect how we reason in Mathematics, moreover they are handy and neat, clear, simple and distinct. So let us use them as tools and see how far we can get with them. I was also very happy with the constructive reals and later on made a sport of teaching my calculus classes constructively. But I certainly was aware of the fact that this discrete Cauchy-Weierstrass type of a net could not suffice to fathom the "labyrinth of the continuum" (to use Leibniz' apt expression, cf. Russell [1900] p. 245 G.II 379[7]).

English winters have a bleakness all of their own. No storms, no disasters, natural or otherwise, just a clammy dreariness. The moods of a coal stove is what keeps you in touch with reality. Those long suffering Brussels' sprouts in the backyard are surviving ostentatiously, shaming you into apologies for having been born without the "stiff upper lip". And the loneliness of total absorption into Mutual Fulfillment !

But there was Kleene's *Introduction*, a great book for someone who prefers to bite herself through rather than be taught. Kreisel did most of his work at home, thinking aloud, writing, rewriting and rewriting again, notes as well as long letters. In particular there were daily essays going back and forth between Reading and Oxford. Often Hao Wang would be calling while Kreisel was writing to him, or vice versa. But, probably the most intense and important correspondence was continuing the exchange of ideas with Gödel initiated in Princeton during the previous couple of years.

After Kreisel and I had parted for what later turned out to be only a few months, I went to Switzerland. Kreisel having encouraged me to get in touch with Bernays, I participated in his seminar in spring 58. But, by that time I was much too preoccupied with serious personal decisions to devote myself actively to the solace of Mathematics. Yet I have a vivid memory of Bernays' gentle personality, and if anything gave me a ray of confidence during that very difficult spring time, it was the knowledge of Bernays' existence. Whether he was playing the piano at Oberwolfach, fussing over technicalities in a seminar or presenting a conceptual analysis, his approach was always illuminating, poised on accomplished insight and presented with respect for the subject at hand as well as for his audience. How much longer would our understanding of Gödel's second theorem have been delayed had not the second volume of Hilbert-Bernays [1939] appeared in 1939! Has it ever been translated?

Stanford

By fall 58 however, via a six week sojourn to Reno, I had regained my civil freedom and Kreisel had accepted an appointment at Stanford. The day after I came down from Nevada, which, incidentally

[7]I am quoting from Russell's selection rather than Leibniz' collected works, because it gives a quick access to Leibniz' view of time, space, the continuum and infinity. But I fervently hope an intuitionist will delve into the original Leibniz sources and write an essay on Leibniz' Labyrinth and choice sequences.

coincided with Dyson's second wedding, Kreisel, who used to shun social gatherings, found a pretext for a party in his house in the Los Altos Hills, so that he could introduce me as "my wife Mrs. Dyson". Whatever the intended effects on Trinity College and its Alumni,[8] that simple statement opened up the world of logicians to me, in the Bay area as well as on the Peninsula. I was generally accepted as part of Kreisel's life and attended his seminar on constructive functionals at Stanford. We even had a bit of a social life, with the Taits and the Fefermans and also the Suppes'. That academic year was a good one for Kreisel too, I believe. He had a home base, a shelter of domesticity that fitted in with the community in which he was engaged with his vocational endeavors. Beyond these confines we had some lively discussions of politics and cultural issues with Paul Baran, the noted and controversial Stanford economist. Once in a while my friend Erika Hublitz, who was engaged in translating the topologist Reidemeister's literary, philosophical and humanistic essays, would drop by, although my phone conversations with her were monitored (by G.K. of course) and limited to three minutes. We would catch our breath in short walks through fragrant neighbor-hoods, enjoy the creations of culinary skills in which Kreisel took considerable pride and conclude the days by drives over golden hills and through olive groves that invariably filled me with nostalgia for the Greece of my childhood.

In order to show Andrzej Mostowski the "other part of California", once on a wintry day we even ventured on a drive to Half Moon Bay where I was swept under by the surf with all my clothes on, (except for my coat) but, without blinking, still drove on to Berkeley as planned. There, in one of the Bay area's most refined establishments (maybe it was Trader Vic's on Jack London square, Tarski's favorite), Kreisel outdid me in *sang froid* by refusing a bottle after sniffing its cork. Mostowski was obviously non plussed and duly impressed, but what Tarski had to say to it all after we had delivered home his Polish disciple I never learned. (No need for M. or me to sniff too, "G.K. knows best". But I have a creeping suspicion that the display was intended to be reported in Berkeley).

It was a good and relaxed climate for work and for collaboration. Generalizations and categorizations don't reach very deeply, but as a first approximation one might sort mathematicians into deep thinkers and craftsmen, puzzlers and puzzle solvers, conceptual innovators and

[8]How ruthless that competition at Trinity must have been began to dawn on me only later when I started puzzling over might have attracted Kreisel to me in the first place and became bewildered by the devastating rigidity of Dyson's self-righteousness.

technical virtuosi. Kreisel would certainly never loose sight of the woods, though he might on occasion trip over an individual tree trunk. He is deeply and tenaciously preoccupied with the mathematical articulation of philosophical and methodological problems in the foundations. His entire personality is ruled by a tendency to introspection. It would be incorrect to say: to the exclusion of taking notice of other matters. On the contrary, he is prone to extend his analytic curiosity to fellow psyches as well as to the working of his and their various inner organs. If not mitigate, this may at least explain his repetitious, uncalled for remarks on Gödel's state of health.[9]

Completeness Problems

One of the most basic problems in the foundations of mathematics is the clarification of the interrelation between Form and Content, Theorem and Proof, Extension and Intension. A formula without an interpretation is an idle tool, a theorem detached from its proof encodes a truth inaccessible to understanding and a set remains a disembodied ghost until it finds its place among the values of some functor.

This century started with the great successes in the art of formalization. Hilbert [1899] distilled Euclid's work into a form ready for Tarski [1948] to dot the i's and prove the formal completeness, and with it a fortiori the decidability, of elementary geometry and algebra. By 1930 the formalization of what we now call first order Logic had crystallized to the point where Gödel was able to prove its completeness. The same year Heyting [1930] formalized intuitionistic predicate logic and articulated the distinction between thought and result in mathematics quoted above. On September 7 of 1930 Gödel declared ([1986], p. 200):

> *man kann von keinem formalen System mit Sicherheit behaupten, dass alle inhaltlichen Überlegungen in ihm darstellbar sind.*

Gödel's remark was meant to justify the following *Gedankenexperiment*. Suppose we were faced with a finitary predicate Φ, for which:

1. $\exists\xi\Phi$ is provable by means of the transfinite methods of classical mathematics using the *tertium non datur*, while

[9] I am using the present tense here, because I am talking about phenomena that were not restricted to the years we spent together, but are manifest throughout his published work. Incidentally, I never called him by any other name than "Kreisel". To my knowledge Kleene is the only person who ever called him "Georg".

2. intensional deliberations ("inhaltliche äberlegungen", translated by "contentual considerations" in [1986]) show that each natural number satisfies $\neg\Phi$.

Such a situation, he points out, would be compatible with the existence of a formal consistency proof of classical Mathematics, for, *no formal system can be claimed with certainty to represent all mathematically meaningful trains of reasoning*. In other words, formal consistency is no guarantee for soundness, unless completeness with respect to content is established. At this point Gödel switched from the hypothetical mode of the *Gedankenexperiment* to the factual report of the existence of recursive properties Φ for which (2) holds, while $\exists\xi\Phi$ is consistent with the formal system of classical mathematics,[10] which shows that, assuming its formal consistency, the formal system of classical mathematics does have formally complete and consistent extensions satisfying the conditions of the *Gedankenexperiment*.

Indeed, from the intuitively informal point of view it makes perfectly good sense to consider absurd the claim that all mathematical problems have a solution, while maintaining, for each individual problem, the absurdity of the presumption that it will never have a solution.[11] At this point of the story I am talking about hunches. The pre scientific era of a field is permeated by psychological hankerings and epistemological hunches. Their antithesis is what motivates the invention of concepts that allow the articulation of tractable problems and provides the driving force to scientific inquiry. The situation just envisaged is a perfect illustration of this fruitful tension: to claim that every problem has a solution, as Hilbert did, presupposes a method, a uniform method for solving all problems. Caution, based on experience, dictates skepticism. Gödel isolated the key concepts and gave them precise form on the basis of which he was able to construct mathematically, to every method a problem, that the method is incapable of solving. Hence the assumption of a universal method for solving all problems is absurd and so then is the claim of the solvability of all problems. However, our desperate need for hope boggles at the thought of a problem that would be absolutely unsolvable. And, within the confines of a definite formal system, we can argue objectively that there is no need to accept such an atrocity as a possibility. Following Gödel again, restrict the

[10] There is no need for the editor (cf. [1986] p. 197) to explain Gödel's correct use of grammar by hypotheses about the psychological effect of von Neumann's remarks on Gödel's hypothetical shyness.

[11] The theory of free choice sequences struck me as a revelation. Not only did it illuminate Leibniz' Labyrinth, but here I met a natural occurrence of the classically taboo conjunction $\neg(\alpha)\exists x\Phi$ & $(\alpha)\neg\neg\exists x\Phi$! Compare with (M) of the next section.

problems to the questions whether or not there exists a natural number having a given "finitary" property Φ. If there were any evidence, in the form of some method of proof, that such a problem is insoluble, then this method would establish the solvability of the problem by the answer "no". For, if there existed a natural number with property Φ, then this fact would be provable in Peano Arithmetic. So, a proof of absolute insolubility and, a fortiori, of Peano insolubility would entail the absurdity of $\exists\xi\Phi$ and with it solve our problem.[12]

During the late fifties Kreisel was keen to stress the distinctions between various, more or less well understood concepts of completeness of a formalism (see [1961]). First of all there is the question whether the formalism is intended to describe one single standard structure or an entire class of variable structures satisfying certain conditions, if any. In the former case the concepts are as clear as the underlying concept of truth in that structure. *Soundness* means that all theorems, i.e., provable propositions, are true, *consistency* claims the unprovability of absurdity – or of $(P \mathbin{\&} \neg P)$, for any proposition P – while *semantic completeness* is the converse of soundness and *formal completeness* asserts the provability of P or of $\neg P$ for each P. Thus, a formally complete and consistent theory is *maximal* in the sense that it has no proper consistent extensions in the same language and *decidable* in view of the generally accepted assumption that a formalism worthy of its name has an effective proof procedure. For a Boolean concept of truth a formally complete and consistent theory is semantically complete if and only if it is sound, and semantic completeness entails formal completeness, while soundness implies consistency for just about all notions of truth.

While consistency as defined above is applicable just as well to formalisms intended for classes of inequivalent structures, what is left of the concept of formal completeness, expressed as a disjunction, is only the decidability. Classical cases of quantifier elimination can be seen in this light. For the theory $\mathcal{A}$ of Abelian groups, for instance, there exists a recursively enumerable set of sentences, B_n, each consistent with $\mathcal{A}$, such that, for every sentence S of the elementary language of Abelian groups, there exists $n \in \omega$ for which either S or else $B_n \Rightarrow \neg S$ is a theorem of $\mathcal{A}$. This is the formal version of Szmielew's disjunctive (semantic)[13] completeness theorem [1955]. She exhibits an effective

[12] This is a familiar argument, but often ignored by popular writers who love to ponder whether Fermat's theorem might turn out to be one of those eternally unsolvables. Alas it has been solved since this essay was submitted. But they still have Goldbach to resort to.

[13] "Completeness" commonly names this or a similar property of formal systems in relation to a concept of interpretation, to be distinguished from purely formal properties, like maximal consistency.

sequence $\{G_m\}_{m\in\omega}$ of recursively axiomatizable Abelian groups such that: every sentence of the language of Abelian groups is either a theorem of $\mathcal{A}$ or else fails in one of the groups G_m. Then, if $n = \langle n_1, n_2\rangle$, using some pairing, choose for B_n the conjunction of the first n_2 axioms for G_{n_1}.

In the more general, undecidable case it is misleading, although classically acceptable, to phrase completeness in terms of a disjunction: all sentences are either theorems of the formal system, e.g., of classical first order logic, or else refutable in a countable model. The proof is a non constructive derivation of refutability from the assumption of non provability, which, for the jump to the entailment of provability by validity, requires the law of double negation.

With the Boolean notion of a set, the concept of an interpretation is ready at hand for the classical first order predicate calculus. Bluntly speaking, sets – and elementary models are the objects that are meant to behave according to the rules of classical logic. Gödel's completeness proof leaves us with a sense of reassurance, but hardly of surprise: we have not forgotten anything in our formal setup, but what have we learned? Its extension to higher order logic is more revealing: some contortions of what the woman in the street would call a model[14] are needed to ensure the desired completeness (Henkin [1950]). Soundness becomes a redundant question when the objects of investigation are defined as formally complete and consistent extensions of an elementary axiom system, and formal consistency will suffice. Thus for instance in the case of the elementary theory of Abelian groups. But then, what the mathematician working with these objects has in mind is much more concrete and formally obscure. So much for this digression.

Tree Semantics

Turning now to intuitionistic first order logic, the question of soundness does not seem to arise because its rules and axioms are carefully chosen to reflect no more than the process of constructive mathematical reasoning. At first glance the question of completeness, on the other hand, does not seem to make sense. As Heyting was quick to stress, when he made the first proposal for a formalization "des intuitionistischen Schliessens", the "possibilities of thinking cannot be reduced to a set of rules constructed in advance". Thinking is a living process open to growth, new insight and uncharted methods. And yet, in the

[14]She would, for instance, call the Peano axioms together with the second order formulation of mathematical induction categorical.

fifties E.W. Beth, well acquainted with intuitionist mathematics, but of broader philosophical tendencies and humanist tradition, started the search for an intuitionistically meaningful concept of a semantics apparently with hopes for a proof of completeness of Heyting's system at the back of his mind. He was looking for interpretations of first order languages on models that would be intuitionistically meaningful structures. Note that, as soon as such concepts are fixed by definitions, the question of soundness does arise. But that gives no trouble when definitions are rigged up to fit the purpose.[15]

Kreisel pointed out in [1961] that, if one insisted on completeness in the run of the mill form of "truth" under all "interpretations" implying the existence of a formal proof, the best one could hope for would be a classical tour de force. However, he thought there might be hope for an intuitionistically acceptable proof of what he called "weak completeness", (WC), namely the double negation of this implication. And that we set out to look for, following Beth's lead closely and tidying up his argument. We succeeded, [1961], up to *Markov's Principle* in the following form: for primitive recursive relations A between choice sequences α and natural numbers n,

$$\forall\alpha\neg\neg\exists n A(\alpha, n) \;\Rightarrow\; \neg\neg\forall\alpha\exists n A(\alpha, n) \qquad (M)$$

The quantifier equivalence $\neg\exists\xi\Phi \Leftrightarrow (\xi)\neg\Phi$ is intuitionistically correct as well as a theorem of HPC, Heyting's Precidate Calculus, while $\neg(\xi)\Phi$ is weaker than $\exists\xi\neg\Phi$. So, (M) is equivalent to

$$\neg\exists\alpha\forall n\neg A(\alpha, n) \;\Rightarrow\; \neg\neg\forall\alpha\exists n A(\alpha, n) \qquad (M')$$

whose consequent is logically stronger than its antecedent.

Now, for a suitable formulation of HPC in terms of sequents, an analysis of the structure of any given sequent S leads to the construction of a binary tree dressed by sequents and to a primitive recursive relation A that holds between a free choice sequence α and a natural number n iff the node $\alpha(n)$ is wearing an instance of an axiom. The dressing of the tree follows a systematic procedure reversing appropriate instances of the basic transformation rules of HPC for sequents along branchings so that, with careful attention to detail, one arrives at the hoped for conclusion:

$$\forall\alpha\exists n A(\alpha, n) \;\Rightarrow\; HPC \vdash S \qquad (*)$$

Nowadays every student of elementary logic is familiar with the method of decomposing a sequent which terminates with the construction of

[15]To put it grotesquely: truth is construed to fit the rules.

a proof, if there exists one. But, what does it have to do with semantic completeness? Well, somewhere a concept of validity must be brought to light. Validity means truth under all interpretations, so, what is a mathematically tractable and conceptually meaningful notion of interpretation for the language of first order logic? The *No-Counter-Example Interpretation* was one of Kreisel's many fruitful insights (*Gedankenblitze*). Classically only a logical equivalence transformation is needed to arrive at the characterization of validity by the non existence of a counter example. Yet Kreisel's fresh outlook lead to new methods, and for our intutitionist project it did just what we needed.

A tree naturally bears a topology. The topological interpretation had already been studied soon after Heyting's formalization, in Tarski [1938] and Mostowski [1948], and seemed natural enough. Restricting attention to trees of choice sequences dressed with formulae we are stripping away all inessentials and availing ourselves of the theory of choice sequences. Truth of a wff B on a choice sequence α is defined inductively according to the formation of B, starting with true atoms occurring at a node of α. Truth is inherited by the entire subtree consisting of all paths through a node $\alpha(n)$, once it is secured at that node. This interpretation is easily shown to be sound for HPC. Moreover, α is a counter example to the sequent S, if all the wffae that occur on the left of the turn style are true on a while at least one wff on the right is fails. Now, the relation A has the property that

$$\alpha \text{ is a counterexample to } S \quad \text{iff} \quad \forall n \neg A(\alpha, n).$$

Validity of S, entailing the non existence of a counter model, then implies $\forall\alpha\neg\neg\exists n A(\alpha, n)$. This is why (M) is wanted to obtain the double negation of the entailment from validity to provability, what Kreisel calls *weak completeness*.

The result of [1961] does not at all hang on the adoption of (M). The articulation of (M) simply brings home the following situation:

$$\begin{array}{lcrcl} (M) & : & \neg\forall\alpha\exists n A(\alpha, n) & \Rightarrow & \neg\neg\exists\alpha\forall n \neg A(\alpha, n) \\ (*) & : & \neg\forall\alpha\exists n A(\alpha, n) & \Rightarrow & HPC \vdash S \\ (WC)? & : & HPC \vdash S & \Rightarrow & \neg\neg\exists\alpha\forall n \neg A(\alpha, n). \end{array}$$

The arrow adorned with question a mark has spurned further research. For instance weak completeness, (WC), fails when the interpretations are restricted to recursive refutation trees (KreiselDyson [1961], theorem 3, p. 57). It is important to observe that concepts, which classically all get lumped together, reveal subtle distinctions under the natural illumination by intuitionism: the existence of a counter model is stronger than invalidity, why should we expect their double negations to be equivalent? Is there any reason to expect non theoremhood to coincide with non validity? And even if it did, why should the existence of a sentence that is neither valid nor refutable by counterexample be inconceivable? We know that intuitionism reaches far deeper than the process of double negating classical results. Indeed, during the fifties the nascent theory of free choice sequences started shedding light (cf. Troelstra [1977]). Nowadays developments in the theory of *topoi* is opening up insights, confirming hunches and reinforcing convictions (e.g., Troelstravan Dalen [1982]).

On a more mundane plane, attention to the fussy details in the proof of KreiselDyson [1961] yields a constructive proof of completeness in its strongest form for Heyting's predicate calculus restricted to prenex formulae: the procedure of KreiselDyson [1961] either terminates with a proof or else produces a recursive counterexample, a nice marker on the watershed between intuitionism and classical logic![16]

But the value of a paper like KreiselDyson [1961] does not lie in its conclusion, which is admittedly inconclusive, but rather in the questions it has stirred up, the research it has spawned. Not that anyone set hopes on a weak completeness proof for Heyting's formalism. Such a result would have been rather baffling and flying in the face of the dynamic, open ended character of intuitionism. But the question what kind of semantics would support approximations to completeness set in motion a vigorous search for interpretations of intuitionistic reasoning. The situation was exactly opposite to that in classical foundations. There we thought we knew what truth is in terms of sets! Here we have an undefined notion of the kind of insight one arrives at subjectively through reasoning guided by content, meaning and intuition, by introspection, contemplation, rumination – by arduous meditation. Finally, rational reflection in the *light of nature* – to use one of Kreisel's favorite expressions – leads us back to *results, that in turn are applicable to concrete situations* (cf. Bernays [1955]).

This is the point at which the intuitionist becomes a constructivist.

[16]It would be instructive to program the procedure and apply it to intuitionistically invalid theorems of classical predicate logic. Monadic sentences like $\exists y \forall x(Fx \Rightarrow Fy)$ are quickly reduced to propositional invalidity. I would like to see an infinite counter example to a classically valid prenex formula. Moreover, with the versatility of computer software one might be able to watch what happens to the sequent M:

$$\forall x \forall y(Rxy \vee \neg Rxy), \forall x \neg\neg\exists y Rxy \vdash \neg\neg\forall x \exists y Rxy?$$

Through mathematical analysis of his mental constructs, he arrives at manifestations and interpretations in the form of concrete, identifiable procedures about which he is able to communicate. Bewilderment over the escalation of formal techniques in the wake of path breaking proofs of their inherent limitations is uncalled for: Formalization and Poetry are the two most direct means we have of externalizing our thoughts.

Concluding Remarks

In my experience communication and feedback within one's realm of preoccupation, no matter how esoteric, is absolutely necessary for anyones sanity. I am sure it was his lively exchange of ideas with a great variety of logicians that kept Kreisel's tendency to depression at bay, at least during the time I knew him. He was a deeply troubled person, haunted and unhappy. We talked quite openly about our psychological conflicts. Karen Horney's astute down to Earth observations in her *Neurosis and Human growth* opened up my perspective and helped me back to reality after the Reading winter. (It was Chris Fernau in his subtle wisdom who had recommended this book to me). I want to stress my conviction that all thinking human beings are prone to depressions and that it is as hypocritical to deny this fact as it is crude to make light of it.

Kreisel's way of work, life and coping took[17] the form of talking to himself and others, mulling things over this way and that, with footnotes and asides and repetitions and not even stopping to leave things alone once they were out in print, but rather churning out appendices and appendices to the appendices. While Gödel had a tendency to shrink his work by overhaul, pare it down to bare essentials and strive for final form, Kreisel's output would keep expanding like a "*Gugelhopf*". His immediate environment could be a gerbil's paradise of snippets of scribbling, notes, letters to and by him, typescripts, reprints and an occasional book. But once all this bulk had been processed it would be discarded, or given away. Heyting's *Introduction*, the Amsterdam *Proceedings* and Tarski, Mostowski and Robinson's *Undecidable Theories* ended up in my hands, not to forget Henry Miller's *Tropics*. Thank you Kreisel, they have all served me well!

Although he proclaimed great respect for grand scale affluence, Kreisel did not believe in possessions. The only objects steadily attached to him that I recall were a change of clothes and a set of blackout

[17]Caution prompts me to use the past tense here, although I cannot imagine an essentially transformed Kreisel.

curtains, that went everywhere that Kreisel went. Instead of owning earplugs he would unplug all appliances before retiring, including the fridge. In contrast, Gödel preserved the most ephemeral scraps of paper. A mere glance into the "Nachlass" box that contains the collection of his notes on Leibniz is overwhelming and terrifying: a wealth of material, the records of a near life time of deep ruminating thought, all cramped into an over size card board box. An image emerges of Gödel being choked from the inside by all those insights that wanted out, but could not be let out because, presumably, he had not yet arrived at their final formulation. Anyone familiar with the paralyzing effects of finding one's self irretrievably trapped inside a world of one's own creation will understand that under such circumstances there is no need for the diagnosis of paranoia to explain an inability to function in the service of self preservation. To be sure, before long some scholars will work through all that repressed material and salvage for posterity what posterity is capable of assimilating ...

I have never known Kreisel as a sociable person.[18] He had no use for small talk, nor for hobbies, sports, entertainment, the outdoors or whatever else people like to fill a vacuum with. In Los Altos Hills we had a grand piano, but I knew better than to play when he was present, which was practically around the clock, unless we were together at Stanford. His occasional flirtations with the life style and the company of jet setters were, I believe, a sham, maybe meant as a provocation. His interests were single mindedly foundational, but not restricted to Mathematics. He was tenaciously inquisitive about the workings of the human psyche, mind and body. Whatever attention he would pay to literary works was guided and restricted by this tunnel vision. He may dismiss these remarks as euphemistic or evasive. Still, I maintain that all his preoccupations – or obsessions, if that is what they deserve to be called – stemmed from an irrepressible curiosity about what makes human beings tick. He could needle, daunt and take you apart till he found what he was looking for. If you were squeamish or got flustered he would ridicule you without bothering to put the pieces back together. It took courage to live with Kreisel, the courage of love, despair and a

[18]He used to retire by 9 p.m. or before, no matter what the occasion. At one of the customary after dinner bull sessions in Oberwolfach a debate sprang up: was Kreisel a misanthrope or simply shy? Bill Boone, always a staunch supporter of the friendlier of two hypotheses, ventured to make a bet, announced he was going to bring Kreisel down for a bottle of wine and ambled upstairs, only to return speedily and red faced: apparently Kreisel's door had opened a mere crack upon Bill's timid knocking and Bill had got around to saying no more than "I hope I am not ..." before, with the utterance "yes you are', the door was politely but definitely closed again.

touch of recklessness. He did respect that, I believe. He was not really a cynic, although he may have liked to think of himself as one.[19]

Kreisel was anything but a hypocrite, in contrast to some less controversial men in my life. In fact, he tried to make himself out a worse person than he really was. I shared whole heartedly his intolerance of hypocrisy and his impatience with phony pretensions and always enjoyed his sarcastic treatment of pomposity. But, I do think, he had more difficulty than ordinary people tolerating feelings, traits and outlooks that he was not familiar with by first hand experience. His friendships would go through a steep upward surge of exploration, level off into lively communication and solid collaboration and then, more often than not, they seemed to fade out of Kreisel's life.[20] Some fled before they were dismissed, used up or destroyed, others came to a clash. No promises – no betrayals, but some disappointments.

Kreisel's association with Gödel certainly outlasted many others. I have allowed myself many digressions on Gödel in this essay not only because he obviously had a deeper influence on Kreisel's thinking on the Foundations than any other scholar, but also because of the similarity of their cultural heritage. I would even venture to suggest that in terms of psychological initial conditions there was a certain kinship between them. To my knowledge Hao Wang and possibly Stan Tennenbaum, whenever his Brownian motions would allow it, were the only logicians in touch with Gödel during the last few years of his life. It is a distressing, if understandable, phenomenon, that towards the end of their life persons of great stature become progressively lonely, whether they choose to withdraw or not. Their disciples', acquaintances' and even friends' respect and admiration turns into a dark form of awe that frightens them into avoiding the person who needs their existential support more than ever before.

Ludwig Wittgenstein's ghost kept haunting us throughout our time together.[21] In the early days of our acquaintance at the Institute during

[19]For one thing, he took himself too seriously for that. I guess he did like to present himself as an "enfant terrible". Enfant-yes, terrible? I would not say so.

[20]I am looking forward to seeing that statement refuted by this *Festschrift*.

[21]How many times would I quietly and stealthily be shedding bitter tears, because I was told that I could never make a respectable philosopher nor do meaningful work in the foundations with my mental block to Wittgenstein's *Bemerkungen* and *Investigations*. Invariably after trying to read Wittgenstein I would end up having nightmares of a Kafkaesque millipede that woke up one nice Sunday morning, and lingering in bed having nothing more urgent to do began to muse over which one of his legs he was meant to be moving first – he was paralyzed for the rest of his life. Later on, teaching in a Philosophy department I learned to appreciate Wittgenstein's aphoristic skills. He never failed me when I found myself in need of a quote for underscoring some point or other.

the late forties Freeman Dyson used to reminisce about his walks and discussions with L. W. of the Trinity days with an amusement kindred to the spirit in which he was initiating me to *Alice in Wonderland*. Kreisel, however, would time and again get tangled up in an uncanny wrestling match of biblical proportions with the late Wittgenstein. To this day I do not quite understand what was happening. I wonder whether Kreisel has ever been able to lay that ghost to rest and sort out Wittgenstein's worries over the Foundations to his own satisfaction. In a way both Wittgenstein and Gödel were family to Kreisel. So was I, as a matter of fact.

At the end of summer 1959 Kreisel returned to Europe and I started teaching at San Jose State. We parted, unceremoniously. Of course it was very painful. We both knew there was no alternative. Nor had there been any alternatives earlier on. Once in a while I would become aware of opinions expressed on the academic cocktail party circuit that I could have had "the best of both worlds", if only I had "played by the rules". I never had it in me to play by rules. I know I would not have been able to split myself up into pieces nor would I have survived the charades of those best possible worlds. The real one is hard and demanding enough. One might say we were too close. Certainly there were times when I felt I was suffocating. Kreisel sported an enormous astral body taking up a vast chunk of space-time and permeating every aspect of my life. Even now, when I write something – not straight forward mathematics – but things like this essay, a letter on "Women in Mathematics" or philosophical ruminations I have the uncomfortable feeling that Kreisel is looking over my shoulder. And when I read Kreisel in print, especially some scathing review, he springs to life like an apparition.

Neither of us ever envisaged an entire life time together. But I believe that my total commitment, free of all reservations, hesitations and goals was as good for Kreisel at that juncture of his life, as it came natural to me. There is no need to say any more at this point. We parted when our relation had run its course.

Professionally I turned back closer to home, to the borderline between group theory and logic. Let me admit in closing that I do not think I basically have the tenacity and the moral stamina needed for coping with the tortures of work in the foundations proper.

Dear Kreisel; I don't think I ever baked a cake for you. So, here is a *Gugelhopf* for you with 70 candles on it burning bright:

$$\chi\rho o\nu\iota\alpha\ \ \pi o\lambda\lambda\alpha\ \ \kappa\alpha\iota\ \ \nu\alpha\ \ \tau\alpha\ \varepsilon\kappa\alpha\tau o\sigma\tau\varepsilon\iota\sigma!$$

Addendum

As an incorrigible disciple of Kreisel's I cannot forego the chance of adding a commentary after having submitted the essay. Moreover, at the end of the introduction (also after completion, but availing myself of the new magic of a Mac), I rashly promised some of the readers to divulge my ID if not my IQ. But, I must admit that rereading the last page takes the bureaucratic winds out of my sails. If you people only knew what it feels like to be confronting eight months after your own 70-th birthday the turning point of your life, a turning point that, had it not been for the sheer tenacity and common sense of your love of life, would have proved your undoing once and for all. (A poem of my childhood springs to mind: *der Reiter über den Bodensee*). So, the above effusion was not written for anyone's entertainment. Nor was it meant as a cathartic exercise. It was, in fact, carried along by the same good humored spirit in which Stan Tennenbaum and I would invariably join forces together against general Kreisel bashing at casual get togethers of logicians during the sixties. Most vividly I remember occasions during my time at Chicago Circle.

Anyway, here is what I deem a sufficient collection of data:

I was born in Naples, raised in Greece, attended the *Deutsche Schule Athen*, throughout the thirties all the time aware of my Swiss origins, and obtained a Ph.D. from Zürich with a thesis on the structure theory of finite groups. I met Freeman Dyson when we both were post doctoral fellows at the Institute for Advanced Studies in Princeton during the year 1948-49. We married in summer 1950, spent a year in Birmingham and two at Cornell, where Freeman decided that teaching was not exactly what he wanted to spend his life on and accepted Robert Oppenheimer's standing invitation to join the permanent staff at the IAS. I had known Kreisel as an intriguing companion for long discussions on all sorts of social and cultural topics for an entire year when we suddenly became lovers. It happened at a time when my marriage had been shaken by a natural catastrophe that had not been any one's choice and need not be discussed here. Nor do I wish to talk here about my children, whose loyalty has survived all.

I believe that Kreisel was taken by surprise by the depth of our involvement on both sides. The original intention may have been a mere extension of the Trinity rivalries. I have never been able to stay "casual" in any relation, no matter how ephemeral. That I immediately followed my impulse to tell Dyson was probably my first infringement of social rules in the entire tragicomedy.

Addendum to the addendum (true Kreisel style)

Rereading Kreisel's letters of the fall 1959 I am moved by their simplicity, depth of feeling and sincerity. My essay is a two dimensional account of an experience too complex and too subtle to be conveyed fully without resorting to a transformation into good literature.

References

Bernays, P.

[1955] Die Mathematik als ein Vertrautes und zugleich Unbekanntes, *Synthese*, 9 (1955) 456-471.

Finsler, P.

[1926] Formale Beweise und Entscheidbarkeit, *Math. Zeitschrift*, 1926, repr. in J. van Heijenoort, *From Frege to Gödel*, 1971.

Gödel, K.

[1986] *Collected Works*, Vol. I, Oxford, 1986.

Henkin, L.

[1950] Completeness in the theory of types, *J. Symbolic Logic*, 15 (1950) 81-91.

Heyting, A.

[1930] Die formalen Regeln des Intuitionistischen Schliessens, *Preuss. AK. Wiss. phys. math.*, 1930, pp. 42-71, 158-169.

[1956] *Intuitionism: an Introduction*, Springer, 1956.

[1959] *Constructivity in Mathematics*. Proceedings of the 1957 colloquium held at Amsterdam, North Holland, 1959.

Hilbert, D.

[1899] *Grundlagen der Geometrie*, 1899, repr. by Teubner, 1962.

Hilbert, D., and Bernays, P.

[1939] *Grundlagen der Mathematik*, vol. II, 1939.

Kleene, S.

[1952] *Introduction to Metamathematics*, van Nostrand, 1952.

Kreisel, G.

[1961] On weak completeness of intuitionistic predicate logic, *Stanford Technical Report*, 3, January 1961, pp. 1-38.

[1980] Kurt Gödel 19061978, *Biographical memoirs of the Royal Society*, 26, 1980.

Kreisel, G, and Dyson, V.

[1961] Analysis of Beth's semantic construction of intuitionistic logic, in Kreisel [1961], chapter 2, pp. 1–65.

Mostowski, A.

[1948] Proofs of non-deducibility in intuitionistic functional calculus, *J. Symbolic Logic*, 13 (1948) 204–207.

Rucker, R.

[1982] *Infinity and the Mind*, Bantam Books, 1982.

Russell, B.

[1900] *A critical exposition of the philosophy of Leibniz*, 1900.

Szmielew, W.

[1955] Elementary properties of Abelian groups, *Fund. Math.*, 41 (1955).

Tarski, A.

[1938] Der Aussagenkalkül und die Topologie, *Fund. Math.*, 31 (1938).

[1948] A decision method for elementary algebra and geometry. *Rand Co.*, 1948.

Troelstra, A.

[1977] *Choice Sequences*, Oxford, 1977.

Troelstra, A., and van Dalen, D.

[1982] *The L.E.J. Brouwer Symposium*, North Holland, 1982.

Wang, H.

[1987] *Reflections on Kurt Gödel*, MIT Press, 1987.

Addendum

by Freeman Dyson

In a letter of June 29, 1993 to the editor, Freeman Dyson has made the following remark (reproduced here with his permission) about notes 6 and 8 of Ms. Huber-Dyson's piece (pp. 54 and 59):

> Concerning my friendship with Kreisel, I never felt myself to be in competition with him. As a result of Hardy's successful reform, the notorious "Tripos Spirit" ceased to exist in Cambridge long before we arrived there. The Tripos had been trivialized so that we did not need to worry about it.
>
> As mathematicians, Kreisel was a deep thinker and I was a craftsman. His life was dominated by Wittgenstein and Gödel, mine was dominated by Besicovitch and Davenport. We lived in different worlds. So we could be friends without being rivals.

A Letter from Professor Kreisel

by Carl G. Jockusch, Jr.

G. Kreisel and I had an extensive mathematical correspondence in the 1970's and early 1980's. We corresponded on effective versions of Ramsey's theorem, Π^0_1-classes and models of arithmetic, and paths through Kleene's $\mathcal{O}$, as well as many other topics. I was astounded by the erudition of his letters and the vigor and promptness of his replies. I will never forget that once, while vacationing on Cape Cod, I received three letters from him in a single day! Nevertheless, he has undoubtedly had an even more active correspondence with many other people.

Kreisel's letters had a great influence on my work. This influence can be seen (however imperfectly) both in the topics I have pursued and in specific technical points. In this short note I will not even attempt to describe the correspondence and its influence on me as a whole, but only illustrate it with a particular example.

In the late 60's and early 70's I was very interested in the extent to which Ramsey's theorem holds effectively, and indeed I remain fascinated by this topic. In 1970 I had shown (extending a result of Specker [1971]) that there is a recursive 2-coloring of the unordered pairs of natural numbers such that there is no infinite Σ^0_2 homogeneous set, and I was attempting to find a corresponding result for colorings of n-tuples. Kreisel suggested that it might be possible to show that for each $n \geq 2$ there is a recursive coloring of n-tuples with no infinite Σ^0_n homogeneous set by relativizing the proof for the case $n = 2$. I carried out his suggestion and thus obtained the best possible result with respect to the arithmetical hierarchy (see [1972], Theorem 5.1). When I sent him a preprint stating this as a joint result, he replied in part as follows:

> 30.7.70
>
> Dear Professor Jockusch,
>
> It seems to me that you should stress more, in fact already in the introduction, the interesting discovery of section 4 that different proofs of Ramsey's theorem lead to different effective versions. ...
>
> It seems to me quite wrong to attribute (part of) 5.1 to me. Remember what happened! You sketched the proof for $n = 2$ and asked what happened at $n > 2$. A rapid study of your sketch suggested[a] to me that something like it would relativize.
>
> Now suppose I had overlooked something, and the answer had turned out to be different. Would you have said:
>
> The following theorem refutes an assertion of K.?
>
> (As I once said at a meeting to somebody who was talking about some unpublished results and called his remarks 'historical': 'Tell me, Prof. X, if your unpublished thoughts are of historical interest, who don't you tell us your unpublished errors?')
>
> As ever
>
> G. Kreisel
>
> [a] Though it may not look like it, I always have proofs in mind for <u>assertions</u> in, say, reviews.

All I can say is "Thank you, Professor Kreisel!"

References

Jockusch, C.

[1972] Ramsey's theorem and recursion theory, *J. Symbolic Logic*, 37 (1972) 268–280.

Specker, E.

[1971] Ramsey's theorem does not hold in recursive set theory, in *Studies in Logic and the Foundations of Mathematics*, North Holland, Amsterdam, 1971, pp. 443–451.

Two Insights

by Michael Morley

Georg Kreisel is known for often telling people what is the crucial heart of a subject. Often it was something one had known already but had not appreciated its significance. I will give two examples from my own life, one mathematical and one pedagogical.

The Cantor-Bendixon Theorem

The Cantor-Bendixon derivative of a compact, separable space is the set of non-isolated points. If not empty this is again a compact, separable space and therefore the operation may be repeated, possibly transfinitely, until there are no further isolated points. Points removed at the αth stage are said to have Cantor-Bendixon rank α. In [1958] Kreisel gave an example of a recursively presented space having points whose ranks range over all the recursive ordinals. I learned of this a few years later. I can still remember how *illuminating* I found this. It is not mathematically difficult. Once you know it exists, the construction is only an exercise. But for me at least, it gave a geometric picture of the hyperarithmetic hierarchy and I began to *understand* it and its relationship to model theory. It led to my results [1967] on ω-logic, and helped me to understand admissible sets.

The Completeness Theorem

I received my degree in 1962 and in the next few years I taught at several universities. Usually I taught an undergraduate course in symbolic logic. These courses were much alike. One introduced the propositional and predicate calculus; the grand climax of the course was the Completeness Theorem. You can imagine how astonished I was when Kreisel told me: Don't teach the Completeness Theorem.

His point was that the fact that a particular proof schema is complete is not the important thing; the mathematical content is that consequences are recursively enumerable. Of course, I knew this. But it had not been reflected in the way I, and most other instructors, taught their courses. I took his advice and changed the way I taught.

References

Kreisel, G.

[1958] The Cantor-Bendixson theorem in hyperarithmetic analysis, *J. Symbolic Logic*, 23 (1958) 460-461.

Morley, M.

[1967] The Hanf number for ω-logic, *J. Symbolic Logic*, 32 (1967) 437.

An Appreciation of Kreisel

by Anil Nerode

Georg Kreisel has been one of my favorite people and favorite logicians for forty-five years. In youth, when we all need approval and advice, he gave it. I found him warm, supportive, and insightful.

I have probably read all of his mathematical logic papers. I started to read them, as they came out, when I was a graduate student in mathematics at the University of Chicago in 1950. I was then much under the influence of Rudolf Carnap in Philosophy, Saunders MacLane,[1] Paul Halmos, and to a lesser extent (Bourbaki as represented by) Andre Weil in Mathematics. The strong Chicago trend was to work on the fundamental, preferably global, theory of almost everything in mathematics, ranging from topological groups and algebras and functional analysis, to statistics, probability, and logic.

Those of us in logic had established idols. Of course there was Gödel. In recursion theory we had Steven Kleene and Emil Post and Alan Turing: Post pointed the way toward mathematical structures in sets and degrees, Kleene toward higher order recursion theory. In model theory we had Alfred Tarski and Abraham Robinson and Leon Henkin, developing the first deep universal algebra. In set theory we had Gödel himself.

But what about proof theory? There we had the papers of Gentzen and Herbrand and Schütte. They were easy to read, but their global mathematical structure and significance were difficult to make out.

[1] I eventually did my thesis with MacLane, in 1956.

What do you gain by reducing consistency of the first order theory of natural numbers to correctness of a primitive recursion up to the first ϵ number? For me, Kreisel was a shining light trying to figure out the structure and significance of proof theory.

One import of Kreisel's papers was that there was beautiful mathematical content and structure in proof theory. We now know this well due to the Curry-Howard isomorphism and the works of Girard and Martin-Löf, but they have all been conditioned by Kreisel's viewpoint.

Another import of Kreisel's papers was that proof theory could be sharpened to understand and establish the constructive content of classical mathematical theorems.

I use the phrase "import of Kreisel's papers" advisedly. For the weakness of Kreisel's papers is in mathematical follow-through. He is extremely insightful and original, and he has new ideas in many papers. But they have often been still-born because they are not developed completely or precisely enough (*Satz-Beweis* style) in the technical sense. Usually the climb upward for secure mathematical knowledge is based on exact proofs and definitions: they build a trail that others can quickly climb to reach a plateau, and continue on from there. In Kreisel's case, one has to rebuild and move and strengthen every detail of the trail, always reaching a slightly different plateau, before being able to continue to climb to greater heights.

This is not altogether bad. The 17th and 18th centuries are replete with famous men whose ideas and methods outstripped their definitions and proofs. In their time they could have published little if held to a higher standard. It is ironic that Kreisel, in shedding light on the mathematical structure of proof theory, did it in the style of that earlier carefree period, without accurate statements or proofs. I often have thought that his contributions would have been greater had he been subjected to the rigorous training of a Chicago or Princeton graduate school in mathematics, where one really has to write down definitions and proofs completely and correctly, instead of having somewhat diffuse training in mathematics, and the standard of exactness of much of the philosophy of the time.

From long experience, all working mathematicians know that there is a preliminary period of rapid advancement in ideas without worrying about exact definitions and proofs, after which there is very hard work to go from that level of accuracy to finished mathematics, where the bugs in definitions and proofs are gone, and concepts are quite clear. A lot of things change in the process. This is the essence of finishing mathematical work. This latter finished aspect is totally absent in Kreisel's single-authored papers.

But, on the other hand, he was able at least twice to secure the collaboration of first class technical mathematicians to finish the technical phase, and establish a plateau to which others could rapidly ascend as a base for a further climb.

A first instance is his work with Troelstra, who made sure all details were finished and in place. That work is tedious to read due to infelicitous notation, but everything is fully written out. This gave later proof theorists a place to look for definitions and proofs, and an excuse to curse the notation.

The second instance is Kreisel's collaboration with Gerald Sacks on metarecursion theory. The Sacks-Kreisel paper was Kreisel's revelation, made manifest by Sacks. This led to an immense outpouring of higher recursion theory, through Sack's Cornell, MIT, and Harvard students, and then through their students to the present day. Of this Kreisel can be proud. It is unlikely that this would have happened without him.

Meetings with remarkable men

I met Kreisel first at the great 1957 AMS *Summer Institute in Symbolic Logic*, organized by J. Barkley Rosser at Cornell. There had never previously been a world-wide mathematical logic conference, because bringing the people together is expensive, and mathematicians and philosophers were paupers. Rosser was well-placed with all government funding agencies, including NSF, and got the money for it pretty easily. It was heady stuff for me as a new Ph.D. when Halmos told me to attend and meet all the great names (except Gödel), from Tarski and Kleene to Quine and Church, and to spend five weeks talking to them, and to deliver my own work.

What a conference! There has been nothing else in logic remotely comparable. I proposed a similar one at Cornell 25 years later, but AMS and NSF said logic was now too large and five weeks too long, and so we limited the repeat to recursion theory.

At the 1957 conference even Gödel was indirectly present, represented by Kreisel, who brought along from Gödel a paper on primitive recursive functionals of higher type and the consistency of arithmetic. Kreisel also gave another talk, on countable functionals, somewhat derivative of Kleene's work. I gave one on general topology and partial recursive functionals, also somewhat derivative of Kleene. Both of us gave other talks as well, but we both noticed the coincidence of interests. Kreisel was constantly at tea after and before sessions, talking about everything, with a truly commanding knowledge of proof theory and constructive methods. I went on after that meeting to an NSF postdoc at the Institute in Princeton, where I had lots of time with Gödel and filled in the remaining gap in my acquaintance with the major logicians. There I was much intrigued with Myhill's new theory of combinatorial functions, and I began to extend it as soon as I had tried isolic arithmetic.

In 1959, after Princeton, I joined Cornell. Sacks was in process of completing preparation for starting his thesis under Barkley Rosser. Barkley was gone that year 1959-60, to set up the Communications Research Division of IDA in Princeton. I was left with instructions from him to prep Sacks as much as possible. Sacks and I spent a lot of time together devoted to understanding the current mathematical content and future outlines of recursion theory and the rest of logic.[2] He got a good start on his thesis, finished it the next year with Rosser in Princeton, and returned in 1962 to Cornell as an Assistant Professor. In many respects, the work of his and my students has been completely intertwined for the succeeding thirty years. Though we never did a joint paper, we often have references acknowledging debts to each other.

A complaint on isols

In 1962–3 I was again in Princeton, on my first sabbatical, and so was Kreisel, living in the countryside. He knew I was working on the Dekker-Myhill theory of recursive equivalence types, in which the isols represent a non-standard definition of finite set. He had mentioned that what he did not like about the recursive equivalence types was that an *uncountable* collection of subsets of the integers was being compared by a *countable* set of 1–1 partial recursive functions. He thought this unfair because it was too easy to do diagonalizations.

I answered this complaint in 1965 by giving Louise Hay the thesis problem of establishing similar results in the *countable* domain of co-simple sets, which she did. This was the beginning of her brilliant career, cut short by cancer.

Twenty years later, McCarty showed that the theory that Dekker, Myhill, and I developed for isols is metamathematically the theory of stable Dedekind finite cardinals in a realizability model of Intuitionistic Zermelo-Fraenkel set theory, itself a generalization of Kreisel-Troelstra realizability (more on this below). A crude summary of the relation is that what is proved intuitionistically about such cardinals in IZF gives rise to a corresponding theorem on isols, while the counterexamples in the isols give independence proofs for IZF. (There is an equally

[2] Although Sacks and I are about the same age, I got an earlier start because he spent four years in the Army, while I was in school (and also doing classified work).

interesting relation between recursive algebra and analysis and IZF, which no one has worked out as yet.)

In the 1960's Kreisel hoped for, but did not see, a real relation of the isols to intuitionistic mathematics, or to his own work, as was later provided by IZF. (But then, neither did I.) Thus the isols turned out to be somewhat closer to his own interests in intuitionistic mathematics than he realized.

To me this is similar to what happened with Saunders MacLane, my advisor. I heard him say, many years ago and in many contexts, quite unkind things about intuitionistic mathematics. But then it turned out that the theory of *topos*, derivative of his work in category theory, is just another language, albeit more elegant, for intuitionistic analysis.

The mystery of how to unveil interrelations between mathematical subjects can only be dimly penetrated. Very often I think that what we all think is individual creativity is mostly determined by the historical epoch in which we work. Thus, subjects that looked quite different thirty years ago seem much the same thirty years later.

More (possible) complaints

Kreisel was quite even handed: he had the same complaint about the theory of Turing degrees of arbitrary sets as he had for isols, namely that an uncountable collection of subsets of the integers is compared by a countable set of partial recursive functionals.

This time it was Sacks who answered the complaint by initiating the detailed theory of r.e. degrees in his famous monograph, which his and my students followed up tenaciously for generations. (The best of mine in this area were Lerman and Soare; he had Harrington, Simpson, Shore, and Slaman.) In the opposite direction, when in 1967 Errett Bishop extended to functional analysis the ideas of recursive analysis, on which Kreisel had worked in the 1950's, Kreisel expressed a lot of admiration. A look at Errett's work reveals, as Kreisel well knew, that every theorem can be construed as a recursive analysis theorem, in which the continuum on which one is operating is the classical one, and only *recursive* maps (partial recursive functionals) between continua are used. (This is not an interpretation that Errett liked. The fact that this is possible is because he regarded the continuum as perpetually unfinished, and finishing it with the classical continuum harms no argument.)

The whole theory of recursive analysis, construed in this way, is subject to the same criticism Kreisel had for isols and degrees: a very large number of sets are being mapped by a very small number of constructive functions. The funny thing is that this seems to be the best way to understand recursive analysis. The theory based on just the recursive reals is horrible and does not resemble classical analysis at all. It has been much pursued in Leningrad by the Shanin school and consists, like recursion theory, mostly of counterexamples.

Perhaps the reason for the success of the theory with arbitrary real numbers and constructive functions, written up by Pour El and Richards in their well-known book, is that analysis is used to describe the world. The real numbers that show up as physical measurements are not presented as recursive real numbers. This is well modelled by the classical continuum. But when we try to do physics and predict using these measurements, we use algorithms. This means we manipulate *arbitrary* reals using *recursive* functionals. The disparity in cardinality between the objects being manipulated and the constructive functions that map them is explained.

I would not be surprised if other such disparities, such as the original ones Kreisel mentioned, in the end may similarly be explained as being reasonable in a context.

Here is one. Look at a distributed computing processes reacting perpetually to the environment, that is, subject to a perpetual disturbance from the environment. The disturbance is an infinite sequence. The distributed system operates on a disturbance to produce its own state sequence. Here again we have countably many operators describing constructive distributed systems which map a continuum of disturbances to a continuum of state sequences.

Two main contributions

As said at the beginning, I regard two of Kreisel's contributions to logic as his mathematical inheritance.

The first is the Kreisel-Troelstra realizability interpretation for higher order intuitionistic type theory, of the late 1950's. The basic idea is that evidence for a higher order universal quantifier is the *same* evidence for all instantiations, rather than different evidence for different instantiations. Twenty years later, at Dana Scotts' instigation, his student McCarty extended this interpretation to IZF, Intuitionistic Zermelo-Fraenkel set theory.

This interpretation explains my work on recursive equivalence types and isols in intuitionistic terms, as I already said. In addition, Kreisel's HEO interpretation of intuitionistic analysis led to the PER interpretation, and is of value in understanding Girard's system F and its generalizations.

Kreisel's other contribution, recursion theory on the ordinals, has no known tight relation to intuitionistic mathematics. But one might in the future come out of investigation of the recursion theoretic complexity of realizability interpretations of IZF or a similar intuitionistic set theory.

This is how it happened that I was *not* involved in recursion theory on the ordinals. In September of 1962 I was interested in Kreisel's opinions on problems in which we might share a common interest. Kreisel always had a terrible problem with disturbed sleep, in case there were any sounds whatsoever at night.[3] The house he had chosen near Princeton was quite remote, and I had trouble finding it. As I entered, he offered me exquisite cookies and a quite exotic array of teas. He was always an excellent if slow cook. We had an afternoon tea. Then we discussed his preference for a different non-standard notion of finiteness than that offered by Myhill-Dekker, what we now call metafinite sets (hyperarithmetic sets of recursive ordinals). We discussed the possible structure of such a subject. I had been referee for several of Clifford Spector's papers, and knew Kleene's and Addison's and Spector's work completely, so this was a quite informative conversation. But I was at the time more interested in finishing the theory of recursive equivalence types and combinatorial functions, and did not take this up. Within a year or so, Sacks did, which led to their joint work and recursion theories on the ordinals.

Anecdotes and myths

Kreisel liked a form of self-advertisement which detracted from his professional reputation. Whether this should be judged as a fault or merely dismissed as an eccentricity, we leave to the reader.

His self-advertisement had nothing in common with the one of mathematicians and philosophers with an exaggerated sense of their historical importance. We all accept that there are those who have spent their lives in the cloistered professoriate, removed from the real world, and who have never lost their sense of youthful omnipotence. They may be forgiven for their unwarranted sense of self-importance. But Kreisel was scientifically very modest, and deferred to the great men of mathematics with as much reverence as a groupie has for the latest

[3] In about 1964 I invited Kreisel to Cornell and put him up in an empty apartment building away from everyone and every noise, I thought.The next morning I asked whether he had got a good night's sleep. He said that several rooms away the building thermostat's clicking had kept him up. He finally found it and ripped it out completely. The rest of the night was fine.

rock star or basketball player.

His form of self-advertisement was different. Simply put, he really enjoyed being outrageous. His conversation was laced with references to rich and famous friends on the Riviera, and orgies in European castles with titled ladies.

One of his anecdotes was about picking up sleeping companions on the Riviera. He walked down the beach striking up a conversation with every unaccompanied uncovered beautiful woman taking a suntan, and asked each to sleep with him. He said the success ratio was one in ten. When Verena Dyson left Freeman Dyson to live with him, he delighted in introducing her as "This is my wife, Mrs. Dyson". I have some doubts as to how comfortable this made Verena.

Once he was asked at a party what was the age of the youngest he had slept with. The answer was instant: "girls 11, boys 9". True, or for effect? And why answer at all?

Mathematicians and philosophers are generally mild and conventional. They regarded this whole series of behaviors as unsavory. But they do like gossip, and Kreisel liked to provide it. He liked to mythologize himself. This was certainly entertaining in a close circle of friends, unsavory or not. But the self-created gossip over the years has tended to make him unwelcome in many places, and to overpower his scientific reputation. He was too successful for his own good in projecting his image of decadence. But he has enjoyed it.

Final judgement

Kreisel could see structures that pure technicians missed, and so was often very helpful. He had much talent, but not enough concentration to be more than an influence on others. His permanent contribution to the edifice of mathematical logic has been assured by his collaborators' industry. He cannot outlive his self-engendered notoriety, but the latter will be buried by the sands of time.

Some Reminiscences of Kreisel

by Rohit Parikh

Since these are personal reminiscences, they will be about myself as well as Kreisel. I knew Kreisel fairly well for a period of about 18 years before our professional lives drifted apart, partly because of his move to Austria but even more due to my increasing involvement with logical problems arising in Computer Science.

Kreisel was one of my three 'advisors'.[1] As a Math student at Harvard interested in Logic (1959-61) I had been dismayed to find that there were no logicians in the Mathematics department. I decided to ask Burton Dreben in Philosophy, whose infectious enthusiasm had been responsible for my going into logic in the first place. He was willing but unfortunately he had a sabbatical coming up very soon. Thus it was that a number of us students at Harvard ended up working with Hartley Rogers, who was already then at MIT.[2] He very kindly agreed to take us on and grounded us thoroughly in recursive function theory. Hartley was a superb teacher and I at least acquired the feeling that there was nothing in recursive function theory that I could not attack. He suggested that I work on developing a hierarchy for recursive functions similar to the Kleene hierarchy for hyperarithmetic functions.

[1] The other two were Hartley Rogers, who actually worked with me and Burton Dreben, who signed my thesis. All three helped me quite a bit, but in different ways.

[2] Somehow we were too mathematically inclined to think of working with Quine himself.

There were several people working on such hierarchies. Kleene and Axt (see Axt [1965]) were working on one which had a sort of primitive recursive jump operator and Feferman ([1962]) was working out a hierarchy of theories based on Turing's notion of ordinal logics. It turned out that my fellow student David Luckham had a copy of a Stanford report containing Feferman's paper and I borrowed it from him.

However, this report also had several other papers in it, including one by Kreisel. In his paper [1959] he showed that hierarchies for classifying low level objects (like r.e. theories or recursive functions) were likely to break down, i.e. suffer from the problem of non-uniqueness.

Hierarchies in recursive function theory are not defined directly on ordinals but on Church-Kleene $\mathcal{O}$, a set of *notations* for recursive ordinals. Thus one associates one's objects (theories or classes of functions) with notations for ordinals. Since the same ordinal may have many notations, one needs to show uniqueness, i.e. that the same object will be associated with all notations for the same ordinal. If the hierarchy associates different objects with two notations for the same ordinal, then we say that such a hierarchy suffers from the problem of non-uniqueness.

Now the problem of deciding whether two notations in $\mathcal{O}$ represent the same ordinal is Π^1_1-complete. Suppose one did have an arithmetical hierarchy which did not suffer from non-uniqueness, then one could use this hierarchy to test whether two ordinal notations represented the same ordinal. Thus one would be able to reduce a Π^1_1-complete problem to an arithmetic one, which is impossible. Thus any arithmetically defined hierarchy must suffer from non-uniqueness from some ordinal onwards. All the hierarchies then being considered were of this kind.

Kreisel considered two popular kinds of hierarchies, of sets and of sets of functions, which were then being studied. He showed ([1959] that such hierarchies must always exhibit non-uniqueness by ordinals ω^3 and ω^4 respectively. However, he left open the question whether these were the lowest possible ordinals. I managed to show ([1967]) that in fact such hierarchies would always break down at ordinals ω^2 and ω^2+1 and gave examples to show that these two upper bounds were the best possible. I communicated my results to Kreisel and eventually I received an offer of an instructorship at Stanford where both Kreisel and Feferman were. I was offered the 'princely' salary of \$6,000 a year and actually it was quite adequate.

This result of Kreisel on hierarchies is a very typical example of the insight that he had into logic, an insight that enabled him to guess when a certain 'obvious' answer was wrong. Some years later he showed a similar insight by proving that even though maximal paths through $\mathcal{O}$ were at best Π^1_1-complete, (as shown by Feferman and Spector [1962]) one could in fact *extract* very little information from them with many-one reducibility. This observation of Kreisel's [1972] led to three other papers: Friedman [1976], Jockusch [1975] and Parikh [1973a]. There are of course more dramatic examples, where his philosophical views led him to see certain connections, but even very technical results like the two I have described showed the same touch.[3]

I actually met Kreisel just after I left Harvard, in summer 1961 at the Institute for Advanced Study at Princeton. I had been working at Bell Labs that summer and so I went up to Princeton to see him about my thesis and also to tell him about another problem that I had solved – a problem about many-one degrees of recursive well orderings, due to Kreisel, Shoenfield and Wang [1960]. Kreisel told me that he had also received a letter from Shih Chao Liu, a student of Kleene's, outlining a solution to the same problem, but that he had not read the proof yet and thought I should go ahead and publish an abstract of my results [1961] in the *Notices* of the AMS. As it turned out, Liu's solution [1963] was correct and also more comprehensive than mine and I never published my results in a paper.

During my first year at Stanford, (1961-62) Kreisel was not there and I talked mostly to Feferman and Myhill and sometimes to Paul Cohen. Paul Cohen lent me his car to look for a place to stay, but he did want to make sure that I had a driver's license. This care seemed superfluous to me since I had just driven to Palo Alto from the East Coast! Kreisel returned during 1962-63 and became for the next few years, the most important single influence in my research life. Many of the papers that I wrote during the next dozen years came out of Kreisel's questions. But it *could* be hazardous listening to him. When Kreisel gave his opinion on something, it invariably turned out that he was right about two thirds of the time and it did not seem to matter whether he said that he was just guessing or whether he said that he had *proved* it.

He asked me a question involving primitive recursive functions and ordinals[4] and told me that he had verified that the answer was positive for ordinals up to ω^n for any n. He asked me to find a counter example at ω^ω. He was convinced that ω^ω was some sort of a natural limit for primitive recursion. However, no matter how hard I tried, I could not bring ω^ω into the picture at all. Finally I discovered ([1966]) what was

[3]Kreisel's result was based on some previous work of Feferman and Spector, and I myself proved a stronger result using Kreisel's insight, but Kreisel's own work had a quality of mystery about it that my paper at least did not have.

[4]The question was, if R is a recursive linear order with no primitive recursive descending sequences, then show that the same holds for R^α where α is an ordinal.

wrong; the smallest ordinal for which the answer was negative was not ω^ω but 2!

However, things were quite different with the Kreisel conjecture on the length of proofs. Kreisel asked Myhill and me if we thought it was likely that if some formula $A(x)$ of Peano arithmetic was such that all its numerical instances $A(n)$ could be proved in k lines where k was a fixed integer, then in fact $(\forall x)A(x)$ would also be provable. Myhill and I were quite contemptuous of such an implausible conjecture and immediately set out to find counter examples. When several days passed without any progress, my disdain turned to a grudging respect and I started to work in the positive direction of the problem.[5]

I once had occasion to witness a remarkable astuteness on Kreisel's part. Bill Tait and I were sitting in a seminar which was just about to begin and I asked him a technical question about something. Before he could answer, Kreisel who had just entered the room answered my question. We were amazed since he could not have heard more than the last half of the question.[6] "What else could you have asked?", said Kreisel to me, smiling. This ability to guess what would be some sort of natural limit to someone else's knowledge begins to approach magic! It was common for Kreisel to tell you a proof of a theorem and then, if you did not like it, offer another. He seemed to believe, a belief apparently also shared by Dijkstra, though in his case about programs, that proofs did not exist only to establish theorems, but that giving aesthetic pleasure was also an important function of theirs. So you had every right to ask for a proof that was pleasant to you.

Kreisel has a habit of writing in footnotes. In his papers, while discussing a point, he may wander off into a side-issue and thence to a still further side issue. By the time Kreisel returns to the main point, the reader has often forgotten what it was in the first place. However, his habit of being cryptic and leaving his t's uncrossed and his i's undotted *can* be helpful. Once, when I was visiting Caltech, Kreisel wrote to me about a conservation result regarding non-standard analysis and standard analysis which he was planning to present at a meeting, also at Caltech, later on. As usual with him, he did not say very explicitly what he had in mind and I could not ask him, because he would think I was too stupid to figure out what he *must* have

[5] I eventually did get a positive answer ([1973]) for a version of Peano arithmetic where plus and times were taken to be ternary predicates, but I never did show the result for the case where Peano arithmetic is axiomatized with these two as function symbols. A solution to the latter problem was eventually announced by M. Baaz.

[6] Tait seemed unwilling to altogether rule out the possibility that Kreisel had been listening at the door.

meant. What is worse, he might not keep his thought to himself! So I said nothing to Kreisel, but soon, on a visit to Buffalo next week I told John Myhill about this over dinner. Myhill said at once, "Kreisel probably meant this ...". I explained that that wasn't possible since there was an obvious counter example. "I conjecture that that is the *only* counter example", said Myhill, and it turned out that he was right. I found out later that Kreisel's result [1967] as completely different and so we had a new theorem (Parikh [1967a]). Myhill very generously insisted on giving me the entire credit and so his name appeared only in an acknowledgement in a footnote.

When I decided to return to India after two years at Stanford, (1961-63) Dana Scott sent me a message through Kreisel that I was making a mistake and there was an opening at UCLA that I should consider. While conveying this message, Kreisel himself did not endorse it. His own view was that if one was isolated in India and did not have to listen to other people's ideas, that could only be a benefit. It turned out either that he was wrong in this or more likely that I did not yet have the strength of character to 'benefit' from my isolation. In any case, in 1965, Kreisel helped me to move from Panjab University where I had been teaching in India, to Bristol where Shepherdson was building up a strong logic group.

Eventually I returned to the US in 1967, ending up at Boston University, where in due course, a letter from Kreisel almost cost me my tenure. According to the chairman, Kreisel's letter had in fact been quite positive, but the committee was expecting something sugar-coated and was taken aback to see that some of Kreisel's reservations about me were also in the letter.

During this period, (the 70's) Kreisel often visited me in Boston and liked to stay at my house which was relatively quiet. However, the windows of the attic where he stayed had to be covered with blinds (he carried these with him) and we were not permitted to flush the toilet once he went to bed. Since the attic had no toilet, he himself carried an empty milk carton up to bed. My wife still recalls with amusement the sight of Kreisel coming down in the morning with his milk carton and asking her what to do with it.

For some reason he had decided that my wife was a midwestern farmer's daughter and therefore a farmgirl.[7] He spoke to her quite confidentially about himself, more than he ever did to me, but neglected

[7] Actually Carol's father had owned a printing business and she herself had degrees from Chicago and Washington universities, but Kreisel preferred his own image to the reality. Later on, he did drop his fantasy and took to discussing *Middlemarch* with her.

to ever thank her for the work she had to do when he stayed with us.[8] Eventually, when I brought up this issue and reproached him, he immediately went out and bought her two dozen roses.

Kreisel seemed to move in society circles which academics did not usually frequent. He seemed to be well acquainted with Yul Brynner's movements (whose niece was at Stanford) and he very proudly introduced my whole family to the Romanovs who lived in the Los Altos Hills. Not all logicians admired him for this worldly side of his. I remember telling a French logician that Kreisel seemed to know a lot about wine and women and hearing the response, "In that case why don't they give him a professorship in wine and women?". However, this European sarcasm, which was Kreisel's own metier, did not come easily to Americans and the occasional complaints about him that I heard from American logicians were always more soberly expressed.

Kreisel does seem to have had much more than his fair share of spats with people, and the other people were not always at fault. At the Buffalo meeting in 1968, organised by Kino, Myhill and Vesley, Kreisel asked me if Americans would know the phrase "What do they of England know, who only England know". I said that I thought logicians would, but wondered why Kreisel needed to know. Then later I understood when Kreisel launched an attack on Burt Dreben and John Denton (who were presenting the paper [1970] on Herbrand style consistency proofs) with the theme, "What do they of Herbrand know, who only Herbrand know". Dreben has never known that I owe him an apology.

Once when he and I were both at Stanford, Kreisel asked me to go to the San Francisco airport to pick up a visiting logician. When I protested that I had not met the man in question and would not recognise him, Kreisel said, "It is very simple, he will be the strangest looking person to walk off the plane"! I somehow managed to get out of it, but in fact Kreisel's description was not at all accurate. The visitor in question – when I finally met him – was no stranger looking than logicians are apt to look anyway. But this remark does show a certain tendency on Kreisel's part to regard as objective certain opinions of his, which were not even inter-subjective!

As for my own relations with him, I have to admit that even though I was always a little afraid of him, he never did harm me in any way. John Myhill was astonished that I would let my children play with

[8]Once, when he visited us in Buffalo, he stayed with us and we gave a party in his honor. At ten sharp, he headed up to bed, leaving the guests to entertain themselves, and his words of thanks to my wife were, "you made too much food, didn't you?".

Kreisel, but in fact he was always courteous to them. They always got chocolates or truffles when he visited us and my son's first water pistol was a gift from Kreisel.[9]

I learned a lot from Kreisel and while my association with him was not always to my liking, this was far more my fault than his.[10] At that time I did not fully grasp the grief he must have felt during his own growing up in England while war was raging in Europe, and that emotional support was one thing which I could not expect from him. I was also more used to the much milder manner of American teachers who were more tolerant of their students acting like children. However, Kreisel was extremely reliable in many concrete ways. He almost always responded to letters by return mail and when I was in India, he meticulously paid my dues to the AMS and the ASL out of funds I had left with him for the purpose. I owe him quite a lot, even though I am perfectly aware that his capacity for mischief often exceeded the patience of his friends and colleagues.

In the *Mahabharata* by Peter Brook, he presents Krishna as a person both Machiavellian and philosophical, somewhat bored with the world. I don't think that the actual Krishna was at all like that but Kreisel did in fact seem like that to me. He did say once to me that when people thought he was angry with them, in fact he was just tired. I wonder if he recognises himself in this little portrait of mine.

References

Axt, P.

[1965] Iteration of primitive recursive functions, *Zeit. Math. Logik*, 11 (1965) 253–255.

Crossley, J., and Parikh, R.

[1963] On isomorphisms of recursive well-orderings", *J. Symbolic Logic*, 28 (1963) 308.

Denton, J., and Dreben, B.

[9] My son Vikram was only three when he met Kreisel. Soon after that, we hired a man from the state unemployment agency to wash our walls before we painted them. Unfortunately the man stepped on Vikram's electric train and broke it. Vikram ran to his mother: "Mummy, mummy, Professor Fooey broke my train!" Somehow, Vikram had acquired the belief that *all* men were to be addressed as 'professor'.

[10] Other research of mine influenced by Kreisel but not mentioned in the text is in Crossley and Parikh [1963], and Parikh [1972].

[1970] Herbrand-style consistency proofs, in *Intuitionism and Proof Theory*, eds. Kino, Myhill and Vesley, North Holland (1970), pp. 419–433.

Feferman, S.

[1962] Transfinite recursive progressions of axiomatic theories, *J. Symbolic Logic*, 27 (1962) 259–316.

Feferman, S., and Spector, C.

[1962] Incompleteness along paths in progressions of theories, *J. Symbolic Logic*, 27 (1962) 383–390.

Friedman, H.

[1976] Recursiveness in P_1^1 paths through $\mathcal{O}$, *Proc. Amer. Math. Soc.*, 54 (1976) 311–315.

Jockusch, C.

[1975] Recursiveness of initial segments of Kleene's $\mathcal{O}$, *Fundamenta Mathematicae*, 87 (1975) 161–167.

Kreisel, G.

[1959] Non-uniqueness results for transfinite progressions, *Bull. Acad. Pol. Sci.*, 7 (1959) 621–626.

[1967] Axiomatizations of nonstandard analysis that are conservative extensions of formal systems for classical standard analysis, in *Applications of Model Theory*, ed. W.A.J. Luxemburg, Holt, Rhinehart and Winston, 1967, 93–106.

[1972] Which number-theoretic problems can be solved in recursive progressions on Π_1^1 paths through $\mathcal{O}$?", *J. Symbolic Logic*, 37 (1972) 311–334.

Kreisel, G., Shoenfield, J., and Wang, H.

[1960] Number-theoretic concepts and recursive well-orderings, *Arch. Math. Logik*, 5 (1960) 42–64.

Liu, S.

[1963] On many-one reducibility, *J. Symbolic Logic*, 28 (1963) 35–42 and 143–153.

Parikh, R.

[1961] Many-one degrees of certain sets of recursive well orderings, *Notices Amer. Math. Soc.*, 8 (1961) 495.

[1966] Some generalisations of the notion of well ordering, *Zeit. Math. Logik Grund. Math.*, 12 (1966) 333–340.

[1967] Non-uniqueness in transfinite progressions, *J. Indian Math. Soc.*, 31 (1967) 23–32.

[1967a] A conservation result, in *Applications of Model Theory*, ed. W.A.J. Luxemburg, Holt, Rhinehart and Winston, 1967, pp. 107–108.

[1972] A note on rigid substructures, *Proc. Amer. Math. Soc.*, 33 (1972) 520–522.

[1973] On the length of proofs, *Trans. Amer. Math. Soc.*, 177 (1973) 29–36.

[1973a] A note on paths through $\mathcal{O}$, *Proc. Amer. Math. Soc.*, 39 (1973) 178–180.

Kreisel, Generalized Recursion Theory, Stanford and Me

by Richard A. Platek

Kreisel played a formative role in my early work in Generalized Recursion Theory; he also played a formative role in my subsequent disillusionment with the life of academic research. In this note I will try to make clear the circumstances surrounding these events. The remarks are personal and perhaps out of place in a collection of scientific papers. I apologize.

After an undergraduate education at M.I.T. I accepted a graduate fellowship in the Stanford Math department in the Fall of 1961. Although a Mathematics major at M.I.T. I had first been attracted to logic by the lectures of the linguist Noam Chomsky. During my senior year I had attended classes at Harvard given by the philosophers W.V. Quine and Burt Dreben. Thus, although receiving a strong mathematics and scientific education, I had never heard a mathematician lecture on logic before attending Stanford! This peculiarity led me to believe that foundational studies where not of the same kind as other mathematical studies; a logician learned and applied mathematics in order to develop and explore philosophical hypotheses in a manner analogous to the way a theoretical physicist applied mathematics to develop and explore physical hypotheses. For me at that time, logic was a branch of philosophy. This conditioning made me a ripe candidate for Kreisel's influence.

The Stanford graduate program in Logic and the Foundations of Math was a new initiative financed by the National Defense Education Act. It aimed at building a center of excellence in Logic at Stanford. For graduate school I had to choose between Berkeley and Stanford. The former had a world famous, well developed group under the direction of Alfred Tarski; the latter was a start-up. Since my entering class would be the first in the new program it appeared that I would be able to have closer contact with faculty than if I attended Berkeley. While the contact at Stanford turned out to be a little too close at times it was a good choice. There were four incoming graduate students in the program with about an equal number of faculty. Berkeley graduate students, on the other hand, were notoriously unnoticed and ill treated. This contributed to the various radical outbursts which began there in the mid sixties; as I write this in December, 1992, I read that Berkeley graduate students are again on strike for better working conditions. At Stanford, on the other hand, we did have our faculty murder at the hands of a disgruntled graduate student; perhaps demonstrations and strikes are better.

Stanford was a logician's paradise. During my stay (1961–65) the faculty included John Myhill, Sol Feferman, William Tait, Dana Scott, and others like Joe Shoenfield and Bill Howard who were visitors. One should of course also mention Paul Cohen who during this period turned his attention to logic with lambent achievements. I had known Cohen at M.I.T. where he taught me linear algebra in a rather abstract manner way before such an approach had entered the curriculum. In addition, we both came from the same section of Brooklyn having lived only a few blocks apart. This shared ethnicity led to an easy going relationship. But of all the people at Stanford the one who made the greatest impression on me, for good and for ill, was Georg Kreisel.

In graduate school I discovered that I was a competent mathematician. Indeed, I was so good that Paul Cohen, before he turned his attention to logic, tried to convince me to leave foundations and focus on real mathematics. He was convinced I was wasting my time with logic; it was not fruitful to worry about what logicians seemed to be preoccupied with. While I really loved mathematics (perhaps too much; my preoccupation with studies played no small role in my divorce while a graduate student) I was irresistibly drawn to logic because of my philosophic tendencies. Since it was this philosophic interest which subsequently led me away from active research in mathematics perhaps Cohen was right. At Stanford, Kreisel was the most "philosophic" of the logicians. Through a continual insistence which amounted to almost a propaganda campaign he inculcated into both colleagues and students the need to reflect on mathematical activity and to subject the activity itself to mathematical analysis.

It is not normal to be philosophic. Normal people are involved in the living of their life. They haven't the time or inclination to reflect much on the meaning of things. If anything is the antithesis of reflection it is passion. And mathematicians in their mathematics are passionate people (despite how they might appear in other contexts). They are too enamored of the beauty and power of their discoveries and pursuits to stop and ask "What's happening?" To ask a mathematician in the throws of creativity to stop and reflect on the process is rather like *coitus interruptus* (to speak this way is one of things I learned from Kreisel; such sophistication was not known in Brooklyn). So the philosophically inclined are a bit abnormal, a bit eccentric, a term many would use for Kreisel who studiously cultivated his eccentricity. Kreisel's mathematical interests were always philosophically motivated. He was, for example, interested in intuitionism and issues of constructivity at a time when hardly anyone mainstream was. Today, constructive logics with their accompanying type theories are a major growth industry in theoretical computer science. Kreisel knew that they should be for philosophical reasons and constantly pointed researchers like Scott and Girard in that direction. I recall his trying to interest me in applying my results in higher order, ordinal and set recursion theory to proof theory. I didn't take the advice but several of the possibilities he mentioned I later found worked out in Girard's theory of Π^1_2–logic.

I didn't take his advice because I didn't really understand what he was talking about. The truth was I never (well, hardly ever) understood what Kreisel was talking about. He was not the greatest lecturer. Nor was he the greatest writer; I could hardly ever read his papers. Nor could many other people. But I maintain that his influence was immense. Many students and colleagues who couldn't follow him basically dismissed Kreisel as unimportant. They were wrong. For example, I believe Kreisel played a pivotal role in Cohen's set-theoretic independence work even though the paper trail might be weak. Cohen's original discovery of forcing was motivated by an attempt to prove analysis consistent. The idea was that statements which seemed to involve infinities could be reduced to pieces of finite information. He gave lectures on his ideas to the logic group at Stanford and through their criticism his ideas matured into the independence results. While in no way denigrating Cohen's achievement I strongly feel that had Kreisel not dominated the logic culture at Stanford it would not have been natural for Cohen to have taken up Hilbert's programme or to have attacked it in the manner he did. My opinion in this matter is worth something, I was present at the creation and knew all parties. A rift opened between Cohen and Kreisel which I believe was caused by Kreisel's insistence that he had anticipated the notion of forcing in some work of his on the logic of intuitionist free choice sequences. Cohen being a mathematician could dismiss such claims; I am sure he had never read much of Kreisel. Kreisel certainly made things like this personal; Cohen told me that Kreisel had said that from now on they would only communicate in writing. People in Brooklyn hardly ever said that to one another.

Although others had their difficulties both intellectually and personally with Kreisel I was fascinated with him. He was also extremely kind to me; dining with me and teaching me European manners. Since he had a joint appointment with Stanford and Paris I did not see him all that often. All my creative work was done while he was away and was presented to him as a fait accompli when he returned. I was not aware of his work with Sacks on metarecursion theory when I worked out the theory of recursive functions over admissible ordinals. I had seen something of his on the definability of predicates over the hyperarithmetic sets which matured into metarecursion theory (see Sacks' article) but it was a definability result not a recursion theoretic result. I think it became the latter through his collaboration with Gerald. At the time I did my work I was also not aware of Kripke's work. My inspiration came from Takeuti's papers on ordinal recursion functions. Takeuti developed a theory of recursive functions over an initial ordinal and I just posed for myself the question of what properties an ordinal had to have in order for Takeuti's theory to go through. This is the kind of question one is taught to ask as a mathematician. I called the ordinals so defined "recursively regular" because the main property in question was similar to the notion of regularity in ordinal theory, namely the sup of the image of a smaller ("finite") ordinal under an ordinal recursive function should be bounded. My thinking was prompted by the conceptions of smallness underlying the axioms of ZF. That is why it is natural for me to formulate what is now known as KP set theory and to use this theory in the development of the ordinal theory. One of my first theorems was that Church-Kleene ω_1, ω_1^{CK}, was the next recursively regular ordinal after ω and that the ω_1^{CK} recursive and recursively enumerable sets when restricted to ω where the hyperarithmetic and Π_1^1 sets respectively. It came as a complete but welcome surprise to me to learn that this case had been also discovered by Kreisel–Sacks and that much of my work was independently discovered by Kripke. The pleasure came from knowing that what I had discovered must be important if others, much more famous than I, were doing it at the same time. I started the project of relating all this to the theory of recursive functionals of finite type which I had learned from Kleene's papers. Kreisel had me present these papers in logic seminar before leaving for Paris and all my investigations were prompted by the question I formulated as to which ordinals were involved in Kleene's inductive definitions of "is defined." Because of Kreisel's emphasis on proof theory it was very natural to pose to oneself questions of characterizing and identifying ordinals.

The amount of material I discovered during that period was so immense I couldn't fit it into a normal thesis. Nor did I publish the results but over the years others have. Why I did not publish is a story which remains to be told.

So I would like to make it perfectly clear. While Kreisel never saw any intermediate versions of my work and while he gave me no specific directions or advice I feel that everything I accomplished was an extension of his interests and teaching. I was carrying out his research program and he deserves credit for promoting the integration of the various "branches" of logic; recursion theory, proof theory, model theory and set theory. Before I began my research he generously talked to me for hours on general philosophical points of view which I half understood but obviously internalized. He planted the seeds. He was my teacher and I am grateful.

But there was a personal side too. As I mentioned Kreisel was very open to me. We met socially and he introduced me to the European style. This was very intoxicating for a graduate student. He arranged for me to have a post-doctoral position in Paris and he introduced me to his various French associates. And then for reasons which were never quite clear there was a falling out. To repeat the details would be spiteful. I will just say that I took it quite hard; in fact, I was traumatized. Unlike the usual superficial relationship between thesis advisor and student there was a deep bonding between us, at least on my side. As odd as it might seem to some (love always seems odd), Kreisel had become my ego ideal. I lost all interest in certain aspects of academic life and turned to other areas of philosophy. Kreisel and I have not communicated for over twenty five years. It is time for amends.

When I was young I believed that the pursuit of scientific Truth was spiritually ennobling; that virtue was a simple side effect of intellectual activity. People might scoff at my adolescent dream but it was mine and Kreisel (and the other "professionals") broke my heart. But one's first love always does and you recover and move on. I was seeking the "Grand Cause" which would engage my total being. Academic life with all its little peculiarities wasn't it. I am still seeking.

I have had a richer, better life thanks to Kreisel. He was one of my teachers and I am grateful.

Kreisel, Generalized Recursion Theory and Me

by Gerald E. Sacks

> "What do they know of recursion theory, who only recursion theory know?"
>
> G.K.

Kreisel first tried to lure me into his shadowy world of generalized recursion theory (GRT) in 1961. We were lolling in the empty, grassy fields that bound the Institute for Advanced Study (IAS) in Princeton. I was silently occupied with some problems in classical recursion theory (CRT). He whispered through the tranquil air. "Better to reign in the hell of GRT than serve in the heaven of CRT." (Thirty years later I find it plausible he did in fact say that.) His words sounded vaguely familiar. I asked myself, as every friend of Kreisel does: "What is the old devil talking about?"

In retrospect I suppose he meant that GRT was hellish because it allowed infinitely long, convergent computations that cast a long shadow on any attempt to lift up the more intricate arguments of CRT. The search for a proof in CRT could be painful, but after Friedberg and Muchnick discovered the priority method, the rest of us could continue in a spirit of heavenly optimism. Not so in GRT. In that domain only syntactical arguments succeeded. Kleene's enumeration theorem lifted up easily. There was no hope of similarly raising the solution to Post's problem. Or so it seemed to me on the deceptively tranquil green fields of the IAS.

In 1963 Kreisel and I were lazying in the gardens of the Beverly Hills hotel. The smell of decaying flowers wafted by. Not unpleasingly. The old devil sniffed appreciatively. He was taking a day off from the ruling passion of his life to try again to lure me into his dark kingdom. This time he hooked me. I was beginning to understand, as many of his friends eventually do, what he was talking about. My newfound comprehension may have been induced by a waning interest in CRT. Kreisel had found the principle missing in all previous attempts to lift CRT to a place where infinitely long computations were lawful. No useful purpose would be served by expressing the principle in his own words. More simply put: CRT is about the interaction of partial recursive functions and finite sets. A fine example is Friedberg's proof, in his solution of Post's problem, that each negative requirement is injured only finitely often. The old devil realized long before anyone else that it was not enough to generalize the notion of partial recursive. In addition the notion of finite had to be generalized!

Easier said than done. But the old devil, despite his well-documented, self-documenting tendency to confuse free association with philosophizing, had a concrete proposal. Replace the natural numbers by hyperarithmetic sets of numbers, and finite by hyperarithmetic. Call the new notion metafinite. Let Q be a Π^1_1 set of unique notations for hyperarithmetic sets of numbers. Then a metafinite set is a hyerarithmetic subset of Q. And a partial metarecursive function f is simply a partial map from Q into Q whose graph is Π^1_1. Later the proposal was lightened. The individuals became the recursive ordinals. Q became a Π^1_1 set of unique notations for the recursive ordinals. Then a set of ordinals is metafinite if it corresponds to a hyperarithmetic set of notations. Suppose f is a partial function from the recursive ordinals into itself. Then f is partial metarecursive if its graph, ordinals replaced by notations in Q, is Π^1_1.

We presented an abstract and a short talk at a meeting of the ASL in 1963. Kleene was the session chairman. I gave the talk and Kreisel fielded the questions. Kleene asked, "Why meta rather than hyper?" Kreisel replied, "You used up hyper". In the talk I mentioned two applications. The existence of a maximal Π^1_1 set, and Friedberg splitting of Π^1_1 sets. Moschovakis, who was in the audience, observed soon afterward that splitting could be done by conventional means.

At first Kreisel's proposal seemed obscure, but he had come to it via powerful model-theoretic reasoning sketched below. He had already verified that the restriction of a partial metarecursive function to a metafinite set on which it is defined is metafinite. That fact won me over. I was ready to sign in blood. He suggested a joint paper to be called *Metarecursion Theory*. He rightly pointed out that the paper would be worthy of a sequel only if it included a proof of a positive solution to Post's problem. In the gardens of the Beverly Hills hotel I felt a divine confidence that all would go well. The old devil knew what he was talking about. Looking back, I think his idea had aspects that even he did not anticipate.

Now for his model-theoretic reasoning. My account of it is not likely to win his approval. For one I do not accord syntax and semantics the primary role in logic he does. Recursion theory can stand on its own legs. For another, effective computation strikes me as more fundamental than logical reasoning. To me a proof looks like some kind of computation, in general less effective than other kinds, but some computations strongly resist being recast as proofs, most notably in higher types recursion. But that is another story, more self-serving than this one, in another volume. Back in the 1960's there was considerable interest in generalizing first-order logic. It was easy as pie to invent new languages, less so to find new compactness theorems. Scott made an excellent suggestion when he invented $\mathcal{L}_{\omega_1\omega}$: allow countable conjunctions and disjunctions, but keep quantifier prefixes finite. Now it was possible to lift the axioms and rules of first-order logic and to mimic Henkin's method of proving completeness. What about compactness? Not possible for full $\mathcal{L}_{\omega_1\omega}$, but countable fragments looked promising, and eventually Barwise's compactness theorem for countable admissible sets was found. However the essential insight was found earlier by Kreisel in the context of ω-logic: first-order logic augmented by a distinguished predicate restricted to the standard integers.

Kreisel proved: Let S be a Π^1_1 set of sentences of ω-logic; suppose every hyperarithmetic subset of S has a model; then S has a model. This is one of many beautiful theorems proved by the old devil before he devoted himself to mathematical philosophy. Contrast the result with first-order compactness: Let S be an arbitrary set of sentences; suppose every finite subset of S has a model; then S has a model. (1) "arbitrary" has been replaced by "Π^1_1". (2) "finite" by "hyperarithmetic". (1) suggests that notions of computability are needed to generalize first-order logic. (2) tells us that "finite" has to be generalized appropriately, and that in the context of Π^1_1, "hyperarithmetic" is appropriate. From here it is a short Kreiselesque leap to the above definitions of metarecursion theory.

Back in the Sixties, that still lambent decade of logic, I gave a talk at UCLA on GRT. Afterwards C.C. Chang said the talk had added a great deal to the already substantial legend of Kreisel. I hope the same will be said of this brief memoir.

Kreisel and I

by Gaisi Takeuti

The first time I met Kreisel was in January 1959. I was a newcomer to the Institute for Advanced Study, and he gave a lecture there. After his lecture I tried to discuss something with him. He told me "Not now." He wished me to speak English better.

Though my progress in English was very slow, we started communicating by letters. The subject of our discussion was proof theory. He always asked me something like why I was doing my type of research and whether it is worthwhile. His questions were very sharp and sounded like attacks. On the one hand, I was very much annoyed by his criticisms since I work on some directions or problems simply because I like them and I don't bother explaining my motivation. On the other hand, I welcomed even negative attentions like his since no one except Gödel, Kreisel and Schütte later had any interest in my work. I now forget what kind of discussions we made. But we were seriously debating when I stayed in the Institute for Advanced Study a second time in 1966-68.

Gödel became very much interested in our discussion and wished to read our correspondence. So I started making copy of our correspondence and gave it to Gödel. This certainly made a big change to my style from sharp words to moderate academic tone. Without knowing what happened Kreisel kept his usual nasty style for a while. Nevertheless this change somehow brought our discussion to an end. We agreed to be different.

At that time, Gödel was always there in our relation, though Kreisel was much closer to Gödel than I. Also Kreisel's name often came out in my conversation with Gödel. Once it was mentioned that both Kreisel's papers and my papers are difficult to read. (This was not Gödel's opinion. He told me that he has no difficulty to read my papers.) I told Gödel "Kreisel's paper is very difficult because his English is too good. My paper is very difficult because my English is too poor." Gödel smiled and said "May be so." Somehow the story was told to Kreisel. He replied "This time, Takeuti's English is too good."

Another time there was a meeting to which both Kreisel and I were invited. After the meeting both of us went to Princeton to talk with Gödel. Kreisel's car was in the New Brunswick airport and he kindly took me to Princeton by his car. My hotel was Nassau Inn. It was on the way to the Institute. Nevertheless he didn't stop at the hotel and told me that he would like to go to the Institute first and then to take me to the hotel later. When we arrived at the Institute, he said that he had to call Gödel, and asked me to wait. After more than 30 minutes telephone conversation with Gödel, he came to me and said that Gödel was coming to talk with him and asked me to go to the hotel by a taxi. Next day Kreisel and I met in a café to talk. In the middle of the conversation, he asked me to wait there. After a while he came back with a nicely decorated small package and explained to me that it was a present to a girl.

Around this time, we became good friends. His witty brilliant letters have constantly delighted me. Though we still agreed to be different, I was also influenced by him. We published a joint paper *Formally self-referential propositions for cut free classical analysis and related systems*. I was actually trapped by his interesting questions. Also the second part of my book *Two Applications of Logic to Mathematics* resulted from his strong influence. I would never have done that type of work without understanding his view of Mathematics. Still he said "You believe that an idea is expressed in a form of theorems. To me the theorems are byproducts of an idea but not the conclusion of it." During these years we were very busy writing letters. I expected his letter every day.

When I talked about mathematical work of someone else or mine, he always asked what the memorable results in it were. I don't remember any memorable results at that time but I remember memorable Kreisel sayings at that time. For proof theoretic ordinals, he referred to a popular tobacco commercial "It's not how long you make it, but how to make it long." For my attitude to proof theory and/or mathematician's attitude to his subject, he referred to a famous line from French a drama "I love my wife because she is mine."

After Gödel's death we wrote about Gödel very often. Besides we discussed many aspects of mathematics. We were thinking more cooperations. Among many other things, logical analysis of Roth's theorem and Siegel's theorem and application of logic to these theorem were one main subject in our discussion. Then he took an early retirement from Stanford University. This was a very sad happening for me. The reason that our cooperation has not materialized is partly due to the fact that my interest switched from mathematics to physics at that time. But more than that his retirement changed our psychological atmosphere drastically.

We still agree to be different. I am now very much interested in proving $P \neq NP$. He said that he is interested in the subject only if $P = NP$. I am still enjoying our correspondences. Nevertheless I nostalgically remember our fierce discussions on proof theory together with the memory of Gödel.

KREISEL'S MATHEMATICS

Kreisel's Unwinding of Artin's Proof

by Charles N. Delzell[1]

1 Introduction

1.1 Unwinding: A peaceful use of Cold War technology

> *"But even if one were not satisfied with consistency and had further scruples, he would at least have to acknowledge the significance of the consistency proof as a general method of obtaining finitary proofs from proofs of general theorems—say of the character of Fermat's theorem—that are carried out by means of the ε-function."* Hilbert [1928, p. 78; 1967, p. 474].

As we shall summarize in (6.4–5) below, the ε-function (not to be confused with the set-membership symbol $\in$) was Hilbert's formal choice function, strong enough to replace the traditional quantifiers $\forall$ and $\exists$ (in a suitable first-order system), and yet conservative over the original system; i.e., any proof, using the ε-symbol, of any formula not containing it (such as $0 \neq 0$, or the Fermat conjecture), can be transformed, line by line, into a new (longer) proof of that formula, all of whose formulae are also free of the ε-symbol; this is the second ε-theorem. The first ε-theorem assumes that the non-logical axioms and the end-formula of the proof are free not only of ε, but also of the

[1]Partially supported by NSF.

more conventional quantifiers $\forall$ and $\exists$, and arranges for the new proof to be not only ε-free, but quantifier-free, as well. (The hypothesis of the first ε-theorem can be satisfied by eliminating any quantifiers originally appearing in the axioms via Skolem-function symbols; a precise statement of the ε-theorems is given in (6.13–15) below.) By reducing a hypothetical proof of $0 \neq 0$, say, whose formulae may contain quantifiers or the ε-symbol, to one whose formulae are free of those symbols, it would be easier to show (finitarily) that such a proof could not have existed in the first place; i.e., to show (finitarily) the consistency of the system. (Cf. (8.1) below for an example of such a consistency proof for real closed fields.)

In the 20's and 30's, this was Hilbert's main advertisement for his ε-theorems. It is a mathematical parallel to the Mutually Assured Destruction doctrine in nuclear war: each side in the conflict builds up an arsenal of sophisticated weapons, capable of destroying the opponent; but in fact both sides are deterred from actually using their weapons, and life goes on. Here, the ε-theorems are like the Strategic Defense Initiative: intended to neutralize an "incoming" proof of $0 \neq 0$, before it can destroy the formal system; like SDI, the ε-theorems have yet to be used in a worst-case scenario, but only as a deterrent.

Back then, the threat of inconsistency (say, in arithmetic) seemed plausible enough to justify a defensive build-up (though critics could call it "Cold War hysteria"). By 1950, however, the trepidation in foundations had largely died down, and the time was ripe to look for possible peaceful, commercial uses of the formidable ε-theorems, as suggested by Hilbert in the opening quote above. But even in this quote, Hilbert's proposed example of a new application was still farfetched: while the originally intended application of his theorems was to sanitize *non-existent* proofs of *ridiculous* formulae (such as $0 \neq 0$), the new application, proposed in the quote, was to sanitize *still*[2] unknown proofs of *serious* formulae, such as Fermat's. Moreover, even if a (non-finitary) proof of the Fermat conjecture had been known, it is not clear what additional mathematical (as opposed to merely foundational) information we could derive from a finitary proof: after all, the Fermat conjecture is a purely universal, not existential, statement (cf. also (1.2) below). It was not until the 50's that anyone applied the ε-theorems (or Herbrand's theorem, or Gentzen's cut elimination procedure, also intended for consistency proofs) to a *known* proof (of a non-universal, or even an interesting, theorem); that someone was G. Kreisel, beginning with his papers [1951–52] and [1958a].

[2]At least, until Wiles' recent announcements.

1.2 Statement of (the first stage of) Kreisel's unwinding program

In [1958a] Kreisel listed a couple of defects of Hilbert's consistency program, and then stated (p. 155):

> "There is a different general program which does not seem to suffer the defects of the consistency program: *To determine the constructive (recursive) content or the constructive equivalent of the non-constructive concepts and theorems used in mathematics,* particularly arithmetic and analysis."

In Kreisel's hands this "different general program" would evolve through three stages over the following 30 years, and the colorless description above would eventually be replaced by the name "unwinding."[3] In [1981, p. 136ff] he included a section entitled, "3. Unwinding of proofs: synthesis and verification of programs," saying:

> "This section contains the—so far—most successful uses of mathematical logic, in particular, proof theory for mechanizing (routine) reasoning."

The following paragraphs in subsection 3.1 of that paper give some background on Kreisel's work on Artin's theorem:

> "3.1 *Three Stages of Unwinding*
>
> The first stage goes back to the literature on so-called non-constructive existence proofs. The latter have been hotly debated for nearly 100 years: criticized for 'hiding' explicit realizations in indirect *ad absurdum* arguments ... which can be difficult to unwind. The first concern was to show that nontrivial realizations can be extracted at all, from proofs of so-called ∀∃ theorems generally.... In the thirties and forties various transformations of proofs, known as cut-elimination and later as normalization, and various functional interpretations were developed to carry out the unwinding. This aim replaced the earlier consistency program of Hilbert, which differed from unwinding in two principal respects:

[3]The first printed use of this word that I can find is the expression, the "mechanical unwinding [of] ... indirect proofs," in Kreisel [1977a, p. 113]. "Unwinding" conveys that a*routine* job is meant, while alternative words, such as "decipher" (favored by some Constructivists) conveys the opposite—some kind of *problem* requiring ingenuity.

> 1. Only purely universal, not existential statements were considered.
> 2. The main emphasis was on the (elementary) proofs used, not on the operations used in unwinding.
>
> For the latter, it is immaterial whether the proof establishing the *correctness* of the unwinding procedure is or is not elementary (provided the proof is sound). Actually, the unwinding involved could be read off from earlier consistency proofs: this possibility had been neglected in accordance with then-current foundational preoccupations with restrictions on metamathematical methods to be used in consistency proofs—preoccupations which had caught the imagination despite the fact that it was not very clear what more we knew from an elementary proof of a universal proposition, for example, of an identity than from a civilized proof.... During the first stage just described, there were some successes, in particular the unwinding of Littlewood's proof ... on the oscillation of $\pi(x) - \mathrm{li}(x)$ or of Artin's theorem on sums of squares (without the need for new ideas)."[4]

We leave the first of these two "successes" (unwinding Littlewood) to Feferman's contribution to this volume, and proceed to the second success:

1.3 Unwinding Artin

("A case of a *knotty unwinding,*" Kreisel [1977a].) The simplest formulation of Artin's theorem [1926] is that positive semidefinite ("psd"; cf. (2.4) below) rational functions $f \in \mathbf{Q}(X_1, \ldots, X_n)$ can be represented as sums of squares of rational functions: $f = \sum_i r_i^2$, for some $r_i \in \mathbf{Q}(X_1, \ldots, X_n)$ (this solved Hilbert's 17th problem [1900]; cf. (2.4) below for details). Many refinements to Artin's theorem, found by others mainly in the fifties, are described in §5 below. While Littlewood's theorem (1.2) had already been constructivized (by Skewes) before Kreisel's work, no completely constructive version of Artin's theorem (in particular, no primitive recursive bound on the number and

[4]The two successes above were mentioned by Kreisel also in [1958b, p. 220 of the 1983 edition], [1968, pp. 361–2], [1977a, pp. 113–6], [1987, p. 395], [1992, pp. 27 and 33], Kreisel, Macintyre [1982, pp. 237–8], and elsewhere. Here we omit further paragraphs from [1981], on the second and third stages of unwinding; cf. Feferman's contribution to this volume for a fine survey of unwinding (exclusive of the unwinding of Artin's proof, treated in this paper).

degrees of the r_i), whether from Artin's original proof or from some new one, had been known prior to Kreisel's attack in 1955.

In [1960a], Kreisel wrote:

> "Artin asked me (Oct. 55) whether his proof could be made constructive. This was done (Nov. 55) by using the ideas in the proof of the first ε-theorem to analyze the transfinite machinery of Artin's proof. Afterwards a simpler, but less informative proof of a sharpened form of Artin's theorem was obtained by a systematic use of the first ε-theorem, and published a year later in *BAMS,* 63 [1957b]."

The rest of [1960a] consisted of sketches of the two methods mentioned above; I shall refer to them as "*Kreisel's first sketch*" and "*Kreisel's second sketch.*" Thus the first sketch (summarized even more briefly in [1957b, p. 100], like the unwinding of Littlewood (1.2), used only the ideas in the *proofs* of the ε-theorems, and followed Artin's original proof. The second sketch (also described in [1958a, pp. 164–8], and summarized in [1957b, p. 99] used the *statement* of the ε-theorems (and of Tarski's quantifier-elimination theorem, (4.1) below), and did not follow Artin's or any other known proof; thus, the second sketch is not properly called an unwinding (strictly speaking, one does not unwind a *theorem*; one unwinds a *proof*). The sense in which the second sketch is less informative than the first is that it gives no *explicit* information on the relative weight of the parameters n and $d := \deg f$ in the complexity of the construction of the r_i, while the first sketch does ((9.1–2) below).

1.4 Unpublished expositions of Kreisel's sketches

Kreisel's first sketch was the basis for the unpublished thesis of Daykin [1961] (cf. also (9.2) below). Kreisel's second sketch was the subject of seminar talks at Yale, Stanford, and the ETH (Zürich), by Macintyre, van den Dries, and Prestel, respectively (and probably others). It was the subject of John Brown's Master's thesis [1965] and D. Isaacson's senior undergraduate thesis [1967]. It also took up pages 11–14 of an unpublished historical survey [1990] by H. Lombardi. All of these written expositions were of high quality, all the more so considering how early in the authors' careers they were written.

Most, if not all, of these unpublished expositions agree that Kreisel's sketch(es) are essentially correct, but in need of more detailed elaboration. For example, Daykin wrote (p. iv):

> "In 1955 G. Kreisel showed by means of symbolic logic [1957b], [1960a], [1958a] that there must exist a constructive solution of Hilbert's problem. Then in 1958 he suggested that the author might try to find such a solution, and the results subsequently obtained form the body of this thesis."

Brown wrote (p. 1):

> "[I]n 1955 Artin asked Kreisel if the proof could be made constructive. Kreisel asserted that this was possible, and published indications of his method in two places, [1960a] and [1958a]. Although [1960a] gives more detail than [1958a], both papers are rather cryptic, neither giving a complete connected account: moreover, for the logical part of the argument, Kreisel uses the Hilbert-Bernays ε-theorems."

Isaacson wrote (p. 3):

> "Kreisel ... published indications of his method in several places [1957b], [1960a], [1958a], but they are quite sketchy. In this thesis my objective is to present in detail a constructive solution to Hilbert's seventeenth problem developed from the ideas of Kreisel."

Finally, Lombardi wrote (p. 2):

> "In 1957 Kreisel publishes an idea of a proof of [Artin's theorem], which ... furnishes a primitive recursive algorithm, and which is constructive (unless there is a gap in the proof)."

And on p. 11 he continued:

> "Here we give an account of the article 'Sums of squares' [1960a], simple notes given at a conference."

This is followed by a footnote:

> "Notes that I could not begin to decipher until after resolving the question of the effective [real] *Nullstellensatz* by a method that, whenall is said and done, is related."

Lombardi's 3-page account of the second sketch is itself not much more detailed than Kreisel's original account, and at the end (p. 13) he wrote:

> "Even after this explicit treatment, Kreisel's proof still seems to us a little obscure."

(In footnote 44 of (6.40) below, I shall identify the one specific source of obscurity that he mentioned.)

From the above comments, it is not surprising that Kreisel's sketches leave the reader with some latitude as to how to fill in the details. In fact, Brown and Isaacson gave two somewhat different approaches to Kreisel's second sketch. In order to explain the difference between these two approaches, we first give:

1.5 An oversimplified sketch of the second sketch[5]

If $f \in \mathbf{Z}[x_1, \ldots, x_n]$ is positive semidefinite (the x_i being variables), then, by the (constructive) completeness of the theory of real closed (= maximally ordered) fields ((4.6); based on Tarski's quantifier-elimination theorem (4.1)), we get a formal proof of $\exists u\,(f(x_1, \ldots, x_n) = u^2)$ from:

(i) the field axioms,

(ii) $a^2 + b^2 \neq -1$ (formal reality),

(iii) $\exists y\,(a = y^2 \;\vee\; -a = y^2)$, and

(iv)$_q$ $\exists z_q\,(z_q^{2q+1} + a_{2q} z_q^{2q} + \cdots + a_0 = 0)$, for finitely many $q = 1, 2, \ldots$ depending on f;

here, $a, b, a_0, \ldots, a_{2q}$ are free variables. In (i), the field axiom $a \neq 0 \;\rightarrow\; \exists v\,(av = 1)$ is made quantifier-free by the new function symbol a^{-1}: $a \neq 0 \;\rightarrow\; aa^{-1} = 1$. Similarly, (iii) and (iv)$_q$ are made quantifier-free by replacing y and z_q by the new function symbols $\rho(a)$ and $\rho_q(a_0, \ldots, a_{2q})$. Since $\exists w_1\,\exists w_2\,(f(x_1, \ldots, x_n) = w_1^2 + w_2^2)$ follows from (i), (iii), and the negation of (ii), it follows from (i), (iii), and (iv)$_q$ alone. By Hilbert's extended first ε-theorem ((6.14) below; or, by Herbrand's theorem), we can prove a disjunction $\bigvee_k (f(x_1, \ldots, x_n) = \sum_{j=1}^{2} t_{kj}^2)$, where the t_{kj} are finitely many terms built up from $0, 1, x_1, \ldots, x_n$ by the field operations and the function symbols ρ and ρ_q. The ρ-symbols are eliminated from this disjunction one at a time, by interpreting a subterm of the form $\rho(s)$ (where s is a term), first as $\sqrt{s}$ and then as $\sqrt{-s}$; both interpretations are consistent with the axioms (i), (iii), and (iv)$_q$, because the latter are satisfied by the algebraic numbers, the

[5] Our complete exposition of this sketch will be postponed until §6 below.

reality condition having been eliminated. The resulting two (disjunctions of) identities are no longer sum-of-squares identities, but they can be contracted, by simple algebraic techniques reminiscent of Galois theory, to a single (disjunction of) sum-of-squares identities, free of either $\sqrt{s}$ or $\sqrt{-s}$. Likewise, each ρ_q-symbol is eliminated via algebra (e.g., the Euclidean algorithm), by an induction on q; here it is important to note that $\rho_q(a_0, \ldots, a_{2q})$ can be interpreted as *any*[6] root z_q of $z_q^{2q+1} + a_{2q}z_q^{2q} + \cdots + a_0$. (As the ρ- and ρ_q-symbols are eliminated, we increase the number of square summands t_{kj}^2—more precisely, the number of values of both k and j; we also introduce additional disjuncts of the form $-1 = \sum_j s_j^2$.) Finally, incorporating an insight of Henkin [1960] (cf. (5.7–8) below), we introduce new constant symbols (instead of integers) for the coefficients of f, in order to extend the result to arbitrary ordered ground fields $(K, \geq)$ (instead of just $\mathbf{Q}$); now we must require f to be nonnegative over the real closure of $(K, \geq)$, and we must allow nonnegative constant "weights" on the squares. $\square$

1.6 Syntactic versus algebraic expositions of the second sketch

One difference between Brown's, Isaacson's, and my treatments of the details of the above sketch centers on the (essentially algebraic) methods of eliminating the ρ- and ρ_q-symbols. Brown introduces (p. 26) notation such as $D_0 : T_0$, $D_1 : T_1$, ..., for the fields $\mathbf{Q}$, $\mathbf{Q}(X_1, \ldots, X_n)$, and, from here on, finitely many branching towers of simple algebraic extensions obtained by successively adjoining roots y of $y^2 - a$ or $y^2 + a$, and/or roots z_q of $z_q^{2q+1} + a_{2q}z_q^{2q} + \cdots + a_0$. He describes some syntactic manipulations on terms that mirror the operations on "actual" fields obtained by adjoining roots of the above polynomials,

> "except that we have made no use of the reducibility or irreducibility of the polynomial[s] ..." (p. 30).

Isaacson, on the other hand, wrote (pp. 3–4):

> "In the outline of this thesis I have sometimes followed Brown [1965]. However, I feel his basically syntactic approach has obscured the algebraic structure, and I have attempted to elucidate it, especially in §5"

(where the branching towers of field extensions are constructed). On pp. 23–4 Isaacson defines maps $\sigma_i : S_i \to K$, where S_i is a certain

[6] Kreisel did not mention this in the fifties, but he did in Kreisel, Macintyre [1982, p. 238]. Cf. also footnote 39 at the end of (6.32) below.

infinite set of terms constructed at the i'th stage, and K is a previously constructed field. At one point he extends σ_i to $\sigma_{i+1} : S_{i+1} \to K(z_q)$, where z_q is "a" root of $z_q^{2q+1} + \sigma_i(t_{2q})z_q^{2q} + \cdots + \sigma_i(t_1)z_q + \sigma_i(t_0)$; here the t_l are terms in S_i. He does this by mapping the term $\rho_q := \rho_q(t_0, \ldots, t_{2q})$ to z_q. It is not clear to me *which* root z_q was to be chosen (or whether *all* of them were, somehow, to be chosen). For example, consider the polynomial $z_1^3 + z_1$, i.e., $z_1(z_1^2 + 1)$ (suggested by Prestel). This has roots 0 and $\pm\sqrt{-1}$. So the field $K(z_1)$ is either K itself, or $K(\sqrt{-1})$ (which, of course, might also be K itself, in case K already contains $\sqrt{-1}$). He wrote (except for trivial changes of notation):

> "It should be clear that σ_{i+1} induces a well defined truth assignment on S_{i+1}, and that under it '$\rho_q^{2q+1} + t_{2q}\rho_q^{2q} + \cdots + t_0 = 0$' is assigned 'true'."

It is not clear to me that σ_{i+1} is well defined at $\rho_1(0, 1, 0)$, although I do see that whether we define $\sigma_{i+1}(\rho_1)$ to be 0 or $\pm\sqrt{-1}$, we shall get "true." So there is no problem here.

On pp. 28–9 Isaacson took up the problem of "constructing" the field $K(\alpha)$ generated over K by "a" root α of a polynomial $p \in K[T]$ (T an indeterminate). Unlike Brown, Isaacson uses the traditional Kroneckerian construction $K[T]/(p)$ if p is irreducible in $K[T]$. If it isn't, then $K[T]/(p)$ is not even an integral domain:

> "Thus we are done if we can effectively find an irreducible factor $q(T)$ of $p(T)$ (since then any root of q is a root of p)."

For this he appeals to van der Waerden's extension (to transcendental or (separable) algebraic extensions K of $\mathbf{Q}$) of the polynomial factorization algorithm commonly attributed to Kronecker[7] for polynomials in $\mathbf{Q}[T]$. He did not say, however, what to do if there are several different such factors: are we to choose any one of them? Or all of them? Or what? (My impression is: all of them.)

1.7 Pros and cons of the syntactic and the algebraic approaches

In the syntactic approach, one works with terms that *resemble* polynomials (cf. (6.18) below for details). With such a polynomial term, one

[7]This algorithm was first discovered by the astronomer Friedrich Theodor von Schubert in [1798], extending a method of Isaac Newton for finding linear and quadratic factors of polynomials $\in \mathbf{Z}[T]$ in his *Arithmetica Universalis* (1707). Cf. Cajori [1908, pp. 136–7]. Kronecker rediscovered von Schubert's method independently in [1882, pp. 256–7 of the 1968 reprint]).

does not even know its *degree,* since one does not know the “values” of the “coefficients” (which themselves are terms, capable of lots of interpretations). So when performing algebraic operations (such as division) on those “polynomials,” one must consider the case where the leading “coefficient” is, or is not, 0; and in case it is 0, one must consider the case where the next “coefficient” is, or is not, 0, etc. (This is so even if the intended ground field is *discrete* (4.4), so that we can determine whether any given field element is 0 or not.) All these case distinctions complicate the syntactic approach.

In the algebraic approach, one works with “actual” polynomials, over good old-fashioned fields and their extensions (such as $K[T]/(p)$, where p is irreducible in $K[T]$, without equivocation). As Isaacson said, this can be arranged via the von Schubert/Kronecker/van der Waerden factorization algorithm. There are two disadvantages to appealing to this algorithm:

1. It is lengthy, in terms of practical computation (but as we shall see from the bounds in §9, practically everything in this business is computationally unfeasible, including the case-distinctions mentioned in the previous paragraph).
2. It works only over special ground fields, such as finite extensions of $\mathbf{Q}$ (or of $\mathbf{Q}(X_1, \ldots, X_n)$). van der Waerden showed that there can be no factorization algorithm over arbitrary discrete fields; it depends on subtle aspects of how the ground field is presented.

Thus the use of factorization forces one to start Kreisel's sketch with $\mathbf{Q}$, and then do a second step, in which one extends to arbitrary real closed (or even arbitrary ordered) discrete ground fields. Actually, even Brown, and not only Isaacson, made this 2-step maneuver from $\mathbf{Q}$ to the general case, even though Brown did not use factorization. Kreisel did not mention factorization; nor did he use the 2-step maneuver from $\mathbf{Q}$ to more general ground fields. This leads me to conclude that he did not *intend* for factorization to be used in his sketch.

As an algebraist (and not a logician), I feel a little uncomfortable with “constructing” an extension field such as $K(z_1)$, where z_1 is a root of, say, $z_1^3 + z_1$, as in (1.6). It would be one thing if we already had at our disposal some algebraically closed extension of K (such as its algebraic closure) in which to locate the various possible values of z_1; but in Kreisel's sketch, we don't (at least, not at first glance).

1. On the one hand, he wrote [1960a, p. 317] that

   > “[Artin's] proof is stripped of unnecessary transfinite machinery,”

allowing us to

> "keep control over the (finite) extensions needed to carry out the computations."

And in [1977a] he wrote (p. 116):

> "Artin uses an infinite tower of field extensions; determining a bound on the relevant finite portion of it is tantamount to finding [a bound for Artin's theorem]."

So if one means to give a *genuinely* algebraic (and finitary) exposition of Kreisel's sketch, then it would seem that one is really supposed to "construct" these finite extension fields; in the absence of factorization, one could use a method such as that in Bridges, Richman [1987, §4.3], which does not depend on factorization, but which does depend on countability.

2. On the other hand, the above quotes on finite extensions were mainly in the context of Kreisel's first sketch, and perhaps were not meant to apply to the second sketch. Indeed, recall the passage from [1981] quoted in (1.2) above:

> "[I]t is immaterial whether the proof establishing the *correctness* of the unwinding procedure is or is not elementary (provided the proof is sound)."

More generally,

> "and this seems worth stressing, ... *we are not rejecting any part of non-constructive work ...*" [1958a, p. 160] (his italics).

On the contrary, the only thing that Kreisel's unwinding program *does* reject is the

> "*central dogma* of Constructivism, cdC, [which] requires constructive principles of proof" Kreisel [1990b].

Specifically, in unwinding, we may shift freely between syntactic and algebraic arguments, constructive or not (e.g., the existence of the algebraic closure of an arbitrary field), for the purpose of establishing the correctness of a new sum-of-squares identity that was obtained from a previous correct identity by certain finitary, syntactic manipulations (such as replacing the term $\rho(s)$ by $\sqrt{s}$ and then $\sqrt{-s}$, and then simplifying, as in (1.5)). In this way, pseudo-algebraic notation such as $K(z_q)$ can be justified.

1.8 New elements in my (unauthorized) exposition

In §6 below I present yet another variant of Kreisel's second sketch. After writing up the first draft of it, I received copies of Brown's and Isaacson's theses, and saw that my presentation is closer to the former than the latter; but I remain on the syntactical level of terms, even more fastidiously than Brown did. Rather, my style is actually closer to that of Hilbert, Bernays [1939, 1970], which I had recently read (and which Kreisel studied in 1942 at Cambridge University; cf. Kreisel [1987, p. 395]); I have also followed Tarski [1951]. My exposition never mentions any actual field or field-element at all, but only field *axioms,* and *terms* in the language of fields. In particular, the question of how to construct "the" extension $K(w_q)$ of K (where w_q is a root of a possibly reducible polynomial) does not arise; for that matter, one need not know even how to "construct" $\mathbf{Q}$ itself. The only things that I construct are *formal proofs* (of sum-of-squares identities). Furthermore, I avoid Brown's and Isaacson's 2-step process of starting with the ground field $\mathbf{Q}$, and later generalizing to arbitrary discrete ordered fields; I treat (the axioms for) the latter right from the beginning (just as Kreisel did, except that he considered only *uniquely orderable* discrete fields with *finite* Pythagoras number (5.6), and thereby got an *unweighted* sum of squares; cf. §5).

Another new element in my exposition is the use of both the statement, and the ideas behind the proofs, of the ε-theorems. Recall that Brown ((1.4) above) had listed Kreisel's use of the ε-theorems as a *disadvantage,* presumably because already in the sixties, they were no longer widely known (in contrast to the closely related Herbrand theorem). As mentioned in (1.5), Kreisel used the (statement of the) extended first ε-theorem to produce the preliminary disjunction of sum-of-squares representations (as do I); Brown and Isaacson simply replaced the extended first ε-theorem with Herbrand's theorem at this step. This disjunction is actually the *end-formula* of a formal proof produced by the extended first ε-theorem. Brown and Isaacson ignored this formal proof, and concentrated instead on the end-formula, from which they eliminated the unwanted ρ- and ρ_q-symbols, via Kreisel's algebraic techniques. This obeys Kreisel's dictum ((1.2) above) that one should not go to extra trouble to seek an "elementary" proof (e.g., one that is free of ρ- and ρ_q-symbols) of a universal formula (such as our sum-of-squares identity), when a "civilized" proof of it is already at hand.

But *quod licet Jovi non licet bovi,* for Kreisel himself seemed to violate his dictum at the end of *his* presentation, where he wrote:

> "By the second ε-theorem the use of the symbols ρ and ρ_q can be eliminated" [1960a, p. 316]. (1.8.1)

In [1958a, pp. 167–8] he elaborated on this:

> "In our proof we used not only the predicate calculus but also the elimination of existence symbols by means of function symbols (ρ and ρ_q). Now, by the completeness theorem, granted that such a use of function symbols is correct, there exists a proof of $f = \sum g_i^2$ (from the axioms above) in the predicate calculus. But if one wants explicit primitive recursive instructions for obtaining such a proof one uses the second ε-theorem. One has to appeal to the ε-theorems *with* equality because one uses e.g. the rule $x = y \to \rho(x) = \rho(y)$. (Actually this refinement is not necessary if one is interested in constructive theorems rather than in constructive proofs, and the explicit form $f = \sum g_i^2$ is the constructive theorem which we are after.) However, the use of the first ε-theorem is quite essential for our purpose."

In the statement in parentheses above, "this refinement" seems to refer not to the preceding sentence (on the need for the equality axioms), but to the sentence before that, on the second ε-theorem. And his reference to a possible interest in "constructive proofs" seems to refer not to a possible interest in getting

> "constructive, formal proofs of $f = \sum g_i^2$ in the predicate calculus"

(whatever that may mean; cf. next paragraph), but to a possible interest in getting

> "constructive proofs of the fact that $f = \sum g_i^2$ is provable in the predicate calculus,"

i.e., an interest in constructing a formal proof in the predicate calculus of $f = \sum g_i^2$.

As summarized at the beginning of (1.1), and presented in detail in (6.15) below, the second ε-theorem does, indeed, accomplish the latter goal, given a suitable formal proof using ρ and ρ_q: it leaves untouched our universal end-formula (which by now has already been cleansed of the unwanted ρ- and ρ_q-symbols), and instead purges those symbols (which are disguised ε-symbols (6.5)) from the rest of the given formal proof; the price we pay is the re-introduction of the more conventional quantifiers $\exists$ and $\forall$, which had been eliminated earlier in the sketch.

The first question is whether, at this stage, Kreisel really still *has* a (first-order) formal proof (using as non-logical symbols only 0, 1, $+$, $-$, $\cdot$, ${}^{-1}$, ρ, ρ_q, and perhaps a few constants), to which to apply the second ε-theorem: indeed, while eliminating $\rho(s)$ from the end-formula (1.5), Kreisel resorted to apparently informal arguments, such as appealing to the existence of the algebraic numbers (to prove consistency), and using symbols such as $\sqrt{s}$ and $\sqrt{-s}$ that do not belong to the formal language apparently being used at this stage. Even if the details of the formalizations of these steps are filled in, there remains a second question, concerning Kreisel's *motive* in appealing to the second ε-theorem: was this supposed to be a step toward an "elementary" proof of an identity ($f = \sum g_i^2$)? If so, it falls between two stools:

1. On the one hand, the second ε-theorem is *not needed* by the silent majority of mathematicians, who were *in*formally introducing functions corresponding to ε-terms (or Skolem-function symbols) such as $\rho(s)$, even before logicians gave formal, "elementary" justifications for such reasoning (cf. (6.4) below). Just as Kreisel wondered ((1.2) above) what more we know from Hilbert's elementary proofs of universal theorems than from civilized proofs, these working mathematicians could wonder what more we know from Kreisel's $\exists$ and $\forall$ symbols than from ρ and ρ_q (even granting that the former really are more "elementary" than the latter).

2. On the other hand, the second ε-theorem is *not enough* to appease militant Finitists, according to Kreisel's own description of their non-negotiable demand:

   > "*Our thesis is: we do not regard an argument as finitist if it contains formulae with bound variables.*" [1951, p. 242]; his italics.

   Since the second ε-theorem re-introduces $\exists$ and $\forall$ (i.e., bound variables) into our formal proof, there is little hope that the second ε-theorem will construct a *finitist* formal proof of $f = \sum g_i^2$ (or of anything else).

My attitude is that as long as we are going to commit the thought-crime of seeking an elementary proof of our sum-of-squares identity, we might as well do it right, and not only eliminate the ρ- and ρ_q-symbols from the formal proof, but also take care to eliminate them *without* re-introducing the quantifiers $\exists$ and $\forall$. For this I follow the ideas in the *proofs* of the ε-theorems; i.e., I adapt Hilbert's ε-substitution method to the theory of real closed fields. In this construction, the main step that was not present in Kreisel's original sketches is my Theorem 6.33, whose proof is, unfortunately, lengthy, in contrast to the unrivalled economy of Kreisel's 15-word statement (1.8.1). The result of my extra effort is a formal proof, using only the quantifier-free fragment of the theory of fields (of characteristic 0, or at least of sufficiently high characteristic), of a representation of f as a weighted sum of squares of rational functions in $K(X_1, \ldots, X_n)$. The fact that this identity holdsover all fields of characteristic 0 (and not only over all real closed and/or algebraically closed fields) has meaning for straight mathematics (as opposed to metamathematics), and does not seem to follow from Kreisel's sketches—unless one appeals to the existence of the algebraic closure of an arbitrary discrete field (say, of characterstic 0), in which case the argument is easy. But as we said above, part of Kreisel's idea was that, while not forbidden, such "transfinite machinery [is] unnecessary."

A thought-crime at least as heinous (according to Kreisel) as the pursuit of elementary proofs of universal theorems in general, is the pursuit of finitary, syntactic proofs of consistency, in particular. We perpetrate this, easily, in (8.1) below, for the theory $\mathbf{RCF}(K, \geq)$ of real closed fields augmented by the diagram of a given ordered field $(K, \geq)$. Specifically, we use the machinery that we have developed for Kreisel's sketch, to reduce any proof of $-1 = \sum_j s_j^2$ in $\mathbf{RCF}(K, \geq)$, to a proof of another (nonnegatively weighted) sum-of-squares representation of -1 in $\mathbf{F}^{\mathrm{qf}}(K)$, the quantifier-free fragment of the theory of fields augmented by the diagram of K. Lombardi had given a similar consistency proof in [1988–91], using his constructive proof of the real *Nullstellensatz* in place of Kreisel's constructive version of Artin's original proof.

While Kreisel wrote in [1958a] of the "mathematical significance of consistency *proofs,*" we shall illustrate (in Part II of this paper) some mathematical significance of our consistency *result,* by using it to give a simple, finitary construction of the real closure of an arbitrary discrete ordered field $(K, \geq)$ (a construction first carried out, with considerable difficulty, by Artin's and Zassenhaus' student Hollkott [1941]). This reverses the traditional sequence, of first constructing a model, and then using it to prove consistency. The fact that this reversal is possible in this case will clarify a speculation of Lombardi, Roy [1991]:

> "Let us note also that a direct ['purely metamathematical'] proof of the consistency and completeness of the formal intuitionistic theory considered would not give a method for constructing the real closure of K, as we can see in the example of the theory of discrete algebraically closed fields

> (the 'completeness theorem' is not valid constructively ...)" (p. 261 of the abridged English version; p. 20 of the French version);

our proof in Part II will show that *very little* besides consistency and completeness is required for this construction (as hinted by Lombardi himself in [1988–91]).

1.9 The "horror" of new (mathematical) ideas

One further aim in unwinding is to obtain suitable bounds *without introducing "further ideas," beyond those in the original non-finitist proof.* Tacitly, new *meta*mathematical ideas (especially those occurring in consistency proofs) are allowed; it is only new *mathematical* ideas that are forbidden. Kreisel first adumbrated this canon in [1951, p. 247], in connection with the unwinding of Littlewood's proof ((1.2) above): both Skewes and Littlewood had stated that a "further idea" (beyond those in Littlewood's original proof, or in Ingham's 1932 exposition) was necessary for finding bound(s). In [1977a], Kreisel claimed (p. 114):

> "This impression was refuted, beyond a shadow of a doubt, by the proof theoretic analysis Kreisel [1951–52] which makes a *routine* application (to Littlewood's original proof for π) of a *general* (logical) method."

He hinted at this rejection of new ideas again in [1958a, p. 170], where he used it in an apparent attempt[8] to marginalize Skewes' obviously prior discovery of the bounds that Kreisel claimed his obscure sketch would also achieve (specifically, Kreisel dismissed Skewes' work as "*ad hoc*"; for more on priority acknowledgments, cf. footnote 21 of (5.4) below). In Kreisel's many subsequent references (e.g., [1968, p. 362], [1977a, p. 114], [1981, p. 137], and [1992, p. 27]) to the discovery of bounds for Littlewood, he dropped Skewes' name altogether, but kept the reference to his own abstinence from new ideas. This virtue is also touted in his "[u]nwinding [of] a proof of Roth" (on approximating algebraic irrationals by rationals; cf. Kreisel [1977a, p. 114]), and in his unwinding of Artin's proof ([1981, p. 137], quoted in (1.2) above).

[8]If this was the aim, it was not achieved: all others who have published on the subject, e.g., Littlewood-Skewes (1955), Lehman (1966), te Riele (1986), and Vaughan [1992] (on p. ix the latter gives references to the former papers) give full credit to Skewes, and do not even mention Kreisel; some of these authors were simply unaware of Kreisel's claims, while the other authors may have found his published work on this to be too sketchy and inconclusive.

The above taboo on new ideas seems to be a novel kind of "purity of method." Impure methods were well known in other areas: one had analytic or topological proofs of arithmetic theorems, and algebraic methods in plane geometry; some people had tried to avoid these impure methods ("the peroration of Hilbert's *Grundlagen der Geometrie* [1899] is a panegyric concerning this ideal"). This led Kreisel to speak of "The Horror of Impure Methods" [1981, p. 134]. But Kreisel seemed to express his own horror at the "further ideas," or "*ad hoc*" (= impure) methods, used by Skewes and others when finding bounds, as seen in the previous paragraph. I have seen no independent evidence that Littlewood, Artin, or anyone else whose proofs Kreisel unwound, shared Kreisel's disapproval of the use of new ideas when seeking constructivity. On the contrary, on p. 49 of Skewes [1955] (a paper that acknowledged Littlewood's help in preparing it for publication), Skewes and Littlewood *re-affirmed* their impression that one needed their "further ideas," several years after the appearance of Kreisel [1951], which had supposedly refuted that impression "beyond a shadow of a doubt." Anyway, the individual reader can decide for himself

(a) how much importance he should give to the avoidance of new ideas while seeking bounds (or the avoidance of impure methods generally), and

(b) whether to pursue one form of purity of method while ridiculing others.

1.10 The purpose of this paper

Our purpose is to work out (in print, finally) Kreisel's two sketches of constructive proofs of Artin's theorem, thereby showing (for those beyond Kreisel's inner circle) that those sketches are essentially correct. Whatever additional results I establish here will also be based only on knowledge (e.g., of Hilbert's ε-substitution method) available up to the time of those sketches (1955). The point of this is to show that after Kreisel's contribution, most results of Artin-Schreier theory (including some of its more recent ones, such as Stengle's *Positivstellensatz* [1974]) either were, or *easily* could have been, satisfactorily constructivized (cf. also §10 below); this historical spin is in contrast to that of Lombardi [1994], who stated that Artin-Schreier theory was, at least until recently, in need of "epistemological rehabilitation." Lombardi and Kreisel are at opposite poles here: Lombardi, being a Constructivist (with a capital 'C'), considers practically *all* of mathematics (even

Artin-Schreier theory) to require constructive rehash before its validity can be secured, while Kreisel holds that nowadays, at least, most foundational questions are either obsolete, crude, dubious, or otherwise "discredited"; and most questions in constructivity, in particular, have answers that are trivial or routine, and in any case their answers are notworth publishing *in detail,* but only as cryptic sketches (or, better yet, as mere claims—e.g., (9.1) below). On the other hand, Kreisel has always *seen* constructive aspects of classical mathematics particularly vividly:

> "I realise that there are other points of view [footnote], but for [arithmetic and analysis], I still see the mathematical core in the combinatorial or constructive aspect of the proof."

The footnote read,

> "E.g. a purely abstract point of view: J.P. Serre once told me that he saw the mathematical core of the Chinese remainder theorem in a certain result of cohomology theory" [1958c, p. 158].

Feferman's contribution to this volume takes a broader look than my paper does at Kreisel's unwinding program (exclusive of the unwinding of Artin, covered here). He discusses, in particular, the unwindings by Kreisel, Girard, Luckhardt, and others, of proofs of Littlewood, van der Waerden, Roth, and others. He raises legitimate questions, and points out some shortcomings. For example, in the case of Littlewood, the obscurity of Kreisel's publications makes it hard for the mathematical community (e.g., contemporary analytic number theorists) even to confirm the essential correctness of that particular unwinding. I hope that my exposition in §6 below will put to rest any similar question about the unwinding of Artin. On the other hand, while Kreisel'sexpositions are obscure, mine is long; both defects are *potential* sources of error:

> "Certainly, in cases of long routine computations mechanization genuinely reduces th[e] probability [of applying correct principles incorrectly], and so in this area Frege's formal rules contribute to rigor in a realistic, not only ethereal sense. But in the bulk of mathematics a more significant safeguard against incorrect applications is to require that *proofs be easy to take in and to remember,* over and above the validity of the principles (intended to be) used" (Kreisel [1984, p. 83]).

As for *actual* errors, however, we shall deal ((7.2) below) with Kreisel's (equally obscure) sketch [1957a, 1958a] of a similar unwinding of Hilbert's *Nullstellensatz*; despite being "easy to take in and remember" (at least, it is brief), that sketch contains an apparently essential gap (also overlooked by P. Lorenzen, K. Schütte, and A. Robinson, in their reviews of [1958a]; cf. the list of references at the end).

Due to limitations of time and space, this paper is only the first of two installments of my exposition of Kreisel's unwinding of Artin. This paper (which I call "Part I"), deals with Kreisel's *second* sketch of Artin's theorem, while in Part II I intend to deal with his (more laborious) first sketch, and extend it to Stengle's [1974] *Positivstellensatz* (5.14); only then will it be possible for me to discuss the *continuous* versions of Artin's theorem found in the 80's. Part II is still in preparation.

Acknowledgments

I thank Drs. Burton Dreben, Dan Isaacson, Ulrich Kohlenbach, Tomas Sander, Philip Scowcroft, Wilfried Sieg, and especially David Marker and Alexander Prestel, for helpful conversations and correspondence. I particularly wish to thank Lou van den Dries, Solomon Feferman, and Henri Lombardi, for extensive criticisms of earlier drafts of this paper. Finally, I thank Professor Kreisel; while this paper takes him to task on some minor points (for the sake of thoroughness), I would like to acknowledge my great debt to him for his correspondence, from 1978 to 1994.

2 Review of Artin-Schreier theory

2.1 Formally real fields[9]

Artin and Schreier [1926] called a field K *(formally) real* if -1 is not a sum of squares in K (thus the characteristic of such a field is 0, so that its prime subfield is—isomorphic to—$\mathbf{Q}$). If K has at least one ordering (compatible with the field operations, tacitly), then K is formally real. ($K := \mathbf{Q}(\sqrt{2})$, for example, has two orderings.) Artin and Schreier proved the converse; in their proof, the well-ordering principle can be replaced by the somewhat weaker (Boolean) prime ideal theorem, as shown by Tarski [1954]. More generally, Artin proved that *any* element

[9] Artin, Schreier [1926, 1927], Artin [1926], Bochnak, et al [1987], Bourbaki [1952–64], Jacobson [1964–89], Lam [1973], Lang [1965–93], Prestel [1975–84], Prieß-Crampe [1983], van der Waerden [1931–70], etc.

of a field K (and not just -1), is a sum of squares in K if and only if it is *positive in every ordering* of K ($\leftrightarrow$: "totally positive"). Thus K is *uniquely* orderable if and only if every element of K, or its negative, is a sum of squares in K. Somewhat more generally, if K is a field and L is a subfield of K endowed with an ordering $\geq$, then he called an element $a \in K$ *totally positive with respect to* $(L, \geq)$ if $a \geq' 0$ relative to every field ordering $\geq'$ of K that extends $\geq$; then he proved that $a \in K$ is totally positive with respect to $(L, \geq)$ iff $a = \sum_i p_i b_i^2$, for some $b_i \in K$ and $p_i \in L$ with $p_i \geq 0$.

2.2 Real closed fields[9]

Artin and Schreier called a field R *real closed* if it is formally real and no proper algebraic extension of R is formally real. They characterized real closed fields as those fields R that satisfy the following three—sets of—elementary (= first-order) axioms:

(a) R is formally real, every element of R or its negative is a square in R (implying unique orderability), and every odd-degree $g \in R[X_1]$ has a root in R.

Also: a field R is real closed iff

(b) $R(\sqrt{-1})$ but not R is algebraically closed.

Upon "enlarging the language" (of rings, or fields) by the symbol $\geq$, Dieudonné [1946], Serre [1949], Bourbaki [1952, 1964] and Prestel [1975–84] call an ordered field $(R, \geq)$ *maximally* ordered if it has no proper algebraic order-extension $(R', \geq')$. If R is real closed, then its unique ordering $\geq$ makes $(R, \geq)$ maximally ordered, while if $(R, \geq)$ is maximally ordered, then R is real closed. The above characterization (a) becomes: an ordered field $(R, \geq)$ is maximally ordered iff

(a′) every positive element of R is a square in R, and every odd-degree $g \in R[X_1]$ has a root in R.

One key step in Artin and Schreier's proof that $R(\sqrt{-1})$ is algebraically closed if R is real closed, is to show that $f \in R[X_1]$ (with $d := \deg f$ even) has a root in $R(\sqrt{-1})$ provided that a certain $g \in R[X_1]$ (with $\deg g = \binom{d}{2}$) does; this reduces the power of 2 that divides d. They attributed this method to one of Gauss' 4 proofs of the fundamental theorem of algebra (specifically, it is the second proof, in [1816]). Bourbaki [1952–64, p. 39] attributed the entirety of these characterizations: maximally ordered $\Leftrightarrow$ (a′) $\Leftrightarrow$ (b), to Euler (1749) and Lagrange

(1772), who had also given (attempted) proofs of the fundamental theorem. However, on p. 150 of [1952] (and p. 163 of [1964]), Bourbaki admitted that Gauss [1799] had pointed out gaps in Euler's and Lagrange's proofs (which, anyway, bear little resemblance, in my opinion, to that of Gauss-Artin-Schreier), and that the proof on p. 39 was "essentially" that of Gauss [1816]. If theArtin-Schreier method must be attributed to someone other than Gauss (and Artin and Schreier themselves), then I (like Bourbaki [1952, p. 150], and [1964, p. 163]) would attribute it to a suggestion of de Foncenex [1759, pp. 120–1], which was carried out more correctly by Laplace [1812, pp. 63–5]; in [1815] Gauss had pointed out a small gap even in Laplace's proof (namely, the use of a splitting field, 100 years before Kronecker had legitimized such reasoning), but by Artin and Schreier's time, this was no longer a problem.

Anyway, Artin and Schreier proved, in addition [1926], that if $(R, \geq)$ is maximally ordered, then

(c) the intermediate-value property for polynomials holds with respect to $\geq$;

conversely, (c) obviously implies (a′). Finally, in [1927], Artin and Schreier went on to prove one more characterization: a field R is real closed iff (d) it has finite co-dimension in its algebraic closure $\overline{R}$ (i.e., $\left[\overline{R} : R\right] < \infty$).

Every ordered field $(K, \geq)$ has a *real closure* $\overline{(K, \geq)}$, i.e., a real closed field that is algebraic over K, and whose unique order extends $\geq$; any two real closures are (uniquely) K-(order-)isomorphic. Every real closed order-extension R of an ordered field $(K, \leq)$ contains—some (order-)isomorphic copy of—$\overline{(K, \geq)}$.

Serre [1949] introduced the useful concept of a *preordering* of a field K with respect to an ordered subfield $(L, \geq)$: it is defined to be a subset $P \subset K$ containing every positive element of L, and such that $0 \notin P$, $P + P \subseteq P$, $\;P \cdot P \subseteq P$, and $K^2 \subseteq P$. A maximal preordering turns out to be—the "positive cone" of—an ordering of K extending $\geq$. He showed how preorderings provide an alternative way to prove most of the above results of Artin-Schreier. They also simplify Dieudonné's generalization of Artin-Schreier theory, where a field K is called *A-orderable* (or *A-real*) if no element of $A \subset K$ is a sum of squares in K (and char $K \neq 2$); and K is called *A-real closed* if it is A-real but no proper algebraic extension is. (Thus the case $A = \{-1\}$ is that of Artin-Schreier.)

2.3 Archimedean orderings[9]

Artin and Schreier, generalizing Hilbert [1899], called an ordered field $(K, \geq)$ *Archimedean* over a subfield L if $(\forall a \in K)\ (\exists b \in L)\ b \geq a$; if $L = \mathbf{Q}$, then they called such a $(K, \geq)$ simply *Archimedean*. $(K, \geq)$ is Archimedean iff $(K, \geq)$ is imbeddable in $(\mathbf{R}, \geq)$, the reals.[10] Thus an Archimedean-ordered field $(K, \geq)$ is (order-)dense in its real closure, while a non-Archimedean-ordered $(K, \geq)$ need not be, as Artin and Schreier showed [1926, p. 99] for the field $K := \mathbf{Q}(X_1)$, ordered so that X_1 is infinitesimally small (i.e., $0 < X_1 < s$ whenever $0 < s \in \mathbf{Q}$). The *complement* of an arbitrary (formally real) field K, however, *is* dense in *any* ordered (proper) extension field $(R, \geq)$ of K, whether or not R is real closed, or R is Archimedean over K, or K is Archimedean over $\mathbf{Q}$; this is easy (cf. Erdös, et al [1955, p. 544]). Finally, even if the ordered field $(K, \geq)$ is not dense in its real closure R, R is Archimedean over K, by the bound on the (real) roots of a polynomial.

2.4 Artin's theorem [1926]

For the rest of this paper, we shall write $X = (X_1, \ldots, X_n)$ for n indeterminates. If $(R, \geq)$ is an ordered field, we call an $f \in R(X)$ *positive semidefinite (psd) over* R if for all $x := (x_1, \ldots, x_n) \in R^n$ such that $f(x)$ is defined, $f(x) \geq 0$.

Artin's theorem begins with the hypothesis that K be a uniquely orderable subfield of $\mathbf{R}$ (i.e., uniquely orderable and Archimedean). Since $K \subseteq \mathbf{R}$, for every $f \in K(X)$, f is psd over K if and only if f is psd over $\mathbf{R}$, or (by Tarski's model-completeness result (4.6) below) over any ordered extension of K. Thus he was justified in using the simpler expression "f is definite" to describe those f to which his theorem applied. The conclusion is:

$$\text{if } f \in K(X) \text{ is psd over } K, \text{ then } f = \sum_i r_i^2, \tag{2.4.1}$$

for some $r_i \in K(X)$. (Cf. also Jacobson [1964, 1980, 1989], or Lang [1971, 1993].) Dropping the unique orderability hypothesis on $K \subseteq \mathbf{R}$,

[10] Hilbert [1899, §17]; his proof applies as well to Archimedean-ordered skew fields, proving that they are actually commutative; this fact was also proved in Artin [1957, p. 47], as H. Lombardi kindly pointed out to me. Cf. also Prestel [1975–84, p. 11]. The embeddability of Archimedean-ordered, hence Abelian, *groups* in $(\mathbf{R}, +, \geq)$ was proved by O. Hölder in 1901; for a survey of similar results for rings, cf. Prieß-Crampe [1983].

he also proved that

$$\text{if } f \in K(X) \text{ is psd over } K, \text{ then } f = \sum_i p_i r_i^2, \tag{2.4.2}$$

for some $r_i \in K(X)$ and $p_i \in K$ with $p_i \geq 0$.

3 A formal system for real closed (= maximally ordered) fields

3.1 Our first formal system, $\mathbf{RCF}_{\geq}$

$\mathbf{RCF}_{\geq}$ will have the (first-order) language of ordered rings with equality, with constants 0 and 1; binary function symbols $+$ and $\cdot$; a unary function symbol $-$; and the binary relation symbol $\geq$ (and $=$); for now we omit the symbol for multiplicative inverse, $^{-1}$. Among the (individual) variables are a, b, $x_1, x_2, \ldots$, $y_1, y_2, \ldots$, and z. Terms are built up from the constants and variables by the function symbols, in the usual way; they look like integer polynomials in the variables. The atomic formulae are $t = s$ and $t \geq s$ for terms t and s; arbitrary formulae are built up from the atomic ones by $\neg$ (negation), $\vee$, $\wedge$, $\rightarrow$, $\leftrightarrow$, $\exists$, $\forall$, (, and), in the usual way; we also allow $\bigwedge_i F_i$ and $\bigvee_i F_i$ for the iterated conjunction and disjunction, respectively, of a finite collection of formulae F_i. We write $t \neq s$ as an abbreviation for $\neg(t = s)$, as usual. We do not use the convention (e.g., Hilbert, Bernays) that variables are classified in advance as being free or bound, and must be renamed if they fall within the scope of a quantifier that is being introduced or eliminated; rather, we use the other convention (e.g., Kleene [1952], Kreisel, Krivine [1967–72], Shoenfield [1967]) that only their occurrences are so classified, so that we may not substitute a term t for the free occurrences of a variable x in a formula $A(x)$ if any such occurrence is in a part of $A(x)$ of the form $\exists y\, B(x, y)$ or $\forall y\, B(x, y)$, for some variable y occurring in t. Besides any of the equivalent sets of axioms and rules of inference forthe classical, first-order predicate calculus with equality (e.g., Hilbert, Bernays [1939, 1970, Supplement I]; Kleene [1952]; or Shoenfield [1967]), we take the following non-logical axioms for real closed (= maximally ordered) fields (Kreisel's [1957b, 1958a, 1960a] version of (2.2)(a)):

the field axioms, (3.1.1)

the reality condition(s): $\forall y_1 \cdots \forall y_s\, (y_1^2 + \cdots + y_s^2 \neq -1)$, $(3.1.2)_s$

$\exists z\, (a = z^2 \ \vee\ -a = z^2)$, (3.1.3)

$$\exists z\,(z^{2q+1} + a_{2q}z^{2q} + \cdots + a_0 = 0), \text{ and} \qquad (3.1.4)_q$$
$$a \geq b \;\leftrightarrow\; \exists z\,(a - b = z^2), \qquad (3.1.5)$$

where $s = 1, 2, \ldots$ in $(3.1.2)_s$, and $q = 1, 2, \ldots$ for $(3.1.4)_q$.

We have already abused (i.e., expanded) the language in standard ways: First, we write ts instead of $t \cdot s$, for terms t and s. Second, we drop superfluous parentheses by associativity (3.1.1). Third, using also commutativity, the class of terms is expanded by allowing symbols for iterated addition and multiplication: $\sum_{i=1}^{g} t_i$, $\prod_{i=1}^{g} t_i$, and t^g abbreviate, say, $t_1 + t_2 + \cdots + t_g$, $t_1 t_2 \cdots t_g$, and $\prod_{i=1}^{g} t$, respectively, where $t, t_1, t_2, \ldots, t_g$ are terms. Similarly, t^0 abbreviates 1, and, more generally, $\prod_{i=1}^{0} t$ "abbreviates" 1, as does $\prod_{i\in\emptyset} t$. We shall also write $t - s$ instead of $t + (-s)$. Likewise, we shall feel free to introduce the relation symbols $\leq$, $>$, and $<$, defined explicitly as Boolean combinations of $=$ and $\geq$, in the usual way. Finally, 2, 3, and 4 will abbreviate the terms $1 + 1$, $1 + 1 + 1$, and $1 + 1 + 1 + 1$, respectively.

3.2 Axiomatic economy

Actually, of the axioms $(3.1.2)_1$, $(3.1.2)_2$, $(3.1.2)_3$, $\ldots$, all we need is[11]

$$\forall y_1\,\forall y_2\,(y_1^2 + y_2^2 \neq -1), \qquad (3.1.2)_2$$

and we shall continue to write $\mathbf{RCF}_{\geq}$ for the system obtained by dropping $(3.1.2)_1$, $(3.1.2)_3$, $(3.1.2)_4$, $\ldots$. To see this, we first show

$$\mathbf{RCF}_{\geq} \;\vdash\; \exists z\,(y_1^2 + y_2^2 = z^2), \qquad (3.2.1)_2$$

where $\vdash$ means that the formula to its right can be proved in the system to its left. First, adding $-(y_1^2 + y_2^2) = z^2$ and $zy = 1$ to the axioms of $\mathbf{RCF}_{\geq}$, we could prove $(y_1 y)^2 + (y_2 y)^2 = -1$ (using (3.1.1)), whence $\exists y_1\,\exists y_2\,(y_1^2 + y_2^2 = -1)$, contradicting $(3.1.2)_2$; thus by the deduction theorem,[12] $\mathbf{RCF}_{\geq}$ plus $-(y_1^2 + y_2^2) = z^2$ proves $zy \neq 1$, whence in turn, $\forall y\,(zy \neq 1)$, $z = 0$ (by (3.1.1)), $y_1^2 + y_2^2 = 0^2$, and $\exists z\,(y_1^2 + y_2^2 = z^2)$. By the deduction theorem again, $\mathbf{RCF}_{\geq}$ alone proves

$$-(y_1^2 + y_2^2) = z^2 \;\rightarrow\; \exists z\,(y_1^2 + y_2^2 = z^2).$$

[11] I am grateful to A. Prestel for this observation.

[12] Cf., e.g., Hilbert, Bernays [1939, 1970, Supplement I, §C], or Kleene [1952, p. 97]. Note that the (strong form of the) deduction theorem applies, because the free—occurrences of the—variables z, y in $zy = 1$ are "held constant." The formulation of the deduction theorem is sometimes simplified by stating it only for *closed* formulae (e.g., Shoenfield [1967, p. 33]); that formulation cannot be applied here until we replace z, y with constant symbols, to be changed back to z, y later by the theorem on constants; the result would be the same.

On the other hand, $y_1^2 + y_2^2 = z^2 \rightarrow \exists z\,(y_1^2 + y_2^2 = z^2)$ is a substitution axiom of $\mathbf{RCF}_\geq$. Thus

$$\mathbf{RCF}_\geq \;\vdash\; [y_1^2 + y_2^2 = z^2 \;\vee\; -(y_1^2 + y_2^2) = z^2] \;\rightarrow\; \exists z\,(y_1^2 + y_2^2 = z^2),$$

which gives $(3.2.1)_2$, using $\exists$-introduction and (3.1.3). Before we can derive $(3.1.2)_3$, $(3.1.2)_4$, $\ldots$ from $(3.1.2)_2$, we need

$$\mathbf{RCF}_\geq \;\vdash\; \exists z\,(y_1^2 + \cdots + y_s^2 = z^2), \qquad (3.2.1)_s$$

for $s = 3, 4, \ldots$, which follows by iterating $(3.2.1)_2$. Finally, to prove $(3.1.2)_s$ for any $s \geq 3$, add $y_1^2 + \cdots + y_s^2 = -1$ to the axioms of $\mathbf{RCF}_\geq$; then $(3.2.1)_{s-1}$ implies $\exists z\,(z^2 + y_s^2 = -1)$, contradicting $(3.1.2)_2$; then apply the deduction theorem. $\square$

Note also that we do not need to include any standard set of order "axioms," such as

$$a \geq b \leftrightarrow a - b \geq 0, \qquad (3.2.2)$$
$$a \geq 0 \;\wedge\; b \geq 0 \;\rightarrow\; a + b \geq 0, \qquad (3.2.3)$$
$$a \geq 0 \;\wedge\; b \geq 0 \;\rightarrow\; ab \geq 0, \qquad (3.2.4)$$
$$a \geq 0 \;\vee\; -a \geq 0, \text{ and} \qquad (3.2.5)$$
$$a \geq 0 \;\wedge\; -a \geq 0 \;\rightarrow\; a = 0; \qquad (3.2.6)$$

to prove each of these, we need (3.1.5)[13] (as well as (3.1.1)); in addition, (3.2.3) uses $(3.2.1)_2$; (3.2.5) uses (3.1.3); and (3.2.6) uses $(3.1.2)_2$.

3.3 The diagram of $(K, \geq)$

Let $(K, \geq)$ be an ordered field, and let $\mathbf{RCF}(K, \geq)$ be the system obtained from $\mathbf{RCF}_\geq$ by adding the diagram of $(K, \geq)$; i.e., to the language we add a constant symbol a_r for each $r \in K$, and to the list of axioms we add the atomic sentences $a_0 = 0$, $a_1 = 1$, and (for each $r, s \in K$) $a_r + a_s = a_{r+s}$, $a_r a_s = a_{rs}$, and (whenever $r > s$) $a_r > a_s$. (Often the diagram is defined to include *all* variable-free atomic formulae involving the a_r that are true in the model $(K, \geq)$, and the negations of those that are false in $(K, \geq)$; but these additional formulae are deducible from our smaller set of formulae.) The models of $\mathbf{RCF}(K, \geq)$ are the real closed order-extensions of $(K, \geq)$.

[13] Prestel noticed that (3.1.5) was missing from Kreisel's list in [1957b], [1958a], and [1960a].

4 Tarski's quantifier-elimination theorem (QE), and its consequences

4.1 Tarski's theorem [1930–67][14]

To any formula $A(y_1, \ldots, y_m)$ in the system $\mathbf{RCF}_{\geq}$ *(§3), we can, in a primitive recursive way, associate two objects:*

1. *a* **quantifier-free** *formula $B(y_1, \ldots, y_m)$ in* $\mathbf{RCF}_{\geq}$ *and*
2. *a* **proof** *in* $\mathbf{RCF}_{\geq}$ *of the equivalence $A \leftrightarrow B$.*

Here, our notation means that no variables besides $y_1, \ldots, y_m$ may occur free in A, and that, likewise, none besides $y_1, \ldots, y_m$ will occur (free) in B, either. And when we say primitive recursive, we are identifying formulae and proofs with their Gödel numbers (or we can use some other natural enumeration of all formulae).

Primitive recursiveness follows from the fact that the main tool in Tarski's construction is a generalized Sturm-chain $h_0, \ldots, h_m \in K[X_1]$ of a pair of polynomials $f, g \in K[X_1]$: $h_0 = f, \quad h_1 = g$, and, for $i \geq 1$, h_{i+1} is the negative of the remainder obtained by dividing h_{i-1} by h_i. If $(K, \geq)$ is an ordered field and R its real closure, and if $a, b \in R$ (with $a < b$ and $fg \neq 0$ at a, b), then the difference δ between the number of sign-changes at a, and the number of sign-changes at b, in this sequence, gives information about the number of points $x \in (a, b) \subset R$ at which f vanishes to a higher order than g (and by an odd integer), and such that f and g have the same sign just to the left of x. As Tarski pointed out in his Note 12 [1951, pp. 50–1], Sturm himself (1829, 1835) considered two special cases of this. The first (and well-known) case is where $g = f'$, when δ gives the number of real roots of f in (a, b). The second case is where f is square-free in $K[X_1]$, when δ gives the difference between

[14] Chapter 4 of Herbrand's thesis [1929] described a weak system of arithmetic and a method of QE for it, using which he reduced the proof of consistency for that system to the (easier) proof of the consistency of its quantifier-free fragment. At the end of Chapter 4, in §8.6, he wrote:

> "It seems probable to us that this [that is, the method of QE] would permit us equally to arrive at the consistency of the theory of real fields and of 'real closed' fields ('reell abgeschlossenen'; cf. *Artin and Schreier 1926*). But the methods of the next chapter lead us to that goal more easily."

Herbrand wrote this in [1929], shortly *before* the appearance of Tarski's [1930] announcement; it seems that the above intriguing quote is all that Herbrand ever wrote about real closed fields; he died in an accident in 1931. Another curiosity about Tarski's theorem is that Kreisel has said that Gödel also discovered it around 1930, but never published it.

the number of roots of f in (a, b) at which $f'g > 0$, and the number of those at which $f'g < 0$. (In 1853 Sylvester replaced g by $f'g$, thereby counting the difference between the number of roots of f in (a, b) at which $g > 0$, and the number of those at which $g < 0$; cf. Ben-Or, Kozen, Reif [1986, p. 255], and Coste, Roy [1988].)

Any atomic formula is provably equivalent in $\mathbf{RCF}_{\geq}$ (or even in the subtheory of ordered rings) to a Boolean combination of formulae of the form $p \geq 0$, for some terms p. If we make such changes to all the atomic formulae in a given quantifier-free formula B, we get a formula tautologically equivalent to a formula of the form $\bigvee_i \bigwedge_j (p_{ij} \geq 0 \wedge q_{ij} > 0)$, for finitely many terms p_{ij} and q_{ij} of $\mathbf{RCF}_{\geq}$; other "normal" forms are also possible, such as $\bigvee_i \left[p_i = 0 \ \wedge \ \bigwedge_j (q_{ij} > 0)\right]$.

Note that the theory of real closed fields does *not* admit QE, that is, if we omit the order-symbol $\geq$. Indeed, if $A(y_1)$ is $\exists x_1 \ (y_1 = x_1^2)$, then for no quantifier-free $B(y_1)$ in the language of rings alone can we prove $A \leftrightarrow B$ on the basis of (3.1.1–4).

4.2 Other proofs of (4.1)

Proofs by P.J. Cohen, Hörmander, Łojasiewicz [1965], A. Robinson, Seidenberg [1954],[15] and others, are also in Bochnak, et al [1987], Brumfiel [1979], Jacobson [1964, 1974, 1975, 1985], Kreisel, Krivine [1967–72], Prestel [1975–84], and Shoenfield [1967]; for a fine historical survey, cf. van den Dries [1988]. The primitive recursiveness of (almost all of) these presentations is obvious. Single- and double-exponential *lower* bounds on the worst-case complexity of any algorithm doing what Tarski's does were found by Fischer/Rabin in 1974, and by Davenport/Heintz and (independently) Weispfenning in 1988, respectively. An upper bound for (4.1)(1) that is double-exponential in the number q of quantifiers in A, was established in 1975, by Collins and L. Monk, independently (and "well-parallelized" by Fitchas/Galligo/Morgenstern and Ben-Or/Kozen/Reif in the late 80's); previous methods required an *un*bounded stack of exponentials, and so were not "elementary" recursive. Recently, q was reduced to q' := the number of alternating *blocks* of like quantifiers in A, by Grigor′ev (in the special case

[15] Hörmander [1955, pp. 163, 225–6] and Łojasiewicz [1965, p. 110] referred to Tarski's theorem as "Seidenberg's theorem." Gorin [1961, pp. 93–4, and §2] and Thom [1964, p. 256] called it "the Seidenberg-Tarski theorem." Brumfiel [1979], Bochnak, Coste, and Roy [1987], and many others (including me, in the early 80's) have called it "the Tarski-Seidenberg theorem." Prestel has pointed out that Seidenberg himself (in the very first sentence of [1954]) credited the theorem entirely to Tarski, and (correctly) claimed only to have given a new, simpler proof (of (4.1)(1), at least).

that A is a sentence, i.e., $m = 0$; 1988), and for arbitrary $m \geq 0$, by Heintz/Roy/Solernó and, independently, Renegar, in 1989; cf. Heintz, et al [1991] and Renegar [1991]; for further recent refinements, see Basu, Pollack, Roy [in preparation].

4.3 Formal proofs in $\mathbf{RCF}_{\geq}$ of $A \leftrightarrow B$ (4.1)(2)

These formal proofs are not emphasized in any of the numerous treatments mentioned above. But Tarski's proof comes closest to doing so, both explicitly, in Notes 9 and 15 [1951, pp. 48–50 and 53–4], and implicitly, inasmuch as he does not deal either with "polynomials in an indeterminate X_1" in the algebraist's sense, or with semialgebraic functions on the real algebraic numbers, but rather with *terms in the formal language* (some of which he calls "polynomials in τ," where τ is a sub-term), and formulae, respectively. Thus, given any single A, if we follow Tarski's original, semi-formal proof step by small step, it becomes only a series of easy exercises in the predicate calculus to write down the formal proof promised in (4.1)(2) (much as we did in our semi-formal proof above that $(3.1.2)_2$ suffices, and as we shall do when we come to Kreisel's second sketch (§6)). Although no one does this exercise, we must emphasize that the possibility of doing it is essential to Kreisel's second sketch (6.1.1); in fact, among the dozens of important applications of Tarski's theorem in the past 65 years (to model theory, semialgebraic geometry and topology, differential equations, etc.), Kreisel's second sketch was, for a long time, the only application actually to use the (effectiveness of the construction of the) formal proof in $\mathbf{RCF}_{\geq}$ constructed in (4.1)(2): he plugged that formal proof directly into Hilbert's first epsilon-theorem. Subsequent applications of the effectiveness of the construction of such formal proofs were along the same lines: Friedman [1981] (cf. (9.3) below) and Friedman, Simpson, Smith [1983] (cf. (5.13) below) gave new effective versions of Artin's theorem; and Scowcroft [1988] and Lombardi [1988–91] gave effective versions of Stengle's improvement (cf. (5.14) below) of Artin's theorem; all of these applications appealed to (the effectiveness of) QE.

We now carry out such a proof for the simple case in which A (or $A(a,b,c)$) is $\exists x(ax^2+bx+c=0)$, where a, b, and c are free(ly occurring) variables. The equivalent, quantifier-free formula B (or $B(a,b,c)$) that Tarski recursively constructs contains > 1000 symbols (give or take a few hundred), after unravelling all of his recursive abbreviations. Most of the atomic subformulae in B are repetitions of formulae such as $0=0$, $0=1$, $a^2<0$, $a^2>0$, $4a^3(b^2-4ac)<0$, etc. After numerous but easy simplifications, we get an equivalent formula such as B':

$$(a \neq 0 \wedge b^2 - 4ac \geq 0) \vee (a = 0 \wedge b \neq 0) \vee (a = 0 \wedge b = 0 \wedge c = 0). \quad (4.3.1)$$

Since Tarski dealt with a, b, c as *terms*, he could not assume that they had definite, determinate values (in some real closed field); he had to allow for all possibilities, e.g., $a < 0$, $a = 0$, $a > 0$, etc. (This is so even when dealing with a discrete ordered field $(K, \geq)$, where $\geq$ is decidable.) Thus when dividing $ax^2 + bx + c$ by its formal derivative $2ax + b$, we first consider the case $a \neq 0$, where we get

$$4a(ax^2 + bx + c) = (2ax + b)(2ax + b) - (b^2 - 4ac) \qquad (4.3.2)$$

(we multiply the dividend by $4a$ since our language does not (yet) have the symbol for reciprocals, e.g., $(2a)^{-1}$). Actually, something stronger than (4.3.2) can be proved in $\mathbf{RCF}_{\geq}$: after expanding out both sides as formal polynomials in x, the coefficient of x^2 on the left and the right hand sides are equal, and likewise for the coefficients of x, and the constant terms. Anyway, the second case we consider is where $a = 0$ but $b \neq 0$; then the remainder from dividing $ax^2 + bx + c$ by $2ax + b$ is *not* obtained by replacing a by 0 in the remainder $4ac - b^2$ computed in (4.3.2). (But since the degree of the divisor, $2ax + b$, drops to 0 when a is replaced by 0, we can already see that the correct remainder term is 0 in this second case.) The third case is $a = b = 0 \neq c$. In summary, we get the following three Sturm-chains:

Case 2: $a = 0 \neq b$:

$bx + c$
b

Case 3: $a = b = 0$:

c (4.3.3)*NotethattheB'in(4.3.1)containsadisjunctforeachofthesethre*

conditiononthebottomelementoftheSturm−chainforthatcase.(Tarski'soriginalBactu
conditionsonafewmorepolynomialtermsina,b,cbesidesthese;butinthissimpleexample,

We now sketch a proof in $\mathbf{RCF}_{\geq}$ that $B' \leftrightarrow \exists x\, A(x)$. We first prove $B' \rightarrow \exists x\, A(x)$. First suppose $a \neq 0$ and $b^2 - 4a \geq 0$. By (3.1.5), $\exists y\, (y^2 = b^2 - 4ac)$. By the field axiom for inverses, for this y, $\exists x (2ax = y - b)$. By (4.3.2), for this$x$, $4a(ax^2 + bx + c) = 0$; by another use of the axiom for inverses, $ax^2 + bx + c = 0$. Second, suppose $a = 0$

and $b \neq 0$; then by the axiom for inverses again, $\exists x\,(bx = -c)$, and this x makes $ax^2+bx+c=0$. The third case is even more trivial: just take $x = 0$. Thus $B' \to \exists x\, A(x)$. To prove the converse, let x be such that $ax^2+bx+c=0$; for this x, (4.3.2) implies $b^2-4ac = (2ax+b)(2ax+b)$, whence $b^2-4ac \geq 0$, by (3.1.5). Considering this same x, we can prove $a = 0 \wedge b = 0 \to c = 0$. These two conclusions amount to B'.

For more complicated formulae A, the main additional idea needed to prove $A \leftrightarrow B$ is the formal factorization of a polynomial into linear and irreducible quadratic factors over a real closed field, for this leads to the intermediate-value theorem for polynomials, the mean-value theorem (including Rolle's theorem), and related facts about the sign of the derivative; we shall formalize Artin and Schreier's proof of this in Part II of this paper, as part of Kreisel's first sketch.

4.4 The distraction of intuitionistic logic when applied to discrete, maximally ordered fields

Our purpose in the above formal proof is to indicate that the only possibly non-intuitionistic steps, i.e., the only applications of the law of excluded middle ($B \vee \neg B$, or, equivalently, $\neg\neg B \to B$), were to *atomic* formulae B. Such applications are justified provided we work over ordered fields $(K, \geq)$ that are *discrete*, i.e., such that the field operations of K are computable, and $\geq$ is decidable in a finite number of steps—in particular, such ordered fields satisfy the trichotomy axiom $a < 0 \vee a = 0 \vee a > 0$ intuitionistically. They include $(\mathbf{Q}, \geq)$ and its real closure, and suitably "explicitly" presented finitely generated extensions of these fields, since here questions of order can be effectively decided (say, by the interval-arithmetic of Kronecker [1882] or Vandiver [1936], to say nothing of Tarski's theorem).

Write $\mathbf{RCF}_{\geq}^{i,d}$ for the system $\mathbf{RCF}_{\geq}$, modified so as to use intuitionistic instead of classical logic, and to which we add the above trichotomy, or discreteness, axiom. In $\mathbf{RCF}_{\geq}^{i,d}$ one can prove the law of the excluded middle, first for quantifier-free formulae B, by induction on the complexity of B (using standard formulae, such as those on pp. 118–9 of Kleene [1952]; cf. also p. 101); and, second, for arbitrary formulae A, by the above observation that the $\mathbf{RCF}_{\geq}$-proof of $A \leftrightarrow B$, at which Tarski generously hinted, is already valid in $\mathbf{RCF}_{\geq}^{i,d}$ (or, if one is fussy, it becomes so with *very* obvious modifications, say, using, in addition, pp. 162–3 of Kleene [1952]).

Conclusion: $\mathbf{RCF}_{\geq}$ *and* $\mathbf{RCF}_{\geq}^{i,d}$ *prove the same theorems*; thus this matter is a *distraction,* and Tarski, and later Kreisel, were perfectly justified in using classical logic right from the beginning, tacitly for discrete ordered fields (and the grandiose name, "the formal intuitionistic theory of real closed discrete fields," introduced at the end of Lombardi, Roy [1991], can be replaced by "the theory of real closed fields"). Although neither Tarski nor Kreisel mentioned this small point in print, it would have been easy for them to do so using facts known at the time, given their familiarity with Tarski's original exposition. Moreover, in the early 80's Kreisel had called my attention to this point (cf. Delzell [1984, p. 371]), as well as the easy fact that the equivalence asserted in theorem 4.1 is intuitionistically false in general: take A to be $\forall x\,(x \geq 0 \,\vee\, x \leq 0)$; (4.1)(1) produces a quantifier-free, provable B, such as $0 = 0$, since A is provable classically but not intuitionistically. Indeed, A is intuitionistically false for $x \in \mathbf{R}$, simply because $(\mathbf{R}, \geq)$ is not discrete.[16] Even earlier, it was again Kreisel who, in the preface to the *second* [1971] (Engish) printing of Kreisel, Krivine, p. XII, wrote:

> "[T]he traditional procedures [of quantifier-elimination] are deceptive in another way too; for example, in the case of real algebra, they simply do not begin to touch the problems which are really significant for a constructive theory, namely the case of an *undecidable* ordering, as for the field of real numbers."

And already in the early 1960's, Kreisel had asked whether Artin's theorem admitted an intuitionistic proof or refutation—roughly, whether Artin's theorem admits continuous versions over $\mathbf{R}$ (cf. Part II, in preparation). So he was well aware of the suitability of classical and intuitionistic logic for discrete and non-discrete ordered fields, respectively.

The later, more widely read proofs of Tarski's theorem do not make this fact so obvious; on the contrary, they leave more to the constructivist's imagination than Tarski did on how to get *any* formal proof, even in $\mathbf{RCF}_{\geq}$, of $A \leftrightarrow B$. Lombardi and Roy were the first to publish (at the end of [1991]) a *proof* of the conclusion in the above paragraph. Their paper begins:

> "The classical theory of ordered fields (Artin-Schreier theory) makes intensive use of non-constructive methods, in particular the axiom of choice. However since Tarski (and

[16] Incidentally, this latter fact has been known for $>$ 100 years (even before Brouwer); cf. Troelstra, van Dalen [1988, Vol. I, pp. 21–2]. But Blum, Shub, and Smale's popular model of computation, purporting to deal with *all* real numbers, "conceives" of $(\mathbf{R}, \geq)$ as discrete. For discussions of this, cf., e.g., Kreisel [1990a, p. 228] (and my review, *MR* 93f:14029, of Heintz, et al [1991]), and Smale [1992, pp. 65–6].

> even since Sturm and Sylvester) one knows how to compute in the real closure of an ordered field K solely by computations in K. This apparent contradiction is solved in this paper. We give here a constructive proof of the first results of the theory of ordered fields, including the existence of the real closure."

They gave some interesting extensions of Sturm's theorem, the mean-value theorem, Thom's lemma, etc., from discrete real closed fields to arbitrary discrete ordered "d-closed" fields (i.e., those satisfying the intermediate-value property for polynomials of degree $\leq d$); in these extensions they preserved the (well-known) constructivity of the original proofs. After providing these tools, they stated the Tarski"-Seidenberg"[15] principle, and proved it as follows:

> "As in Bochnak, Coste, Roy [1987] by using Hörmander's method since all the tools needed are available."

All this could give the impression that they considered earlier proofs, such as Tarski's own, as not *already* constructive, and that Tarski's theorem needed a "correct" proof. In fact, however, their intention here was only to give a *simpler* proof of Tarski's theorem. It was at the end that they pointed out that their proofs work also in $\mathbf{RCF}_{\geq}^{i,d}$. (The main result of their paper, in my opinion, is their new construction of the real closure of a discrete ordered field; in Part II we shall show how this can be done in a different way, with pre-1956 methods from logic, due to Tarski and Kreisel.)

4.5 More on formal proofs in RCF$_{\geq}$ of $A \leftrightarrow B$ (4.1)(2)

(a) As to the complexity of the formal proof of $A \leftrightarrow B$, Friedman said [1981, pp. 2.1–2.2]:

> "A detailed examination of [Collins' and Monk's] proofs in these terms is very tedious, and is best left to someone else."

After giving some hints, he states (p. 2.3) that "a very safe bet" is that the number of symbols in the formal proof provided by Collins' and Monk's accelerated versions of Tarski's algorithm, is at most $2^{2^{2^{2^{2^p}}}}$, where p is the number of symbols in A (actually, his estimate is based on a formalization in a Gentzen-style sequent-calculus, and not the Hilbert-style system underlying our $\mathbf{RCF}_{\geq}$; and he points out that the bound depends on whether one uses the axioms in (2.2)(c), or those in (2.2)(a) and (3.1)).

I have not attempted to master Friedman's sophisticated techniques, and so I cannot judge the reliability of his claimed results. In the early nineties, however, H. Lombardi and M.-F. Roy began seeking (and, apparently, finding) their own elementary recursive bound on the size of this formal proof. More recently they have compared their methods to Friedman's. Lombardi tells me that as a result of their intensive study of the subject, they are skeptical that Friedman's cryptic hints really lead to any elementary recursive bound on the formal proof—i.e., they believe that the best bound obtainable from any straightforward analysis of Cohen's, Collins', or Monk's proofs is a stack of exponentials of (*non*constant) height q, where q is the number of variables in the given formula.

(b) Speaking of formal proofs, we note, finally, that while the equivalence $A \leftrightarrow B$ is provable in $\mathbf{RCF}_{\geq}$, Tarski's theorem itself (4.1) is not provable or even expressible in $\mathbf{RCF}_{\geq}$, since (4.1) involves quantification over—Gödel numbers of—formulae and proofs; when interpreting $\mathbf{RCF}_{\geq}$ in a model, quantified variables range only over some real closed field, and not over the power set of such a field, or over certain subsets such as $\mathbf{Z}$, $\mathbf{Q}$, or $\mathbf{N} := \{0, 1, 2, \ldots\}$. (4.1) can be proved, instead, in **PRA**, primitive recursive arithmetic, since the latter contains, by definition, a term for every primitive recursive function. But we shall not need this observation in this paper.

4.6 Logical consequences of QE

Tarski denoted the mapping $A \mapsto B$ promised in (4.1)(1) by U. He also constructed a function W mapping a formula B without any variables to one of the formulae $0 = 0$ or $0 = 1$ (by an obvious recursion on the complexity of B; cf. [1951, Definition 34, pp. 41–2]), in such a way that $\mathbf{RCF}_{\geq} \vdash [B \leftrightarrow W(B)]$. It follows that $\mathbf{RCF}_{\geq}$ is *complete*, i.e., for every sentence (= closed formula) A (i.e., for every formula with no free—occurrences of—variables), either $\mathbf{RCF}_{\geq} \vdash A$ or $\mathbf{RCF}_{\geq} \vdash \neg A$.

There is the following self-strengthening of (4.1): Let $(K, \geq)$ be an ordered field, and let $\mathbf{RCF}(K, \geq)$ be the system obtained from $\mathbf{RCF}_{\geq}$ by adding the diagram of $(K, \geq)$, as in (3.3). Then *every formula* $A(y_1, \ldots, y_m)$ *in* $\mathbf{RCF}(K, \geq)$ *is provably equivalent in* $\mathbf{RCF}(K, \geq)$ *to a quantifier-free formula* $B(y_1, \ldots, y_m)$ *in* $\mathbf{RCF}(K, \geq)$. This follows from (4.1) by extending the mapping U to formulae A of $\mathbf{RCF}(K, \geq)$, as follows. Replace the (finitely many) constants a_r occurring in A by new variables y'_r, yielding a formula $A'(y_1, \ldots, y_m, y'_{r_1}, \ldots, y'_{r_{m'}})$ in

$\mathbf{RCF}_{\geq}$; denote the quantifier-free, $\mathbf{RCF}_{\geq}$-equivalent formulae $U(A')$ by $B'(y_1, \ldots, y_m, y'_{r_1}, \ldots, y'_{r_{m'}})$; now replace the y'_r in B' by a_r; denote the resulting formula B of $\mathbf{RCF}(K, \geq)$ by $U(A)$. Finally, replace the y'_r by a_r throughout the $\mathbf{RCF}_{\geq}$-proof of the equivalence $A' \leftrightarrow B'$, as well.

$\mathbf{RCF}_{\geq}$ is also *model-complete*, which means that for every model $(R, \geq)$ of $\mathbf{RCF}_{\geq}$, $\mathbf{RCF}(R, \geq)$ is complete. Equivalently, if $(R_1, \geq_1)$ is a submodel of the model $(R_2, \geq_2)$ of $\mathbf{RCF}_{\geq}$, then a sentence of $\mathbf{RCF}(R_1, \geq_1)$ is true in $(R_1, \geq_1)$ iff it is true in $(R_2, \geq_2)$; the fact that $\mathbf{RCF}_{\geq}$ is model-complete is, essentially, Tarski's *transfer principle*. A syntactic characterization of model-completeness is that every formula in $\mathbf{RCF}_{\geq}$ is equivalent to an existential formula, i.e., a formula in prenex form with no universal quantifiers; this is just a weakening of (4.1), sufficient for some applications (such as A. Robinson's (5.4) and Henkin's (5.7–8), but not Kreisel's (6.50), which relies on (6.1.1)).

Tarski went on to argue (in Note 15, p. 53) that $\mathbf{RCF}_{\geq}$ is *consistent*, by a proof that "has what is called a constructive character," as follows. First, one must note that most of his proof was a recursive definition of U (24 pages long!), based on the complexity of A; one of the clauses (p. 38) was that for A atomic, $U(A)$ is defined to be A; similarly, one of the clauses (p. 41) in the (short) definition of W is that $W(s = t)$ is defined to be $0 = 0$, provided that the variable-free terms s and t have equal "value" (defined in an obvious way by recursion on the complexity of s and t); in particular, $WU(0 = 0)$ is $0 = 0$. Second, he said that we can "easily" show, for every sentence A and every sentence B, that

1. $WU(\neg B)$ is $0 = 1$ iff $WU(B)$ is $0 = 0$, and
2. if $\mathbf{RCF}_{\geq} \vdash A$, then $WU(A)$ is $0 = 0$;

taking B to be $0 = 0$ and A to be $0 \neq 0$, (1) and (2) would imply that $\mathbf{RCF}_{\geq} \nvdash 0 \neq 0$, i.e., consistency.

To show (1), one must note that another of the clauses in the definition of U was that $U(\neg B)$ is defined to be $\neg U(B)$; similarly, another clause in the definition of W was that for any variable-free formula C (such as $U(B)$, where B is a sentence), $W(\neg C)$ is defined to be $0 = 1$ or $0 = 0$, according as $W(C)$ is $0 = 0$ or $0 = 1$; (1) follows.

But (2) escapes me: his definition of $U(A)$ had nothing obvious to do with the existence or non-existence of any hypothetical proof of A; and indeed, $U(A)$ is defined even when A is neither a sentence nor provable, such as $x = 0$ or $0 \neq 0$ (for which we know there are no proofs at all, once the argument is completed). As mentioned in footnote 14 above, in 1929 Herbrand [1971, p. 124] used a much simpler quantifier-elimination mapping U' for a weak system of arithmetic to reduce the proof of the consistency of that system to the consistency of its quantifier-free fragment. He used induction on proofs: first, those logical and non-logical *axioms* A involving quantifiers are mapped by U' to *theorems* of the quantifier-free system, and, second, if $U'(A)$ and $U'(A')$ are theorems of the quantifier-free system, then so is $U'(B)$, for any formula B obtained from A and A' by any rule of inference. In the last sentence of Chapter 4 [1971, p. 132], Herbrand expressed the hope that something similar could be done for $\mathbf{RCF}_{\geq}$. Perhaps this is also what Tarski intended: i.e., to reduce, via U, the proof of the consistency of $\mathbf{RCF}_{\geq}$ to that of the quantifier-free theory of ordered fields. But even some of the *axioms* A of $\mathbf{RCF}_{\geq}$, such as $\exists x\ (x^3 + a_2x^2 + a_1x + a_0 = 0)$, are mapped by U to formulae that seem to be thousands of characters long; it does not look "easy" to me to prove $U(A)$ in the theory of ordered fields (i.e., without the use of the axioms for square roots or odd-degree roots), or to show, by purely syntactical means, that the closure $\forall a_0 \forall a_1 \forall a_2\, A(a_0, a_1, a_2)$ of A is mapped by WU to $0 = 0$.

When we say "by purely syntactical means," we mean, in particular, "without presupposing the existence of a real closed field," and, indeed a finitarily constructed real closed field, such as that of the real algebraic numbers $\overline{(\mathbf{Q}, \geq)}$; apparently Tarski was trying to avoid using this well-known fact, which makes the finitary consistency of $\mathbf{RCF}_{\geq}$ obvious even without U. In Part II we shall summarize the history of constructions of the real closure of an ordered field.

Open Question 4.7 *Were Herbrand and Tarski* **both** *mistaken about how to prove the consistency of* $\mathbf{RCF}_{\geq}$*?*

In (8.1) we give a finitary, syntactic proof of the consistency of $\mathbf{RCF}(K, \geq)$, by a variation on Kreisel's second sketch; this consistency will lead easily, in Part II, to a finitary construction of the real closure of $(K, \geq)$.

Once consistency is established (by one means or another), the *decidability* of—the set of theorems of—$\mathbf{RCF}_{\geq}$ follows from the completeness of $\mathbf{RCF}_{\geq}$.[17] (Here, the decision procedure would be general recursive even if the mapping WU were not, for we can systematically search through all possible proofs, until we find one of A or $\neg A$. But since WU is actually primitive recursive, the decision procedure is,

[17] Actually, even an inconsistent system is decidable trivially.

too.)

We have seen that QE leads easily to several other properties, such as decidability, completeness, and model completeness (although apparently not consistency; recall (4.7)). In the opposite direction, A. Robinson proved in 1958 that certain model-complete theories automatically admit QE; some (more useful) variants of Robinson's model-theoretic criterion for QE appeared in Kreisel, Krivine [1967–72], Feferman [1974], and Shoenfield (1970 and 1977); (cf. van den Dries [1988] for details).[18]

5 Refinements of Artin's theorem

> *On the other hand, our proof [of (2.4)] is indirect, and provides no explicit method for the decomposition [(2.4.1)]. However, one may well expect that the proof could be completed in this direction.* Artin [1926, p. 110].

And Kreisel reported [1977a, p. 115]:

> "In seminars Artin had raised the problem of finding bounds, ever since his original proof in the twenties."

First we review some partial results by others.

5.1 Artin-Schreier theory

One source of apparent indirectness in Artin's proof of (2.4) was his reliance on his and Schreier's "construction" of

an ordering $\geq$ of a given formally real field K (2.1), and (5.1.1)

the real closure of an ordered field $(K, \geq)$ (2.2). (5.1.2)

(The *uniqueness* of the real closure is unnecessary for (2.4).) As we stated in (1.8) and (4.4), we shall save until Part II of this paper a

[18] Sacks gave another proof of Tarski's theorem in [1972]. van den Dries informed me of a gap (on request, Sacks was aware of it, too): on p. 93 it was assumed, implicitly, that for an arbitrary ordered field $(K, \geq)$, if x and y in two ordered extension fields of $(K, \geq)$ are transcendental over K and make the same Dedekind cut in K, then the field isomorphism $K(x) \to K(y)$ sending x to y (and fixing K) preserves order; this is equivalent to the assumption that an *arbitrary* ordered field $(K, \geq)$ is dense in its real closure (which is false: recall (2.3) above). (Lang had already made a similar error in [1965], by relying on the weaker, but still false, assumption that if a field is *uniquely* orderable, then it is dense in its real closure—cf. (5.10) below.)

discussion of the history of efforts (notably by Zassenhaus' and Artin's student Hollkott in [1941]) to make (5.1.2) finitary. One reason is that the improvements to Artin's main theorem, listed below, do not directly address (5.1.1–2); and indeed one of Kreisel's observations was that it is not *necessary* to address (5.1.1–2) directly in order to answer Artin's main question (although some of Kreisel's methods lead to an easy construction of (5.1.2), at least; cf. Part II).

5.2 Habicht

Habicht [1940] considered *homogeneous* $F \in \mathbf{R}[X]$ that are *strictly* definite (tacitly, over $\mathbf{R}^n \setminus \{(0,\ldots,0)\}$); upon de-homogenizing, say by setting $X_n = 1$, we get an $f \in \mathbf{R}[X_1,\ldots,X_{n-1}]$ such that

$$f(x_1,\ldots,x_{n-1}) \geq \epsilon \cdot (x_1^2 + \cdots + x_{n-1}^2 + 1)^{d/2},$$

where $\epsilon = \min\{\, F(x) \mid \sum_{i=1}^n x_i^2 = 1 \,\} > 0$ and $d = \deg F$ ($= \deg f$, since F is assumed to be strictly definite); conversely, if

$$f(x_1,\ldots,x_{n-1}) \geq \epsilon \cdot (x_1^2 + \cdots + x_{n-1}^2 + 1)^{d/2}$$

for $d = \deg f$ and for some $\epsilon > 0$, then f homogenizes to a strictly definite F.[19] (By QE, ϵ is algebraic over the subfield of $\mathbf{R}$ generated by the coefficients of f.) Habicht gave a straightforward algorithm for representing such F as $\sum_j r_j^2$, with homogeneous $r_j \in \mathbf{R}(X)$; and if the coefficients of F are rational, then we can arrange the same for those of the r_j. He used the following result of Pólya (1928): if $G \in \mathbf{R}[X]$ is homogeneous and takes positive values whenever each $x_i \geq 0$ (except possibly at the origin), then $(X_1^2 + \cdots + X_n^2)^e G(X)$ is a form with positive coefficients, for some $e \geq 0$. (Pólya's theorem had been proved by Poincaré in 1888 for $n = 2$, and by Meissner in 1911 for $n = 3$; Meissner's method applies to $n > 3$, but does not lead to so simple a result; cf. Hardy, et al [1934–91] for references.) In [1982b] we stated that Habicht arranged for the common denominator of his r_j to be $(X_1^2 + \cdots + X_n^2)^e$. This is not quite correct: he *began* with such a

[19] Caution: (a) E.A. Gorin's [1961] inhomogeneous quartic $f_0 := X_1^2 + (1 - X_1X_2)^2$ is (strictly) definite over $\mathbf{R}^2$, but has infimum 0, and so is not considered by Habicht; f_0 has a zero on the line at infinity, as we see from its homogenization $F_0 := X_1^2X_3^2 + (X_3^2 - X_1X_2)^2$, which has the non-trivial zero $(0,1,0)$.

(b) Even if $f \in \mathbf{R}[X_1,\ldots,X_{n-1}]$ has *positive* infimum ϵ on $\mathbf{R}^{n-1}$, its homogenization $F \in \mathbf{R}[X]$ need not be strictly definite on $\mathbf{R}^n \setminus \{(0,\ldots,0)\}$. For example, take $f = \epsilon + f_0$ (with f_0 as in (a)), which has infimum ϵ; then $F = \epsilon X_3^4 + F_0$, which still has the zero $(0,1,0)$. Thus not even all $f \in \mathbf{R}[X_1,\ldots,X_{n-1}]$ with positive infimum are handled by Habicht's method, contrary to p. 109 of—the original French version of—Bochnak, et al [1987]; I thank H. Lombardi for this reference.

denominator, but transformed it several times before getting the final denominator. However, one can still verify that his final denonimator is strictly definite on $\mathbf{R}^n \setminus \{(0,\ldots,0)\}$ (although Habicht did not mention it); this partially answers a question, on the "bad points" of F, that was considered repeatedly from 1955 to 1977; cf. (9.3)(b)(ii) below. For recent improvements of Habicht's result, cf. Reznick [1995] and de Loera, Santos [1995].

5.3 Lang

Lang [1953, p. 387] weakened Artin's hypothesis (2.4.1) that K be Archimedean and uniquely orderable. Now K may be an arbitrary formally real field, and $f \in K(X)$; but we must now assume that for all "zero-dimensional real places" ϕ of $K(X)/K$, either $\phi(f) = \infty$ or, for *any* ordering $\geq$ of $\phi(K(X))$, $\phi(f) \geq 0$; then the conclusion is that f is a sum of squares in $K(X)$. Translated into the vernacular, suppose that K is formally real, and, for every algebraic, formally real extension L of K, and for every $x \in L^n$, either $f(x)$ is undefined, or, for any ordering $\geq$ of L, $f(x) \geq 0$; then f is a sum of squares in $K(X)$. Equivalently, recalling the terminology "positive semidefinite" (psd) of (2.4): for K formally real[20]:

$$\begin{gathered}\text{if, for } \textit{every} \text{ ordering } \geq \text{ of } K,\ f \text{ is psd over the real closure of } (K, \geq),\\ \text{then } f \text{ is a sum of squares in } K(X).\end{gathered} \tag{5.3.1}$$

Corollary:

$$\begin{gathered}\text{if } K \text{ is } \textit{uniquely} \text{ orderable and } f \text{ is psd over the real closure of } K,\\ \text{then } f \text{ is a sum of squares in } K(X).\end{gathered} \tag{5.3.2}$$

(5.3.2) contains Artin's theorem (2.4.1), for if K is not only uniquely orderable, but Archimedean as well, and if f is psd over K, then, as discussed in (2.3–4), f is already psd over the real closure of $(K, \geq)$. In [1971, 1993] Lang attributed (5.3.2) to Artin, without explicitly mentioning its corollary (2.4.1), which is what Artin actually stated in [1926]. On the other hand, the methods of Artin [1926] were more than adequate to prove (5.3.1–2). For example, (5.3.1) is reminiscent of another one of Artin's sufficient conditions for $f \in K[X]$ to be a sum of squares in $K(X)$: K is an arbitrary (i.e., not necessarily formally real) algebraic number field, and f is "totally definite," i.e., $\forall x \in K^n$,

[20] Lang did not mention it, but (5.3.1) remains true even if K is not formally real, provided that the characteristic of K is not 2, using the well-known identity $f = [(f+1)/2]^2 + (-1)[(f-1)/2]^2$.

$f(x)$ is totally positive ($:\leftrightarrow f(x) \geq 0$ in every ordering of $K \leftrightarrow f(x) =$ a sum of squares in K).

5.4 A. Robinson

Robinson was the first[21] logician to contribute to Artin's theorem. In [1955] (cf. also [1963], and Kreisel, Krivine [1967–72, pp. 74–6]) he showed that in Artin's theorem (2.4.1), the Archimedean hypothesis on the uniquely orderable field $(K, \geq)$ can be replaced by the hypothesis that K be real closed: i.e., if R is real closed, then

$$\text{if } f \in R(X) \text{ is psd over } R, \text{ then } f = \sum_i r_i^2, \text{ for some } r_i \in R(X). \tag{5.4.1}$$

For if the conclusion fails, then $f < 0$ relative to some ordering $\geq$ of $R(X)$ (by (2.1)), whence $(\exists z \in S^n)$ $(f(z) < 0)$, where S is the real closure of $(R(X), \geq)$ (taking $z_i = X_i \in S$); the last (elementary) statement remains true when S is replaced by R, by model-completeness (4.6). Actually, (5.4.1) is a special case of corollary (5.3.2) of Lang's

[21](a) In [1977a], Kreisel wrote (p. 115):

> "[A. Robinson's] proof appeared in Robinson [1955], not long after the proof theoretic solution had been found."

This statement is impeccable, for a legalistic temperament: Kreisel stated (in (1.3) above) that he had found his "proof theoretic solution" in November 1955, while Robinson's paper appeared in the issue dated December 16, 1955. (Robinson's paper was submitted on April 30, 1955.)

(b) The same temperament was even more strongly manifested in Kreisel [1958a, p. 170]:

> "Later Skewes [1955] published an *ad hoc* solution which gave the same bound as the method of Kreisel [1951–52]."

(Recall (1.9) above; cf. also Feferman's contribution to this volume.) Here, the word "Later" (like the word "after" in the quotation in (a) above) is impeccable, legally. On the other hand, while Kreisel (in [1951–52]) had cited Littlewood's [1948] report, he never called attention to the following two statements there:

> "[A]ssuming R.H. [the Riemann hypothesis] ... Dr Skewes found [1933] a new line of approach leading to [the bound] $10^{10^{10^{34}}}$."

And on p. 169:

> "[S]uch a [bound], free of hypotheses, was found by Dr Skewes in 1937; his work has not yet been published, [footnote] but it should be before very long."

The footnote read:

> "It is accessible in a thesis deposited in the Cambridge University Library."

(This footnote was revised in the later reprints of Littlewood's article, e.g., [1968].)

theorem, since real closed fields are uniquely orderable; but logicians in the 50's seem to have overlooked Lang's [1953] paper.

Robinson also got a weighted sum-of-squares representation for each $f \in R(X)$ that is nonnegative or undefined on the set in R^n where given $g_1, \dots, g_t \in R(X)$ are all positive or undefined; this was the first, weak "*Positivstellensatz*"; cf. (5.14) below for modern versions.

5.5 Robinson's parametrizations

Parametrized versions of Artin's theorem were also first considered by Robinson. We introduce the following notation. Fix an integer $d \geq 0$. Let $m = \binom{n+d}{n}$, and let $C := (C_1, \dots, C_m)$ and $C' := (C'_1, \dots, C'_m)$ be additional indeterminates. Let $F \in \mathbf{Z}[C; X]$ be the general polynomial of degree d in X with coefficients C: $F = \sum_{|e| \leq d} C_{g(e)} \prod_{i=1}^n X_i^{e_i}$, where $e := (e_1, \dots, e_n) \in \mathbf{N}^n$ is a multi-index, $|e| = \sum_i e_i$, and g is a bijection from $\{ e : |e| \leq d \}$ to $\{1, 2, \dots, m\}$. For a real closed field R, let c and c' denote elements of R^m with $c' \neq (0, \dots, 0) \in R^m$. Define $P'_{nd} := P'_{n,d,R}$ by

$$P'_{nd} = \{ (c, c') \mid F(c; X)/F(c'; X) \in R(X) \text{ is psd in } X \text{ over } R \}.$$

For $b = 0, 1, \dots$, Robinson considered the formula

$$\exists r_1, \dots, r_b \in R(X) \text{ with } \deg r_i \leq b^{22} \text{ s.t. } F(c; X)/F(c'; X) = \textstyle\sum_i r_i^2;$$

this can be expressed by a formula $A_b(c, c')$ of $\mathbf{RCF}_{\geq}$. Since $(c, c') \notin P'_{n,d,R'}$ holds in every model R' of the axioms for real closed fields augmented by the axioms $\neg A_0, \neg A_1, \dots$ (by (5.4.1)), the compactness theorem implies that, for some $\lambda := \lambda(n, d) \in \mathbf{N}$, for every real closed field R, $(c, c') \notin P'_{n,d,R}$ follows from $\mathbf{RCF}_{\geq} + \neg A_1 + \cdots + \neg A_\lambda$; i.e.,

$$(\exists \lambda \in \mathbf{N})\ (\forall R)\ (\forall (c, c') \in P'_{nd})\ \exists r_1, \dots, r_\lambda \in R(X) \text{ with } \deg r_i \leq \lambda$$
$$\text{such that} \quad F(c; X)/F(c'; X) = \textstyle\sum_i r_i(X)^2. \tag{5.5.1}$$

I.e., the number and degrees of the $r_i \in R(X)$ required in (5.4.1) depend only on n and d, not R or (c, c'). Once $\lambda(n, d)$ is known, we can compute the (coefficients of the) r_i as "semialgebraic" functions of (c, c'). Robinson did not mention it in [1955], but the minimum value of $\lambda(n, d)$ as in (5.5.1) is obviously a *general recursive* function of n and d, since the statement in (5.5.1) (apart from the prefix "$\exists \lambda \in \mathbf{N}$, $\forall R$") is (even primitive) recursively decidable for each λ, by QE. Although this

[22]For $g, h \in R[X]$ relatively prime, $\deg(g/h)$ is defined to be $\max\{\deg g, \deg h\}$, as usual.

means that λ could be found from n and d in a finite number of steps, the only apparent way to find out *how* many steps is to search through all possible values of λ until the first correct one is found; many consider such an algorithm to be not finitary, since the proof of its termination (i.e., (5.5.1)) did not appear to be finitary by the model theory known at the time. (On the other hand, Robinson did not particularly set out to get overtly finitistic results, either for Artin's theorem or in the many other areas of mathematics to which he applied model theory.)

5.6 Kreisel

(Recall (1.3) and footnote 21(a) of (5.4).) Kreisel's 1955 work on Artin's theorem began by considering (after trivial changes of notation) "a commutative [formally] real field K in which every positive element can be represented as the sum of $\leq k$ squares" ([1958a, p. 165]; similar formulations are in Kreisel [1957b, p. 99], and Daykin [1961, p. 124]). The word "positive" means little in "a commutative [formally] real field"; for example, is $\sqrt{2}$ "positive" in the formally real field $K := \mathbf{Q}(\sqrt{2})$? (Answer: It is positive in *one* ordering of $\mathbf{Q}(\sqrt{2})$, and negative in the other; recall (2.1) above.) One *a priori* plausible reading of this condition on K is obtained by replacing "positive" by "totally positive" (recall (2.1)); on this reading, $\mathbf{Q}(\sqrt{2})$ does, indeed, satisfy the condition (with $k = 4$, by a theorem of Siegel (1921)). But Kreisel's formulation ((5.6.1) below) of Artin's theorem does not hold with $K = \mathbf{Q}(\sqrt{2})$ (cf. footnote 23(b) below). So he must have intended to consider "a *uniquely* orderable field K in which every positive element is the sum of $\leq k$ squares"; $\mathbf{Q}(\sqrt{2})$ does *not* satisfy this condition, leaving (5.6.1) intact.[23]

The modern word here is the *Pythagoras number* $P(K)$ of an arbitrary field (or even commutative ring) K, which is defined to be the least number $k \le \infty$ such that every sum of squares in K is the sum of k squares in K. Thus the fields K apparently considered by Kreisel were those that are *uniquely* orderable and have $P(K) \le k$. Letting R be a real closed field[24] whose unique order extends that of K, Kreisel sketched proofs of the following theorem:

$$\text{if } f \in K[X] \text{ is psd over } R, \text{ then } f = \sum_i r_i^2, \tag{5.6.1}$$

for some $r_i \in K(X)$. Evidently Kreisel's (5.6.1), like Robinson's (5.4.1), was a re-discovery of corollary (5.3.2) of Lang's theorem [1953].

While Kreisel's attempt to clarify Artin's hypotheses on K was "a day late and a dollar short," his real aim (to constructivize Artin's proof) succeeded admirably. From either of his two sketches one could obtain a bound $\lambda(n, d, k)$ on the number and degrees of the r_i in (5.6.1), analogous to Robinson's bound $\lambda(n, d)$ for (5.5.1). But in [1960a] Kreisel stressed that while Robinson's (and Henkin's) $\lambda(n, d)$ is *general* recursive, Kreisel's $\lambda(n, d, k)$ is *primitive* recursive (although on p. 167 of [1958a] he pointed out that

---

[23](a) A doubly clumsy formulation occurred in Kreisel [1977a, p. 115]:

> "Artin's solution of Hilbert's problem was stated for archimedean formally real, that is, orderable fields K in which every positive element is a sum of $\le k$ squares."

In this case, it is not only the word "positive" that means little in a "formally real, that is, orderable field"; it is also the word "archimedean" (recall (2.3) above). For example, is the formally real field $K := \mathbf{Q}(X_1)$ "archimedean"? (Answer: some of its *orderings* are, such as the one determined by declaring any $r \in K$ to be "positive" if and only if $r(\pi) > 0$ (where $>$ is the unique ordering on $\mathbf{R}$); and other orderings are not, as remarked in (2.3) above.) As above, Kreisel overlooked the necessary condition (mentioned by Artin) that K be *uniquely* orderable. Correct formulations are in Daykin [1961, p. 138] and Kreisel [1969, p. 113].

(b) Actually, even Hilbert failed to require that the ground field be uniquely orderable when he formulated his 17th problem in [1900]. He seemed to allow K to be an arbitrary subfield of $\mathbf{R}$; in this level of generality, his conjecture is *false*: as usual, take $K = \mathbf{Q}(\sqrt{2})$ (this time ordered so that $\sqrt{2} > 0$), and $f(X) \equiv \sqrt{2} \in K[X]$ (i.e., $d = 0$, and n is irrelevant); then f is positive definite relative to this ordering on K, but is not a sum of squares in $K(X)$.

[24]Here again we see the need for the *unique* orderability of K: In [1958a, p. 165], and elsewhere, Kreisel wrote, "if $f \ge 0$ in some real-closed extension of the field generated by the coefficients" of f, then f is (allegedly) a sum of squares in $K(X)$. However, if K admits more than one ordering, then $f \in K[X]$ could be positive semidefinite over some real closed extensions of K, but not others; this is enough to ruin the desired conclusion, as footnote 23(b) above shows.

> "[this] conventional distinction ... misses the mark (though it is formally correct)[, because] this difference is overstated ...";

cf. (5.12) below). In fact, at least from the *first* of Kreisel's two sketches (Part II of this paper), $\lambda(n, d, k)$ could be calculated explicitly as an iterated-exponential function of its parameters; cf. also (9.1) and (9.2) below. (Of course, the field K to which Kreisel's algorithm is finally applied must be discrete, too.)

Some have wondered whether Kreisel's introduction of the hypothesis $P(K) \leq k$ into Artin's theorem constitutes a forbidden "new idea," in conflict with Kreisel's ideal of—a certain kind of—purity of method ((1.9) above). But while the idea was obviously new for the original, published version of (2.4.1), Artin may well have already observed on his own (i.e., before posing his question to Kreisel in 1955), that this new hypothesis can hardly be forbidden if there is to be the slightest hope of unwinding (2.4.1) at all: there is *no* bound $\lambda'(n, d)$, independent of k or K, on the number of squares in (2.4.1). In fact, even $\lambda'(0, 0)$ could not exist (i.e., for *constant* $f \in K$), since it would obviously have to be $\geq P(K)$ for every uniquely orderable K. However, it was not definitely known at the time that $P(K)$ could be arbitrarily large (or even infinite) for uniquely orderable K (Archimedean or not): until 1977, the only numbers known to occur as $P(K)$ for some uniquely orderable field K were 1, 2, and 4. In [1978] Bröcker constructed *non*-Archimedean uniquely orderable K with $P(K)$ arbitrarily large; and in [1978], Prestel constructed Archimedean examples with these properties; he even arranged for $P(K)$ to be any prescribed number in the set $\{\, 2^n, 2^n+1 \mid n < \infty \,\} \cup \{\infty\}$ (and asked whether other numbers are possible). Anyway, after Henkin's improved formulation ((5.7.1), below), allowing nonnegative weights $p_i \in K$ on the squares r_i^2 in Artin's theorem, $P(K)$ and k become irrelevant to the number of *weighted* squares. And Kreisel's sketches of (5.6.1) are easily adapted to (5.7.1), as we shall show in §6 below; thus, even if this "new idea" is repugnant, it is easily dispensed with after Henkin's insight.

Kreisel stated (5.6.1) only for polynomials $f \in K[X]$, and not for arbitrary rational functions $f \in K(X)$, as Artin, Lang, and A. Robinson had done in (2.4.1), (5.3.1–2), and (5.4.1) and (5.5.1). To extend (5.6.1) to $f \in K(X)$, one could alter Kreisel's proof slightly. However, one can, instead, first deduce directly, for every $f, g \in K[X] \setminus \{0\}$ relatively prime, that if f/g is psd over R, then so are f and g (or $-f$ and $-g$). Indeed, for $x \in R^n$, if either f or g changes sign near x, then so does the other; the "transversal zeros theorem" would then imply that f and g have a common, non-constant factor in $K[X]$, contradiction. (The transversal zeros theorem seems to have been folklore in real algebraic geometry at least until Dubois and Efroymson published the main idea in [1970, Theorem 2.7]; a complete statement appeared in Choi, Lam, Knebusch, and Reznick [1982] (especially Corollary 2.5).) Second, once we know that f and g are psd over R, (5.6.1) gives $f = \sum r_i^2$ and $g = \sum s_j^2 \quad (r_i, s_j \in K(X))$, whence $f/g = (\sum r_i^2)/\sum s_j^2 = (\sum r_i^2 \sum s_j^2)/(\sum s_j^2)^2 = \sum_{i,j}(r_i s_j / \sum s_k^2)^2$.

5.7 Henkin

L. Henkin got interested in Artin's theorem in 1956, after hearing a lecture at Berkeley by Kreisel. In [1960] Henkin gave a new version of Robinson's model-theoretic treatment (5.4) of Artin's theorem. While apparently unaware of Lang's formulation (5.3.2), he improved the latter (and hence also Robinson's and Kreisel's formulations, except for the overtly finitary character of Kreisel's proof), by removing not only the Archimedean hypothesis on K, but also the unique orderability hypothesis: let R be a real closed field, and let K be a subfield, with the inherited order. Then Henkin proved:

$$\text{if } f \in K[X] \text{ is psd over } R, \text{ then } f = \sum_i p_i r_i^2, \tag{5.7.1}$$

for some $r_i \in K(X)$ and $p_i \in K$ with $p_i \geq 0$. The price for this relaxation of the hypotheses on K is the introduction of the nonnegative weights p_i; but this price is small, since if K happens to be uniquely orderable, then each p_i is a sum of squares in K, which can be absorbed into the r_i. Despite the effortless superiority of (5.7.1) over all other formulations, for some reason it is rarely presented in expositions of Artin's theorem (e.g., not in Jacobson [1964, 1975, 1980, 1989], or Lang [1971, 1993], which still give the needlessly restrictive formulations).

Next, Henkin improved Robinson's parametrized version (5.5.1) of Artin's theorem. Robinson had shown that the number and degrees of the rational functions r_i can be bounded in terms of n and d, independent of (R or) the coefficients c, c' of the general rational function $F(c; X)/F(c'; X)$ of degree d in n variables; but in [1955] he said nothing about how, for fixed n and d, the coefficients of the r_i (or the p_i in (5.7.1)) depend on c, c'. Henkin showed that they depend *piecewise-rationally* on c, c'. Actually, Henkin, like Kreisel, considered the general *polynomial* F (rather than the general rational function) of degree d in n variables; so now define

$$P_{nd} := P_{n,d,R} = \{\, c \in R^m \mid F(c; X) \in R[X] \text{ is psd in } X \text{ over } R \,\}.$$

Theorem 5.8 (Henkin [1960, pp. 287–8]) *For each n and d there exists a partition of P_{nd} into subsets $D_1, \ldots, D_l$ that are "$\mathbf{Z}$-semialgebraic" (= definable over $\mathbf{Z}$; cf. (5.9)(a) below), and for each $i \leq l$, there exist $p_{ij} \in \mathbf{Z}[C]$ and $r_{ij} \in \mathbf{Q}(C; X)$ such that, $\forall c \in D_i$,*

$$F(c; X) = \sum_j p_{ij}(c) r_{ij}(c; X)^2, \quad \textit{and} \tag{5.8.1}$$

$$\bigwedge_j [p_{ij}(c) \geq 0 \ \wedge \ \textit{the denominator} \in R[X] \ \textit{of}\ r_{ij}(c; X) \ \textit{is not}\ 0]. \tag{5.8.2}$$

(All this is uniform in R.)

van den Dries [1977, pp. 113 and 131–2] generalized (5.8) to commutative f-rings R that are "von Neumann regular" (i.e., that satisfy $\forall a \exists b \, (a^2 b = a)$).

5.9 Kreisel's premature refinement of (5.8)

(a) Let R be real closed, and K a subdomain of R. A subset S of R^m is called (K-)*subbasic closed semialgebraic* if $S = \{ c \in R^m \mid p(c) \geq 0 \}$, for some $p \in K[C]$, where $C = (C_1, \ldots, C_m)$; a subset of R^m is called *semialgebraic* if it is a Boolean combination of subbasic closed semialgebraic sets. Thus P_{nd} is $\mathbf{Z}$-semialgebraic, for by eliminating the quantifiers (4.1) from the condition $(\forall x \in R^n)\, F(c, x) \geq 0$ on c, we get $P_{nd} = E_1 \cup \cdots \cup E_l$, where each E_i is of the form

$$\left\{ c \in R^m \;\middle|\; \left(\bigwedge_j (p_{ij}(c) \geq 0) \right) \wedge \bigwedge_j (q_{ij}(c) > 0) \right\},$$

for finitely many $p_{ij}, q_{ij} \in \mathbf{Z}[C]$. In [1960a, p. 315], Kreisel claimed more: that no q_{ij}'s (i.e., no $>$ symbols) are necessary here. This fact does not follow from Tarski's theorem alone. It was not until the late 70's that several proofs were found (e.g., Bochnak, et al [1987], Coste, Roy [1979], Delzell [1982a], van den Dries [1982]) of a refinement of Tarski's theorem now known as the "finiteness theorem for closed (K-) semialgebraic sets" (such as P_{nd}, with $K = \mathbf{Z}$): such sets can be written in the form $E_1 \cup \cdots \cup E_l$, where each E_i is a "basic closed (K-) semialgebraic set," i.e., a set of the form $\{ c \in R^m \mid \bigwedge_j (p_{ij}(c) \geq 0) \}$, for finitely many $p_{ij} \in K[C]$. (Łojasiewicz had proved the analog for semi-analytic sets in [1965].) The fact that P_{nd} is closed relative to the usual Euclidean topology on R^m (induced by the order topology on R) amounts to the fact that the limit of a convergent sequence of psd polynomials of degree $\leq d$ in n variables is again psd.

(b) In [1982a,b] I gave a simple proof that for each n and d there exist finitely many $g_{iJ} \in \mathbf{Z}[C]$ and $r_{iJ} \in \mathbf{Q}(C;X)$ such that:

$$\bigwedge_i \quad F(C;X) = \sum_J g_{iJ}(C) r_{iJ}(C;X)^2, \quad \text{and} \tag{5.9.1}$$

$$(\forall c \in P_{nd}) \bigvee_i \bigwedge_J \Big[g_{iJ}(c) \geq 0 \wedge \big[r_{iJ}(c;X)\text{'s denominator } \neq 0 \in R[X] \big] \Big]. \tag{5.9.2}$$

(5.9.1–2), like (5.8), is uniform in R. Here, for fixed i, the g_{iJ} are—not necessarily distinct—products of the p_{ij}, where the latter are found by applying the finiteness theorem for closed semialgebraic sets to the set P_{nd}, as in (a) above. For fixed i, the common denominator of the r_{iJ} is of the form $F^{2e} + \sum_J g_{iJ} h_{iJ}^2$, where $h_{iJ} \in \mathbf{Z}[C;X]$ (using Stengle's *Positivstellensatz* (5.14.2)); this explains why this common denominator does not vanish identically in X so long as $c \in D_i := \{\, c \in R^m \mid \bigwedge_j (p_{ij}(c) \geq 0) \,\}$. With this choice of D_i, we see that (5.9.1–2) contains Henkin's (5.8), with the additional information that the subsets D_i of P_{nd} are not only $\mathbf{Z}$-semialgebraic, but even *basic closed* $\mathbf{Z}$-semialgebraic.

(c) In [1968, pp. 361–2] and [1977a, pp. 115–6], Kreisel had claimed that his [1960b] and/or Daykin's [1961] had already established (5.9.1–2); I repeated his claim in [1982a,b], [1984], and [1994]. But I obtained a copy of Daykin's thesis only recently; in fact, it makes no such claim. And now that I have really studied Kreisel's sketches, I believe that even if one reads them generously, one can find no hint of this claim (or at least no hint of its proof). On the contrary, suppose that g_{iJ} and r_{iJ} have been found satisfying (5.9.1–2). Use those g_{iJ} to define $D_i = \{\, c \in R^m \mid \bigwedge_J (g_{iJ}(c) \geq 0) \,\}$. Then $P_{nd} \subseteq \bigcup_i D_i$ (by (5.9.2)) and $P_{nd} \supseteq \bigcup_i D_i$ (by (5.9.1)). Thus (5.9.1–2) entails $P_{nd} = \bigcup_i D_i$, which is precisely the finiteness theorem for closed semialgebraic sets, at least as it applies to the set P_{nd}. As far as I can tell, it is just as difficult to decompose P_{nd} into suitable D_i as it is to so decompose an *arbitrary* closed semialgebraic set[25]; and I see no evidence in Kreisel's work that he realized that such a decomposition is non-trivial. On the other hand, Kreisel's methods *are* good enough to prove Henkin's (5.8) (constructively) without the finiteness theorem, as we shall show in §6 and in Part II.

[25] Unless $d \leq 2$, when it is easy, by Delzell [1982a, pp. 92–3] and [1982b, pp. 99–100]. But when $d \geq 4$ it is difficult, even for $n = 1$, by Delzell [1982a,b]; in fact, even the simpler problem of eliminating the quantifier from $\forall x(ax^4+bx^3+cx^2+dx+e \geq 0)$ at all (i.e., even if we *allow* both $\geq$ and $>$ relations in the answer) is difficult, by Arnon, Mignotte [1988] and Lazard [1988].

5.10 Dubois, McKenna

Some 10 years *after* Henkin's bull's-eye formulation (5.7.1) for arbitrary ordered K (allowing nonnegative weights), Lang [1965] returned to Artin's original weightless representation (2.4.1), and tried to drop the Archimedean hypothesis on K while retaining unique orderability. (Recall footnote 18 at the end of §4 above, on a similar mistake by Sacks.) Dubois found a counterexample in [1967], with $n = 1$ and $d = 6$. And a memorable result was observed by McKenna [1975] (cf. also Jacobson [1989, pp. 661–2]): if K is an ordered field, then (2.4.1) holds if and only if K is uniquely orderable and dense in its real closure.

5.11 General recursiveness

From the logical form and the mere validity of (5.8) (i.e., without looking at its proof), one can compute (a) the number, degrees, and integer coefficients of the p_{ij} and r_{ij}, and (b) the number, degrees, and coefficients of the polynomials in $\mathbf{Z}[C]$ defining the D_i as semialgebraic sets, by general recursive functions of n and d. Indeed, one simply enumerates all finite sequences $\langle p_{ij}\rangle$ and $\langle r_{ij}\rangle$ of elements of $\mathbf{Z}[C]$ and $\mathbf{Q}(C;X)$, and all finite sequences $\langle D_i\rangle$ of semialgebraic subsets of P_{nd}, checking whether (5.8.1–2) holds (using the completeness (4.6) of $\mathbf{RCF}_{\geq}$); by (5.8), this "systematic trial-and-error" method must terminate. Similarly for (5.9)(b), this time without the need to enumerate the sequences $\langle D_i\rangle$ explicitly. On the other hand, this proof of termination is not finitistic until (5.8) or (5.9)(b) themselves are established finitistically; thus, these general recursive functions cannot yet be considered finitist. (Compare the "*central dogma* of Constructivism" (cdC) Kreisel [1990b], quoted in (1.7) above.)

5.12 Are the non-primitive recursive bounds (5.11) of Robinson/Henkin already finitist?

According to Kreisel, those model theoretic bounds are actually ω^{ω^ω}-recursive, by general model-theoretic results available even then; and by the end of the fifties, after some analysis of the idea(l) of finitist proof ([1960b] and [1965]), Kreisel argued that every function in the class $\mathcal{E}^*$ (= the class of α-recursive functions, for every ordinal $\alpha < \epsilon_0 := \lim_{i\to\infty} \omega^{\cdot^{\cdot^{\omega}}}\}^i$) is finitist. (Cf. Rose [1984] for background.) This argument has received (1) positive, (2) negative, and (3) cautious reactions from the following three experts, respectively.

(1) Rose [1984, p. xi] wrote:

> "Kreisel (1965) ... has put forward powerful arguments which imply that the theory of the extended Grzegorczyk hierarchy for ordinals less than ϵ_0 is a finitist theory."

And on p. 138:

> "[A] case can be made out for claiming that $\mathcal{E}^*$-arithmetic is exactly the finitist part of PA; see Kreisel (1965)."

(2) While Kreisel has often (e.g., in [1987, p. 396]) cited Hilbert's mention (in *Über das Unendliche* [1926]) of Ackermann's function and certain other non-primitive recursive functions, as evidence that those functions were to be accepted as finitist, Tait [1981, p. 545] wrote:

> "But, possibly because of [Kreisel's] misreading of Hilbert's remarks concerning Ackermann's function, Kreisel's analysis completely misses the mark...."

(3) Finally, Gödel [1972, p. 274] wrote:

> "An unobjectionable version is given in *Kreisel 1965,* pp. 168–173, 177–178. Theorem 3.43 on page 172 of these lectures states that ϵ_0 is the limit of this process. Kreisel wants to conclude from this fact that ϵ_0 is the exact limit of idealized concrete intuition. But his arguments would have to be elaborated further in order to be fully convincing."

5.13 Primitive recursiveness via model theory

(a) Over the (very weak) system $\mathbf{RCA}_0$ (named after the recursive comprehension axiom), Friedman, Simpson, and Smith proved [1983, p. 167] that the *countable* version of the Artin-Schreier theorem (2.1) that real fields are orderable, follows from **WKL**, i.e., weak König's lemma; the latter asserts that every infinite binary (directed) tree has an infinite path. (They even "reversed" this, i.e., they also proved the converse.)

(b) **WKL** is conservative over **PRA** (primitive recursive arithmetic, (4.5)) for Π_2-theorems, i.e., those of the form $(\forall y \in \mathbf{N})(\exists z \in \mathbf{N})P(y,z)$, for some quantifier-free formula P in the language of arithmetic. This means that from a proof using **WKL** (and a very basic logical apparatus, such as $\mathbf{RCA}_0$) of a Π_2-formula as above, we can construct a term $\tau(y)$ in **PRA** (representing a primitive recursive function) such that **PRA** proves $P(y,\tau(y))$. This was argued in Mints [1976], using the no-counterexample interpretation. (Mints used a basic system, denoted S^+, that is slightly different from $\mathbf{RCA}_0$.) In contrast to this proof-theoretic work, Friedman [1979] proved this conservation result model-theoretically. In 1981, Sieg found an error in Mints' argument, and in [1985] Sieg gave a correct proof-theoretic demonstration of—a number of extensions of—this conservation result, for both basic systems; cf. his Corollaries 5.9 and 5.7, respectively. (For even stronger conservation results, cf. Kohlenbach [1992].)

(c) At least if $K = \mathbf{Q}$ (the most important example of a computable ordered field), Artin's theorem, and even (5.8) and (5.9)(b), are, indeed, Π_2: e.g., for Artin's theorem, one codes the coefficients of $f \in \mathbf{Q}[X]$ into a single integer y; likewise, one codes a finite list of $r_i \in \mathbf{Q}(X)$ into an integer z. $P(y, z)$ should say that if f is positive semidefinite, then $f = \sum_i r_i^2$, for the f and r_i determined by y and z, respectively. P is expressible in the language of arithmetic, by QE.

From (a)–(c) follows a primitive recursive bound for Artin's theorem.

5.14 Stengle

Stengle's *Positivstellensatz* [1974], the last refinement of Artin's theorem considered here, generalizes (2.4.1) and even (5.7.1), as follows. Again let R be a real closed field, and K a subfield, with the inherited order. Let $f, g_1, \ldots, g_s \in K[X]$, and suppose that $\forall x \in R^n$, if each $g_i(x) \geq 0$, then $f(x) \geq 0$. Then

$$\left(f^{2l} + \sum_I p_I G_I h_I^2\right) f = \sum_I p'_I G_I {h'_I}^2, \tag{5.14.1}$$

where $l \geq 0$, $h_I, h'_I \in K[X]$, $0 \leq p_I, p'_I \in K$, and the G_I are—not necessarily distinct—products of the g_i. He called this theorem a "*Positivstellensatz*." (In the early eighties, this theorem was slightly generalized, allowing also equalities and strict inequalities; we leave the details to Part II.) Taking $s = 0$, we obtain a refinement of (2.4.1) and (5.7.1):

$$\text{if } f \text{ is psd over } R, \text{ then } f = \frac{\sum_I p'_I {h'_I}^2}{f^{2l} + \sum_I p_I h_I^2}. \tag{5.14.2}$$

This representation offers more control over the denominator than in Artin's theorem or its earlier refinements: here, the denominator can vanish only at points $x \in R^n$ where $f(x) = 0$, a fact that is often useful (say, in (5.9.2) above, or for continuity, in Part II).

Stengle's "proof combines a standard proof of the Hilbert *Nullstellensatz* (e.g. Jacobson [1964]) with the most basic arguments of Artin-Schreier theory" [1974, p. 89]. (There are other proofs, such as those in Lam [1984] and Bochnak, et al [1987], that are based on the real spectrum (= the space, or the mere set, of orderings) of the ring $K[X]$; this set goes back to Prestel [1975, (1.4)] and Coste, Roy [1979], independently, and is, like Stengle's proof, also based on Artin-Schreier theory, generalized from fields to commutative rings.) Therefore, one can be confident that it is only an exercise to extend at least most of the model and proof theoretic methods and bounds for Artin's theorem, mentioned in §5 above and in §9 below, to Stengle's theorem. In particular:

(a) His theorem can be parametrized as in (5.8–9) (even continuously, in certain cases: Scowcroft [1989], Delzell [1993], Delzell, González-Vega, Lombardi [1993], González-Vega, Lombardi [1993]; in fact, even r times differentiably, for any finite r (González-Vega, Lombardi [to appear]), but not real analytically (Delzell [1994])).

(b) General recursive bounds follow from its mere validity as in (5.11), although its validity must be established constructively if such bounds are to be constructive (by cdC (1.7)).

(c) Such bounds are arguably automatically finitist, as in (5.12).

(d) Primitive recursive bounds would follow from the conjecture that it, like Artin's theorem, can be proved using **WKL** as in (5.13).

(e) Stengle's or Lam's proofs could be unwound as in Kreisel's *first* sketch (1.3) for Artin's proof.

(f) An explicit iterated-exponential bound could thereby be extracted.

In Part II of this paper I plan to work out the details of (e), as part my exposition of Kreisel's first sketch. The only approach that does not seem to extend to Stengle's theorem is Kreisel's *second* sketch (§6); we shall pinpoint the difficulty in (7.1). Rather than extending (from Artin to Stengle) the above, existing approaches to constructivity (a)–(f), Scowcroft sketched a proof-theoretic approach [1988, pp. 70–1] different from Kreisel's (cf. (9.5) below); and Lombardi [1988–91] provided a direct constructive proof, and an explicit calculation of a bound ([1992]; cf. (9.6) below).

6 Kreisel's second sketch

(Recall the outline in (1.5).)

6.1 Eliminating the quantifiers $\forall x_i$. Recall that in (5.5) we defined the general polynomial $F \in \mathbf{Z}[C;X]$ of degree d in the indeterminates $X := (X_1, \ldots, X_n)$, with indeterminate coefficients $C := (C_1, \ldots, C_m)$ (where $m = \binom{n+d}{n}$), to be $F = \sum_{|e| \le d} C_{g(e)} \cdot \prod_{i=1}^n X_i^{e_i}$, where e and g are as in (5.5). And we wrote $x_1, \ldots, x_n, c_1, \ldots, c_m$ for elements of a real closed field R.

Now, instead, from (6.1) through (6.44), we shall write $x_1, \ldots, x_n$ for (individual) variable symbols in the formal language of $\mathbf{RCF}_{\ge}$ (§3); and now we shall write $c_1, \ldots, c_m$ for new *constant* symbols; technically, they do not belong to the language of $\mathbf{RCF}_{\ge}$ (whose only constants are 0 and 1), so we write $\mathbf{RCF}^{c}_{\ge}$ for the system obtained by adding these constants (and no new axioms). Finally, F will no longer denote the general polynomial $\in \mathbf{Z}[C;X]$ of degree d in X with coefficients C, but rather the corresponding *term,* in the language of $\mathbf{RCF}^{c}_{\ge}$ (in fact, in the language of rings); formally, F is now

$$\sum_{|e| \le d} c_{g(e)} \cdot \prod_{i=1}^{n} x_i^{e_i}.$$

Tarski's theorem (4.1) works just as well in the presence of the new constant symbols $c_1, \ldots, c_m$ as in their absence, since we may, temporarily, treat $c_1, \ldots, c_m$ as variable symbols, rather than constants. Thus we can construct finitely many terms p_{ij} and q_{ij}, built up from 0, 1, the c_i's, $+$, $-$, and $\cdot$ alone (i.e., "polynomials" in the c_i's with "integer" coefficients), such that

$$\mathbf{RCF}^{c}_{\ge} \vdash (\forall x_1 \cdots \forall x_n \, F \ge 0) \leftrightarrow \bigvee_i \bigwedge_j (p_{ij} \ge 0 \wedge q_{ij} > \mathbf{0}). \quad (6.1.1)$$

(Recall (5.9)(c), where we decided not to use the finiteness theorem for closed semialgebraic sets.)

From now until (6.45) below, we fix i (as well as n and d).

6.2 Elimination of $\ge$. Expand the language of $\mathbf{RCF}^{c}_{\ge}$ by introducing, for each j, new constant symbols a_{ij} and b_{ij}, and add the axiom

$$\bigwedge_j (p_{ij} = a_{ij}^2 \wedge q_{ij} b_{ij}^2 = 1). \quad (6.2.1)$$

Then $\mathbf{RCF}^{c}_{\ge} + (6.2.1) \vdash \bigwedge_j (p_{ij} \ge 0 \wedge q_{ij} > 0)$. So

$$\mathbf{RCF}^{c}_{\ge} + (6.2.1) \vdash \exists z \, (F = z^2), \quad (6.2.2)$$

using (6.1.1) and (3.1.5). Write $\mathbf{RCF}_{ndi}$ for the system obtained from $\mathbf{RCF}^c_{\geq}$ by dropping the $\geq$ symbol (in particular, dropping (3.1.5)), and adding the axiom (6.2.1) (in particular, adding the constants a_{ij} and b_{ij} for all j). Furthermore, among the axioms for roots of polynomials of degree $2q+1$ $(3.1.4)_q$, retain in $\mathbf{RCF}_{ndi}$ only those with $q = 1, 2, \ldots, Q$, where Q is the maximum q such that $(3.1.4)_q$ occurs as an initial formula[26] in the proof in (6.1.1). Then $\mathbf{RCF}_{ndi} \vdash \exists z\,(F = z^2)$, since $\exists z\,(F = z^2)$ does not contain the $\geq$ symbol (which was defined explicitly by (3.1.5) and which did not appear in the other nonlogical axioms of $\mathbf{RCF}_{\geq}$).

6.3 Elimination of formal reality. Let $\mathbf{RCF}'$ be the system obtained from $\mathbf{RCF}^c_{\geq}$ by dropping the $\geq$ symbol and the axioms for square roots and odd-degree roots ((3.1.3) and $(3.1.4)_q$), and replacing the axiom $(3.1.2)_2$ (reality) by its negation. Thus the nonlogical axioms of $\mathbf{RCF}'$ are (3.1.1), $\neg(3.1.2)_2$, and (3.1.3). Then

$$\mathbf{RCF}' \vdash \exists z_1\,\exists z_2\,(F = z_1^2 + z_2^2). \tag{6.3.1}$$

To see this, first suppose $-F = z^2$ and $-1 = y_1^2 + y_2^2$. Then we can prove $F = (zy_1)^2 + (zy_2)^2$. On the other hand, from $F = z^2$ would follow $F = z^2 + 0^2$. Either way,

$$\begin{aligned} &-1 = y_1^2 + y_2^2 \rightarrow \\ &\qquad [(F = z^2 \vee -F = z^2) \rightarrow \exists z_1 \exists z_2 (F = z_1^2 + z_2^2)], \text{ whence} \\ &[\exists y_1 \exists y_2 (-1 = y_1^2 + y_2^2)] \rightarrow \\ &\qquad [(F = z^2 \vee -F = z^2) \rightarrow \exists z_1 \exists z_2 (F = z_1^2 + z_2^2)]. \end{aligned}$$

Then $(F = z^2 \vee -F = z^2) \rightarrow \exists z_1\,\exists z_2\,(F = z_1^2 + z_2^2)$, by $\neg(3.1.2)_2$. From this we similarly obtain (6.3.1), this time using (3.1.3). □

Let $\mathbf{RCF}'_{ndi}$ be obtained from $\mathbf{RCF}_{ndi}$ by replacing $(3.1.2)_2$ with one of its consequences, $2 \neq 0$; this will simplify the statement and proof of (6.32) (footnote 37), and both the end of (6.35), and footnote 43 there. Thus the axioms of $\mathbf{RCF}'_{ndi}$ are (3.1.1), (3.1.3), $(3.1.4)_1, \ldots, (3.1.4)_Q$, $2 \neq 0$, and (6.2.1). Then

$$\mathbf{RCF}'_{ndi} \vdash \exists z_1\,\exists z_2\,(F = z_1^2 + z_2^2); \tag{6.3.2}$$

this proof is easily assembled from those in (6.2) and (6.3.1).

[26] An *initial formula* in a proof is one that does not arise from previous formulae in the proof by any rule of inference, but rather is—a substitution-instance of—an axiom; regarding the proof as a directed tree, an initial formula is an initial node.

6.4 Introduction of Skolem-function symbols. From the system $\mathbf{RCF}'_{ndi}$ let us form the new system $\mathbf{RCF}''_{ndi}$, by introducing the function symbols $\rho(a)$, $\rho_q := \rho_q(a_0, \ldots, a_{2q})$, and a^{-1} (where $a, a_0, \ldots, a_{2q}$ are variables), and replacing axioms (3.1.3), $(3.1.4)_q$, and the axiom $a \neq 0 \;\to\; \exists z\,(az = 1)$ for inverses (3.1.1), by the new axioms

$$a = \rho(a)^2 \;\vee\; -a = \rho(a)^2, \qquad (6.4.1)$$

$$\rho_q^{2q+1} + a_{2q}\rho_q^{2q} + \cdots + a_0 = 0, \qquad (6.4.2)_q \text{ and}$$

$$a \neq 0 \;\to\; aa^{-1} = 1. \qquad (6.4.3)$$

Obviously (6.4.1), $(6.4.2)_q$, and (6.4.3) imply (3.1.3), $(3.1.4)_q$, and $a \neq 0 \;\to\; \exists z\,(az = 1)$ (3.1.1), respectively. Thus (6.3.2) leads to

$$\mathbf{RCF}''_{ndi} \;\vdash\; \exists z_1\,\exists z_2\,(F = z_1^2 + z_2^2). \qquad (6.4.4)$$

We now digress to explain, mainly for algebraists, the introduction of such "Skolem-function" symbols. Readers with a logical education can skip ahead to (6.16).

Actually, mathematicians are already familiar with the idea of replacing existential quantifiers by function symbols, as Shoenfield explains via an example from arithmetic [1967, p. 55]:

> "[S]uppose we are discussing natural numbers and have proved that for every x, there is a prime y such that $y > x$. In the course of a later proof we might say: ley y be a prime such that $y > x$. We would then have to keep in mind through the rest of the proof that y depends on x. If we wished to indicate this by the notation, we would say instead: for each x, let $f(x)$ be a prime greater than x. Of course, f would be a new symbol which does not appear in the result we are trying to prove."

In algebra, the original existential statement is often an *axiom* rather than (merely) a theorem. For example, instead of the axiom $\exists z\,(a + z = 0)$, for negatives, one may introduce the unary function symbol '$-$' and the axiom $a + (-a) = 0$ (say, for rings, or Abelian groups); the new axiom obviously implies the old, and the resulting formal system is conservative over the original one, in the sense that any formula in the original language (i.e., not containing $-$) that is provable in the new system (e.g., $a + b = a \;\to\; b = 0$) is provable already in the old one, though not as straightforwardly. An even simpler example is the introduction of the constant 0 and its axiom $a + 0 = a$, replacing the existential axiom $\exists z \forall a\,(a + z = a)$ (we may regard a constant symbol as a function symbol with 0 arguments). In (3.1) we included the three "function" symbols $-$, 0, and 1 in $\mathbf{RCF}_{\geq}$ right from the beginning.

The function symbols ρ and ρ_q introduced in (6.4.1) and $(6.4.2)_q$, however, differ in one respect from $-$, 0, and 1 above and their axioms: these new axioms do not determine ρ and ρ_q uniquely (just as in Shoenfield's illustration above); all we know from the axioms (3.1.3) and $(3.1.4)_q$ is that at least one suitable "value" of $\rho(a)$ and $\rho_q(a_0, \ldots, a_{2q})$ exists. Now (6.4.1) and $(6.4.2)_q$ may seem to assert the existence of choice functions, capable of making infinitely many choices at once. (Actually, we are faced with a choice even in (6.4.3); but there, only one choice has to be made, in the definition of 0^{-1}.) However, if the end-formula does not contain these new function symbols, then this impression that genuine choice functions are involved is deceptive, as Kreisel explains [1958a, p. 165]:

> "The elimination of ρ, which is a kind of choice function, may be understood roughly as follows: In any proof by the predicate calculus, ρ is applied only to a finite number of terms, and a finite number of choices is harmless; the exact meaning of this phrase is provided by the analysis contained in the proof of the second ε-theorem."

Another "official" justification of the use of such function symbols is Skolem's theorem and "the theorem on functional extensions," which Shoenfield's [1967, pp. 55–7] text proves using Herbrand's theorem; Hilbert, Bernays [1939, 1970] proves both Skolem's and Herbrand's theorems from the second ε-theorem ((6.15) below). Another way to view the symbols ρ and ρ_q is this: in any given model of the axioms (here, a suitable field K), no matter how we interpret the constant symbols $a_{ij}, b_{ij}, c_1, \ldots, c_m$ as elements $\in K$ satisfying (6.2.1), and the free variables $x_1, \ldots, x_n$ as arbitrary elements $\in K$, and no matter which of the possible values $\in K$ (satisfying (6.4.1) and $(6.4.2)_q$) that we choose for each term $\rho(s_0)$ and $\rho_q(s_0, \ldots, s_{2q})$ occurring in the formal proof under consideration, the proof becomes a perfectly logical sequence of statements about those (finitely many) elements of K; and if (as we assume) the end-formula contains no ρ- or ρ_q-symbols, then our arbitrary choices for the values of those symbols have no effect on the end-formula. Thus we obtain *many* (not quite formal) proofs of the end-formula, one for each choice of values for the ρ- or ρ_q-symbols. (In particular, the axiom of choice has nothing to do with it.)

6.5 Hilbert's ε-calculus. Instead of introducing Skolem-function symbols one at a time, as we did in (6.4.1), $(6.4.2)_q$, (6.4.3) above, we can introduce all of them in one stroke, using Hilbert's ε-symbol; Bourbaki called this "without doubt the most interesting" modification of the early formal languages ([1957, p. 80], [1966, p. 61]). The ε-calculus will enable us to eliminate the quantifiers $\exists z_l$ in (6.4.4).

Write $\mathcal{A}(z)$ for a "formula-variable" with one argument z; in a proof, $\mathcal{A}(z)$ can be replaced by any formula $A(z)$ of $\mathbf{RCF}''_{ndi}$ containing z free. We expand the language by adding the symbol ε, which is allowed to occur in new terms of the form $\varepsilon_z\mathcal{A}(z)$ (an *ε-term*). Such occurrences of z will now be called *bound* occurrences, just like the occurrences in—the scope of—$\forall z$ and $\exists z$; and the ε-symbol will now be counted as a "quantifier," just like $\forall$ and $\exists$; in particular, a "quantifier-free" formula contains no ε-symbol. We add one new (logical) axiom, the "ε-formula":

$$\mathcal{A}(b) \;\rightarrow\; \mathcal{A}(\varepsilon_z\mathcal{A}(z)) \tag{6.5.1}$$

(where b is a variable, as usual). Denote the new system by $\mathbf{RCF}''_{ndi\varepsilon}$. The idea is that $\varepsilon_z\mathcal{A}(z)$ represents some element z of the universe for which $\mathcal{A}(z)$ holds, *provided there is such an element* (b); otherwise, while it still represents some element, we are unconcerned by the fact that $\mathcal{A}(z)$ does not hold for it. If we replace $\mathcal{A}(z)$ by a specific formula $A(z)$ containing parameters (= additional free variables, besides z), then $\varepsilon_z\mathcal{A}(z)$ represents a function of those parameters.

The fact that the ε-formula (6.5.1) introduces a new, second-order kind of variable (*viz.,* the formula-variable $\mathcal{A}$) to our system should not cause alarm: we shall never *quantify* this variable (i.e., we shall never write $\forall\mathcal{A}$ or $\exists\mathcal{A}$). In fact, we need never write $\mathcal{A}(z)$ at all: any single proof needs only finitely many substitution-instances of the ε-formula, obtained by replacing $\mathcal{A}(z)$ by some formula $A(z)$ of the original language, i.e., a formula not containing any formula-variables. Thus (6.5.1) is really an axiom *schema* of the original language. We also call (6.5.1) an "improper" axiom, because it contains a formula variable. From now on, the words "formula" and "axiom," unless preceded by the word "improper," will mean "proper formula" or "proper axiom." Whenever a substitution-instance of the ε-formula appears as an initial formula in some particular proof, that instance is called a "critical ε-formula" in the proof.

Hilbert, Bernays [1939, 1970] showed, easily, that the usual quantifiers $\forall$ and $\exists$, as well as the corresponding axioms and rules of inference, are superfluous in the presence of the ε-symbol and its formula (6.5.1).

For one can define $\forall$ and $\exists$ explicitly by

$$\forall z\, A(z) \;\leftrightarrow\; A(\varepsilon_z \neg A(z)) \quad \text{and} \quad \exists z\, A(z) \;\leftrightarrow\; A(\varepsilon_z A(z)),$$

respectively, and then *prove* that thereby the axioms and rules of inference for $\forall$ and $\exists$ become tautological consequences of (6.5.1). (For example, Bourbaki's formal language for set theory [1954, 1960] took the ε-symbol as a primitive, in terms of which they defined $\forall$ and $\exists$; actually, instead of the notation $\varepsilon_z A(z)$, they used $\tau_z(A(z))$, which in Hilbert [1923] had corresponded to what is here written as $\varepsilon_z \neg A(z)$.)

In $\mathbf{RCF}''_{ndi\varepsilon}$, for example, we can replace $\mathcal{A}(z)$ by $a = z^2 \vee -a = z^2$, obtaining the ε-term $\varepsilon_z(a = z^2 \;\vee\; -a = z^2)$. Substituting into (6.5.1), we get the following instance of the ε-formula:

$$\begin{aligned}&(a = b^2 \vee -a = b^2) \;\rightarrow\\ &[a = (\varepsilon_z(a = z^2 \vee -a = z^2))^2 \vee -a = (\varepsilon_z(a{=}z^2 \vee -a{=}z^2))^2], \text{ whence}\\ &[\exists z\; (a = z^2 \vee -a = z^2)] \;\rightarrow\\ &[a = (\varepsilon_z(a = z^2 \vee -a = z^2))^2 \vee -a = (\varepsilon_z(a{=}z^2 \vee -a{=}z^2))^2];\end{aligned}$$

by (3.1.3), we get

$$a = (\varepsilon_z(a = z^2 \vee -a = z^2))^2 \vee -a = (\varepsilon_z(a{=}z^2 \vee -a{=}z^2))^2. \quad (6.5.2)$$

Thus we could have defined the function symbol $\rho(a)$ explicitly, by $\rho(a) = \varepsilon_z(a = z^2 \;\vee\; -a = z^2)$, instead of implicitly, by (6.4.1); then (6.4.1) would follow from (6.5.2).

Similarly we obtain the ε-term $\varepsilon_z(z^{2q+1} + a_{2q}z^{2q} + \cdots + a_0 = 0)$, and the instance

$$\begin{aligned}&b^{2q+1} + a_{2q}b^{2q} + \cdots + a_0 = 0 \;\rightarrow\\ &\qquad (\varepsilon_z(z^{2q+1} + a_{2q}z^{2q} + \cdots + a_0 = 0))^{2q+1}\\ &\qquad +a_{2q}(\varepsilon_z(z^{2q+1} + a_{2q}z^{2q} + \cdots + a_0 = 0))^{2q} + \cdots + a0 = 0\end{aligned}$$

of (6.5.1), whence by $(3.1.4)_q$ (as above)

$$\begin{aligned}&(\varepsilon_z(z^{2q+1} + a_{2q}z^{2q} + \cdots + a_0 = 0))^{2q+1}\\ &+a_{2q}(\varepsilon_z(z^{2q+1} + a_{2q}z^{2q} + \cdots + a_0 = 0))^{2q} + \cdots + a0 = 0;\end{aligned}$$

this allows us to define ρ_q explicitly by

$$\rho_q(a_0, \ldots, a_{2q}) \;=\; \varepsilon_z(z^{2q+1} + a_{2q}z^{2q} + \cdots + a_0 = 0).$$

Finally, we could have defined a^{-1} explicitly by $\varepsilon_z(a \neq 0 \rightarrow az = 1)$.

6.6 Russell and Whitehead's ι-symbol. For the predicate calculus with equality, the ι-symbol is a (more familiar) special case of the ε-symbol. For any formula $A(z)$ with free variable z, the ι-rule allows us to introduce the *ι-term* $\iota_z A(z)$, provided we have previously proved the following *existence and uniqueness formulae for* A:

$$\exists z\, A(z) \quad \text{and} \quad \forall y\, \forall z\, [(A(y) \wedge A(z)) \ \rightarrow \ y = z].$$

The rule then allows us also to add the formula $A(\iota_z A(z))$ as an axiom. For example, in our first system $\mathbf{RCF}_{\geq}$ (3.1), we can take $A(z)$ to be $(a \geq 0 \ \wedge \ z = a) \ \vee \ (a \leq 0 \ \wedge \ z = -a)$, since in $\mathbf{RCF}_{\geq}$ we can prove

$$\exists z\, [(a \geq 0 \ \wedge \ z = a) \ \vee \ (a \leq 0 \ \wedge \ z = -a)] \quad \text{and}$$

$$\forall y \forall z \left[\begin{pmatrix} [(a \geq 0 \ \wedge \ y = a) \ \vee \ (a \leq 0 \ \wedge \ y = -a)] \ \wedge \\ [(a \geq 0 \ \wedge \ z = a) \ \vee \ (a \leq 0 \ \wedge \ z = -a)] \end{pmatrix} \ \rightarrow \ y = z \right];$$

furthermore, we are allowed to use as an axiom the formula

$$\begin{aligned} &[a \geq 0 \wedge \iota_z[(a \geq 0 \ \wedge \ z = a) \ \vee \ (a \leq 0 \ \wedge \ z = -a)] = a] \ \vee \\ &[a \leq 0 \wedge \iota_z[(a \geq 0 \ \wedge \ z = a) \ \vee \ (a \leq 0 \ \wedge \ z = -a)] = -a]; \end{aligned}$$

one could also introduce the familiar notation $|a|$, defined explicitly by $\iota_z[(a \geq 0 \ \wedge z = a) \vee (a \leq 0 \ \wedge z = -a)]$. (For an example of an ι-term in arithmetic, recall the μ-operator: $\mu_z A(z)$ is the *least* nonnegative integer z such that $A(z)$.) Hilbert, Bernays [1934, 1968] proved that if one extends a formal system according to the ι-rule, then the resulting system is conservative over the original one: any proof, of any formula A not containing the ι-symbol, can be converted into a proof of A in the original system. However, this result will be superseded by the second ε-theorem (6.15), below.

6.7 Skolem's idea was intermediate between the ι-symbol and the ε-symbol, and can be expressed by a modified ι-rule, which Hilbert, Bernays [1939, 1970] (p. 10) called the *η-rule*, obtained by dropping the requirement that the *uniqueness* formula be provable, but retaining the requirement that the *existence* formula $\exists z\, A(z)$ be provable. Then one may introduce the *η-term* $\eta_z A(z)$ and the axiom $A(\eta_z A(z))$. The fact that this expanded system is conservative over the original one is called "the theorem on functional extensions" (Shoenfield [1967, pp. 55–6]); when it is iterated so as to eliminate *all* existential quantifiers appearing in one or more given nonlogical axioms, it is called *Skolem's theorem* (pp. 56–7; recall also (6.4) above).

We can now see the difference between the η-symbol and the ϵ-symbol: for the latter, we drop not only the requirement that the *uniqueness* formula be provable, but also the requirement that the *existence* formula $\exists z\, A(z)$ be provable.

6.8 The $\boldsymbol{\varepsilon}$-theorems, (6.13–15) below, will apply to any formal system $\mathcal{S}$ arising from the predicate calculus (with or without equality) (1) by adding to the language the ε-symbol and certain individual-, predicate-, and function-symbols, and (2) by adding the ε-formula (6.5.1) to the logical axioms, and by taking as nonlogical axioms certain other ε-*free* formulae (i.e., formulae that do not contain the ε-symbol).[27] For example, our system $\mathbf{RCF}''_{ndi\varepsilon}$ (6.5.1) has these properties, as well as the property hypothesized in (6.13) and (6.14) (but not (6.15)) below that all of its nonlogical axioms be quantifier-free (in particular, ε-free).

6.9 The quantifier-free fragment of $\boldsymbol{\mathcal{S}}$. In case all the nonlogical axioms of $\mathcal{S}$ are quantifier-free, we shall refer to "the quantifier-free fragment $\mathcal{S}^{\mathrm{qf}}$ of $\mathcal{S}$," or "the free-variable calculus extended by the nonlogical axioms of $\mathcal{S}$." This is the system arising from $\mathcal{S}$ by dropping all quantifiers (and hence all bound occurrences of variables) from the language, and dropping those logical axioms containing quantifiers (*viz.*, $\forall z \mathcal{A}(z) \to \mathcal{A}(a)$ and $\mathcal{A}(a) \to \exists z\, \mathcal{A}(z)$), and those rules of inference pertaining to quantifiers (*viz.*, $\dfrac{A \to B(z)}{A \to \forall z\, B(z)}$ and $\dfrac{B(z) \to A}{\exists z\, B(z)\ \to A}$). Thus the only logical axioms are the tautologies (or even some generating subset thereof), and, if the $=$ symbol is allowed, the two equality axioms $a = a$ and $a = b \ \to \ (\mathcal{A}(a) \to \mathcal{A}(b))$. And the only rules of inference are *modus ponens*: $\dfrac{A \quad A \to B}{B}$, and two substitution rules: substitution of terms t for (free) variables a: $\dfrac{A(a)}{A(t)}$, and substitution of formulae A for formula-variables $\mathcal{A}$.

[27] Hilbert and Bernays restricted themselves ([1939, pp. 18, 381]; [1970, pp. 18, 394]) to what they called "first-order" axiom systems: those with *only finitely many* nonlogical axioms (all proper). At least for the purpose of the ε-theorems, it seems to me that one could allow infinitely many (proper) axioms, since the ε-theorems deal with only one proof at a time, and any single proof can use only finitely many axioms anyway. While Kleene's definition [1952] of "first-order theory" (p. 421) kept the requirement that the number of proper axioms be finite, Shoenfield [1967] dropped that requirement (p. 22).

6.10 We can dispense with the two rules of substitution if we allow (a) in addition to the logical and nonlogical axioms, all "substitution-instances" of those axioms, i.e., all formulae arising from the original axioms by a sequence of substitutions of the above two types; and (b) "repetition of a previously derived formula" as a new rule of inference.[28] A proof using such an extended set of axioms (a) and the extra rule of inference (b) is called a proof *in the extended sense.* For example, in addition to the logical axiom $\neg\mathcal{A} \vee \mathcal{A}$, we now allow $a \neq 0 \ \vee \ a = 0$; and in addition to $a + b = b + a$, we now allow $\rho(1+1)+0 = 0+\rho(1+1)$. To see that we may thenceforth restrict the application of the substitution rules to the axioms, note that (1) any proof can be organized into the form of a directed, binary tree, with axioms at the initial (top) nodes, and the end-formula (= the theorem being proved) at the bottom node, and such that any formula in the proof that follows from one or two previous formula(e) in the proof by some rule of inference is placed at the node immediately below the node(s) of the previous formula(e); and (2) any substitution occurring in the proof can be pushed back up through all the "proof-threads" passing through that node, right up to the initial formulae (which are axioms). Cf. Hilbert, Bernays, Vol. II, Supplement I ([1939, pp. 388–91]; [1970, pp. 402–5]), and, for more detail, Vol. I ([1934, pp. 221–8]; [1968, pp. 220–7]). Furthermore, if after this any formula-variables $\mathcal{A}$ or free individual-variables a remain in the proof (besides those that occur in the end-formula), then they can all be replaced by suitable variable-free formulae or terms (such as $0 = 0$ and 0, respectively), without disturbing the connectedness of the proof, and without altering the end-formula; cf. Vol. I ([1934, p. 228]; [1968, p. 227]).

6.11 Of the two equality axioms, $a = b \ \rightarrow \ (\mathcal{A}(a) \rightarrow \mathcal{A}(b))$ and $a = a$, the first one (the "general" equality axiom) is improper. But the following (proper) instances of the first axiom, called "special" equality axioms, suffice (together with the second axiom, $a = a$) for proofs in

[28] Actually, Hilbert and Bernays allowed repetition even in the presence of the two rules of substitution. But that is unnecessary, since if we want to repeat a formula A, we can justify it by saying that we have substituted some term, say 0, for some individual variable, say a, that does not occur (freely) in A; or if we are using a formal system that has no terms (such as the pure predicate calculus), then we can say that we have substituted some formula, say B, for some formula variable, say $\mathcal{A}$, that does not occur in A.

$\mathcal{S}^{\mathrm{qf}}$ of all proper theorems:

$$\begin{aligned}
&a = b \;\rightarrow\; (a = c \;\rightarrow\; b = c),\\
&a = b \;\rightarrow\; \zeta(a) = \zeta(b), \qquad \text{and}\\
&a = b \;\rightarrow\; (P(a) \rightarrow P(b)),
\end{aligned}$$

for all function symbols $\zeta(a)$ of one or more arguments, and for all predicate symbols $P(a)$ of one or more arguments. Cf. Hilbert, Bernays, Vol. II ([1939, pp. 59 and 387–8]; [1970, pp. 59 and 401–2]) for a statement of this fact, and Hilbert, Bernays, Vol. I ([1934, pp. 373–5]; [1968, pp. 382–4]) for a proof. (Taking $P(a)$ to be $a = c$, we see that the first special equality axiom listed above is a special case of the third.) If there are additional arguments in ζ or P, then those positions are filled with free variables in the obvious way; for example, if $\mathcal{S}$ is the system $\mathbf{RCF}''_{ndi}$ of interest to us (6.4) (whose quantifier-free fragment we shall denote by $\mathbf{RCF}^{\mathrm{qf}}_{ndi}$), and if ζ is $\rho_q(a_0, \ldots, a_{2q})$, then we actually have $2q + 1$ special equality axioms for ζ:

$$\begin{aligned}
&a = b \;\rightarrow\; \rho_q(a, a_1, \ldots, a_{2q}) = \rho_q(b, a_1, \ldots, a_{2q}),\\
&a = b \;\rightarrow\; \rho_q(a_0, a, a_2 \ldots, a_{2q}) = \rho_q(a_0, b, a_2 \ldots, a_{2q}),\\
&\quad\vdots \qquad\qquad \vdots \qquad\qquad\qquad\qquad\qquad \vdots\\
&a = b \;\rightarrow\; \rho_q(a_0, \ldots, a_{2q-1}, a) = \rho_q(a_0, \ldots, a_{2q-1}, b).
\end{aligned} \tag{6.11.1$_q$}$$

In fact, we now list the entire set of equality axioms of our particular system $\mathbf{RCF}^{\mathrm{qf}}_{ndi}$: in addition to $a = a$ and $(6.11.1)_1$, …, $(6.11.1)_Q$, we have

$$\begin{aligned}
&a = b \;\rightarrow\; (a = c \;\rightarrow\; b = c), && a = b \;\rightarrow\; ac = bc,\\
&a = b \;\rightarrow\; a + c = b + c, && a = b \;\rightarrow\; a^{-1} = b^{-1}, \qquad \text{and}\\
&a = b \;\rightarrow\; -a = -b, && a = b \;\rightarrow\; \rho(a) = \rho(b).
\end{aligned}$$

6.12 More axiomatic economy. We allow all substitution-instances of the special equality axioms above, as mentioned in (6.10); but we really need only the "proper"—instances of the—special equality axioms, i.e., those in which we substitute *distinct* terms for a and b; the other ("improper") instances, such as $t = t \;\rightarrow\; t^{-1} = t^{-1}$, are derivable from the identity axiom (in this case, $t^{-1} = t^{-1}$), and so may be dropped.[29]

For convenience, we, like Kreisel, *add the axiom* $0^{-1} = 0$ *to* $\mathbf{RCF}^{\mathrm{qf}}_{ndi}$; at the very end (6.46), we (unlike Kreisel) shall eliminate this axiom.

[29]Such formulae resemble the "improper" axioms of ε-equality considered in Hilbert, Bernays, Vol. II, p. 64.

We can dispense with $a = b \to -a = -b$, since it is provable from $a = b \to (a = c \to b = c)$, $a = b \to a + c = b + c$, and the abelian-group axioms for $+$, $-$, and 0.

By a parallel argument, the axiom $a = b \to a^{-1} = b^{-1}$ is dispensable under the extra hypothesis that $a \neq 0$: we can prove the former from $a = b \to (a = c \to b = c)$, $a = b \to ac = bc$, $a \neq 0 \to aa^{-1} = 1$, and the commutative-monoid axioms for $\cdot$ and 1.

Without the hypothesis $a \neq 0$ (e.g., even under the hypothesis $a = 0$), however, $a = b \to a^{-1} = b^{-1}$ is not provable from the remaining axioms. On the other hand, the conjunction $(0^{-1} = 0) \wedge (a = b \to a^{-1} = b^{-1})$ is interdeducible (via the remaining axioms) with $a = 0 \to a^{-1} = 0$. Indeed, the latter follows from an instance of $a = b \to a^{-1} = b^{-1}$ (with b replaced by 0), $0^{-1} = 0$, and the special equality axiom for $=$: $0^{-1} = a^{-1} \to (0^{-1} = 0 \to a^{-1} = 0)$. Conversely, assume $a = 0 \to a^{-1} = 0$. Then, first, $0^{-1} = 0$ follows easily from the instance $0 = 0 \to 0^{-1} = 0$ of the assumption. Second, to prove $a = b \to a^{-1} = b^{-1}$, we may assume $a = 0$, since otherwise the previous paragraph proves it. Therefore $a^{-1} = 0$, by the assumption. Suppose $a = b$. Then $b = 0$. By assumption, $b^{-1} = 0$. Therefore $a^{-1} = b^{-1}$.

We may therefore, whenever convenient, replace the pair, $0^{-1} = 0$ *and* $a = b \to a^{-1} = b^{-1}$, *with* $a = 0 \to a^{-1} = 0$.

Next, *the axiom* $a = b \to ac = bc$ *can be replaced with the theorem* $a = 0 \to ac = 0$. Indeed, assume the latter. To prove $a = b \to ac = bc$, assume $a = b$. Then $a - b = 0$, whence $(a - b)c = 0$ (by the hypothesis), whence $ac - bc = 0$, whence $ac = bc$.

Note: our system has, besides $=$, no predicate symbols (the symbol $\geq$ having been dropped in (6.2)). And our list (6.11) omits the "left-handed" equality axioms $a = b \to c + a = c + b$ and $a = b \to ca = cb$ for $+$ and $\cdot$, since they are derivable from our right-handed versions by the commutativity axioms. Finally, our list omits the "axioms of ε-equality" $a = b \to \varepsilon_z \mathcal{A}(z, a) = \varepsilon_z \mathcal{A}(z, b)$, where $\varepsilon_z \mathcal{A}(z, a)$ is a "Grundtypus" (= "basic type"; cf. Hilbert, Bernays, Vol. II [1939, 1970, pp. 57–8]) of some other ε-term; the reason for this omission (as well as the omission of von Neumann's definition of Grundtypus) is that we shall need to study the special equality axioms only for $\mathbf{RCF}^{\mathrm{qf}}_{ndi}$, which, by definition, contains no ε-symbols, and hence no axioms of ε-equality. On the other hand, our Skolem-function symbols ρ and ρ_q are, actually, ε-symbols in disguise, as we mentioned in (6.5). (They are fairly simple, being of "rank 1" in Ackermann's hierarchy—cf. Hilbert, Bernays, Vol. II [1939, 1970, p. 25].) Thus $\mathbf{RCF}^{\mathrm{qf}}_{ndi}$ includes a certain portion of the ε-calculus.

Theorem 6.13 First ε-theorem[30] (Hilbert, Bernays, Vol. II [1939, 1970, pp. 18, 57, 79–80) *Suppose $\mathcal{S}$ is a first-order formal system as in (6.8), all of whose nonlogical axioms are quantifier-free. Furthermore, let A be any quantifier-free formula of $\mathcal{S}$. Then, given a proof of A in $\mathcal{S}$, we can, in a primitive recursive way, construct a proof of A in the quantifier-free fragment $\mathcal{S}^{\mathrm{qf}}$ of $\mathcal{S}$ (6.9).*

Theorem 6.14 Extended first ε-theorem (Hilbert, Bernays, Vol. II, [1939, 1970, pp. 32–3, 80]) *Suppose $\mathcal{S}$ is a first-order formal system as in (6.8), and $A(z_1, \dots, z_r)$ is a quantifier-free formula of $\mathcal{S}$. Then, given a proof in $\mathcal{S}$ of $\exists z_1 \cdots \exists z_r\, A(z_1, \dots, z_r)$, we can, in a primitive recursive way, construct (1) finitely many terms $t_{1,1}, \dots, t_{1,r}, t_{2,1}, \dots, t_{s,r}$ in $\mathcal{S}$, not containing the ε-symbol,*[31] *and (2) a proof, in the quantifier-free fragment $\mathcal{S}^{\mathrm{qf}}$ of $\mathcal{S}$, of the formula $\bigvee_{k=1}^{s} A(t_{k,1}, \dots, t_{k,r})$.*

Theorem 6.15 Second ε-theorem (Hilbert, Bernays, Vol. II, [1939, 1970, pp. 18, 131, 141]) *Suppose $\mathcal{S}$ is a formal system as in (6.8), and A is an ε-free formula of $\mathcal{S}$. Then, given a proof of A in $\mathcal{S}$, we can, in a primitive recursive way, construct a proof of A that contains no ε-symbol; i.e., the new proof of A is in the predicate calculus (with or without equality, as appropriate), using also the nonlogical axioms of $\mathcal{S}$.*

6.16 We now pick up Kreisel's second sketch where we left off, namely at (6.4.4). Recall that we defined $\mathbf{RCF}''_{ndi\varepsilon}$ by adding to $\mathbf{RCF}''_{ndi}$ the ε-symbol and the ε-formula (6.5.1); then (6.4.4) leads to $\mathbf{RCF}''_{ndi\varepsilon} \vdash \exists z_1 \exists z_2\, (F = z_1^2 + z_2^2)$. The extended first ε-theorem tells us how to construct ε-free terms $t_{1,1}, t_{1,2}, t_{2,1}, \dots, t_{s,2}$ in $\mathbf{RCF}^{\mathrm{qf}}_{ndi}$ (the quantifier-free fragment, above $(6.11.1)_q$, of $\mathbf{RCF}''_{ndi\varepsilon}$), and a proof in $\mathbf{RCF}^{\mathrm{qf}}_{ndi}$ of $\bigvee_{k=1}^{s}(F = t_{k,1}^2 + t_{k,2}^2)$. By (6.11) the only equality axioms

[30] Hilbert supplied the initial ideas of the proof; Ackermann finished the proof. Nowadays most proof theory books present cut-elimination, instead of the ϵ-calculus. One reason is that the latter is a corollary of the former: Ladrière [1951] derived the ε-theorems from Gentzen's Hauptsatz. I am grateful to H. Lombardi for calling my attention to Leisenring [1969], which thoroughly explains the ε-theorems and their connection to cut-elimination, and provides many additional references. Other references include Łoś, Mostowski, Rasiowa [1956]; Smirnov [1971]; Wessels [1977]; and Schütte [1960] (but not [1977]). Leisenring's proof of the second ε-theorem (with equality) was corrected in Flannagan [1975]; the latter was itself corrected in Ferrari [1987, 1989].

[31] Hilbert and Bernays did not mention it here, but we can choose the t's to contain only those functions symbols, constants, and free variables occurring in either the nonlogical axioms or $\exists z_1 \cdots \exists z_r A(z_1, \dots, z_r)$.

used are those listed there. By (6.10) the only rule of inference used is modus ponens, provided we expand our set of axioms to include, in addition to the tautologies, all substitution-instances of the nonlogical axioms and of the special equality axioms. *The terms $t_{k,l}$ are built up by the function symbols $+$, $-$, $\cdot$, $^{-1}$, and $\rho, \rho_1, \ldots, \rho_Q$, from the variables $x_1, \ldots, x_n$, and from the constants 0, 1, $c_1, \ldots, c_m$ (6.1), a_{ij}, and b_{ij} (all j (6.2.1)).* Artin's theorem calls for *rational* terms, so we must eliminate ρ and ρ_q from the $t_{k,p}$, simplify any nested reciprocals, and then arrange for the a_{ij} and b_{ij} (all j) to occur only as squares.

6.17 The kind of formal proofs considered. Let $y_1, \ldots, y_p$ be variables, and let $A(y_1, \ldots, y_p)$ be a formula of $\mathbf{RCF}^{\mathrm{qf}}_{ndi}$ that does not contain $^{-1}$, ρ, or ρ_q. For the purpose of Kreisel's second sketch, for example, we shall, in fact, be taking $A(y_1, \ldots, y_p)$ to be $\bigvee_k \big(F = \sum_l y_{k,l}^2\big)$, where k ranges over a finite index-set, and now l ranges over any finite index-sets. For the purpose of consistency proofs (8.1), on the other hand, we could take A to be simply $-1 = \sum_l y_l^2$.

Suppose we have terms $t_1, \ldots, t_p$, and a proof, in $\mathbf{RCF}^{\mathrm{qf}}_{ndi}$, of $A(t_1, \ldots, t_p)$. We now begin to show how to transform this proof into one with no ρ- or ρ_q-symbols (in particular, with no instances of the axioms (6.4.1) or $(6.4.2)_q$); the new end-formula will be of the form $\bigvee_{k'} A'(t_{k',1}, \ldots, t_{k',p})$, where the $t_{k',e}$ are new terms constructed from the given proof, and A' is quite similar to A (e.g., $F = \sum_{l'} y_{k',l'}^2$). To prepare for these eliminations, we first formalize some standard algebraic manipulations, such as the Euclidean algorithm (6.20).

6.18 Formal polynomials. Let ζ be a term of $\mathbf{RCF}^{\mathrm{qf}}_{ndi}$. We call a term t a *polynomial (term) in* ζ if it is of the form $t_k\zeta^k + \cdots + t_0$, where the t_i are also terms. We could distinguish two kinds of *degree* for t: its *formal* degree is k, while its *algebraic* degree is the integer $d \le k$ such that $t_d \ne 0 \wedge \bigwedge_{l=d+1}^{k}(t_l = 0)$ (provided $\bigvee_l (t_l \ne 0)$). Both degrees are poorly defined. *First,* the term t alone does not determine its formal degree in ζ, since, for example, the entire polynomial could also be construed as a *constant* polynomial t_0', i.e., of formal degree 0 in ζ; so, to be quite correct, we should define a polynomial term of formal degree k in ζ to be a *sequence* $t_k, \ldots, t_0$ of terms, rather than the single term $t_k\zeta^k + \cdots + t_0$ that is determined by that sequence. For ease of exposition, we shall not continue to dwell on this distinction. *Second,* the algebraic degree depends, in general, on how the terms t_l are interpreted (in some model), since in general we cannot prove $t_d \ne 0 \wedge \bigwedge_{l=d+1}^{k}(t_l = 0)$ for *any* d (even if we can prove $\bigvee_l (t_l \ne 0)$). Often we shall consider separately the $k+1$ possible values for d (and even the possibility that $\bigwedge_l(t_l = 0)$). Usually we shall speak of "the degree" of a polynomial term, rather than the formal or algebraic degree; the meaning should be clear from the context.

If s is a polynomial term, say of the form $s_l\zeta^l+\cdots+s_0$, whose formal degree l in ζ is, say, $\leq k$, then we shall write $t \equiv_\zeta s$ as an abbreviation for the formula $\left(\bigwedge_{i=0}^{l}(t_i = s_i)\right) \wedge \left(\bigwedge_{i=l+1}^{k}(t_i = 0)\right)$; the case $l > k$ is handled similarly. We shall write $t \not\equiv_\zeta s$ for $\neg(t \equiv_\zeta s)$. Thus $t \not\equiv_\zeta 0$, for example, is short for $\neg\bigwedge_{i=0}^{k}(t_i = 0)$.

We now have three relationships between polynomial terms t and s: "t is s," "$t \equiv_\zeta s$," and "$t = s$." By the first we shall mean that t and s, considered as terms, are syntactically identical, i.e., they are the same string of symbols in our alphabet. The statement $t \equiv_\zeta s$ corresponds to the algebraist's statement that two conventional polynomials $\in K[\zeta]$ are "equal" (where for them ζ is an indeterminate over some field K). It is clear that if t is s, then $t \equiv_\zeta s$. And from the latter follows, by the ring axioms, not only $t = s$, but also, for any term τ: $t_k\tau^k+\cdots+t_0 \equiv_\tau s_l\tau^l+\cdots+s_0$; in fact, this last formula *is* $t \equiv_\zeta s$. Anyway, this formula would not follow from $t = s$, usually.

Next, we shall write $t \oplus_\zeta s$ as an abbreviation for the polynomial term

$$t_k\zeta^k + \cdots + t_{l+1}\zeta^{l+1} + (t_l + s_l)\zeta^l + \cdots + (t_0 + s_0),$$

in case $l \leq k$; the case $l > k$ is handled similarly. We shall also write $t \otimes_\zeta s$ as an abbreviation for the polynomial term

$$(t_k s_l)\zeta^{k+l} + \cdots + \Bigg(\sum_{\substack{a+b=r\\ 0\leq a\leq k\\ 0\leq b\leq l}} t_a s_b\Bigg)\zeta^r + \cdots + t_0 s_0,$$

where $r = 0, 1, \ldots, k+l$. By the ring axioms, $t \oplus_\zeta s \;=\; t+s$ and $t \otimes_\zeta s \;=\; ts$.

Finally, for $0 \leq l \leq k$, we call $t_l\zeta^l + \cdots + t_0$ the *l'th reductum* of $t_k\zeta^k + \cdots + t_0$.

Lemma 6.19 Formal division of t by (reducta of) s. *Let t and s be the polynomial terms $t_k\zeta^k+\cdots+t_0$ and $s_l\zeta^l+\cdots+s_0$. Then we can construct, for $e = 0, 1, \ldots, l$, terms $q'_{e,0}, \ldots, q'_{e,k-e}$ and $r_{e,0}, \ldots, r_{e,e-1}$, each built up from the t_j and s_j by the field operations, such that for* *each e,*

$$s_e \neq 0 \ \rightarrow \left[\begin{array}{l} t_k\zeta^k + \cdots + t_0 \ \equiv_\zeta \\ \ \big[(s_e\zeta^e + \cdots + s_0) \ \otimes_\zeta \ (q'_{e,k-e}\zeta^{k-e} + \cdots + q'_{e,0})\big] \\ \hfill \oplus_\zeta \ (r_{e,e-1}\zeta^{e-1} + \cdots + r_{e,0}) \end{array} \right]$$

is provable from the field axioms. (Here $q'_{e,k-e}\zeta^{k-e} + \cdots + q'_{e,0}$ denotes the term 0 in case $k - e < 0$; and $r_{e,e-1}\zeta^{e-1} + \cdots + r_{e,0}$ denotes the term 0 in case $e = 0$.) Consequently,

$$s \not\equiv_\zeta 0 \ \rightarrow \bigvee_{e=0}^{l} \left[\begin{array}{l} t_k\zeta^k + \cdots + t_0 \ \equiv_\zeta \\ \ \big[(s_e\zeta^e + \cdots + s_0) \ \otimes_\zeta \ (q'_{e,k-e}\zeta^{k-e} + \cdots + q'_{e,0})\big] \\ \hfill \oplus_\zeta \ (r_{e,e-1}\zeta^{e-1} + \cdots + r_{e,0}) \end{array} \right].$$

Proof. Suppose $\bigvee_{e=0}^{l} (s_e \neq 0)$. Then there are $l+1$ cases, indexed by $e = 0, 1, \ldots, l$:

Case e: $s_e \neq 0 \ \wedge \ \bigwedge_{a=e+1}^{l}(s_a = 0)$ (if $e = l$, this means $s_l \neq 0$). We shall prove by induction on k that for this fixed e, we can construct terms q'_{ej} and r_{ej} for which we can prove

$$t_k\zeta^k + \cdots + t_0 \ \equiv_\zeta$$
$$\big[(s_e\zeta^e + \cdots + s_0) \otimes_\zeta (q'_{e,k-e}\zeta^{k-e} + \cdots + q'_{e,0})\big] \oplus_\zeta (r_{e,e-1}\zeta^{e-1} + \cdots + r_{e,0})$$

from the field axioms. For $k = 0, 1, \ldots, e-1$, we need no q'_{ej}'s, and we may take r_{ej} to be $\begin{cases} t_j & \text{if } 0 \le j \le k, \text{ and} \\ 0 & \text{if } k < j < e. \end{cases}$ For $k \ge e$, we may assume that we have already described how to handle dividends of the form $t_{k-1}\zeta^{k-1} + \cdots + t_0$ in Case e. Take $q'_{e,k-e}$ to be $s_e^{-1}t_k$. Note that we can prove

$$(t_k\zeta^k + \cdots + t_0) \ \oplus_\zeta \ \big[\,(s_e\zeta^e + \cdots + s_0) \otimes_\zeta \ (-q'_{e,k-e})\zeta^{k-e}\big]$$
$$\equiv_\zeta \ 0 \cdot \zeta^k + t'_{k-1}\zeta^{k-1} + \cdots + t'_0,$$

for suitable terms $t'_{k-1}, \ldots, t'_0$ built up from $t_0, \ldots, t_l, s_0, \ldots, s_e, s_e^{-1}$ by the ring operations. By the inductive hypothesis, we have q'_{ej}, r_{ej}, and a proof that

$$\begin{aligned}
&t'_{k-1}\zeta^{k-1} + \cdots + t'_0 \ \equiv_\zeta \\
&\quad \big[(s_e\zeta^e + \cdots + s_0) \ \otimes_\zeta \ (q'_{e,k-1-e}\zeta^{k-1-e} + \cdots + q'_{e,0})\big] \\
&\qquad\qquad\qquad \oplus_\zeta \ (r_{e,e-1}\zeta^{e-1} + \cdots + r_{e,0}); \quad \text{then} \\
&t_k\zeta^k + \cdots + t_0 \ \equiv_\zeta \\
&\big[(s_e\zeta^e + \cdots + s_0) \ \otimes_\zeta \ (q'_{e,k-e}\zeta^{k-e} + q'_{e,k-1-e}\zeta^{k-1-e} + \cdots + q'_{e,0})\big] \\
&\qquad\qquad \oplus_\zeta \ (r_{e,e-1}\zeta^{e-1} + \cdots + r_{e,0}).
\end{aligned}$$

□

Lemma 6.20 Formalized Euclidean Algorithm. *Let t and s be the polynomial terms $t_k\zeta^k + \cdots + t_0$ and $s_u\zeta^u + \cdots + s_0$. Then we can construct a finite index set E, and for each index $e \in E$: (a) a quintuple $a_e, b_e, d_e, t'_e, s'_e$ of polynomial terms in ζ, each of whose ζ-coefficients is a term built up from the t_j, s_j by the field operations; (b) a formula A_e, the terms occurring in which are built up from the t_j, s_j by the field operations; and (c) a proof from the field axioms of*

$$A_e \ \rightarrow \ [t \equiv_\zeta d_e \otimes_\zeta t'_e \ \wedge \ s \equiv_\zeta d_e \otimes_\zeta s'_e \ \wedge \ (t \otimes_\zeta a_e) \oplus_\zeta (s \otimes_\zeta b_e) \equiv_\zeta d_e].$$

Moreover, $\bigvee_e A_e$ is a tautology (and so is $\neg(A_e \wedge A_{e'})$ for e' distinct from e). Thus, at least one of the d_e is a "ζ-GCD" (greatest common divisor) of t and s.

Proof. The first case we consider is $s \equiv_\zeta 0$, which we take for A_e. Then we may take d_e to be t; a_e and t'_e to be 1; and b_e and s'_e to be 0.

For the other cases, we may assume $s \not\equiv_\zeta 0$. First divide t by s, as in (6.19). We get, for $e_0 = 0, 1, \ldots, u$:

$$s_{e_0} \neq 0 \ \rightarrow \ \begin{bmatrix} t_k\zeta^k + \cdots + t_0 \ \equiv_\zeta & \\ & \big[(s_{e_0}\zeta^{e_0} + \cdots + s_0) \otimes_\zeta q_{e_0}\big] \\ & \oplus_\zeta (r_{(e_0),e_0-1}\zeta^{e_0-1} + \cdots + r_{(e_0),0}) \end{bmatrix},$$

for suitable polynomials q_{e_0} and $r_{(e_0)}$ of degrees $\max\{0,\ k-e_0\}$ and $e_0 - 1$ in ζ, respectively. (Exception: if $e_0 = 0$, $r_{(e_0)}$ will be 0.) Second, for $e_0 = 1, 2, \ldots, u$, divide the e_0th reductum of s by $r_{(e_0)}$. We get, for $e_1 = 0, 1, \ldots, e_0 - 1$:

$$\begin{aligned}
&r_{(e_0),e_1} \neq 0 \ \rightarrow \\
&\quad \begin{bmatrix} s_{e_0}\zeta^{e_0} + \cdots + s_0 \ \equiv_\zeta \big[(r_{(e_0),e_1}\zeta^{e_1} + \cdots + r_{(e_0),0}) \ \otimes_\zeta \ q_{e_0,e_1}\big] \\ \oplus_\zeta \ (r_{(e_0,e_1),e_1-1}\zeta^{e_1-1} + \cdots + r_{(e_0,e_1),0}) \end{bmatrix},
\end{aligned}$$

for suitable polynomials $r_{(e_0,e_1)}$ and q_{e_0,e_1} of degrees e_1-1 and $\max\{0, e_0-e_1\}$ in ζ, respectively (except that $r_{(e_0,e_1)}$ denotes 0 if $e_1=0$). (Actually, our use of the notation q_{e_0,e_1} here is only temporary, and has nothing to do with our more permanent use, throughout the rest of §6, of the notation q_{ij} in (6.1.1).)

To continue this iteration, for each sequence $e_0, e_1, \ldots, e_p$ $(p \geq 1)$ such that $u \geq e_0 > e_1 > \cdots > e_p > 0$, divide the e_pth reductum of $r_{(e_0,\ldots,e_{p-1})}$ by $r_{(e_0,\ldots,e_p)}$ (as ζ-polynomials). We get, for $e_{p+1} = 0, 1, \ldots, e_p - 1$:

$$r_{(e_0,\ldots,e_p),e_{p+1}} \neq 0 \;\rightarrow$$
$$\left[\begin{array}{l} r_{(e_0,\ldots,e_{p-1}),e_p}\zeta^{e_p} + \cdots + r_{(e_0,\ldots,e_{p-1}),0} \\ \qquad \equiv_\zeta \;\left[\left(r_{(e_0,\ldots,e_p),e_{p+1}}\zeta^{e_{p+1}} + \cdots + r_{(e_0,\ldots,e_p),0}\right) \otimes_\zeta q_{e_0,\ldots,e_{p+1}}\right] \\ \qquad\quad \oplus_\zeta \left(r_{(e_0,\ldots,e_{p+1}),e_{p+1}-1}\zeta^{e_{p+1}-1} + \cdots + r_{(e_0,\ldots,e_{p+1}),0}\right) \end{array}\right],$$

for suitable polynomials $q_{e_0,\ldots,e_{p+1}}$ and $r_{(e_0,\ldots,e_{p+1})}$ of degrees $\max\{0, e_p - e_{p+1}\}$ and $e_{p+1}-1$ in ζ, respectively (except that $r_{(e_0,\ldots,e_{p+1})}$ denotes 0 in case $e_{p+1}=0$).

By this process we eventually get (finitely) many sequences $e_0, e_1, \ldots, e_{p+1}$ (for various $p \geq 0$) such that $u \geq e_0 > e_1 > \cdots > e_{p+1} > 0$. We shall denote such a sequence by e. The formula A_e may be taken to be

$$s_{e_0} \neq 0 \;\wedge\; \left(\bigwedge_{k=e_0+1}^{u} s_k = 0\right) \wedge$$
$$r_{(e_0),e_1} \neq 0 \;\wedge\; \left(\bigwedge_{k=e_1+1}^{e_0-1} r_{(e_0),k} = 0\right) \wedge$$
$$\vdots \quad \vdots \quad \vdots \quad \vdots$$
$$r_{(e_0,\ldots,e_p),e_{p+1}} \neq 0 \;\wedge\; \left(\bigwedge_{k=e_{p+1}+1}^{e_p-1} r_{(e_0,\ldots,e_p),k} = 0\right) \wedge$$
$$\bigwedge_{k=0}^{e_{p+1}-1} r_{(e_0,\ldots,e_{p+1}),k} = 0.$$

For each sequence e, we have the theorem $A_e \rightarrow B_e$, where B_e is

$$\left[\begin{aligned} t_k\zeta^k + \cdots + t_0 \equiv_\zeta & \left[(s_{e_0}\zeta^{e_0} + \cdots + s_0) \otimes_\zeta q_{e_0}\right] \\ & \oplus_\zeta \ (r_{(e_0),e_0-1}\zeta^{e_0-1} + \cdots + r_{(e_0),0}) \end{aligned} right]\right. \ \wedge$$

$$\left[\begin{aligned} s_{e_0}\zeta^{e_0} + \cdots + s_0 \equiv_\zeta & \left[(r_{(e_0),e_1}\zeta^{e_1} + ots + r_{(e_0),0}) \otimes_\zeta q_{e_0,e_1}\right] \\ & \oplus_\zeta \ (r_{(e_0,e_1),e_1-1}\zeta^{e_1-1} + \cdots + r_{(e_0,e_1),0}) \end{aligned}\right] \wedge$$

$$\left[\begin{aligned} r_{(e_0),e_1}\zeta^{e_1} + \cdots + r_{(e_0),0} \quad 05in & \equiv_\zeta \left[(r_{(e_0,e_1),e_2}\zeta^{e_2} + \cdots + r_{(e_0,e_1),0}) es \zeta q_{e_0,e_1,e_2}\right] \\ & \oplus_\zeta \ (r_{(e_0,e_1,e_2),e_2-1}\zeta^{e_2-1} + \cdots + r_{(e_0,e_1,e_2),0}) \end{aligned}\right] \wedge$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$\left[\begin{aligned} & r_{(e_0,\ldots,e_{p-2}),e_{p-1}}\zeta^{e_{p-1}} + \cdots + r_{(e_0,\ldots,e_{p-2}),0} \\ & \quad \equiv_\zeta \ \left[(r_{(e_0,\ldots,e_{p-1}),e_p}\zeta^{e_p} + \cdots + r_{(e_0,\ldots,e_{p-1}),0}) \otimes_\zeta q_{e_0,\ldots,e_p}\right] \\ & \qquad \oplus_\zeta \ (r_{(e_0,\ldots,e_p),e_p-1}\zeta^{e_p-1} + \cdots + r_{(e_0,\ldots,e_p),0}) \end{aligned}\right] \wedge$$

$$\left[\begin{aligned} & r_{(e_0,\ldots,e_{p-1}),e_p}\zeta^{e_p} + \cdots + r_{(e_0,\ldots,e_{p-1}),0} \\ & \quad \equiv_\zeta \ (r_{(e_0,\ldots,e_p),e_{p+1}}\zeta^{e_{p+1}} + \cdots + r_{(e_0,\ldots,e_p),0}) \otimes_\zeta q_{e_0,\ldots,e_{p+1}} \end{aligned}\right].$$

As in the usual Euclidean algorithm, we may take the GCD d_e of t and s to be $r_{(e_0,\ldots,e_p),e_{p+1}}\zeta^{e_{p+1}} + \cdots + r_{(e_0,\ldots,e_p),0}$ in the eth case. Also as usual, we may construct the required polynomial terms a_e, b_e, t'_e, s'_e by working backwards through the above chain(s) of equations. $\Box$

6.21 The depth of a root-term. Let ζ denote any "root-term," i.e., a term of either of the forms $\rho_q(s_0, \ldots, s_{2q})$ for some $q \geq 1$, or $\rho(s_0)$, for some subterms $s_0, \ldots, s_{2q}$ (the "argument(s)" of that root-term). Define a *root-chain* to be a sequence $\zeta_1, \ldots, \zeta_E$ of root-terms ζ_e such that for $1 < e \leq E$, ζ_e is a proper subterm of ζ_{e-1}; thus, ζ_e occurs within (one of) the argument(s) of ζ_{e-1}. Define the *depth* of a root-term ζ_1 to be the maximum of the lengths of all root-chains beginning with ζ_1.[32] One root-chain may have many "occurrences" in a given proof: e.g., ζ_1 may occur several times in a formula in the proof, and if ζ_1 is $\rho_1(s_0, s_1, s_2)$, then ζ_2 may have several occurrences in the argument s_0, and several more in s_1 and s_2, etc. Let M denote the maximum of the depths of all the root-terms in the proof considered

[32]The depth of a root-term is similar to what Ackermann and Hilbert called the "degree" of an ε-term; cf. Hilbert, Bernays, Vol. II, [1939, 1970, p. 25]. But in our context, the word degree might suggest the formal degree of the defining polynomial of ζ (i.e., either $2q + 1$, or 2), which has nothing to do with what I am calling the depth.

in (6.17); this equals the maximum of the lengths of all root-chains in that proof.

6.22 The overall strategy for eliminating $\boldsymbol{\rho}$ and $\boldsymbol{\rho_q}$. Let ζ be one of those root-terms of depth M. We shall give a (lengthy) procedure for eliminating ζ from the entire proof, without introducing any new root terms of depth $\geq M$.[33] This will reduce the *number* of root-terms of depth M occurring in the proof. By iterating this procedure, we eliminate *all* root-terms of depth M, thereby reducing M itself; once M is 0, all the ρ- and ρ_q-symbols will have been eliminated from the entire proof.

6.23 The defining polynomial s of ζ will be, in the case where ζ is of the form $\rho_q(s_0, \ldots, s_{2q})$, the term $\zeta^{2q+1} + s_{2q}\zeta^{2q} + \cdots + s_0$; in the second case, where ζ is $\rho(s_0)$, the defining polynomial will be either $\zeta^2 - s_0$ or $\zeta^2 + s_0$, depending on which stage of the process we are in. We shall also distinguish between the *original* and the *current* defining polynomial for ζ. The original defining polynomial is the one specified above, denoted by s. As the algorithm proceeds, we shall consider various cases, subcases, etc. In each (sub)case, we shall select certain (monic, non-constant) "factors" s' of the original defining polynomial, for which we prove $s' = 0$ in that (sub)case; the most-recently selected s' will be called the current defining polynomial.

Denote the degree of the current defining polynomial by u.

Definition/Lemma 6.24 Organizing a term into a polynomial in $\boldsymbol{\zeta}$ (compare Tarski [1951, p. 17]) *Let ζ be the target root-term selected in (6.22) above, and let t be any term. We define a certain formal polynomial $P_\zeta(t)$ by recursion on t, as follows. If t is either $0, 1, c_j, x_j, a_{ij}, b_{ij}$, or any root-term other than ζ, then define $P_\zeta(t)$ to be t; if t is ζ, define $P_\zeta(t)$ to be $0+1\cdot\zeta$; and if t_1 and t_2 are terms, define $P_\zeta(t_1 + t_2)$, $P_\zeta(-t_1)$, $P_\zeta(t_1 t_2)$, and $P_\zeta(t_1^{-1})$ to be $P_\zeta(t_1) \oplus_\zeta P_\zeta(t_2)$, $(-1) \otimes_\zeta P_\zeta(t_1)$, $P_\zeta(t_1) \otimes_\zeta P_\zeta(t_2)$, and $(P_\zeta(t_1))^{-1}$, respectively. From the ring axioms we can prove $t = P_\zeta(t)$.*

6.25 Organizing atomic subformulae of the proof. Recall (6.17) that the end-formula of our proof is $\bigvee_k (F = \sum_l y_{k,l}^2)$. Writing $P_\zeta(y_{k,l})$ in the form $\sum_r u_{k,l,r}\zeta^r$ for certain terms $u_{k,l,r}$, we can now prove $\bigvee_k [F = \sum_l (\sum_r u_{k,l,r}\zeta^r)^2]$.

[33] On the other hand, some of the old root-terms may occur more often in the new proof than in the old, and some new root-terms of depth $< M$ may be introduced.

Now replace each atomic subformula $a = b$ occurring in the proof (where a and b are terms), with $P_\zeta(a - b) = 0$. From the ring axioms we can prove $a = b \leftrightarrow P_\zeta(a - b) = 0$. Therefore our proof has been transformed into a(nother) sequence of provable formulae. (However, this sequence is not quite a proof; for example, the axiom $1 \neq 0$ becomes $1 - 0 \neq 0$, which is not an axiom. Rather, our new sequence could be called a "skeleton" of a proof, in the vague sense that we could restore the proof-connectedness of our sequence by inserting small pieces of proof, using only the ring axioms, between consecutive formulae.)

Now each atomic subformula is of the form $u_p\zeta^p + \cdots + u_0 = 0$, for various integers p and terms u_k. In particular, the end-formula is of the form $\bigvee_k \left[P_\zeta\left(F - \sum_l (\sum_r u_{k,l,r}\zeta^r)^2\right) = 0\right]$, for some terms $u_{k,l,r}$. The only way that ζ can occur in the u_k is as a subterm of some reciprocal t^{-1} in u_k (since ζ has maximal depth). In the next proposition, we rearrange these atomic subformulae so that this does not happen.

Proposition 6.26 Eliminating reciprocals containing $\boldsymbol{\zeta}$. *We can transform the above sequence of provable formulae into a sequence of formulae provable in* $\mathbf{RCF}^{\mathrm{qf}}_{ndi} + (s = 0)$ *(where s is the defining polynomial of ζ (6.23)), in which each atomic subformula is of the form $v_p\zeta^p + \cdots + v_0 = 0$, for various integers p and terms v_k* **not containing** *ζ. The end-formula is equivalent (via the ring axioms) to another (bigger) disjunction of the form $\bigvee_{k'}[F = \sum_{l'}(\sum_{r'} v_{k',l',r'}\zeta^{r'})^2]$, where ζ does not occur in the $v_{k',l',r'}$.*

Proof. We call a term of the form t^{-1} a *reciprocal term,* or simply a *reciprocal.* We say that t^{-1} *contains* ζ if ζ occurs in t. A reciprocal t^{-1} containing ζ will be called *innermost*, if t contains no subterm of the form ${t'}^{-1}$ in which ζ occurs. Thus inside every reciprocal u^{-1} containing ζ, we can find an innermost reciprocal t^{-1} containing ζ.

Let t^{-1} be an innermost reciprocal containing ζ, and suppose t^{-1} occurs in our sequence of provable formulae. We shall show how to eliminate t^{-1} from our sequence, without introducing any new reciprocals containing ζ. Upon iterating this procedure, we shall eliminate *all* such reciprocals from our sequence.

Since t^{-1} occurs in our sequence, t has already been organized into a polynomial in ζ, say $t_v\zeta^v + \cdots + t_0$; the t_j do not contain ζ, since t^{-1} is innermost.

Apply the Euclidean algorithm (6.20) to t and s, where s is the current defining polynomial of ζ (6.23); we obtain

$$A_e \;\rightarrow\; [t \equiv_\zeta d_e \otimes_\zeta t'_e \;\wedge\; s \equiv_\zeta d_e \otimes_\zeta s'_e \;\wedge\; (t \otimes_\zeta a_e) \oplus_\zeta (s \otimes_\zeta b_e) \equiv_\zeta d_e],$$

for finitely many values of the index e, as in (6.20). Recall that both the ζ-coefficients of the $a_e, b_e, t'_e, s'_e, d_e$, and the terms occurring in the formulae A_e, are built up by the field operations from the ζ-coefficients of t and s, and therefore do not contain ζ. Also, recall that the A_e represent the various possible (sub)cases occurring in the Euclidean algorithm; we shall consider each case separately.

Thus we fix a value of e, and temporarily add the corresponding formula A_e to the list of axioms of $\mathbf{RCF}^{\mathrm{qf}}_{ndi}$. Then we can prove (from A_e and the field axioms)

$$t \equiv_\zeta d_e \otimes_\zeta t'_e \quad \wedge \quad s \equiv_\zeta d_e \otimes_\zeta s'_e \quad \wedge \quad (t \otimes_\zeta a_e) \oplus_\zeta (s \otimes_\zeta b_e) \equiv_\zeta d_e.$$

Write d_e as $d_{e,v_e}\zeta^{v_e} + \cdots + d_{e,0}$, and write s'_e as $s'_{e,u'_e}\zeta^{u'_e} + \cdots + s'_{e,0}$ (where the $d_{e,j}$ and $s'_{e,j}$ are free of ζ).

Note that $d_e \not\equiv_\zeta 0$, and that $v_e \leq u$ (= the degree of s), since $s \equiv_\zeta d_e \otimes_\zeta s'_e$ and s is monic. Thus one of the following subcases holds:

Subcase *e*.0. The degree of d_e is 0: $d_{e,0} \neq 0 \wedge \bigwedge_{k=1}^{v_e}(d_{e,k} = 0)$. Then $d_e \neq 0$. Therefore $t \neq 0$, using $ta_e + sb_e = d_e$ and $s = 0$. Also $(ta_e)d_e^{-1} + (sb_e)d_e^{-1} = 1$. Using $s = 0$ again, we conclude $t(a_e d_e^{-1}) = 1$; multiplying both sides by t^{-1}, we obtain $t^{-1} = a_e d_e^{-1}$ (using $t \neq 0$). Thus we may replace t^{-1} by $a_e \otimes_\zeta d_e^{-1}$ throughout the sequence of provable formulae, in this subcase.

Subcase *e.u*. The degree of d_e is u: $v_e = u$ and $d_{e,u} \neq 0$. Then s'_e is a constant polynomial: $s'_{e,0} \neq 0 \wedge \bigwedge_{k=1}^{u'_e}(s'_{e,k} = 0)$. Therefore $d_e \equiv_\zeta {s'_{e,0}}^{-1} \otimes_\zeta s$ (since $s \equiv_\zeta d_e \otimes_\zeta s'_e$). Therefore $d_e = 0$ (using $s = 0$). Then $t = 0$ (using $t = d_e t'_e$). Thus $t^{-1} = 0$, using $0^{-1} = 0$. We may therefore replace t^{-1} by 0 throughout our sequence of provable formulae, in this subcase.

Subcase *e.p* . The degree of d_e is p, where $0 < p \leq \min\{v_e, u-1\}$: $d_{e,p} \neq 0 \wedge \bigwedge_{k=p+1}^{v_e}(d_{e,k} = 0)$. (In case $v_e = p < u$, this collapses to $d_{e,p} \neq 0$.) Then $s'_{e,z} \neq 0 \wedge \bigwedge_{k=z+1}^{u'_e}(s'_{e,k} = 0)$, where $0 < z := z_{e,p} = u - p < u$. I.e., neither d_e nor s'_e are constant polynomials.

This subcase splits into two "subsubcases": $t = 0$ and $t \neq 0$. If $t = 0$, then $d_e = 0$ (using $s = 0$). Write d' for $d_{e,p}^{-1} \otimes_\zeta d_e$; note that d' is monic and non-constant, and $d' = 0$. Conversely, $d' = 0$ implies $s = 0$ and $t = 0$, also by the field axioms. We take d' as the new, current defining polynomial for ζ (6.23), instead of s. As in subcase $e.u$ above, we may replace t^{-1} by 0 throughout our sequence of provable formulae.

If, on the other hand, $t \neq 0$, then $d_e \neq 0$ (since $t = d_e t'_e$). Therefore $s'_e = 0$ (using $s = 0$). This time, write s^* for the monic, non-constant polynomial ${s'_{e,z}}^{-1} \otimes_\zeta s'_e$. Then $s^* = 0$. Conversely, from $s^* = 0$, we could prove $s = 0$, in this subsubcase; however, we might not be able to prove $t \neq 0$ just from $s^* = 0$. We take s^* as the new, current defining polynomial for ζ, instead of s. To conclude this subsubcase, we iterate the above procedure, this time applying the Euclidean algorithm to t and s^*, rather than to t and s. This involves yet more cases (indexed by, say, $e.p.e'$ for various e'), and then subcases $e.p.e'.0$, $e.p.e'.z$, and $e.p.e'.p'$, analogous to $e.0$, $e.u$, and $e.p$, respectively. (Actually, the subcase $e.p.e'.z$ analogous to $e.u$ will not arise, since we have assumed $t \neq 0$.) In subcase $e.p.e'.0$, we prove $t^{-1} = u$, for some ζ-polynomial u whose coefficients are free of—reciprocals containing—ζ. In subcase $e.p.e'.p'$, we go through the iteration yet again, leading to yet another current defining polynomial for ζ, and yet another greatest common divisor; at each stage the degree of s^* or d_e (or both) is lowered. Thus, after at most u repetitions of all this, the degree of s^* or d_e will be 0; in fact, using $t \neq 0$, we see that it is d_e, and not s^*, that will end up having degree 0. Conversely, $s^* = 0$ implies the vanishing of all the prior defining polynomials, as well as the non-vanishing of t, using only the assumptions of the present (sub)case. This last s^* will become the final defining polynomial s for ζ. As in subcase $e.0$, we shall have found a suitable replacement polynomial for t^{-1}.

Now carry out this replacement, according to the above (sub)cases. For each resulting atomic subformula $u = 0$, re-organize the term u into a polynomial in ζ—that is, replace u with $P_\zeta(u)$ (6.24).

If the new sequence of provable formulae contains no more reciprocals t'^{-1} containing ζ, then we are done; otherwise, not forgetting which of the above (sub)cases we are in and what the current defining polynomial for ζ is, we select from the new sequence an innermost reciprocal t'^{-1} (other than t^{-1}, which has been eliminated), and introduce yet more cases, subcases, etc., in each of which we construct a replacement polynomial term for t'^{-1} whose coefficients are free of—reciprocals containing—ζ. Eventually, we replace all such reciprocals with polynomial terms.[34] Then we have a sequence of provable formulae with no reciprocals containing ζ. The end-formula is of the required form.

Actually, we have *many* such sequences, one for each case, subcase, etc. And when we say "provable," we now mean "provable using the additional, temporary axioms corresponding to the present case, subcase, etc." On the other hand, since it is a tautology that at least

[34]For any polynomial term v in ζ, we shall, in (6.33), write $Q_e(v)$ for the result of the above transformation, i.e., successively replacing innermost reciprocals containing ζ with polynomials whose coefficients are ζ-free, and then re-applying P_ζ. Thus every atomic subformula $v = 0$ is transformed by (6.26) into $Q_e(v) = 0$ in case e.

one of these (sub)cases must occur, we can assemble from all these sequences of provable formulae a single sequence of provable formulae, the proofs of which do not use as initial formulae any of the temporary axioms corresponding to the various (sub)cases; the new end-formula will be the disjunction of the various end-formulae corresponding to the (sub)cases, and so will still be of the required form.[35] □

6.27 Reducing the end-formula mod s. So far in the present (sub)case, our end-formula is (after use of the ring axioms) of the form $\bigvee_k [F = \sum_l (\sum_r u_{k,l,r}\zeta^r)^2]$, with each $u_{k,l,r}$ free of ζ. Recall that s is the (original) defining polynomial of ζ, i.e., either $\zeta^{2q+1} + s_{2q}\zeta^{2q} + \cdots + s_0$, $\zeta^2 - s_0$, or $\zeta^2 + s_0$. Divide each squared polynomial $\sum_r u_{k,l,r}\zeta^r$ by s: $\sum_r u_{k,l,r}\zeta^r \equiv_\zeta (s \otimes_\zeta q'_{k,l}) \oplus_\zeta r_{k,l}$, for ζ-polynomials $q'_{k,l}$ and $r_{k,l}$; here the degree of $r_{k,l}$ is $u-1$ (where u is the degree of s), and, writing $r_{k,l}$ as $\sum_{j=0}^{u-1} v_{k,l,j}\zeta^j$, the coefficients $v_{k,l,j}$ are free of ζ. Using the axiom $s = 0$, we conclude $\sum_r u_{k,l,r}\zeta^r = r_{k,l}$, whence $\bigvee_k [F = \sum_l r_{k,l}^2]$, our new end-formula.

We shall treat separately the two cases:

(a) where ζ is of the form $\rho_q(s_0, \ldots, s_{2q})$, and

(b) where ζ is of the form $\rho(s_0)$.

6.28 In case (a), we shall show (6.30–32) how to convert this proof into a proof of a disjunction $\bigvee_{k'} \bigvee_k [F = \sum_l t_{k',k,l}^2]$, for new terms $t_{k',k,l}$ not containing ζ; no new root-terms will be added to the end-formula. Actually, we may not need to apply (6.32), since ζ need not necessarily occur in the end-formula; but if we do apply (6.32), it is possible that it will introduce some new root-terms ζ^* of the form $\rho_{q^*}(s'_0, \ldots, s'_{2q^*})$ $(q^* < q)$ into *earlier* parts of the proof; these new root-terms may even have depth $\geq M$. If so, we start over and select one of these root-terms—call it ζ^*—of maximal depth $M' \geq M$ among all root-terms in the new proof. In (6.33) we shall show how to completely eliminate any maximal-depth root-term (such as ζ^*) from any given proof; in particular, the original defining axiom for ζ^* (i.e., an instance of $(6.4.2)_{q^*}$, with $q^* < Q$) will no longer occur. No new root-terms will be introduced by (6.33) anywhere in the proof. The only cost is that the new proof uses a new axiom, denoted by P_p, stating that the characteristic of the field is either 0 or $\geq p$, where p is computable from the original proof. The elimination will be done in such a way

[35]However, in future steps it will be useful to remember the various cases, and the corresponding current defining polynomials for ζ.

as to cause only minor alterations to the end-formula, provided that the latter does not contain ζ^* (which it won't, thanks to (6.32)); in particular, the end-formula will still be a disjunction of sum-of-squares representations of F. Repeated application of (6.33) will eliminate from the proof all other root-terms of depth $\geq M$ that may have been introduced by (6.32). Once they are gone, we apply (6.33) to ζ itself. Now we have eliminated ζ from the entire proof, without introducing any new root-terms of depth $\geq M$ anywhere. Thus the *number* of root-terms of depth M occurring in the proof will have been reduced, completing, in case (a), the strategy set forth in (6.22).

6.29 In case (b) (where ζ is $\rho(s_0)$), we shall begin by carrying out two parallel constructions, one in which we add $s_0 = \zeta^2$ to the list of axioms of $\mathbf{RCF}^{\mathrm{qf}}_{ndi}$, and the other in which we add $-s_0 = \zeta^2$ (either of which, alone, implies the instance $s_0 = \zeta^2 \ \vee\ -s_0 = \zeta^2$ of (6.4.1)). We apply (6.33) to each of the two proofs, eliminating ζ. In this case, the method of elimination consists, essentially, of replacing each occurrence of ζ^2 by s_0 in the first subcase (where we assumed $s_0 = \zeta^2$), and replacing each occurrence of ζ^2 by $-s_0$ in the second subcase. (The exact nature of the alterations will be stated in (6.33).) This time (unlike in case (a)), the fact that neither the reality axiom $a_1^2+a_2^2 \neq -1$, nor any of its instances, occur as initial formulae, is essential, since if ζ is, say, $\rho(-1)$, then instances such as $\rho(-1)^2 + 0^2 \neq -1$ could be transformed into underivable (or even refutable) formulae. Another difference between cases (a) and (b) is that when we enter case (b), the end-formula may well contain ζ (and even ζ^2); a side-effect of (6.33) will be to alter the end-formula in two different ways (i.e., replacing ζ^2 by s_0 and $-s_0$ in the two subcases, respectively). No new root-terms will be introduced anywhere. Thus we shall obtain two proofs, neither of which will contain the temporary axioms $s_0 = \zeta^2$ and $-s_0 = \zeta^2$ (both having been transformed into $0 = 0$ in the appropriate cases); nor will these two proofs contain even the original axiom $s_0 = \zeta^2 \vee -s_0 = \zeta^2$ (6.4.1). Then we shall have a single, ζ-free proof in $\mathbf{RCF}^{\mathrm{qf}}_{ndi} + P_p$ of the *conjunction* of the two new end-formulae. Unfortunately, our replacements of ζ^2 by s_0 and $-s_0$ could ruin the sum-of-squares representation of F; so, in (6.35), we shall derive from this conjunction a new end-formula that does give a sum-of-squares representation, of the form $\bigvee_{k'}[F = \sum_l {t'_{k',l}}^2]$, where the terms $t'_{k',l}$ are free of ζ. This will complete, in case (b), the strategy set forth in (6.22).

6.30 Eliminating odd root-symbols ρ_q from the end-formula. As promised in (6.28), we now consider the case where ζ is $\rho_q(s_0, \ldots,$

s_{2q}), and s is $\zeta^{2q+1} + s_{2q}\zeta^{2q} + \cdots + s_0$. Thus $s = 0$ is an instance of $(6.4.2)_q$. We have a proof of $\bigvee_k \left[F = \sum_l (v_{k,l,2q}\zeta^{2q} + \cdots + v_{k,l,0})^2\right]$, where the $v_{k,l,p}$ are finitely many terms not containing ζ; the proof uses, in addition to the axioms of $\mathbf{RCF}^{\mathrm{qf}}_{ndi}$, the additional, temporary axioms introduced in the present case, subcase, etc. (recall the proof of (6.26), especially footnote 35).

Before eliminating ζ from the end-formula in (6.32), we must eliminate those instances of the special equality axioms for ρ_q $(6.11.1)_q$ that contain ζ.

Lemma 6.31 *In the above situation, we can transform our proof of* $\bigvee_k \left[F = \sum_l (v_{k,l,2q}\zeta^{2q} + \cdots + v_{k,l,0})^2\right]$ *into a proof of*

$$\left(\bigvee_k j\left[F = \sum_l (v_{k,l,2q}\zeta^{2q} + \cdots + v_{k,l,0})^2\right]\right) ll$$
$$\vee \bigvee_{k'} \bigvee_k \left[F = \sum_l (v_{k,l,2q}\zeta_{k'}^{2q} + \cdots + v_{k,l,0})^2\right] .d$$

The new proof does not use as an initial formula any of the special equality axioms for ρ_q that contain ζ:

$$s_j = t \;\rightarrow\; \rho_q(s_0, \ldots, s_j, \ldots, s_{2q}) = \rho_q(s_0, \ldots, t, \ldots, s_{2q}) \quad \text{or } (6.31.1)_j$$
$$t = s_j \;\rightarrow\; \rho_q(s_0, \ldots, t, \ldots, s_{2q}) = \rho_q(s_0, \ldots, s_j, \ldots, s_{2q}), \qquad (6.31.2)_j$$

for any term t and integer j with $0 \le j \le 2q$. The new root-terms $\zeta_{k'}$ are the terms $\rho_q(s_0, \ldots, t, \ldots, s_{2q})$ occurring in (6.31.1–2)$_j$.[36]

Proof. We adapt—a small part of—the Ackermann-Hilbert method of simultaneously eliminating critical ε-formulae and formulae of ε-equality (Hilbert, Bernays, Vol. II, [1939, 1970, pp. 56–79]).

We need not be concerned with the instance of $(6.31.1–2)_j$ in which t is s_j (recall (6.12)). So assume t is not s_j.

On the one hand, if we add the formula $s_j \neq t$ to the axioms, then $(6.31.1)_j$ follows tautologically. We handle $(6.31.2)_j$ similarly, using the derivable formula $s_j = t \leftrightarrow t = s_j$. Thus $(6.31.1–2)_j$ are now derived formulae, rather than initial formulae.

On the other hand, if we add the formula $s_j = t$ instead of $s_j \neq t$ to the axioms, then we eliminate $(6.31.1–2)_j$ as follows. For any term r or

[36]It may appear that the end-formula has lost its previous form, since now some of the disjuncts are representations of F as sums of squares of polynomials in $\zeta_{k'}$, for certain root-terms $\zeta_{k'}$ other than ζ. But those polynomials do not contain ζ, and so can be construed as *constant* polynomials in ζ.

formula B, write r^* or B^* for the result of replacing $\rho_q(s_0, \dots, s_j, \dots, s_{2q})$ by $\rho_q(s_0, \dots, t, \dots, s_{2q})$ throughout r or B, respectively. (Thus ζ^* is $\rho_q(s_0, \dots, t, \dots, s_{2q})$, for example.) Replace each formula B in the proof by B^*.

When B is $(6.31.1)_j$, the resulting formula, $(6.31.1)^*_j$, is $s_j = t \to \zeta^* = \zeta^*$ (the terms s_j and t in the premise are unaffected by this replacement, since they do not contain ζ, by the choice of ζ in (6.22)); this formula follows from $\zeta^* = \zeta^*$ (as in the case of improper axioms of equality treated above). Likewise for $(6.31.2)_j$.

When B is $s_j = t' \to \rho_q(s_0, \dots, s_j, \dots, s_{2q}) = \rho_q(s_0, \dots, t', \dots, s_{2q})$, with t' distinct from t, s_j, B^* is $s_j = t' \to \zeta^* = \rho_q(s_0, \dots, t', \dots, s_{2q})$, which is no longer a special equality axiom. However, it follows from the derivable formula $s_j = t' \to t = t'$ (which uses $s_j = t$), and the axiom $t = t' \to \zeta^* = \rho_q(s_0, \dots, t', \dots, s_{2q})$ (in which ζ does not occur). (Again, we should also consider formulae B of the form $t' = s_j \to \rho_q(s_0, \dots, t', \dots, s_{2q}) = \zeta$; they are handled similarly.)

It remains to examine the effect that our replacement of ζ by ζ^* has on the rest of the proof.

First, the only other special equality axioms in (6.11) that might contain ζ are those for the field operations, such as $r_1 = r_2 \to r_1 + v = r_2 + v$; these become new instances of axioms of the same type (e.g., $r_1^* = r_2^* \to r_1^* + v^* = r_2^* + v^*$).

Second, instances of the field axioms (e.g., $ab = ba$), become new instances of axioms of the same type.

Third, our replacement will transform each application of modus ponens: $\dfrac{B_1 \quad B_1 \to B_2}{B_2}$ into another instance of modus ponens, since $(B_1 \to B_2)^*$ is obviously $B_1^* \to B_2^*$.

Fourth, the axioms $2 \neq 0$ and (6.2.1), as well as the temporary axioms such as A_e that were added in the present (sub)case (6.26), do not contain ζ, and so are unaffected by our replacement.

Fifth, instances of the axiom (6.4.1) (for ρ) also do not contain ζ (since ζ has maximal depth).

Sixth, the tautologies become new tautologies.

Seventh, the end-formula becomes $\bigvee_k \left[F = \sum_l (v_{k,l,2q}{\zeta^*}^{2q} + \cdots + v_{k,l,0})^2\right]$.

Finally (eighth), by the choice of ζ, the only instance of axiom $(6.4.2)_q$ (for ρ_q) that contains ζ is $s = 0$, where s is $\zeta^{2q+1} + s_{2q}\zeta^{2q} + \cdots + s_0$. Replacing ζ by ζ^* gives${\zeta^*}^{2q+1} + s_{2q}{\zeta^*}^{2q} + \cdots + s_j{\zeta^*}^j + \cdots + s_0 = 0$. Unfortunately, this is not an axiom. But it does follow from the three derivable formulae $s_j = t \rightarrow s_j\zeta^{*j} = t\zeta^{*j}$,

$$s_j\zeta^{*j} = t\zeta^{*j} \rightarrow \begin{bmatrix} \zeta^{*2q+1} + s_{2q}\zeta^{*2q} + \cdots + s_j\zeta^{*j} + \cdots + s_0 = \\ \zeta^{*2q+1} + s_{2q}\zeta^{*2q} + \cdots + t\zeta^{*j} + \cdots + s_0 \end{matrix} \quad right],$$

and $\zeta^{*2q+1} + s_{2q}\zeta^{*2q} + \cdots + t\zeta^{*j} + \cdots + s_0 = 0$.

Thus we have proofs of $\bigvee_k \left[F = \sum_l (v_{k,l,2q}\zeta^{2q} + \cdots + v_{k,l,0})^2\right]$ and $\bigvee_k \left[F = \sum_l (v_{k,l,2q}\zeta^{*2q} + \cdots + v_{k,l,0})^2\right]$, using $s_j \neq t$ and using $s_j = t$ respectively, but not using the formulae $(6.31.1\text{–}2)_j$ as initial formulae, and without introducing any new such formulae into the proof. From these two proofs we can easily assemble a proof, using neither $s_j \neq t$ nor $s_j = t$ as initial formulae, of the disjunction of the two above end-formulae.

If more formulae of the form $(6.31.1\text{–}2)_j$ (for different j or different terms t) occur as initial formulae in our new proof, then eliminate those formulae one by one, in the same way, obtaining a bigger disjunction as end-formula, but still of the required form. $\square$

Recall the situation described in (6.30), and apply (6.31). Now we have a proof of $\bigvee_k \left[F = \sum_l (v_{k,l,2q}\zeta^{2q} + \cdots + v_{k,l,0})^2\right]$, where the $v_{k,l,p}$ are finitely many terms not containing ζ; the proof uses, in addition to the axioms of $\mathbf{RCF}^{\mathrm{qf}}_{ndi}$, the additional, temporary axioms A_e introduced in the present case, subcase, etc. by (6.26) (recall footnote 35), and possibly also the axiom P_p, introduced during an earlier application of (6.33) (recall (6.28)). However, the proof does not use $(6.31.1\text{–}2)_j$ as initial formulae. (It is possible that ζ does not occur at all in our end-formula, but only in earlier lines of the proof; if so, we can skip over (6.32) below, and continue the algorithm with (6.33).)

Proposition 6.32 *In the above situation, we can construct a proof of* $\bigvee_{k'} \left[F = \sum_{l'} w^2_{k',l'}\right]$, *where the* $w_{k',l'}$ *are built up by the field operations from* $s_0, \ldots, s_{2q}$, F, *and the* $v_{k,l,p}$ *(and hence do not contain* ζ*). No new root-terms are introduced into the end-formula (although some new root-terms of the form* $\rho_{q^*}(s'_0, \ldots, s'_{2q^*})$, *for* $q^* < q$, *may be introduced into the* rest *of the proof).*

Proof. We assume, by induction on $q \geq 1$, that for $1 \leq q^* < q$, we can transform any proof of $\bigvee_k \left[F = \sum_l (v^*_{k,l,2q}\zeta^{*2q} + \cdots + v^*_{k,l,0})^2\right]$ (for some given terms $v^*_{k,l,0}, \ldots, v^*_{k,l,2q}$ not containing ζ^*, where ζ^* is $\rho_{q^*}(s'_0, \ldots, s'_{2q^*})$ for some given terms $s'_0, \ldots, s'_{2q^*}$) into a proof of $\bigvee_{k'} \left[F = \sum_{l'} (w'_{k',l'})^2\right]$, where the $w'_{k',l'}$ are built up by the field operations from F, $s'_0, \ldots, s'_{2q^*}$, and the $v^*_{k,l,p}$ (and hence are free of ζ^*, and contain no new root-terms).

We may assume $F \neq 0$. Assume the kth disjunct of our hypothesis: $F = \sum_l (v_{k,l,2q}\zeta^{2q} + \cdots + v_{k,l,0})^2$. So for some $N \leq 2q$, we have $\left[\bigvee_l (v_{k,l,N} \neq 0)\right] \wedge \bigwedge_l \bigwedge_{p=N+1}^{2q} (v_{k,l,p} = 0)$. Note: if $N = 0$, we are done; so assume $N > 0$. By the division algorithm (6.19), write

$$\left[\bigoplus_l \left[(v_{k,l,N}\zeta^N + \cdots + v_{k,l,0}) \otimes_\zeta (v_{k,l,N}\zeta^N + \cdots + v_{k,l,0})\right]\right] \oplus_\zeta (-F)$$
$$\equiv_\zeta \left[s \otimes_\zeta (q_z\zeta^z + \cdots + q_0)\right] \oplus_\zeta r, \qquad (6.32.1)$$

where $z = \max\{0, 2N-2q-1\}$; $\bigoplus_l$ denotes iterated use of $\oplus_\zeta$ (indexed by l); s and r are $\zeta^{2q+1} + s_{2q}\zeta^{2q} + \cdots + s_0$ and $r_{2q}\zeta^{2q} + \cdots + r_0$, respectively; and q_p and r_p are terms built up by the field operations from $s_0, \ldots, s_{2q}$, F, and the $v_{k,l,p}$. Note: the summand $-F$ is a "constant" polynomial in ζ in the left-hand side of (6.32.1); so the left-hand side has, for example as itsconstant-term, $\left(\sum_l v_{k,l,0}^2\right) - F$.

We may assume $\sum_l v_{k,l,N}^2 \neq 0$, for otherwise, choosing l_0 so that $v_{k,l_0,N} \neq 0$, we would have $-1 = \sum_{l \neq l_0} (v_{k,l_0,N}^{-1} v_{k,l,N})^2$; this, combined with the identity $F = [2^{-1}(F+1)]^2 + (-1)[2^{-1}(F-1)]^2$ (using $2 \neq 0$, above (6.3.2)), would then lead to $F = \sum_{l'} w_{l'}^2$ for suitable $w_{l'}$.[37]

Case 1. $r \equiv_\zeta 0$. Then $q_z\zeta^z + \cdots + q_0 \not\equiv_\zeta 0$, since $N > 0$. So $z \neq 0$, since otherwise the left- and right-hand sides of (6.32.1) would have even and odd algebraic degrees, respectively; so $z = 2N - 2q - 1$ (odd!). Furthermore, $q_z \neq 0$. If $z \geq 3$, write ζ' for $\rho_{N-q-1}(q_z^{-1}q_0, \ldots, q_z^{-1}q_{z-1})$. Replace ζ by ζ' in (6.32.1), drop r, and then replace $\oplus_\zeta$, $\otimes_\zeta$, and $\equiv_\zeta$ by $\cdot$, $+$, and $=$, respectively, obtaining $\left[\sum_l (v_{k,l,N}{\zeta'}^N + \cdots + v_{k,l,0})^2\right] - F = s'q'$, where s' and q' are ${\zeta'}^{2q+1} + s_{2q}{\zeta'}^{2q} + \cdots + s_0$ and $q_z{\zeta'}^z + \cdots + q_0$, respectively. But $q' = 0$, by $(6.4.2)_{N-q-1}$, making the right-hand side vanish. Then we find within the left-hand side a representation of F as a sum-of-squares of terms involving ζ'. It is possible that the $v_{k,l,p}$ contain ζ', possibly even in subterms of the form t^{-1}; if so, reorganize the terms $v_{k,l,N}{\zeta'}^N + \cdots + v_{k,l,0}$ as polynomials in ζ' whose coefficients are free of ζ', in the same way that we once did for ζ. Since $N - q - 1 \leq 2q - q - 1 = q - 1$, the induction hypothesis applies. (If $e = 1$, then the

[37] Whenever such a representation of -1 as a sum of squares arose as a disjunct in Kreisel's sketch, he just left that equation as one of the disjuncts. At the very end (6.50), he could have restored the reality axioms $(3.1.2)_s$, and then dropped those "bad" disjuncts. Thus, he never had to add the axiom $2 \neq 0$. In our alternative exposition (above), we transform those bad disjuncts into good ones as soon as they arise, with the help of $2 \neq 0$; again, at the very end, we could restore reality, and drop $2 \neq 0$ (which follows from reality).

argument is easier: just take ζ' to be $-q_1^{-1}q_0$. Then there is no need for an induction hypothesis.)[38]

Case 2. $r \not\equiv_\zeta 0$. Nevertheless, $r = 0$, as we see by replacing $\oplus_\zeta$, $\otimes_\zeta$, and $\equiv_\zeta$ in (6.32.1) by $+$, $\cdot$, and $=$, respectively (as in Case 1 above): the left-hand side vanishes by hypothesis, and $s = 0$ by $(6.4.2)_q$. Next, apply the Euclidean algorithm (6.20) to s and r:

$$\bigvee_e [s \equiv_\zeta d_e \otimes_\zeta s'_e \quad \wedge \quad r \equiv_\zeta d_e \otimes_\zeta r'_e \quad \wedge \quad (s \otimes_\zeta a_e) \oplus_\zeta (r \otimes_\zeta b_e) \equiv_\zeta d_e];$$

here, $a_e, b_e, d_e, s'_e, r'_e$ are finitely many quintuples of polynomial terms in ζ, each of whose ζ-coefficients is free of ζ, since each is built up from the s_p, r_p by the field operations. Assume the eth disjunct:

$$s \equiv_\zeta d_e \otimes_\zeta s'_e \quad \wedge \quad r \equiv_\zeta d_e \otimes_\zeta r'_e \quad \wedge \quad (s \otimes_\zeta a_e) \oplus_\zeta (r \otimes_\zeta b_e) \equiv_\zeta d_e.$$

The last of these three equations implies $d_e = 0$; and the first implies $d_e \not\equiv_\zeta 0$; therefore, writing d_e as $d_{e,a_e}\zeta^{a_e} + \cdots + d_{e,0}$, we see that $D > 0$, where D is the algebraic degree of d_e, i.e., the integer such that $d_{e,D} \neq 0 \wedge \bigwedge_{p=D+1}^{a_e}(d_{e,p} = 0)$ (6.18). Also, $D \leq 2q$, by the second equation. Similarly, write s'_e as $s'_{e,c_e}\zeta^{c_e} + \cdots + s'_{e,0}$; since $s'_e \not\equiv_\zeta 0$ (by the first equation again), the algebraic degree c of s'_e is well-defined: $s'_{e,c} \neq 0 \wedge \bigwedge_{p=c+1}^{c_e}(s'_{e,p} = 0)$. Then $c + D = 2q + 1$, by the first equation, whence $1 \leq c \leq 2q$.

Subcase 2a. c is odd, say $2q^* + 1$ (so that $q^* < q$). If $c \geq 3$, let ζ^* be $\rho_{q^*}({s'_{e,c}}^{-1}s'_{e,0}, \ldots, {s'_{e,c}}^{-1}s'_{e,c-1})$; if $c = 1$, let ζ^* be $-{s'_{e,1}}^{-1}s'_{e,0}$. For any term t or formula B, write t^* or B^* for the result of replacing ζ by ζ^* throughout t or B, respectively. Replace each formula B by B^* throughout the given proof of $\bigvee_k \left[F = \sum_l (v_{k,l,2q}\zeta^{2q} + \cdots + v_{k,l,0})^2\right]$. Our replacement will transform each application of modus ponens: $\dfrac{B \quad B \to C}{C}$ into another instance of modus ponens, since $(B \to C)^*$ is $B^* \to C^*$. The tautologies become new tautologies. Instances of the field axioms become new instances of axioms of the same type. Since

[38]The argument in case 1 is a "descent" that is common in both the theory of Diophantine equations and the algebraic theory of quadratic forms. For the former, cf., e.g., Mordell's [1969] exposition (pp. 74–5) of a result of Heegner [1952]. For the latter, cf., e.g., Lam [1973] and Prestel [1975–84], who attributed it to Springer [1952]. But Springer attributed it to Artin and Schreier [1926]; they, in turn, did not attempt to attribute it to anyone. But it goes back at least to the original, [1899] edition of Hilbert's *Grundlagen der Geometrie*, where he used it to solve the case $n = 1$ of—what he would in [1900] call—his 17th problem. (The later, post-Artin, editions omitted it, since Artin's [1926] solution to the 17th problem superseded the result Hilbert was proving.)

ζ does not occur in (6.2.1), $2 \neq 0$, or even P_p, nor in the temporary axioms such as A_e that were added in the present (sub)case (6.26), these are all unaffected by our replacement. Instances of the axiom (6.4.1) (for ρ) also lack ζ, since ζ has maximal depth. For the same reason, the only instance of $(6.4.2)_q$ (for ρ_q) that contains ζ is $s = 0$, which becomes $s^* = 0$. While this is no longer an instance of $(6.4.2)_q$ (or even of $(6.4.2)_{q^*}$), it can be derived, as follows. From $s \equiv_\zeta d_e \otimes_\zeta s'_e$ above, we derive $s^* = d_e^*(s'_e)^*$, whose right-hand side vanishes since $(s'_e)^*$ does (by $(6.4.2)_{q^*}$ if $c \geq 3$, and by (6.4.3) and the other field axioms if $c = 1$). Instances of the special equality axioms (6.11) are transformed into new instances of axioms of the same type, with the possible exception of $(6.31.1\text{–}2)_j$, for various integers $j \leq 2q$ and various terms t other than s_j; such an instance of $(6.31.1)_j$, for example, would become $s_j = t \to \zeta^* = \rho_q(s_0, \ldots, t, \ldots, s_{2q})$, which is not an axiom for most j and t. But part of our hypothesis was that such instances of $(6.31.1\text{–}2)_j$ do not occur as initial formulae in our given proof. Finally, the end-formula becomes $\bigvee_k \left[F = \sum_l (v_{k,l,2q}(\zeta^*)^{2q} + \cdots + v_{k,l,0})^2\right]$, since F and the $v_{k,l,p}$ do not contain ζ. It is possible that the $v_{k,l,p}$ contain ζ^*, possibly in subterms of the form t^{-1}; if so, reorganize the terms $v_{k,l,2q}(\zeta^*)^{2q} + \cdots + v_{k,l,0}$ as polynomials in ζ^* whose coefficients are free of ζ^*, in the same way that we once did for ζ. If $c = 1$, we're done; if $c \geq 3$, apply the inductive hypothesis.[39]

[39] Kreisel's [1960a] exposition of (6.32) (his p. 314) was too informal here:

> "If ξ is an indeterminate, $p(\xi) = \xi^{2N+1} + a_1\xi^{2N} + \cdots + a_{2N+1}$, $p(\alpha) = 0$, $f = \sum q_m(\alpha)^2$, q_m polynomial [*sic*] with coefficients in [a given field] K, then either f [or -1 is a sum of squares in K]."

In 1993, Prestel pointed out a counterexample (recall (1.6)): take $N = 1$, $K = \mathbf{Q}$, $p(\xi) = \xi^3 + \xi$ $(= \xi(\xi^2 + 1))$, $\alpha = \sqrt{-1}$, $f = -1$, and $q_1(\xi) = \xi$; then $f = \alpha^2$, but neither f (nor, of course, -1) is a sum of squares in K.

As if Kreisel had anticipated some such problem with his [1960a] exposition, he wrote in Kreisel, Macintyre [1982, p. 238] that the symbol $\rho_q(s_0, \ldots, s_{2q})$ is to be interpreted as *any* root z of $z^{2q+1} + s_{2q}z^{2q} + \cdots + s_0$ (recall footnote 6). This clarification has the effect of stengthening the hypothesis of Kreisel's lemma enough to exclude Prestel's example: for while $f = q_1(\alpha)^2$ holds with $\alpha = \sqrt{-1}$, it fails with $\alpha = 0$ (the other root of $\xi(\xi^2 + 1)$).

Our exposition of (6.32) above also avoids Prestel's counterexample, but not by Kreisel's "algebraic" hypothesis that $f = \sum q_m(\rho_q(s_0, \ldots, s_{2q}))^2$ hold for all possible interpretations of $\rho_q(s_0, \ldots, s_{2q})$; rather, we have appealed to the "syntactic" hypothesis that this equation be *provable*. These two hypotheses are equivalent, by the completeness theorem: for every formula A, *if A is valid, then it is provable,* where *valid* means that A holds in all models of the axioms, and under all possible interpretations of its (free) variables and its other symbols (such as ρ and ρ_q). The completeness theorem is, in general, not constructive. But it does have constructive meaning at least when applied to the earlier system $\mathbf{RCF}_\geq$, which is complete in Tarski's sense: for every *closed* formula B (in particular, for the universal closure of A), either $\mathbf{RCF}_\geq \vdash B$ or $\mathbf{RCF}_\geq \vdash \neg B$ —and we can decide which by calculating

Subcase 2b. c is even. Then D is odd (say, $D = 2q' + 1$), since $c + D = 2q + 1$. Recall, moreover, that $D \leq 2q$. This time let ζ' be $\rho_{q'}(d_{e,D}^{-1}d_{e,0}, \ldots, d_{e,D}^{-1}d_{e,D-1})$ if $D \geq 3$, and $-d_{e,1}^{-1}d_{e,0}$ if $D = 1$. Now argue as in subcase 2a.

Finally, glue together the proofs obtained for all k and e. $\square$

As promised in (6.28–9), we now show how to eliminate our maximal-depth root-term ζ from any proof in $\mathbf{RCF}_{ndi}^{\mathrm{qf}}$.

Theorem 6.33 Cleansing an entire proof, of $\boldsymbol{\zeta}$. *We can transform any proof in $\mathbf{RCF}_{ndi}^{\mathrm{qf}} + (s = 0)$ into a new proof in $\mathbf{RCF}_{ndi}^{\mathrm{qf}}$, in which the maximal-depth root-term ζ does not occur anywhere. Here, s is a monic, non-constant polynomial in ζ (= our maximal-depth root-term). In particular, the new proof does not use $s = 0$ as an initial formula. On the other hand, it may need to use the extra hypothesis that the characteristic of the field is either 0 or $> p$, where p can be computed primitive-recursively from the given proof; this hypothesis may be expressed by the formula P_p:*

$$(1 + 1 \neq 0) \ \wedge \ (1 + 1 + 1 \neq 0) \ \wedge \ \cdots \ \wedge \ (1 + \cdots + 1 \neq 0),$$

where the number of 1's in the last inequation is p.[40] *This elimination procedure will cause only minor alterations to the end-formula (to be described at the end of the proof).*

Proof. Part 1. In case ζ is of the form $\rho_q(s_0, \ldots, s_{2q})$, we may, by (6.31), transform the given proof into a new proof in which no special

---

$WU(B)$, where WU is Tarski's function ((4.6) above), mapping B to one of the sentences $0 = 0$ or $0 = 1$; then (4.1)(2) constructs a proof in $\mathbf{RCF}_{\geq}$ of $B \leftrightarrow WU(B)$; we conclude that if B is valid, then $WU(B)$ is $0 = 0$, whence $\mathbf{RCF}_{\geq} \vdash B$. However, while $\mathbf{RCF}_{\geq}$ is complete in Tarski's sense, $\mathbf{RCF}_{ndi}^{\mathrm{qf}}$ (the system with which we are now dealing) is not: if B is $\rho(1)^2 = 1$ (closed!), then neither B nor $\neg B$ is valid (or provable); what *is* valid (and provable) is $\rho(1)^2 = 1 \vee \rho(1)^2 = -1$. Thus a little extra work would seem to be necessary in order to get a constructive strengthening of the completeness theorem for $\mathbf{RCF}_{ndi}^{\mathrm{qf}}$, *if that is desired.* (I have not tried; perhaps one could construct a suitable translation of formulae of $\mathbf{RCF}_{ndi}^{\mathrm{qf}}$ into formulae of $\mathbf{RCF}_{\geq}$.) On the other hand, for the purpose of unwinding, recall that

> "it is immaterial whether the proof establishing the *correctness* of the unwinding procedure is or is not elementary (provided the proof is sound)" ([1981], quoted in (1.2) above).

[40] Actually, all we need to assume is that the ground field have *cardinality* $> p$; thus instead of introducing the axiom P_p, we could introduce new constant symbols $a_0, \ldots, a_p$ and the new axiom $\wedge_{j<k}(a_j \neq a_k)$. Then our new end-formula would, in general, contain some of the a_j's.

equality axiom for ρ_q $(6.11.1)_q$ that contains ζ (6.31.1–2) occurs as an initial formula; this changes the end-formula C into $\bigvee_k C_k$, where the C_k are obtained from C by replacing ζ by certain other root-terms ζ_k in the original proof. And in case ζ is of the form $\rho(s_0)$, a similar argument shows that we may assume that no special equality axiom for ρ (6.11) that contains ζ (e.g., $a = s_0 \;\rightarrow\; \rho(a) = \rho(s_0)$, for some term a free of root-terms of depth $\geq M$) occurs as an initial formula in the proof.

Part 2. As in the proof of (6.26), we call s the "original defining polynomial of ζ"; presumably, s is either $\zeta^{2q+1} + s_{2q}\zeta^{2q} + \cdots + s_0$, or $\zeta^2 - s_0$, or $\zeta^2 + s_0$. Transform the proof, by the method in (6.26), into a sequence of provable formulae, all of whose atomic subformulae are of the form $u_b\zeta^b + \cdots + u_0 = 0$, where ζ does not occur in the u_p. Actually, (6.26) produces *many* sequences of provable formulae, one for each case A_e introduced there. Specifically, recall that from each formula B of the original proof, we obtained a new formula B_e by first replacing each atomic subformula $a = b$ with $P_\zeta(a - b) = 0$, and then, in case e, successively replacing any innermost reciprocals containing ζ that occur in—the coefficients of—$P_\zeta(a - b)$, by various polynomial terms (possibly just 0) whose ζ-coefficients are free of ζ, and then re-organizing into polynomials by P_ζ. As in footnote 34, we denote the resulting atomic subformula by $Q_e(P_\zeta(a - b)) = 0$. Recall that in this process, we constructed, in case e, a new monic, non-constant polynomial (which we shall call s'_e), which we took as "the current defining polynomial of ζ"; from A_e and the field axioms we proved that s'_e is a "factor" of s, and, moreover, that $s = 0 \leftrightarrow s'_e = 0$.

Fix e.

We shall need a few facts about Q_e. First, while we can, in case e, prove $t = P_\zeta(t) = Q_e(P_\zeta(t))$ for any term t (using $s = 0$), one cannot, for each t, prove $P_\zeta(t) \equiv_\zeta Q_e(P_\zeta(t))$. For example, suppose s is $\zeta^2 - 1$ and t is ζ^{-1}; then $Q_e(\zeta^{-1})$ is ζ (at least if e refers to the most interesting case), and yet $\zeta^{-1} \not\equiv_\zeta \zeta$ (since these polynomials have degrees 0 and 1, respectively).

On the other hand, for any polynomials $r_a\zeta^a + \cdots + r_0$ and $t_b\zeta^b + \cdots + t_0$ (written as r and t, respectively), we can prove

$$Q_e(r \oplus_\zeta t) \equiv_\zeta Q_e(r) \oplus_\zeta Q_e(t) \quad \text{and} \quad Q_e(r \otimes_\zeta t) \equiv_\zeta Q_e(r) \otimes_\zeta Q_e(t). \tag{6.33.1}$$

Indeed, let v^{-1} be the *first* innermost reciprocal containing ζ that Q_e replaces with a polynomial term whose coefficients are ζ-free. For any term u, write u^* for the result of carrying out this one replacement in u. Note: $(u + w)^*$ is $u^* + w^*$, and $(uw)^*$ is u^*w^*. At this first stage in the construction of $Q_e(r)$ and $Q_e(t)$, we have $r_a^*\zeta^a + \cdots + r_0^*$ and $t_b^*\zeta^b + \cdots + t_0^*$ (ζ^* is ζ, since ζ does not contain v^{-1}; on the contrary, v contains ζ). Then we apply P_ζ: $P_\zeta(r_a^*\zeta^a + \cdots + r_0^*)$ and $P_\zeta(t_b^*\zeta^b + \cdots + t_0^*)$. Recalling the recursive definition of P_ζ, we see that at this early stage of the construction of $Q_e(r) \oplus_\zeta Q_e(t)$, we have

$$[(P_\zeta(r_a^*)\otimes_\zeta P_\zeta(\zeta^a))\oplus_\zeta\cdots\oplus_\zeta P_\zeta(r_0^*)]\oplus_\zeta[(P_\zeta(t_b^*)\otimes_\zeta P_\zeta(\zeta^a))\oplus_\zeta\cdots\oplus_\zeta P_\zeta(t_0^*)]. \tag{6.33.2}$$

Meanwhile, at this stage of the construction of $Q_e(r \oplus_\zeta t)$, we have

$$[(P_\zeta(r_a^*) \oplus_\zeta P_\zeta(t_a^*)) \otimes_\zeta P_\zeta(\zeta^a)] \oplus_\zeta \cdots \oplus_\zeta (P_\zeta(r_0^*) \oplus_\zeta P_\zeta(t_0^*)). \tag{6.33.3}$$

(In (6.33.3) we have assumed $a = b$, for simplicity.) By the ring axioms we can prove that (6.33.2) and (6.33.3) have equal coefficients. This situation continues through all the other replacements made during the execution of Q_e. The proof of $Q_e(r \otimes_\zeta t) \equiv_\zeta Q_e(r) \otimes_\zeta Q_e(t)$ is similar, recalling the definition of $r \otimes_\zeta t$ (6.18), and using the fact that $(r_k t_l)^*$ is $r_k^* t_l^*$.

Part 3. Before we can eliminate ζ, we must analyze those initial formulae in the original proof that are of the forms $t = 0 \rightarrow tr = 0$ or $t = 0 \rightarrow t^{-1} = 0$ (i.e., variations on the special equality axioms for $\cdot$ and $^{-1}$ introduced in (6.12); here r and t are terms). They become, first, $Q_e(P_\zeta(t)) = 0 \rightarrow Q_e(P_\zeta(tr)) = 0$ and $Q_e(P_\zeta(t)) = 0 \rightarrow Q_e(P_\zeta(t^{-1})) = 0$. In case t does not contain ζ (i.e., t is a constant polynomial), we shall not need to analyze these formulae further; otherwise, we do.

We shall continue our analysis of these special equality axioms by considering various cases (indexed by e'), in each of which we shall construct a certain (monic, non-constant) factor $s_{e,e'}$ of s'_e, which we shall take as the new current defining polynomial of ζ. We shall write u_e and $u_{e,e'}$ for the degrees of s'_e and $s_{e,e'}$, respectively. (Our analysis will imitate the proof of (6.26), where we analyzed reciprocals t^{-1} according to the cases $t = 0$ and $t \neq 0$.)

Apply the Euclidean algorithm (6.20) to t and s'_e, where s'_e is the current defining polynomial of ζ. As in the proof of (6.26), we obtain cases $A_{e,e'}$ involving the coefficients of t and s'_e; in each case, we obtain a ζ-GCD $d_{e,e'}$ of t and s'_e. We then obtain subcases corresponding to the degree p of $d_{e,e'}$. As in (6.26), when $p = 0$, we prove $t \neq 0$, and leave the current defining polynomial of ζ unchanged. In case $p = u_e$, we prove $t = 0$, and again leave the current defining polynomial of ζ unchanged. Finally, in case $0 < p < u_e$, we consider two possibilities, $t = 0$ and $t \neq 0$, which we prove equivalent to $d_{e,e'} = 0$ and $d_{e,e'} \neq 0$, respectively. In case $t = 0$, we take $d_{e,e'}$ as the current defining polynomial of ζ (actually, we take the *monic* constant-multiple of $d_{e,e'}$). In case $t \neq 0$, we take (the monic constant-multiple of) $s'_{e,e'}$ as the current defining polynomial of ζ, where $s'_{e,e'}$ occurs in the identity $s'_e \equiv_\zeta d_{e,e'} \otimes_\zeta s'_{e,e'}$ given by Euclid; then we repeat the above, applying the Euclidean algorithm to the pair $t, s'_{e,e'}$ instead of t, s'_e; we iterate this until $p = 0$, whereupon, writing s' for the most current defining polynomial, we find that $s' = 0 \leftrightarrow [t \neq 0 \wedge s = 0]$ is provable from the field axioms in this (sub)case.

If the original proof contains other initial formulae $t' = 0 \to t'r' = 0$ or $t' = 0 \to (t')^{-1} = 0$ of the same form, and if ζ occurs in t', then, not forgetting which case we are in at the moment, and what the current defining polynomial s' for ζ is, we apply the above procedure to t' and s', leading to yet more cases, in each of which we obtain yet another current defining polynomial s'' that is a certain (monic, non-constant) factor of s'. Eventually, all initial formulae of the forms $t = 0 \to tr = 0$ or $t = 0 \to t^{-1} = 0$ are analyzed in this way.

For future reference, let w be the number of terms t such that $t = 0 \to tr = 0$ or $t = 0 \to t^{-1} = 0$ occurs in the above analysis (for this count, we survey *all* cases e, e') .

Part 4. After this preparation, we are now ready for the **key step**: still not forgetting which case we are in, we replace each atomic subformula $u = 0$, in our current sequence of provable formulae, with $\mathrm{Rem}_{s_{e,e'}}(u) = 0$, where $\mathrm{Rem}_v(u)$, for any polynomials u and v, is the remainder upon dividing u by v: thus $u \equiv_\zeta (v \otimes_\zeta q) \oplus_\zeta \mathrm{Rem}_v(u)$, and the formal degree of $\mathrm{Rem}_v(u)$ is less than that of v (6.19). *We claim that the resulting formulae are provable without the use of $s = 0$,* but possibly with the use of the following new axiom (which we shall denote by $P_{(u-1)w}$)[40]: $(1+1 \neq 0) \wedge (1+1+1 \neq 0) \wedge \cdots \wedge (1+\cdots+1 \neq 0)$; here, the number of 1's in the last inequation is $(u-1)w$, where u is the degree of the original defining polynomial s of ζ, and where w is as in the previous paragraph.

To see this, we use induction on the original proof.

First, each application $\dfrac{B_1 \quad B_1 \to B_2}{B_2}$ of modus ponens in the original proof becomes another instance of modus ponens, since we have changed only the atomic subformulae of B_1 and B_2, and in a uniform way (throughout the present case).

Second, for the same reason, the tautologies become new tautologies.

Third, we examine the ring axioms. For example, $r + t = t + r$ becomes, first, $P_\zeta((r + t) - (t + r)) = 0$. The left-hand side of this is $[P_\zeta(r) \oplus_\zeta P_\zeta(t)] \oplus_\zeta [(-1) \otimes_\zeta (P_\zeta(t) \oplus_\zeta P_\zeta(r))]$. Applying Q_e, we get, by

(6.33.1),

$$[Q_e(P_\zeta(r)) \oplus_\zeta Q_e(P_\zeta(t))] \oplus_\zeta [(-1) \otimes_\zeta (Q_e(P_\zeta(t)) \oplus_\zeta Q_e(P_\zeta(r)))];$$

denoting this by u, we can then prove $u \equiv_\zeta 0$ using the ring axioms. Therefore $\mathrm{Rem}_{s_{e,e'}}(u) \equiv_\zeta 0$, by the field axioms (in particular, not using $s = 0$). By the ring axioms, $\mathrm{Rem}_{s_{e,e'}}(u) = 0$ (still not using $s = 0$). This is what we claimed to be able to prove, as far as $r+t = t+r$ is concerned. Similar arguments work for the other ring axioms: $r+(t+u) = (r+t)+u$, $r + 0 = r$, $r + (-r) = 0$, $rt = tr$, $r(tu) = (rt)u$, $1 \cdot r = r$, and $r(t + u) = rt + ru$.

Fourth, we examine the (other) field axioms, $r \neq 0 \rightarrow rr^{-1} = 1$ and $1 \neq 0$. The latter becomes $1 - 0 \neq 0$, which is, of course, provable from the field axioms.[41]

To analyze the other field axiom, $r \neq 0 \rightarrow rr^{-1} = 1$, we may suppose that r is already in polynomial form. We study $Q_e(r^{-1})$. Let v^{-1} be the first innermost reciprocal containing ζ that Q_e selects for replacement. There are three possibilities:

1. v^{-1} does not occur in r^{-1}. Then $(r^{-1})^*$ is just r^{-1} (which is $(r^*)^{-1}$). Then, at the first stage of the execution of Q_e, we get

$$P_\zeta(r) \neq 0 \rightarrow (P_\zeta(r) \otimes_\zeta (P_\zeta(r)^{-1})) \oplus_\zeta ((-1) \otimes_\zeta 1) = 0.$$

2. v^{-1} occurs in r. Then $(r^{-1})^*$ is (again) $(r^*)^{-1}$, leading to

$$P_\zeta(r^*) \neq 0 \rightarrow (P_\zeta(r^*) \otimes_\zeta (P_\zeta(r^*)^{-1})) \oplus_\zeta ((-1) \otimes_\zeta 1) = 0.$$

3. Finally, v^{-1} is precisely r^{-1}. (This possibility is inevitable; and it is the last stage in the execution of Q_e that affects this axiom.)

   3.a $r = 0$. Then r^* is r, and $(r^{-1})^*$ is 0; our field axiom is, finally, transformed into

$$\mathrm{Rem}_{s_{e,e'}}(r) \neq 0 \rightarrow \mathrm{Rem}_{s_{e,e'}}[(r \otimes_\zeta 0) \oplus_\zeta ((-1) \otimes_\zeta 1)] = 0. \tag{6.33.4}$$

   Recall that while executing Q_e, we chose the then-current defining polynomial s'_e to be a formal factor of r, in this case $(r = 0)$; i.e., we can prove $r \equiv_\zeta s'_e \otimes_\zeta q_e$ (some q_e) from A_e and the field axioms; and in the current subcase e', we

[41] Here we see the need for arranging for $s_{e,e'}$ to be nonconstant: if $s_{e,e'}$ were constant, then $\mathrm{Rem}_{s_{e,e'}}(1)$ would be 0, and the field axiom $1 \neq 0$ would be transformed into $0 \neq 0$, which is unprovable in some cases A_e (namely, in the consistent cases).

arranged for $s_{e,e'}$, in turn, to be a factor of s'_e, and hence of r. Thus $\mathrm{Rem}_{s_{e,e'}}(r) \equiv_\zeta 0$, from which (6.33.4) follows via the ring axioms.

3.b $r \neq 0$. Then Q_e replaces r^{-1} with u_e, where u_e is a polynomial with ζ-free coefficients, and for which we can prove, using A_e and the field axioms, $(r \otimes_\zeta u_e) \oplus_\zeta (s'_e \otimes_\zeta q_e) \equiv_\zeta 1$, for some polynomial q_e. Our field axiom becomes, first, $r \neq 0 \;\rightarrow\; (r \otimes_\zeta u_e) \oplus_\zeta ((-1) \otimes_\zeta 1) = 0$, and then

$$\mathrm{Rem}_{s_{e,e'}}(r) \neq 0 \;\rightarrow\; \mathrm{Rem}_{s_{e,e'}}[(r \otimes_\zeta u_e) \oplus_\zeta ((-1) \otimes_\zeta 1)] = 0. \tag{6.33.5}$$

We have

$$\mathrm{Rem}_{s_{e,e'}}[(r \otimes_\zeta u_e) \oplus_\zeta ((-1) \otimes_\zeta 1)] \equiv_\zeta \mathrm{Rem}_{s_{e,e'}}(s'_e \otimes_\zeta q_e),$$

by the choice of u_e; and $\mathrm{Rem}_{s_{e,e'}}(s'_e \otimes_\zeta q_e) \equiv_\zeta 0$, by the field axioms. We now establish (6.33.5) by noting that, for any polynomials u and v, we can prove $u \equiv_\zeta v \rightarrow \mathrm{Rem}_{s_{e,e'}}(u) \equiv_\zeta \mathrm{Rem}_{s_{e,e'}}(v)$ from the field axioms.

Thus in both subcases of case 3, and hence in all three cases, we have established our claim as far as the axiom $r \neq 0 \;\rightarrow\; rr^{-1} = 1$.

Fifth, of the instances of axioms (6.4.1) and $(6.4.2)_q$ (for ρ and ρ_q), those that do not contain ζ are only trivially affected by P_ζ, Q_e, and division by $s_{e,e'}$, and so remain provable without $s = 0$. Since ζ has maximal depth, the only such axiom that contains ζ is the defining formula for ζ (i.e., $\zeta^{2q+1} + s_{2q}\zeta^{2q} + \cdots + s_0 = 0$ in case ζ is $\rho_q(s_0, \ldots, s_{2q})$, and $s_0 = \rho(s_0)^2 \vee -s_0 = \rho(s_0)^2$ in case ζ is $\rho(s_0)$). But we may assume that the instance of (6.4.1) or $(6.4.2)_q$ corresponding to ζ does not occur as an initial formula in the given proof, unless that instance happens to be $s = 0$. Indeed, if this instance does so occur, then we can replace s by the GCD of s and either $\zeta^{2q+1} + s_{2q}\zeta^{2q} + \cdots + s_0 = 0$, or $\rho(s_0)^2 - s_0$, or $\rho(s_0)^2 + s_0$, as appropriate. Anyway, we must also consider $s_{e,e'} = 0$, where $s_{e,e'}$ is the *current* defining polynomial for ζ. Since $s_{e,e'}$ is a formal factor of s (using $A_{e,e'}$), our axiom $s = 0$ is transformed into $0 = 0$ upon modding out by $s_{e,e'}$; this is, obviously, provable without using $s = 0$.

Sixth, ζ does not occur in the axioms (6.2.1), $2 \neq 0$ (or even any formula P_p introduced as an axiom by a prior application of (6.33)), nor in the axioms defining the present (sub)case; therefore they are only trivially affected.

Seventh, the identity axiom $a = a$ is transformed, first, into

$$Q_e(P_\zeta(a)) \oplus_\zeta ((-1) \otimes_\zeta Q_e(P_\zeta(a))) = 0;$$

apply $\mathrm{Rem}_{s_{e,e'}}$, and proceed as we did above for the ring axioms.

Eighth (finally), we have the special equality axioms (6.11). The easiest one of these is $a = b \rightarrow (a = c \rightarrow b = c)$, which we leave to the reader. Next comes $a = b \rightarrow a + c = b + c$, which becomes, first,

$$\begin{aligned} &Q_e(P_\zeta(a)) \oplus_\zeta ((-1) \otimes_\zeta Q_e(P_\zeta(b))) \equiv_\zeta 0 \rightarrow \\ &[Q_e(P_\zeta(a)) \oplus_\zeta Q_e(P_\zeta(c))] \oplus_\zeta [(-1) \otimes_\zeta (Q_e(P_\zeta(b)) \oplus_\zeta Q_e(P_\zeta(c)))] \equiv_\zeta 0. \end{aligned} \tag{6.33.6}$$

By the ring axioms,

$$\begin{aligned} &[Q_e(P_\zeta(a)) \oplus_\zeta Q_e(P_\zeta(c))] \oplus_\zeta [(-1) \otimes_\zeta (Q_e(P_\zeta(b)) \oplus_\zeta Q_e(P_\zeta(c)))] \\ &\qquad\qquad \equiv_\zeta Q_e(P_\zeta(a)) \oplus_\zeta ((-1) \otimes_\zeta Q_e(P_\zeta(b))). \end{aligned}$$

Thus, the application of $\mathrm{Rem}_{s_{e,e'}}$ to (6.33.6) will yield a formula provable from the field axioms.

Recall that we arranged in Part 1 for no ζ-containing instances of the special equality axioms for ρ or ρ_q to occur as initial formulae in our proof. Those instances in which ζ does not occur are only trivially affected by our transformation.

The last special equality axioms to consider are $t = 0 \rightarrow tr = 0$ and $t = 0 \rightarrow t^{-1} = 0$ (6.12), which become, first,

$$Q_e(P_\zeta(t)) = 0 \rightarrow Q_e(P_\zeta(t)) \otimes_\zeta Q_e(P_\zeta(r)) = 0$$

and $Q_e(P_\zeta(t)) = 0 \rightarrow Q_e(P_\zeta(t^{-1})) = 0$, respectively.

Case 1. $\mathrm{Rem}_{s_{e,e'}}(Q_e(P_\zeta(t))) \equiv_\zeta 0$. Then $\mathrm{Rem}_{s_{e,e'}}(Q_e(P_\zeta(t)) \otimes_\zeta Q_e(P_\zeta(r))) \equiv_\zeta 0$, by field theory, which establishes the transform of the first of our two special equality axioms. To establish that of the second, observe that $Q_e(P_\zeta(t^{-1}))$ is 0 in this case.

Case 2. $\mathrm{Rem}_{s_{e,e'}}(Q_e(P_\zeta(t))) \not\equiv_\zeta 0$. In this case, our two special equality axioms are transformed into formulae that are not necessarily provable in $\mathbf{RCF}^{\mathrm{qf}}_{ndi}$ alone, and *without* the use of $s = 0$. For example, suppose s is $\zeta^2 - 1$, t is $\zeta - 2$, and r is ζ, so that our two special equality axioms are $\zeta - 2 = 0 \rightarrow \zeta^2 - 2\zeta = 0$ and $\zeta - 2 = 0 \rightarrow (\zeta - 2)^{-1} = 0$, respectively. Note: $Q_e(P_\zeta((\zeta - 2)^{-1}))$ is $-3^{-1}(\zeta + 2)$. Since Rem_s has no effect on $\zeta - 2$ and $-3^{-1}(\zeta + 2)$, and since $\mathrm{Rem}_s(\zeta^2 - 2\zeta)$ is $-2\zeta + 1$, our axioms are transformed into $\zeta - 2 = 0 \rightarrow -2\zeta + 1 = 0$ and $\zeta - 2 = 0 \rightarrow -3^{-1}(\zeta + 2) = 0$, respectively, neither of which is provable without $s = 0$ (or something like it).

To overcome such problems, write $t_1, \ldots, t_w$ for the list of all the terms t that occur in this way in the proof. Denote $\mathrm{Rem}_{s_{e,e'}}(Q_e(P_\zeta(t_l)))$ by r_l, a polynomial with ζ-free coefficients. Write $T(\zeta)$ for the formula $\bigwedge_{l=1}^{w} [r_l \not\equiv_\zeta 0 \rightarrow r_l \neq 0]$.

So far we have shown that for each formula $B(\zeta)$ (possibly containing ζ) that arises from our transformation of the given proof, $T(\zeta) \to B(\zeta)$ is provable in $\mathbf{RCF}^{\mathrm{qf}}_{ndi}$, and without using $s = 0$ as an initial formula. In these new proofs, we may replace ζ by any term w, since both the axiom $s = 0$, and the special equality axiom(s) governing ζ, no longer appear as initial formulae; the result will be proofs in $\mathbf{RCF}^{\mathrm{qf}}_{ndi}$ (without $s = 0$) of $T(w) \to B(w)$, where $T(w)$ and $B(w)$ result from $T(\zeta)$ and $B(\zeta)$, respectively, by replacing ζ by w.

To eliminate $T(\zeta)$ (or $T(w)$), we first recall that a non-zero polynomial of degree $< u$ over a field can have at most $u-1$ roots in that field; therefore, if the field contains $\geq u$ elements, then the polynomial cannot vanish at every element of the field. More generally, given w non-zero polynomials of degrees $< u$, and $(u-1)w+1$ distinct elements of the field, there must be at least one element at which all the polynomials are non-zero. One way[40] to ensure that a field has $> (u-1)w$ elements is to impose the condition that its characteristic be either 0 or $> (u-1)w$: i.e., we shall work in $\mathbf{RCF}^{\mathrm{qf}}_{ndi} + P_{(u-1)w}$, where $P_{(u-1)w}$ was defined at the beginning of Part 4.

Introduce the abbreviation $\overline{k}$ for the term $1 + \cdots + 1$, with k 1's ($k \geq 1$); we also write $\overline{0}$ for 0. Then the above paragraph is formalized by proving, from $P_{(u-1)w}$ and the field axioms: $\bigvee_{k=0}^{(u-1)w} T\big(\overline{k}\big)$. Combining this with $T\big(\overline{k}\big) \to B\big(\overline{k}\big)$ (proved above from $\mathbf{RCF}^{\mathrm{qf}}_{ndi}$ without $s = 0$), we prove $\bigvee_{k=0}^{(u-1)w} B\big(\overline{k}\big)$ from $\mathbf{RCF}^{\mathrm{qf}}_{ndi} + P_{(u-1)w}$ (without using $s = 0$).

Part 5. The end-formula $C(\zeta)$ in the given proof has undergone the following alterations during our transformation. In Part 1, it became $\bigvee_k C(t_k)$ (which we shall denote by C'), where the t_k are certain root-terms in the original proof. In Part 2, C' became $\bigvee_e C'_e$ (which we shall denote by C''), where C'_e arises from C' by replacing the atomic subformulae $a = b$ in C' by $Q_e(P_\zeta(a-b)) = 0$. In Part 4, C'' became $\bigvee_{e,e'} C'_{e,e'}$ (which we shall denote by $C'''(\zeta)$), where $C'_{e,e'}$ arises from C'_e by replacing the atomic subformulae $Q_e(P_\zeta(a-b)) = 0$ in C'_e by $\mathrm{Rem}_{s_{e,e'}}(Q_e(P_\zeta(a-b))) = 0$. Finally, in Part 5, $C'''(\zeta)$ became $\bigvee_{k=0}^{(u-1)w} C'''\big(\overline{k}\big)$.[42] □

6.34 Recapitulation. In (6.22) we set forth the overall strategy for eliminating the ρ- and ρ_q-symbols from the proof of $\bigvee_k (F = \sum_l t_{k,l}^2)$: we chose ζ to be a root-term of maximal depth among those occurring

[42] I thank Prestel for helpful conversations on the above proof, during which he pinpointed a gap in an earlier, more naive version of this algorithm.

in the given proof, and set out to eliminate it. The elimination procedure split into two cases, one in which ζ is of the form $\rho_q(s_0, \ldots, s_{2q})$, and the other in which ζ is the form $\rho(s_0)$. In (6.28) and (6.29) we gave more detail on the strategy in these two cases. Now that (6.32) (on eliminating $\rho_q(s_0, \ldots, s_{2q})$ from the end-formula) and (6.33) (on eliminating either $\rho_q(s_0, \ldots, s_{2q})$ or $\rho(s_0)$ from the entire proof) have been proved, the stategy in (6.28) for $\rho_q(s_0, \ldots, s_{2q})$ can be carried out completely. The plan in (6.29) for $\rho(s_0)$, however, has not quite been completed.

6.35 Complete elimination of $\rho(s_0)$. So far, we have done the following.

First, in (6.16) we obtained a proof of $\bigvee_k (F = \sum_l t_{k,l}^2)$ in $\mathbf{RCF}_{ndi}^{\mathrm{qf}}$, for some terms $t_{k,l}$ in the language.

Second, we selected ζ of the form $\rho(s_0)$ (of maximal depth), and in (6.25) we transformed this proof into a new sequence of provable formulae, by replacing all atomic subformulae $a = b$ by $P_\zeta(a - b) = 0$, where for any term u, $P_\zeta(u)$ is a certain polynomial in ζ such that $u = P_\zeta(u)$ is provable from the ring axioms.

Third, we considered the systems $\mathbf{RCF}_{ndi}^{\mathrm{qf}} + (s^+ = 0)$ and $\mathbf{RCF}_{ndi}^{\mathrm{qf}} + (s^- = 0)$, where s^+ and s^- (the original defining polynomials of $\rho(s_0)$ in the two systems), are taken to be $\rho(s_0)^2 - s_0$ and $\rho(s_0)^2 + s_0$, respectively.

Fourth, regarding our proof as a proof in $\mathbf{RCF}_{ndi}^{\mathrm{qf}} + (s^+ = 0)$ and in $\mathbf{RCF}_{ndi}^{\mathrm{qf}} + (s^- = 0)$ (one at a time), we transformed this proof (6.26) into proofs of $\bigvee_e \bigvee_k \big(F = \sum_l \big(t_{e,k,l}^+\big)^2\big)$ and $\bigvee_e \bigvee_k \big(F = \sum_l \big(t_{e,k,l}^-\big)^2\big)$, respectively, where the $t_{e,k,l}^{\pm}$ are polynomials in ζ *with coefficients that are free of—reciprocals containing—ζ*. (Technically, we are supposed to move all terms to the left-hand side of each equal-sign; we shall not do this for these two end-formulae, in order to improve readability.) Each e represents a certain *case*, in which we selected a "current" defining polynomial $s_e^{\pm}$ of $\rho(s_0)$; $s_e^{\pm}$ is some monic, non-constant formal factor of $s^{\pm}$.

Fifth, in (6.27) we replaced $t_{e,k,l}^{\pm}$ with $\mathrm{Rem}_{s^{\pm}}(t_{e,k,l}^{\pm})$ ("remainder") in the end-formula, obtaining $\bigvee_e \bigvee_k \big[F = \sum_l \big(t_{e,k,l,1}^{\pm}\rho(s_0) + t_{e,k,l,0}^{\pm}\big)^2\big]$, for some terms $t_{e,k,l,s}^{\pm}$ free of ζ.

Sixth, denote the term $\sum_l t_{e,k,l,0}^{\pm} t_{e,k,l,1}^{\pm}$ by $v_{e,k}^{\pm}$, and note that if $F = \sum_l \big(t_{e,k,l,1}^{\pm}\rho(s_0) + t_{e,k,l,0}^{\pm}\big)^2$, then[43] $v_{e,k}^{\pm} \neq 0 \rightarrow \rho(s_0) = r_{e,k}^{\pm}$, where

[43] Here again we use the axiom $2 \neq 0$.

$r^{\pm}_{e,k}$ is

$$(2v^{\pm}_{e,k})^{-1}\Bigl(F - \rho(s_0)^2 \sum_l \bigl(t^{\pm}_{e,k,l,1}\bigr)^2 - \sum_l \bigl(t^{\pm}_{e,k,l,0}\bigr)^2\Bigr).$$

Therefore, in $\mathbf{RCF}^{\mathrm{qf}}_{ndi} + (s^{\pm} = 0)$ we can prove

$$\bigvee_e \bigvee_k \Bigl[\Bigl[v^{\pm}_{e,k} = 0 \to F = \rho(s_0)^2 \sum_l \bigl(t^{\pm}_{e,k,l,1}\bigr)^2 + \sum_l \bigl(t^{\pm}_{e,k,l,0}\bigr)^2\Bigr] \wedge \Bigl[v^{\pm}_{e,k} \neq 0 \to F = \sum_l \bigl(t^{\pm}_{e,k,l,1} r^{\pm}_{e,k} + t^{\pm}_{e,k,l,0}\bigr)^2\Bigr]\Bigr],$$ or, more simply,

$$\bigvee_e \bigvee_k \Bigl[\Bigl(F = \rho(s_0)^2 \sum_l \bigl(t^{\pm}_{e,k,l,1}\bigr)^2 + \sum_l \bigl(t^{\pm}_{e,k,l,0}\bigr)^2\Bigr) \vee \Bigl(F = \sum_l \bigl(t^{\pm}_{e,k,l,1} r^{\pm}_{e,k} + t^{\pm}_{e,k,l,0}\bigr)^2\Bigr)\Bigr]. \tag{6.35.1$^{\pm}$}$$

Seventh, recall that (6.33) applied $\mathrm{Rem}_{s^{\pm}_{e,e'}}$ to each atomic subformula $u = 0$ in the eth disjunct of $(6.35.1)^{\pm}$, where $s^{\pm}_{e,e'}$ (for various e') are certain monic, non-constant factors of $s^{\pm}_e$. Since $s^{\pm}$ is of degree 2 in ζ, each $s^{\pm}_{e,e'}$ is either $s^{\pm}$ itself, or a *linear* polynomial of the form $\rho(s_0) + u^{\pm}_{e,e'}$, where $u^{\pm}_{e,e'}$ does not contain $\rho(s_0)$. For those e, e' with $s^{\pm}_{e,e'}$ being $s^{\pm}$, the application of $\mathrm{Rem}_{s^{\pm}_{e,e'}}$ will amount to replacing $\rho(s_0)^2$ with $\pm s_0$; this transforms $r^{\pm}_{e,k}$ into a term in which $\rho(s_0)$ does not occur. Therefore our new end-formula will contain a disjunct, for each such e, e', of the form

$$\bigvee_k \Bigl[\Bigl(F = \pm s_0 \sum_l \bigl(t^{\pm}_{e,k,l,1}\bigr)^2 + \sum_l \bigl(t^{\pm}_{e,k,l,0}\bigr)^2\Bigr) \vee \Bigl(F = \sum_l \bigl(w^{\pm}_{e,k,l}\bigr)^2\Bigr)\Bigr],$$

where the terms $w^{\pm}_{e,k,l}$ do not contain $\rho(s_0)$. For those e, e' with $s^{\pm}_{e,e'}$ being $\rho(s_0) + u^{\pm}_{e,e'}$, the application of $\mathrm{Rem}_{s^{\pm}_{e,e'}}$ will amount to replacing each occurrence of $\rho(s_0)$ with $-u^{\pm}_{e,e'}$: $\bigvee_{k'} \bigl[F = \sum_l \bigl(w^{\pm}_{e,e',k',l}\bigr)^2\bigr]$, where the $w^{\pm}_{e,e',k,l}$ are terms free of $\rho(s_0)$. Combining these formulae for all e, e', the application of Rem has transformed our end-formula into

$$\bigvee_g \Bigl(F = \pm s_0 \sum_l \bigl(s^{\pm}_{g,l,1}\bigr)^2 + \sum_l \bigl(s^{\pm}_{g,l,0}\bigr)^2\Bigr) \vee \bigvee_h \Bigl(F = \sum_l \bigl(z^{\pm}_{h,l}\bigr)^2\Bigr), \tag{6.35.2^{p}m}$$

where all terms are free of $\rho(s_0)$.

After applying Rem, the last act that (6.33) performs on the end-formula (and the entire proof) is to replace any remaining occurrences of $\rho(s_0)$ with $0,\ 1,\ 1+1, \ldots,\ (u-1)w$, where $u = 2$ in this case, and w was computed by analyzing the earlier proof; then we would have a $\rho(s_0)$-free proof in $\mathbf{RCF}^{\mathrm{qf}}_{ndi} + P_w$ of the disjunction of the resulting end-formulae (recall, P_w asserts that the characteristic of the field is either 0 or $> w$). But $(6.35.2)^{\pm}$ has no occurrences of $\rho(s_0)$. So (6.33) asserts that $(6.35.2)^{+}$ itself is provable in $\mathbf{RCF}^{\mathrm{qf}}_{ndi} + P_w$ without the use of $s^{+} = 0$ (or any occurrence of $\rho(s_0)$ at all), and that $(6.35.2)^{-}$ is provable in $\mathbf{RCF}^{\mathrm{qf}}_{ndi} + P_w$ without the use of $s^{-} = 0$. I.e., the *conjunction* of $(6.35.2)^{+}$ and $(6.35.2)^{-}$ is provable without any use of $\rho(s_0)$. Using the distributivity of $\vee$ and $\wedge$, this conjunction is of the form

$$\bigvee_g \bigvee_{g'} \left[\left(F = s_0 \sum_l \left(s^{+}_{g,l,1}\right)^2 + \sum_l \left(s^{+}_{g,l,0}\right)^2 \right) \wedge \right. \tag{6.35.3}$$
$$\left. \left(F = -s_0 \sum_l \left(s^{-}_{g',l,1}\right)^2 + \sum_l \left(s^{-}_{g',l,0}\right)^2 \right) \right] \vee \bigvee_a \left(F = \sum_l y^2_{a,l} \right).$$

Therefore, to complete our task of proving a disjunction of sum-of-squares representations of F without the use of $\rho(s_0)$, we fix g, g', and assume

$$F = s_0 \sum_l \left(s^{+}_{g,l,1}\right)^2 + \sum_l \left(s^{+}_{g,l,0}\right)^2 \wedge F = -s_0 \sum_l \left(s^{-}_{g',l,1}\right)^2 + \sum_l \left(s^{-}_{g',l,0} gr\right)^2;$$

from these two weighted sum-of-squares representations we seek a single, unweighted sum-of-squares representation. **Case 1**. $\sum_l \left(s^{+}_{g,l,1}\right)^2 + \sum_l \left(s^{-}_{g',l,1}\right)^2 \neq 0$. Then we can prove the following equations, in turn:

$$F = \left[\left(s_0 \sum_l \left(s^{+}_{g,l,1}\right)^2 + \sum_l \left(s^{+}_{g,l,0}\right)^2 \right) \left(\sum_l \left(s^{+}_{g,l,1}\right)^2 + \sum_l \left(s^{-}_{g',l,1}\right)^2 \right) \right]$$
$$\left(\sum_l \left(s^{+}_{g,l,1}\right)^2 + \sum_l \left(s^{-}_{g',l,1}\right)^2 \right)^{-1},$$

$$F = \left[s_0 \sum_l \left(s^{+}_{g,l,1}\right)^2 \sum_l \left(s^{+}_{g,l,1}\right)^2 + s_0 \sum_l \left(s^{+}_{g,l,1}\right)^2 \sum_l \left(s^{-}_{g',l,1}\right)^2 + \right.$$
$$\left. \sum_l \left(s^{+}_{g,l,0}\right)^2 \sum_l \left(s^{+}_{g,l,1}\right)^2 + \sum_l \left(s^{+}_{g,l,0}\right)^2 \sum_l \left(s^{-}_{g',l,1}\right)^2 \right]$$
$$\left(\sum_l \left(s^{+}_{g,l,1}\right)^2 + \sum_l \left(s^{-}_{g',l,1}\right)^2 \right)^{-1},$$

$$F=\Big[\sum_l\big(s^+_{g,l,1}\big)^2\Big(s_0\sum_l l(s^+_{g,l,1})^2+s_0\sum_l\big(s^-_{g',l,1}\big)^2+\sum_l\big(s^+_{g,l,0}\big)^2\Big)+$$
$$\sum_l\big(s^+_{g,l,0}\big)^2\sum_l\big(s^-_{g',l,1}\big)^2\Big]\Big(\sum_l\big(s^+_{g,l,1}\big)^2+\sum_l\big(s^-_{g',l,1}\big)^2\Big)^{-1},$$
$$F=\Big[\sum_l\big(s^+_{g,l,1}\big)^2\Big(s_0\sum_l\big(s^+_{g,l,1}\big)^2+\Big(\sum_l\big(s^-_{g',l,0}\big)^2-F\Big)+\sum_l\big(s^+_{g,l,0}\big)^2\Big)$$
$$+\sum_l\big(s^+_{g,l,0}\big)^2\sum_l\big(s^-_{g',l,1}\big)^2\Big]\Big(\sum_l\big(s^+_{g,l,1}\big)^2+\sum_l\big(s^-_{g',l,1}\big)^2\Big)^{-1},\text{ and}$$
$$F=\Big[\sum_l\big(s^+_{g,l,1}\big)^2\Big(\sum_l\big(s^-_{g',l,0}\big)^2\Big)+\sum_l\big(s^+_{g',l,0}\big)^2\sum_l\big(s^-_{g',l,1}\big)^2\Big]$$
$$\Big(\sum_l\big(s^+_{g,l,1}\big)^2+\sum_l\big(s^-_{g',l,1}\big)^2\Big)^{-1};$$

the last expression is easily transformed into the required form.

Case 2. $\sum_l\big(s^+_{g,l,1}\big)^2+\sum_l\big(s^-_{g',l,1}\big)^2=0$. The desired conclusion is already at hand if $\sum_l\big(s^+_{g,l,1}\big)^2=0$; otherwise, in this case we have $-1=\big(\sum_l\big(s^+_{g,l,1}\big)^2\big)^{-1}\sum_l\big(s^-_{g',l,1}\big)^2$, whence $-1=\sum_l w_l^2$ for terms w_l built up by the field operations from the $s^+_{g,l,1}$, $s^-_{g',l,1}$. This, combined with the identity $F=[2^{-1}(F+1)]^2+(-1)[2^{-1}(F-1)]^2$ (using $2\neq 0$, above (6.3.2)), leads to the desired sum-of-squares expression.[37]

As promised in (6.17), we have now proved

Theorem 6.36 Elimination of all ρ- and ρ_q-symbols from an entire proof. *From a proof in $\mathbf{RCF}^{\mathrm{qf}}_{ndi}$ of $\bigvee_k(F=\sum_l t^2_{k,l})$, for some terms $t_{k,l}$ in the language of $\mathbf{RCF}^{\mathrm{qf}}_{ndi}$, we can construct a (still quantifier-free) proof of $\bigvee_{k'}(F=\sum_{l'}{t'_{k',l'}}^2)$, for finitely many terms $t'_{k',l'}$ built up by the field operations from the variables $x_1,\dots,x_n$ and the constants $0,1,c_1,\dots,c_m,a_{i,1},\dots,a_{i,J_i},b_{i,1},\dots,b_{i,J'_i}$, for our (still fixed!) n,d,i; the new proof uses as initial formulae only the field axioms, $0^{-1}=0$, P_p (some $p\geq 2$), and (6.2.1).*

Proposition 6.37 Eliminating a_{ij} and b_{ij}. *Suppose we have, for fixed i, a quantifier-free proof, using as initial formulae only the logical axioms, the field axioms, $0^{-1}=0$, P_p, and (6.2.1), and with end-formula $\bigvee_k(F=\sum_l t^2_{k,l})$, for some terms $t_{k,l}$ built up by the field operations from the variables $x_1,\dots,x_n$ and the constants $0,1,c_1,\dots,c_m$, $a_{i,1},\dots,a_{i,J_i},b_{i,1},\dots,b_{i,J'_i}$. Then we can construct a quantifier-free proof not containing the constants a_{ij} and b_{ij} (in particular not con-*

taining the axiom (6.2.1)), whose initial formulae are otherwise as above, together with $\bigwedge_j (q_{ij} \neq 0)$ *(cf. (6.1.1)); the new end-formula is* $\bigvee_{k'}[F = \sum_{l'} w_{k',l'}(t'_{k',l'})^2]$*, where the* $w_{k',l'}$ *are products of the* p_{ij} *and* q_{ij}^{-1}*.*

Proof. First prove $\bigvee_k (F = \sum_l 1 \cdot t_{k,l}^2)$. Next, let ζ be any one of the symbols a_{ij} or b_{ij}; denote ζ by a_{i,j_0} or b_{i,j_0}, respectively. It suffices to show how to convert a proof of $\bigvee_k (F = \sum_l w_{k,l} t_{k,l}^2)$, where the $w_{k,l}$ are arbitrary terms built up by the field operations from the constants $0, 1, c_1, \ldots, c_m$, into a ζ-free proof of $\bigvee_{k'}[F = \sum_{l'} w'_{k',l'}(t'_{k',l'})^2]$, where the $w'_{k',l'}$ are products of the $w_{k,l}$ and either p_{i,j_0} or q_{i,j_0}^{-1} (according as ζ is a_{i,j_0} or b_{i,j_0}).

Our main tool is—the method of—(6.33). In previous applications of (6.33), ζ was a root-term, i.e., either $\rho(s_0)$ or $\rho_q(s_0, \ldots, s_{2q})$; ζ was "defined" by the presence of an extra axiom $s = 0$, where s was a monic, non-constant polynomial in ζ with coefficients that are free of ζ. Now, as we said, ζ is either a_{i,j_0} or b_{i,j_0}. If ζ is a_{i,j_0}, then we take s to be $a_{i,j_0}^2 - p_{i,j_0}$. Likewise, if ζ is b_{i,j_0}, we take s to be $b_{i,j_0}^2 - q_{i,j_0}^{-1}$; in this case, the conjunct $q_{i,j_0} b_{i,j_0}^2 = 1$ appearing in axiom (6.2.1) entails $s = 0 \wedge q_{i,j_0} \neq 0$, and conversely. All the essential hypotheses of (6.33) are satisfied. The conclusion is that we can construct a new proof in which ζ does not occur anywhere. The only axioms needed in the new proof are the field axioms, $0^{-1} = 0$, $\quad P_t \quad (t = \max\{p, (u-1)w\}$, where u and w are as in the proof of (6.33)) and the *rest* of (6.2.1) (i.e., the conjunction of those equations not containing ζ); but in case ζ is b_{i,j_0}, then we must add one more axiom: $q_{i,j_0} \neq 0$.

The new end-formula arises from the old one, $\bigvee_k (F = \sum_l w_{k,l} t_{k,l}^2)$, in much the same way as described in (6.35). First we apply P_ζ to the atomic subformulae, organizing the terms into polynomials in ζ. Unfortunately, some of the coefficients of these polynomials may contain reciprocals in which ζ occurs; so we introduce several cases e, in each of which we replace those reciprocals with polynomials in ζ whose coefficients are free of ζ. Now our end-formula is $\bigvee_e \bigvee_k [F = \sum_l w_{e,k,l} v_{e,k,l}^2]$, where the $v_{e,k,l}$ are polynomials in ζ whose coefficients are ζ-free, and where the $w_{e,k,l}$ are re-indexed copies of the $w_{k,l}$. Reduce the $v_{e,k,l}$ mod s (using $s = 0$); i.e., roughly, replace ζ^2 by p_{i,j_0} or q_{i,j_0}^{-1}, as appropriate. We obtain $\bigvee_e \bigvee_k [F = \sum_l w_{e,k,l}(t_{e,k,l,1}\zeta + t_{e,k,l,0})^2]$, where the $t_{e,k,l,s}$ do not contain ζ. Next, denote the term $\sum_l w_{e,k,l} t_{e,k,l,0} t_{e,k,l,1}$ by $v_{e,k}$, and note that if $F = \sum_l w_{e,k,l}(t_{e,k,l,1}\zeta + t_{e,k,l,0})^2$, then[43] $v_{e,k} \neq 0 \rightarrow \zeta = r_{e,k}$, where $r_{e,k}$ is

$$(2v_{e,k})^{-1}\Big(F - \zeta^2 \sum_l w_{e,k,l} t_{e,k,l,1}^2 - \sum_l w_{e,k,l} t_{e,k,l,0}^2\Big).$$

Therefore,

$$\bigvee_e \bigvee_k \Big[\Big(F = \zeta^2 \sum_l w_{e,k,l} t_{e,k,l,1}^2 + \sum_l w_{e,k,l} t_{e,k,l,0}^2\Big) ill \\ \vee \Big(F = \sum_l w_{e,k,l}(t_{e,k,l,1} r_{e,k} + t_{e,k,l,0})^2\Big)\Big].$$

Next, we apply Rem, i.e., we take the remainder upon dividing by s and by certain monic, non-constant factors of s. If the factor is quadratic (i.e., s itself), then Rem replaces ζ^2 with p_{i,j_0} or q_{i,j_0}^{-1}; if the factor is linear, then Rem replaces ζ with some term not containing ζ. Either way, the defining equation $s = 0$ becomes $0 = 0$, which is provable without the use of ζ; the other axioms are only trivially affected. In case ζ is a_{i,j_0}, we get

$$\bigvee_g \Big(F = p_{i,j_0} \sum_l w_{e,k,l} s_{g,l,1}^2 + \sum_l w_{e,k,l} s_{g,l,0}^2\Big) \vee \bigvee_h \Big(F = \sum_l w_{e,k,l} z_{h,l}^2\Big);$$

in case ζ is b_{i,j_0}, we get q_{i,j_0}^{-1} instead of p_{i,j_0} above, and $z'_{h,l}$ instead of $z_{h,l}$. Either way, all terms in this end-formula are free of ζ. If ζ still occurs in the earlier parts of the proof, just replace those occurrences by 0. $\square$

6.38 What is a rational function? The result of (6.37) is a proof of $\bigvee_k (F = \sum_l w_{k,l} t_{k,l}^2)$; here the $t_{k,l}$ are terms built up by the field operations from $x_1, \ldots, x_n, 0, 1, c_1, \ldots, c_m$; the $w_{k,l}$ are products of the p_{ij} and q_{ij}^{-1} (for various j); and the p_{ij} and q_{ij}, in turn, are built up by the ring operations from $0, 1, c_1, \ldots, c_m$. This is essentially what Hilbert's 17th problem asks for—a representation of F as a (suitably weighted) sum of squares of rational functions. But are terms such as the $t_{k,l}$ properly called "rational functions"? More precisely, do they correctly *formalize* the notion of rational function?

In—this part of—algebra, one usually starts with a field K and some indeterminates, such as $C_1, \ldots, C_m, X_1, \ldots, X_n$, from which one forms the ring $K[C; X]$ of polynomials. Rational functions are then defined to be *equivalence classes of ordered pairs* (f, g), where $f, g \in K[C, X]$ and $g \neq 0$, subject to the usual equivalence relation. The classes are usually denoted by f/g. Strings of symbols such as $f/((g/h) \cdot k)$ or $(f^{-1}g + 1)^{-1}$ are not included in this definition.

On the other hand, it is often convenient to consider such strings as legitimate rational functions. For example, in [1958a] Kreisel wrote (p. 166):

> "The disjunction can be contracted to $f = \sum T_i^2$ since, e.g., $(f = t_1^2 \vee f = t_2^2) \leftrightarrow (f - t_1^2)(f - t_2^2) = 0$, i.e., $f^2 - f(t_1^2 + t_1^2) + t_1^2 t_2^2 = 0$, i.e., $f = (f^2 + t_1^2 t_2^2)/(t_1^2 + t_2^2)$."

(One must also consider the possibility (mentioned in [1960a]) that $t_1^2 + t_2^2 = 0$, when $-1 = (t_1/t_2)^2$; and in this last equation, one must *also* consider the possibility (not mentioned in [1960a]) that $t_2 = 0$, etc.) In fact, this method works as well for *weighted* sums of squares. The problem with this trick is that even if t_1 and t_2 are rational functions in the sense of mainstream algebra (previous paragraph), $(f^2 + t_1^2 t_2^2)/(t_1^2 + t_2^2)$ need not be; it is usually only a *compound* fraction. When it comes time to simplify and/or evaluate such a compound fraction, one is led inevitably back to a *disjunction* of cases: either $t_1^2 + t_2^2 = 0$ or $t_1^2 + t_2^2 \neq 0$; and in the former case, we have the subcases $t_1 = 0$ or $t_1 \neq 0$; and in these subcases there can be yet further cases, depending on how many reciprocals are nested inside of t_1 and t_2.

Once we recognize that such nested reciprocals are just *disjunctions in disguise,* one may ask what this trick of "contracting" disjunctions achieves. The answer, as far as I can tell, is, "(Less than) nothing": i.e., performing this maneuver actually *sets us back,* since it introduces *new* nested reciprocals; recall the need in our purely syntactic presentation (e.g., in (6.32) and (6.35)) for successively *eliminating* (6.26) reciprocals containing certainunwanted terms.

In this exposition, we seek to simplify the compound fractions $t_{k,l}$ so that they look like mainstream rational functions, even if this means *increasing* the number of disjuncts. We begin with

Lemma 6.39 Putting terms into $\sum\prod$-form, or "sum-product form". *Given any term u built up by the field operations from $0, 1, c_1, \ldots, c_m, x_1, \ldots, x_n$, we can construct terms u_{gh}, and a proof using only the ring axioms, of $u = \sum_g \prod_h u_{gh}$. Moreover, we can arrange for each u_{gh} to be either -1, or a subterm of u not of the forms $s + s'$, $-s$, or ss' (where s and s' are subterms of u); thus each u_{gh}* is *either -1 or one of the "basic polynomial subterms" 0, 1, $c_1, \ldots, c_m$, $x_1, \ldots, x_n$, or s^{-1} (where s is a further subterm of u).*

Proof. We define the u_{gh} by recursion on terms. If u is already -1 or one of the basic polynomial subterms, we are done (or, to be pedantic, we take u_{11} to be u itself). If u is $\left(\sum_k \prod_l s_{kl}\right) + \left(\sum_{k'} \prod_{l'} s'_{k'l'}\right)$

(where the two big summands are in $\sum\prod$-form), then we are again done, by associativity of $+$ (recalling that $\sum_k$ is just an abbreviation for successive $+$ symbols with suitable parentheses); at most we must re-label the indices k, k' as suitable g. Next, if u is $-\sum_g \prod_{h=1}^{m_g} u_{gh}$, then take $u_{g,0}$ to be -1, so that $t = \sum_g \prod_{h=0}^{m_g} u_{gh}$. Finally, if u is $\left(\sum_k \prod_l s_{kl}\right)\left(\sum_{k'} \prod_{l'} s'_{k'l'}\right)$, then $u = \sum_{k,k'} \left(\prod_l s_{kl}\right)\left(\prod_{l'} s'_{k'l'}\right)$, so we can take u_{gh} to be s_{kl} or $s'_{k'l'}$, for suitable indices depending on g, h (guided by distributivity). $\square$

Proposition 6.40 Putting terms into "rational form". *Let t be a term built up by the field operations from $x_1, \ldots, x_n, 0, 1, c_1, \ldots, c_m$. Then we can construct (a) terms p_j, q_j (for finitely many j) that are built up by the ring operations from $x_1, \ldots, x_n, 0, 1, c_1, \ldots, c_m$ (in fact, the p_j, q_j may be taken to be in $\sum\prod$-form); and (b) a proof, from the field axioms and $0^{-1} = 0$, of $\bigvee_j (t = p_j q_j^{-1})$.*

Proof. Organize t into $\sum\prod$-form: $t = \sum_g \prod_h t_{g,h}$, by (6.39). Trivially, $t = \left(\sum_g \prod_h t_{l,g,h}\right) \cdot 1^{-1}$.

Thus we may suppose that we have constructed (a) terms p_j, q_j that are built up by the *field* operations from $x_1, \ldots, x_n, 0, 1, c_1, \ldots, c_m$, in $\sum\prod$-form, and (b) a proof, from the field axioms and $0^{-1} = 0$, of $\bigvee_j (t = p_j q_j^{-1})$. But there is no assurance as yet that the p_j, q_j are free of $^{-1}$.

Define the *depth* of a reciprocal term s_1^{-1} by analogy with the depth of a root-term (6.21): it is the maximum of the lengths E of all "chains of reciprocals" $s_1^{-1}, \ldots, s_E^{-1}$ beginning with s_1^{-1} (i.e., for $1 < e \leq E$, s_e^{-1} is a subterm of s_{e-1}, and hence a proper subterm of s_{e-1}^{-1}). Let M denote the maximum of the depths of all reciprocals occurring in any of the p_j, q_j (that is, we do not consider the depth of q_j^{-1}, which is 1 greater than the depth of q_j). Thus M equals the maximum of the lengths of all chains of reciprocals occurring in the p_j, q_j.

Let s^{-1} be one of those reciprocals of depth M occurring in at least one of the p_j, q_j. We shall construct (a) new terms $p'_{j'}, q'_{j'}$ in which s^{-1} does not occur, and into which no new reciprocals have been introduced, and (b) a proof, from the field axioms and $0^{-1} = 0$, of $\bigvee_{j'} (t = p'_{j'} (q'_{j'})^{-1})$. Then the *number* of reciprocals of depth M occurring in the p_j, q_j will have been reduced; as in (6.22), we can iterate this procedure until *all* reciprocals of depth M have been eliminated, thereby reducing M itself. Once M is 0, we shall be done.

We have written p_j and q_j as $\sum_g \prod_h p_{j,g,h}$ and $\sum_g \prod_h q_{j,g,h}$, respectively, where each $p_{j,g,h}$ and $q_{j,g,h}$ is either $0, 1, x_1, \ldots, x_n, c_1, \ldots, c_m$, or one of the reciprocals r^{-1} occurring in the p_j, q_j.

1. $s = 0$. Then replace each occurrence of s^{-1} in every p_j, q_j with 0; the resulting terms p'_j, q'_j still have $\bigvee_j \left(t = p'_j\left(q'_j\right)^{-1}\right)$, using $0^{-1} = 0$.

2. $s \neq 0$. Fix j, and suppose $t = p_j q_j^{-1}$.

   2a Our maximal-depth reciprocal s^{-1} does not occur in either p_j or q_j. Then take p''_j, q''_j to be p_j, q_j, respectively, and go the last paragraph of the proof.

   2b s^{-1} occurs in p_j. Then for some values g_0, h_0 of g, h, p_{j,g_0,h_0} must be s^{-1}, since s^{-1} has maximal depth. Thus we can write p_j in the form

$$s^{-1} \prod_{h \neq h_0} p_{j,g_0,h} + \sum_{g \neq g_0} \prod_h p_{j,g,h}.$$

   For any terms a, b, s, we can prove $s \neq 0 \rightarrow ab^{-1} = (sa)(sb)^{-1}$, using the field axioms and (for the case in which $b = 0$) $0^{-1} = 0$.[44] Then

$$t = \left(\prod_{h \neq h_0} p_{j,g_0,h} + \sum_{g \neq g_0} s \prod_h p_{j,g,h}\right)\left(\sum_g s \prod_h q_{j,g,h}\right)^{-1}.$$

   For those products $\prod_h p_{j,g,h}$ (for $g \neq g_0$) or $\prod_h q_{j,g,h}$ (for any g) that still contain occurrences of s^{-1}, we may cancel one such occurrence in each of those products, together with the factor s added to the left of each of those products; leave the other products alone. If s^{-1} still occurs in the "numerator" p_j, repeat subcase 2b; if s^{-1} (still) occurs in the "denominator" q_j, go to subase 2c.

[44] As we mentioned in (1.4) above, in [1990, p. 13] Lombardi wrote (after trivial changes of notation):

> "Even after this explicit treatment, Kreisel's proof still seems to us a little obscure. Notably, it is not altogether clear to know how one manages with the symbol $^{-1}$ (which represents the inverse of an element, with $0^{-1} = 0$ by convention): the reduction of a representation comprising only the function symbols $+$, $-$, $\cdot$, and $^{-1}$ to a 'rational fraction' form poses in fact a problem on account of uv^{-1} not being equal to $uw(vw)^{-1}$ unless w is not 0 or uv is 0."

I believe that my exposition in (6.38–40) above clarifies this (legitimate) question: namely, we simply consider *all cases separately*: $w = 0$, $w \neq 0$, $v = 0$, $v \neq 0$, etc., and for each case we add a new disjunct (i.e., a new sum-of-squares representation of F) to the end-formula of our proof.

2c s^{-1} occurs in q_j. Similar to subcase 2b.

Denote the terms arising, upon completion of subcases 2a–c, by p''_j, q''_j. Combining this for all j, we get $\bigvee_j \left(t = p''_j \left(q''_j\right)^{-1}\right)$.

Combining cases 1 and 2, we have $\bigvee_j \left[\left(t = p'_j\left(q'_j\right)^{-1}\right) \vee \left(t = p''_j\left(q''_j\right)^{-1}\right)\right]$, in which s^{-1} no longer occurs. $\square$

Theorem 6.41 *Using as alphabet only the symbols* $0, 1, x_1, \ldots, x_n, c_1, \ldots, c_m, +, -, \cdot, ^{-1}, =, \vee, \wedge, \neg, \rightarrow, (,)$ *(in particular, without quantifiers, the ε-symbol, or—other—root-symbols ρ or ρ_q), we can, for fixed n, d, and i, construct a proof of*

$$\bigvee_k \left[F = \sum_l w_{k,l} \left[\left(\sum_g \prod_h p_{k,l,g,h}\right)\left(\sum_g \prod_h q_{k,l,g,h}\right)^{-1}\right]^2\right], \quad (6.41.1)$$

where each $p_{k,l,g,h}$ and $q_{k,l,g,h}$ is among $0, 1, x_1, \ldots, x_n, c_1, \ldots, c_m$, and each $w_{k,l}$ is a product of the p_{ij}, q_{ij}^{-1} (for various j; the polynomials p_{ij}, q_{ij} were introduced in (6.1.1)).[45] *The only initial formulae in this proof are the tautologies and the special equality axioms (6.12) for $+$, $\cdot$, and $^{-1}$ (the "logical" axioms), and—instances of—the field axioms, $0^{-1} = 0$, P_p (asserting that the characteristic of the field is either 0 or $> p$, for some sufficiently large p),*[40] *and*[45] $\bigwedge_j (q_{ij} \neq 0)$ *(the "non-logical" axioms); the only rule of inference is modus ponens.*

Proof. In (6.37) we proved $\bigvee_k(F = \sum_l w_{k,l} t_{k,l}^2)$, for some terms $t_{k,l}$ built up by the field operations from $0, 1, c_1, \ldots, c_m, x_1, \ldots, x_n$. Apply (6.40) to each $t_{k,l}$, one at a time: we get—proofs of—disjunctions of representations of $t_{k,l}$ as "mainstream" rational functions as in (6.38). Plugging these representations into $\bigvee_k(F = \sum_l w_{k,l} t_{k,l}^2)$, we get a (bigger) disjunction, of the required type. $\square$

6.42 Equality of rational functions. There is still some mopping up to do in order to derive Henkin's (5.8) from (6.41) above. In (6.38) we noted that the informal conception of "rational function" differs from the strict conception in ordinary algebra, by allowing *nested* reciprocals; we have obtained "mainstream" rational functions (without nested reciprocals) in (6.41.1). But this is not all: when an algebraist says that two rational functions f_1/g_1 and f_2/g_2 $(f_i, g_i \in K[X]$,

[45] Recall (6.1) that the q_{ij} are actually unnecessary, by the finiteness theorem; but this was not known until the 70s. And even if the q_{ij} are kept, we can replace each occurrence of q_{ij}^{-1} in the $w_{k,l}$ with q_{ij}, upon multiplying the corresponding "denominator" $\sum_g \prod_h q_{k,l,g,h}$ by q_{ij}.

$X = (X_1, \ldots, X_n)$, $g_i \neq 0$, K a field) are "equal," he does *not* mean that $\forall x := (x_1, \ldots, x_n) \in K^n$, $(f_1/g_1)(x) = (f_2/g_2)(x)$ (although this is essentially what (6.41.1) says). Rather, he means that $f_1 g_2 = f_2 g_1$; and even this equation does not mean that $\forall x \in K^n$, $(f_1 g_2)(x) = (f_2 g_1)(x)$, but rather that *corresponding X-coefficients* ($\in K$) *of the left- and right-hand sides agree.* For example, if K is finite, say $K = \mathbf{Z}/3\mathbf{Z}$, and $h(X_1) = X_1^3 - X_1 \in K[X_1]$, then $h \neq 0$, even though $\forall x_1 \in K$, $h(x_1) = 0$. However, it is an easy theorem of field theory that if K has sufficiently many elements, then the two definitions of equality coincide: precisely, for $h_1, h_2 \in K[X]$,

$$[|K| > D \ \wedge \ (\forall x \in K^n)\,(h_1(x) = h_2(x))] \ \rightarrow \ h_1 = h_2; \tag{6.42.1}$$

here, $D = \max_{1 \le j \le n}\{\deg_{X_j} h_1, \ \deg_{X_j} h_2\}$. For the proof, note that when $n = 1$, (6.42.1) just says that a polynomial $h \in K[X_1] \setminus \{0\}$ has $\le D := \deg h$ roots in K; for $n > 1$, use induction, applying the 1-variable case to the variable X_n, and then applying the $(n-1)$-variable case to the X_n-coefficients.

To improve (6.41.1), we shall need a small extension of (6.42.1): for $h_{k,1}, h_{k,2} \in K[X]$ $(k = 1, 2, \ldots, M)$,

$$\left[|K| > MD \wedge (\forall x \in K^n) \bigvee_{k=1}^{M} (h_{k,1}(x) = h_{k,2}(x))\right] \rightarrow \bigvee_{k=1}^{M} (h_{k,1} = h_{k,2}); \tag{6.42.2}$$

this time, $D = \max\limits_{\substack{1 \le k \le M \\ 1 \le j \le n}} \{\deg_{X_j} h_{k,1}, \ \deg_{X_j} h_{k,2}\}$. For the proof, note that when $n = 1$, (6.42.2) follows from (6.42.1) by the pigeonhole principle: for some k, there are $> D$ elements $x_1 \in K$ at which $h_{k,1}(x_1) = h_{k,2}(x_1)$, whence for that k, $h_{k,1} = h_{k,2}$. For $n > 1$, use induction as in the previous paragraph.

6.43 Notation for formal polynomials in $\boldsymbol{x_1, \ldots, x_n}$. In (6.18) we introduced notation for polynomial terms in any single term ζ. Now we handle polynomial terms in $x_1, \ldots, x_n$, by which we mean terms t of the form

$$\sum_{i_n=0}^{I} \left[\sum_{i_{n-1}=0}^{I_{i_n}} \left[\cdots \left[\sum_{i_1=0}^{I_{i_2,\ldots,i_n}} t_{i_1,\ldots,i_n} x_1^{i_1} gr \right] \cdots \right] x_{n-1}^{i_{n-1}} \right] x_n^{i_n}; \tag{6.43.1}$$

here (unlike in (6.18)), the coefficient terms $t_{i_1,\ldots,i_n}$ are built up by the ring operations from $0, 1, c_1, \ldots, c_m$ alone (i.e., they do not involve the x_j). (Thus here, unlike in (6.18), we do not need to clarify the definition of "polynomial term" in $x_1, \ldots, x_n$ by saying that it is really the *array of coefficient terms* $t_{i_1,\ldots,i_n}$ themselves.) Anyway, the $t_{i_1,\ldots,i_n}$ can be constructed from any given t by first constructing $P_{x_n}(t)$ (6.24), a polynomial term $\sum_{i_n=0}^{I} t_{i_n} x_n^{i_n}$ in x_n with coefficient terms t_{i_n} that are free of x_n; then, for each i_n, construct $P_{x_{n-1}}(t_{i_n})$, etc.

Suppose that t' is another polynomial term in $x_1, \ldots, x_n$, say

$$\sum_{i_n=0}^{I'} \left[\sum_{i_{n-1}=0}^{I'_{i_n}} \left[\cdots \left[\sum_{i_1=0}^{I'_{i_2,\ldots,i_n}} t'_{i_1,\ldots,i_n} x_1^{i_1} ggr] \cdots \right] x_{n-1}^{i_{n-1}} \right] x_n^{i_n}, \tag{6.43.2}$$

where the terms $t'_{i_1,\ldots,i_n}$ are also free of $x_1, \ldots, x_n$. Then we shall write $t \equiv_x t'$ as an abbreviation for the formula consisting of the conjunction of:

(a) the equations $t_{i_1,\ldots,i_n} = t'_{i_1,\ldots,i_n}$, for all multi-indices $i_1, \ldots, i_n$ for which such terms exist,

(b) the equations $t_{i_1,\ldots,i_n} = 0$, for all $i_1, \ldots, i_n$ for which $t_{i_1,\ldots,i_n}$ but not $t'_{i_1,\ldots,i_n}$ exists, and

(c) the equations $0 = t'_{i_1,\ldots,i_n}$, for all $i_1, \ldots, i_n$ for which $t'_{i_1,\ldots,i_n}$ but not $t_{i_1,\ldots,i_n}$ exists.

Next, we shall write $t \oplus_x t'$ for the term of the form (6.43.1–2) in which the coefficient terms are:

(a) $t_{i_1,\ldots,i_n} + t'_{i_1,\ldots,i_n}$, for all multi-indices $i_1, \ldots, i_n$ for which such terms exist,

(b) $t_{i_1,\ldots,i_n}$, for all $i_1, \ldots, i_n$ for which $t_{i_1,\ldots,i_n}$ but not $t'_{i_1,\ldots,i_n}$ exists, and

(c) $t'_{i_1,\ldots,i_n}$, for all $i_1, \ldots, i_n$ for which $t'_{i_1,\ldots,i_n}$ but not $t_{i_1,\ldots,i_n}$ exists.

Finally, we shall write $t \otimes_x t'$ for the term of the form (6.43.1–2) in which the coefficient terms are suitable polynomial combinations of the $t_{i_1,\ldots,i_n}$ and $t'_{i'_1,\ldots,i'_n}$, guided by the conventional algebraic definition of the product of polynomials in $x_1, \ldots, x_n$ (generalizing that in (6.18) in the case $n = 1$).

Note that the formula $t \equiv_x t'$ and the polynomial terms $t \oplus_x t'$ and $t \otimes_x t'$ are all free of $x_1, \ldots, x_n$; they contain only $0, 1, c_1, \ldots, c_m$ (and the ring-operation symbols).

6.44 We now return to the task of deriving Henkin's (5.8) from (6.41.1). For each k and l in (6.41.1), write $r_{k,l}$ and $s_{k,l}$ for the terms $\sum_g \prod_h p_{k,l,g,h}$ and $\sum_g \prod_h q_{k,l,g,h}$, respectively. By (6.43), we may assume that each $r_{k,l}$ and $s_{k,l}$ has already been organized into the form of a polynomial in $x_1, \ldots, x_n$, as in (6.43.1). Thus (6.41.1) becomes

$$\bigvee_{k=1}^{M} \left[F = \sum_{l \in L_k} w_{k,l} \left(r_{k,l} s_{k,l}^{-1} \right)^2 \right], \tag{6.44.1}$$

for some $M \in \mathbf{N}$ and suitable index-sets L_k. There is, as yet, no assurance that $s_{k,l} \not\equiv_x 0$, for each k, l, as Henkin's (5.8.2) requires. So, for each k, we consider all possible subsets $L' \subseteq L_k$, and write $A_{k,L'}(c)$ (where c denotes $(c_1, \ldots, c_m)$) for the formula

$$\left[\bigwedge_{l \in L'} (s_{k,l} \not\equiv_x 0) \right] \wedge \bigwedge_{l \in L_k \setminus L'} (s_{k,l} \equiv_x 0).$$

Note that no x_j's occur in $A_{k,L'}(c)$, and that for every k, the formula $\bigvee_{L' \subseteq L_k} A_{k,L'}(c)$ is a tautology (so is the formula $\neg(A_{k,L'}(c) \wedge A_{k,L''}(c))$, whenever $L' \neq L'' \subseteq L_k$).

Next, we "clear denominators" in each disjunct of (6.44.1), and for each $L' \subseteq L_k$, we write $B_{k,L'}$ for the formula

$$F \cdot \prod_{l \in L'} (s_{k,l} \cdot s_{k,l}) = \sum_{l \in L'} \left[w_{k,l} (r_{k,l} \cdot r_{k,l}) \cdot \prod_{l \neq l' \in L'} (s_{k,l'} \cdot s_{k,l'}) \right].$$

By (6.44.1) and the ring axioms, we have $\bigvee_{k=1}^{M} B_{k,L_k}$. Applying (6.42.2) we get a proof of $\bigvee_{k=1}^{M} C_{k,L_k}(c)$, where, for each $L' \subseteq L_k$, $C_{k,L'}(c)$ denotes the formula

$$F \otimes_x \bigotimes_{l \in L'} (s_{k,l} \otimes_x s_{k,l}) \equiv_x \bigoplus_{l \in L'} \left[w_{k,l} \otimes_x (r_{k,l} \otimes_x r_{k,l}) \otimes_x \bigotimes_{l \neq l' \in L'} (s_{k,l'} \otimes_x s_{k,l'}) \right];$$

here $\bigoplus_{l \in L'}$, $\bigotimes_{l \in L'}$, etc., indicate iterated use of $\oplus_x$, $\otimes_x$, etc., and we regard $w_{k,l}$ as a "constant" polynomial term in $x_1, \ldots, x_n$. (As usual, $\bigoplus_{l \in \emptyset} t_l$ and $\bigotimes_{l \in \emptyset} t_l$ indicate 0 and 1, respectively, for any terms t_l.) $C_{k,L'}(c)$ (like $A_{k,L'}(c)$) is free of $x_1, \ldots, x_n$. The proof of $\bigvee_{k=1}^{M} C_{k,L_k}(c)$ requires the hypothesis that the ground field contain $> MD$ elements, where D is the maximum, for all $k = 1, \ldots, M$ and all $j = 1, \ldots, n$, of the formal degree in x_j of the right- and left-hand sides of B_{k,L_k}. For this, we add to our system the axiom P_{MD}: $1{+}1 \neq 0 \wedge \cdots \wedge 1{+}\cdots{+}1 \neq 0$, where the last inequation contains MD 1's.

For each k and each $L' \subseteq L_k$, $C_{k,L_k}(c) \to (A_{k,L'}(c) \to C_{k,L'}(c))$, using the ring axioms and $0^{-1} = 0$.

6.45 For the first time since the end of (6.1), we "unfix" *i*. Looking once again at Henkin's (5.8), we see that we are to construct a certain semialgebraic partition $D_1, \dots, D_l$ of P_{nd}. So far, we already have, for any real closed field R, the semialgebraic sets

$$E_i := \left\{ c \in R^m \;\middle|\; \bigwedge_j [p_{ij}(c) \geq 0 \,\wedge\, q_{ij}(c) > 0] \right\},$$

where $c := (c_1, \dots, c_m)$ is now a list of actual elements of R (and not a list of variable symbols), and where the p_{ij}, q_{ij} now represent the actual polynomials $\in \mathbf{Z}[C]$ in indeterminates C that correspond to the terms formerly denoted by p_{ij} and q_{ij} in (6.1.1); the latter says that $P_{nd} = \bigcup_i E_i$, although the E_i need not be pairwise-disjoint. To get our (preliminary) partition $D_1, \dots, D_l$, let $D_i = E_i \setminus \bigcup_{j=1}^{i-1} E_j$, which is still semialgebraic. For each $c \in E_i$ (and hence for each $c \in D_i$), we have established, formally, that $\bigvee_k C_{k,L_k}(c)$ holds (since the characteristic of R is 0).

This partition, however, is not quite good enough for (5.8). For each i, k, write

$$D_{i,k} = \{ c \in D_i \mid \neg C_{1,L_1}(c) \,\wedge\, \cdots \,\wedge \neg C_{k-1,L_{k-1}}(c) \,\wedge\, C_{k,L_k}(c) \}.$$

The equation $D_i = \bigcup_{k=1}^{M} D_{i,k}$ follows from $\bigvee_{k=1}^{M} C_{k,L_k}(c)$. For each $c \in D_{i,k}$ we have $C_{k,L_k}(c)$.

Finally, for each $L' \subseteq L_k$, write $D_{i,k,L'} = \{ c \in D_{i,k} \mid A_{k,L'}(c) \}$. Then $D_{i,k} = \bigcup_{L' \subseteq L_k} D_{i,k,L'}$ (disjoint, semialgebraic), and for $c \in D_{i,k,L'}$ we have $C_{k,L'}(c)$ and $A_{k,L'}(c)$, which, together, express Henkin's (5.8.1) and (5.8.2), respectively. $\square$

Proposition 6.46 *We can eliminate the temporary axiom* $0^{-1} = 0$ *from the proof constructed by (6.41) (refined by (6.44)) of the end-formula* $\bigvee_k C_{k,L_k}(c)$.

Proof. First suppose that no instances $s = t \rightarrow s^{-1} = t^{-1}$ (for terms s and t) of the special equality axiom for $^{-1}$ occur asinitial formulae in the given proof. Then we replace each occurrence of 0^{-1} with 0 throughout the proof. We claim that the resulting sequence of formulae is still a proof (after trivial adjustments), and, indeed, one in which $0^{-1} = 0$ no longer occurs as an initial formula. Indeed, applications of modus ponens are (as usual) converted into new such applications. Tautologies and ring axioms are converted into new instances of same. P_p, $1 \neq 0$, and $\bigwedge_j (q_{ij} \neq 0)$ are not changed at all. The (other) field axiom $s \neq 0 \rightarrow ss^{-1} = 1$ is unchanged unless s is 0, in which case we get $0 \neq 0 \rightarrow 0 \cdot 0 = 1$, which follows tautologically from the identity axiom $0 = 0$. The special equality axioms for $+$ and $\cdot$ become other such axioms. The temporary axiom $0^{-1} = 0$ becomes $0 = 0$, which is an identity axiom. Thus we have a proof in which $0^{-1} = 0$ no longer occurs as an initial formula. The end-formula $\bigvee_k C_{k,L_k}(c)$ (6.44) contains no reciprocals, and so is unaffected by this replacement.

If, however, the instance $a = 0 \rightarrow a^{-1} = 0^{-1}$ of the special equality axiom for $^{-1}$ had occurred as an initial formula in the given proof, then the replacement described above would have converted it into $a = 0 \rightarrow a^{-1} = 0$, which is not provable without $0^{-1} = 0$ (or something like it).

So (6.46) will be established in general if we can convert the given proof into a new proof in which no instances of the special equality axiom for $^{-1}$ occur as initial formulae, and in which the end-formula is unaffected; $0^{-1} = 0$ may still occur as an initial formula in the new proof.

Let $\mathcal{G}$ be the set of all instances $s = t \rightarrow s^{-1} = t^{-1}$ of the special equality axiom for $^{-1}$ that occur as initial formulae in the given proof. (Here, we may assume that the terms s and t are distinct, by (6.12); on the other hand, while this observation was useful in the proof of (6.31), the reader will notice that one could carry out the argument below just as well without it.) Let $\mathcal{F}$ be the set of all the terms s^{-1} and t^{-1} that occur as above in formulae in $\mathcal{G}$. Fix an arbitrary enumeration of the terms in $\mathcal{F}$, subject to the sole conditions that (a) 0^{-1} occurs first in the enumeration, if at all, and (b) if s^{-1} and t^{-1} in $\mathcal{F}$ have unequal depth,[46] then the term with lower depth comes earlier in the enumeration.

If $\mathcal{F} \subseteq \{0^{-1}\}$, then in fact $\mathcal{F} = \emptyset$, and we are done. Otherwise, let $t^{-1} \in \mathcal{F} \setminus \{0^{-1}\}$ be the last term in this enumeration. It suffices to convert the given proof into a new proof for which the corresponding set $\mathcal{F}'$ is a subset of $\mathcal{F} \setminus \{t^{-1}\}$.

To do this, first temporarily add $t \neq 0$ to the list of axioms. Then each formula in $\mathcal{G}$ of the forms $s = t \rightarrow s^{-1} = t^{-1}$ or $t = s \rightarrow t^{-1} = s^{-1}$ (for some term s) becomes provable from the remaining axioms and the temporary axiom $t \neq 0$, as shown in (6.12). We have therefore removed those formulae from $\mathcal{G}$; then t^{-1} gets removed from $\mathcal{F}$, in this case.

Second, add $t = 0$ instead of $t \neq 0$ to the axioms. Replace every occurrence of t^{-1} with 0 throughout the given proof. We claim that we still have a proof (after trivial adjustments). This is checked as in

[46] Recall the proof of (6.40).

the first paragraph, with only the following differences: (1) This time, the field axiom $s \neq 0 \;\rightarrow\; ss^{-1} = 1$ is unchanged unless s is t, in which case we get $t \neq 0 \;\rightarrow\; t \cdot 0 = 1$, which follows tautologically from the temporary axiom $t = 0$. (2) The temporary axiom $0^{-1} = 0$ is unchanged, since t is not 0. (3) This time, we must contend with instances $r = s \;\rightarrow\; r^{-1} = s^{-1}$ in $\mathcal{G}$ of the special equality axiom for $^{-1}$. Since t has maximal depth, the only such instances that contain t^{-1} are those of the forms $t = s \rightarrow t^{-1} = s^{-1}$ and $s = t \rightarrow s^{-1} = t^{-1}$, for some s distinct from t. The first of these two instances becomes $t = s \rightarrow 0 = s^{-1}$ (note: t^{-1} cannot occur in s or even s^{-1}, for otherwise the depth of s^{-1} would exceed that of t^{-1}, s being distinct from t). Unfortunately, this is no longer a special equality axiom. But we can prove it, as follows: Assume $t = s$. From our temporary axiom $t = 0$ we then conclude $s = 0$. We are at liberty to use $s = 0 \;\rightarrow\; s^{-1} = 0^{-1}$ as an initial formula in the new proof, since s^{-1} and 0^{-1} both occur earlier in the enumeration of $\mathcal{F}$ than t^{-1} does. So $s^{-1} = 0^{-1}$. Using $0^{-1} = 0$ (as we may, at this stage), we conclude $s^{-1} = 0$, as required. (The second of the two instances is similar.) The end-formula is again unchanged, not containing any reciprocals.

Therefore we have proved our end-formula twice, once using $t \neq 0$, and again using $t = 0$, as initial formulae. We can therefore assemble from these two proofs a new proof, using neither $t \neq 0$ nor $t = 0$ as initial formulae. In the process, t^{-1} has been eliminated from $\mathcal{F}$, as required. $\Box$

6.47 Comparison with Ackermann-Hilbert. The above proof is a variation on the Ackermann-Hilbert method of simultaneously eliminating critical ε-formulae and formulae of ε-equality (Hilbert, Bernays, Vol. II [1939, 1970, pp. 56–79]). While the method we used in (6.31) followed the Ackermann-Hilbert method closely, the method used in (6.46) above differs from it as follows: where we introduced the two cases $t \neq 0$ and $t = 0$ in (6.46), they considered $t \neq s$ and $t = s$ (where s occurs earlier in the enumeration of $\mathcal{F}$); and where we replaced t^{-1} with 0, they replaced it with s^{-1}.

6.48 Comparison with Kreisel. Kreisel did not formulate or prove (6.46) in his sketches. Instead, he stated only:

> "By convention, $a^{-1} = 0 \leftrightarrow a = 0$, which is consistent with the field axiom $a \neq 0 \;\rightarrow\; a^{-1}a = 1$."

This is in accordance with his advice (quoted in (1.2) above) to ignore restrictions on methods of proof—advice that we have rejected here, as discussed in (1.8).

6.49 Nonnegativity of the weights. So far, we have said nothing about the signs of the weights $w_{k,l}$ in (6.41) and (6.44); but they are obviously well-defined and nonnegative whenever $\bigwedge_j (p_{ij} \geq 0 \wedge q_{ij} > 0)$ holds. By (6.1.1), the latter will hold for some i, whenever F is positive semidefinite in x. This part of the result cannot be proved or even stated in the same limited formal system as that described in (6.41), since here we need the $\geq$-symbol; furthermore, the very concept of "positive semidefiniteness" (mentioned in the hypothesis) cannot even be defined without mentioning quantifiers at least temporarily.

Theorem 6.50 Summary of §6. *From n and d we can construct not only Henkin's semialgebraic partition $D_1, \ldots, D_l$ of P_{nd} and the identities (5.8.1), where the weights $p_{ij}(c)$ satisfy (5.8.2), but we can also construct formal proofs of (5.8.1) and (5.8.2). The proof of the latter is in $\mathbf{RCF}^c_{\geq}$ (recall (6.1.1)). The proof of the former is as in (6.41) and (6.46): the alphabet is restricted to the symbols $0, 1, x_1, \ldots, x_n, c_1, \ldots, c_m, +, -, \cdot, ^{-1}, =, \vee, \wedge, \neg, \rightarrow, (,)$ (in particular, no quantifiers, no ε-symbol, and no (other) root-symbols ρ or ρ_q); the only initial formulae in this proof are the tautologies and the special equality axioms (6.11–12) for $+, \cdot$, and $^{-1}$ (the "logical" axioms), and—instances of—the field axioms, P_p (which can be replaced with a finite set of the reality conditions $(3.1.2)_s$), and[45] $\bigwedge_j (q_{ij} \neq 0)$ (the "non-logical" axioms); the only rule of inference is modus ponens.*

7 Stellensätze vs. the second sketch

7.1 Stengle's Positivstellensatz

In [1990, p. 11], Lombardi wrote as follows about Kreisel's *second* sketch (§6, above):

> "Today it seems to me that Kreisel's proof should be able to be adapted to the [real] *Nullstellensatz*."

Here, the "real *Nullstellensatz*" is, essentially, the "*Positivstellensatz*" (5.14) of Stengle. *Here I wish to express my skepticism.* True, Stengle's theorem is a straightforward generalization of Artin's theorem, as discussed in (5.14). But there is (only) one problem in adapting Kreisel's second sketch to Stengle: in (5.14.1) we are asserting the existence of *polynomials,* and not arbitrary rational functions as in Artin's theorem.

Syntactically, this will require the elimination, from the end-formula (5.14.1), of all terms of the form t^{-1} such that t contains at least one occurrence of $x_1, \ldots, x_n$; they must be eliminated *without spoiling the special form of the left-hand side*; thus one cannot simply clear denominators, as we did in (6.40). One might hope to arrange that the $^{-1}$ symbols never get applied, during the formal proof, to terms containing $x_1, \ldots, x_n$, but only to terms built up from $0, 1$, and the coefficients of the given polynomials. Unfortunately, the *statement* (4.1) of Tarski's QE theorem gives us no such assurance; a finer examination of almost any *proof* of Tarski's theorem, however, does lead to the desired conclusion; but then we are doing Kreisel's *first* sketch—which I postpone to Part II.

7.2 Hilbert's Nullstellensatz

Kreisel had already encountered the above difficulty with reciprocals (without noticing it) in the analogous case of Hilbert's *Nullstellensatz* (Kreisel [1957a], and [1958a, pp. 164–7]).[47] He hinted at his error in [1992, p. 33] (and probably earlier):

> "*Correction* of a sloppy aside on proof-theoretic aspects of [Hilbert's] *Nullstellensatz* .... In contrast to Kreisel [1960a], with detailed material on sums of squares I never returned to this matter; but cf., for example, Scarpellini [1969]. (Also in contrast to sums of squares there was no interest up-market in such aspects.)"

This typically cryptic statement may give the impression that Scarpellini [1969] fills the gap in Kreisel's sketch for Hilbert's *Nullstellensatz*; but such an impression would be wrong. Rather, on pp. 81–4 of Scarpellini (and also on p. 567 of Whiteley [1991]), one finds a syntactical treatment of a *weakened* form of Hilbert's *Nullstellensatz* not considered by Kreisel in [1957a] or [1958a] (or by anyone else, as far as I know—and rightly so):

Weak Hilbert Nullstellensatz 7.3 *Suppose $f, g_1, \ldots, g_t$ belong to $\mathbf{Z}[X_1, \ldots, X_n]$, and that for every* **integral domain** *D, and for every $x := (x_1, \ldots, x_n) \in D^n$, $f(x) = 0$ whenever each $g_i(x) = 0$. Then $f^e = \sum h_i g_i$, for some $e \in \mathbf{N}$ and $h_i \in \mathbf{Z}[X_1, \ldots, X_n]$.*

[47] P. Lorenzen, K. Schütte, and A. Robinson, in their reviews of Kreisel [1958a], all reported Kreisel's claimed unwinding of the *Nullstellensatz* as if it had been correct.

Then Scarpellini (and Whiteley, independently) proved what he called "the syntactic counterpart" of Hilbert's *Nullstellensatz*:

[Weak] Syntactic Nullstellensatz 7.4[48] *Suppose we have a proof of* $\forall x_1 \cdots \forall x_n \left[\left(\bigwedge_{i=1}^{t}(g_i = 0)\right) \rightarrow f = 0\right]$ *using only the axioms of* **integral domains**. *Then we can, primitive recursively, construct* $e \in \mathbf{N}$ *and terms* h_i *such that* $f^e = \sum h_i g_i$. (Here, f, g_i, h_i are terms in the language of rings with 1.)[49]

The reason I call this form of the Hilbert *Nullstellensatz* "weak" is that its hypothesis is stronger than that of the usual *Nullstellensatz*; the latter's hypothesis is only that $\left(\bigwedge_{i=1}^{t}(g_i = 0)\right) \rightarrow f = 0$ holds over all *algebraically closed fields*; in the "syntactic counterpart," the hypothesis would be that we are given only a proof, from the axioms for *algebraically closed fields,* of $\left(\bigwedge_{i=1}^{t}(g_i = 0)\right) \rightarrow f = 0$. By limiting themselves in (7.3–4) to the weakened form of the *Nullstellensatz*, Scarpellini and Whiteley duck the task of eliminating either root-terms (analogous to Kreisel's $\rho(s_0)$ and $\rho_q(s_0, \ldots, s_{2q})$ in the real case) or reciprocals s^{-1}. As we saw in most of §6, this elimination is the bulk of the work. On the other hand, Scarpellini and Whiteley weren't particularly *trying,* apparently, to constructivize mainstream theorems by unwinding non-finitist proofs, as Kreisel was.

Actually, by reading between the lines of Kreisel's sketch for the *Nullstellensatz*, one can prove the following result, intermediate between the usual and the weak Hilbert *Nullstellensatz*:

Intermediate-Strength Syntactic Nullstellensatz 7.5 *Suppose we have a proof, using only the axioms of* **integrally closed domains with algebraically closed fraction field**,[50] *of the formula*

$$\forall x_1 \cdots \forall x_n \left[\left(\bigwedge_{i=1}^{t}(g_i = 0)\right) \rightarrow f = 0\right].$$

Then we can, primitive recursively, construct $e \in \mathbf{N}$ *and terms* h_i *such that* $f^e = \sum h_i g_i$ *(and we can also construct a* proof *of this equation).*

---

[48]Scarpellini said (p. 71) that this is "a constructive version ... suggested by G. Kreisel."

[49]Scarpellini and Whiteley didn't seem to mention it, but we can also construct a *proof* of this equation.

[50]Here we mean the axioms of integral domains, supplemented by the axioms $\exists x\,(x^d + a_{d-1}x^{d-1} + \cdots + a_0 = 0)$ $(d = 2, 3, \ldots)$ ensuring roots of *monic* nonconstant polynomials. The only field axiom that is missing here is that for inverses.

For a hint of the proof, we note only that in [1957a] Kreisel at least mentioned, correctly, that ordinary elimination theory (for algebraically, not real, closed fields) allows us to eliminate the root-terms that arise from use of the extended first ε-theorem; it was only the reciprocals that he overlooked.

On the other hand, as long as we must get our hands dirty with elimination, we might as well do the entire job with elimination theory *alone,* as Hilbert seems to have done in his *original* proof of the *Nullstellensatz* ([1893a]; cf. pp. 233–40 of the English translation [1978]). Responding to van der Waerden's and Krull's later, more advanced and/or abstract proofs, Zariski (1947), R. Brauer (1948), and Seidenberg [1956] gave a series of progressively more "elementary" proofs of the *Nullstellensatz* (some emphasizing elimination theory). But neither they nor Kreisel seemed to comment on whether there was anything essentially non-constructive in Hilbert's original proof, a question that van den Dries has raised in correspondence; on a preliminary reading, the answer seems to be no. If this impression holds up, then it is unclear what the ε-theorems, cut-elimination, or some of the *ad hoc* methods really contribute to the *Nullstellensatz*. And as Kreisel himself said in [1992, p. 33]:

> "Looking back I see no evidence that logical formulations could have served here (even) as first steps towards such sharp results as (Amoroso 1990), (Kollár 1988), and, quite recently, (Phillipon 1991)."

8 A finitary, syntactic proof of consistency

As in (3.3), let $(K, \geq)$ be a discrete ordered field, and let $\mathbf{RCF}(K, \geq)$ be the system obtained from $\mathbf{RCF}_{\geq}$ (3.1) by adding the diagram of $(K, \geq)$; i.e., to the language we add a constant symbol a_r for each $r \in K$, and to the list of axioms we add the atomic sentences $a_0 = 0$, $a_1 = 1$, and (for each $r, s \in K$) $\quad a_r + a_s = a_{r+s}, \quad a_r a_s = a_{rs}$, and (whenever $r > s$) $\quad a_r > a_s$.

Let $\mathbf{F}^{\mathrm{qf}}(K)$ denote the system obtained from $\mathbf{RCF}(K, \geq)$ by dropping $\geq$ and all quantifiers (in particular, dropping the axioms (3.1.3), $(3.1.4)_q$, (3.1.5), and the field axiom $a \neq 0 \rightarrow \exists z(az = 1)$), but adding the $^{-1}$ symbol together with its defining axiom (6.4.3). Thus the only non-logical axioms in $\mathbf{F}^{\mathrm{qf}}(K)$ are—all instances of—the quantifier-free field axioms and the diagram of K. The logical axioms are the tautologies and the equality axioms; the only rule of inference is modus ponens. The existence of K, together with the fact that it is a field, implies that $\mathbf{F}^{\mathrm{qf}}(K)$ is consistent; moreover, the fact that K has an ordering implies that *for every finite set of nonnegative elements p of K, and every set of terms* t_p *built up by the field operations from (0, 1, and) the* a_r $(r \in K)$*:* $\mathbf{F}^{\mathrm{qf}}(K) \nvdash -1 = \sum_p a_p t_p^2$. As promised in (1.8), in the sentence after (4.7), and in (6.17), we now prove:

Theorem 8.1 *There is a primitive recursive procedure to transform any given proof of* $\exists y_1 \exists y_2(-1 = y_1^2 + y_2^2)$ *in* $\mathbf{RCF}(K, \geq)$*, into a proof of* $-1 = \sum_p a_p t_p^2$ *in* $\mathbf{F}^{\mathrm{qf}}(K)$*, for some* a_p *and* t_p *as above. Since no such proof exists (as mentioned above),* $\mathbf{RCF}(K, \geq)$ *is consistent.*

Remark. This argument is similar to Lombardi's proof in [1988–91]. The main difference is that he used his constructivized proof (9.6) of the *Positivstellensatz* (5.14), while we shall use the constructive proof of Artin's theorem.

Proof. We imitate most of §6, replacing F by -1. One difference: in (6.1–2) we used QE to *construct* a proof of $\exists z\,(F = z^2)$; now we just *assume* (presumably wrongly) that we have a proof of $\exists y_1 \exists y_2\,(-1 = y_1^2 + y_2^2)$ in $\mathbf{RCF}(K, \geq)$. Among the initial formulae of this proof, some belong to the diagram of $(K, \geq)$; as in (6.2.1), replace those of the form $a_r > a_s$ $(r > s)$ with $(a_r - a_s)p_{r,s}^2 = 1$, where the $p_{r,s}$ are new constant symbols. Now the axioms (and not merely the end-formula) are free of the $\geq$-symbol, with the exception of axiom (3.1.5), which serves as an explicit definition of $\geq$, using which we can cleanse the entire proof of that symbol. As in (6.3), it is clear that our proof can be re-arranged so as not to contain the reality axiom(s) $(3.1.2)_s$ as initial formulae; here the argument is even easier than it was in (6.3). Next, introduce the symbols ρ, ρ_q, and $^{-1}$, together with their defining axioms, (6.4.1), $(6.4.2)_q$, and (6.4.3), as in (6.4); remove the quantifiers; denote the resulting system by $\mathbf{RCF}^{\mathrm{qf}}(K)$. By the extended first ε-theorem (6.14) (as in (6.16)), we obtain a proof of a disjunction, possibly containing ρ- and ρ_q-symbols. The overall strategy (6.22) for eliminating ρ- and ρ_q-symbols works exactly the same here as it did in §6, replacing F by -1; the only difference is that now the argument is a little simpler, since there are no variables $x_1, \ldots, x_n$ (which didn't come into play until (6.38) and later). We can stop already at (6.37), when the end-formula is $\bigvee_k(-1 = \sum_l w_{k,l} t_{k,l}^2)$, for some terms $t_{k,l}$ built up by the field operations from the a_r, and where the $w_{k,l}$ are products of terms $a_r - a_s$ such that $r > s \in K$. For each k, either $\mathbf{F}^{\mathrm{qf}}(K) \vdash -1 = \sum_l w_{k,l} t_{k,l}^2$ or $\mathbf{F}^{\mathrm{qf}}(K) \vdash -1 \neq \sum_l w_{k,l} t_{k,l}^2$, and we can decide which, using the diagram of K (since K is discrete). In this way we compute k_0 such that $\mathbf{F}^{\mathrm{qf}}(K) \vdash -1 = \sum_l w_{k_0,l} t_{k_0,l}^2$, since $\mathbf{F}^{\mathrm{qf}}(K) \vdash \bigvee_k (-1 = \sum_l w_{k,l} t_{k,l}^2)$. Finally, replace the $w_{k_0,l}$ by a_p, for suitable $p \geq 0$ in K depending on l, and re-label the $t_{k_0,l}$ as t_p. $\square$

Note that if, in the above argument, we had tried to use the second ε-theorem instead of (6.33), the conclusion would not have been sufficient to establish (8.1).

Note also that our proof did not presuppose the existence of a model of $\mathbf{RCF}(K, \geq)$ (i.e., a real closed order-extension of $(K, \geq)$); nor did we construct such a model; the proof was purely syntactical, in the spirit of Hilbert, Bernays. On the other hand, in Part II we shall use this consistency result to give an *easy,* finitary construction of such a model, i.e., a real closure of $(K, \geq)$. (Hint: it will be an "ι-term model"; recall (6.6).)

9 Underground bounds

In this section we assemble scattered bounds for Artin's theorem; with a few above-ground exceptions, the proofs (and, until now, even some of the statements) are unpublished. We also point out some apparent discrepancies, since even the unpublished bounds below have been cited in publications. Here, R is real closed, and $K \subseteq R$.

9.1 Kreisel's bounds

Of course, Kreisel's *second* sketch (§6 above) gives *some* primitive recursive (roughly, some multiply-iterated exponential) bound, in terms of n and d, on the number, degrees, and integer coefficients, of the rational functions in Henkin's sum-of-nonnegatively-weighted-squares representation (5.8.1). The main obstacle to finding this bound *explicitly* is the lack of any known, explicit bound for the extended first ε-theorem (6.14). The other main meta-theorem used in the second sketch is Tarski's QE (4.1); as stated in (4.5)(a), H. Friedman asserted a bound on the number of symbols in the formal proof produced by QE. For all of the other steps in Kreisel's second sketch it would be an easy exercise to develop explicit bounds.

Kreisel's *first* sketch (Part II of this paper) offers a real chance to get explicit bounds, since it (unlike the second sketch) does not appeal to the statement of the (first) ε-theorem, or even of QE. As stated in (5.6), Kreisel considered a uniquely orderable field K with Pythagoras number $P(K) \leq k$. His [1960a] and [1958a] expositions gave no explicit bounds. But in a couple of abstracts [1957b] he considered a special case of (5.6.1), namely $n = 2$: for an $f \in K[X_1, X_2]$ of degree d that is psd over R, he asserted, without proof, that "a rough calculation shows that" the number and degrees of the $r_i \in K(X_1, X_2)$ in the weightless sum-of-squares representation (5.6.1), as found by his first sketch, is "of order"

$$2^{2^{2^{2^{2^{2^{d}}}}}}. \tag{9.1.1}$$

As big as this is, it is impossibly small, since it overlooks the extra parameter k, which Kreisel himself had introduced—of necessity: as discussed in (5.6) above, even for *constant* $f \in K$, any bound, independent of k, on the number of squares required in (5.6.1), would have to be $\geq P(K)$ for every uniquely orderable K (and $P(K)$ can be arbitrarily large for uniquely orderable K). Since Kreisel offered no clue as to how he derived (9.1.1) from Artin's original proof, it is hard to know where the mistake(s) arose, or how serious they are. The most positive spin would be that he may have intended to treat k as a *constant,* rather than a parameter; then, presumably, k might not affect the "order" of (9.1.1).

Later, others gave detailed proofs of bounds applying to larger values of n, in case K (and not only R) is real closed (i.e., they apply only to Robinson's (5.4.1)). These (by no means underground) proofs were followed, one by one, by more unproved (and cryptic) claims from Kreisel about the bounds that his papers [1958a] or [1960a] "give." These other bounds (unlike Kreisel's) apply only to the *number,* and not the degrees, of the rational functions r_i in the weightless sum-of-squares representation (5.4.1); but they are independent of d.

Specifically, (1) Ax (1966, unpublished) proved that for real closed R, and for $n \leq 3$, 2^n squares suffice in (5.4.1). (For $n = 1$ this says $P(R(X_1)) = 2$, which follows from the fact (2.2) that $R(\sqrt{-1})$ is algebraically closed, combined with the 2-square identity; at least for $R = \mathbf{R}$, this was extended by Witt [1934] to $P(L) = 2$, where $L = R(X_1, \alpha)$, for any α algebraic over $R(X_1)$; i.e., L is now *any* function field of transcendence degree 1 over R. For $n = 2$, Hilbert had proved $P(R(X_1, X_2)) = 4$ in [1893b], at least for $R = \mathbf{R}$.) In [1968, p. 362] Kreisel mentioned Ax's result, and then asserted:

> "Thus, for $n = 3$, Kreisel [1958a] gives essentially $N \leq 2^{2^{2^{d}}}$ ...."

Here, N is a bound on the number, but not, apparently, the degrees, of the r_i. It is not clear (to me) whether Kreisel's bound N was supposed to apply to (a) Henkin's *nonnegatively weighted* sum-of-squares representations (5.7.1) and (5.8) (which hold over arbitrary ordered fields K), (b) Robinson's weightless sum-of-squares representation (5.4.1) (for real closed K), or (c) Kreisel's weightless sum-of-squares representation (5.6.1) (which holds only over *uniquely orderable* K with $P(K) \leq k$). Kreisel stated all 3 versions on pp. 361–2 of [1968]. In fact, Kreisel [1958a], which supposedly "gives" the bound $2^{2^{2^d}}$ "essentially," deals only with the representation in (c); but in case (c), this bound, if read literally, is impossibly small, for the same reason as in (9.1.1): it overlooks Kreisel's new parameter k.

(2) Independently of Ax's work, A. Pfister [1967] (cf. also Lam [1973], Jacobson [1980, 1989], and (9.4) below) proved the above 2^n-bound for arbitrary $n \geq 0$; again, it applies to the number, and not the degrees, of the r_i in (5.4.1), and it requires that the ground field K be real closed. This was followed by Kreisel [1974b]:

> "With the help of the 'logical' analysis of Artin's proof, one obtains a simultaneous bound (for both the *number* and the *degrees* of the squares) of the order of magnitude
>
> $$2^{\cdot^{\cdot^{2^d}}}\Big\}n \text{ .''} \tag{9.1.2}$$

(I repeated the bound (9.1.2) in Delzell [1982a] and [1982b], p. 91 of each.) Again, not only is there no hint of how this bound was derived from Artin's original proof, but even its *statement* is not clear: are weights to be allowed on the squares (as in (1a) above), or not (as in (1b) and (1c))? And what hypotheses apply to the ground field K? As in (1) above, these issues are crucial to the correctness of (9.1.2). Similarly, in [1977a, p. 115], Kreisel began by stating his weightless sum-of-squares representation (5.6.1) (which holds only over uniquely orderable K with $P(K) \leq k$); he then wrote:

> "Model theory supplied a recursive method of determining a bound (by trial and error applied to derivations in predicate logic), proof theory a bound involving n-fold *exponentiation*."

The latter refers to (9.1.2), presumably; but, as in (1) above, this is impossible under these hypotheses.

Another puzzling aspect of Kreisel's bounds is that they differ rather significantly from:

9.2 Daykin's bound

In [1961] G.E. Daykin wrote a Ph.D. thesis under the direction of R. Rado at Reading (and examined also by B.H. Neumann). It presented the details of Kreisel's first sketch (the unwinding, in Part II), including a computation of a bound for (5.6.1). The thesis was never published. While writing the thesis, Daykin was able to consult periodically with Kreisel, who had spent most of the previous decade at Reading; however, Kreisel was in France when the thesis was finally finished, so he never saw it until 1992 (in 1961, it was not easy to copy handwritten manuscripts).

To state his bound, Daykin first defined

$$\Lambda(1,d,k) = k^{2^{(\frac{1}{2}d)^{2d}}}; \tag{9.2.1}$$

he then defined

$$\Lambda(\nu,d,k) = \Lambda\left(\nu-1,\, 7^{\cdot^{\cdot^{7^{d^d}}}}\Big\}3^d, k\right) \qquad (1<\nu\le n); \tag{9.2.2}$$

then the number of squares is at most $\Lambda(n,d,k)$; he did not explicitly state a bound on $\deg r_i$, though it, too, is undoubtedly some stack of exponentials.

Actually, since d, being the degree of f, is a *constant,* it should not be used as a variable during the recursion in (9.2.1–2); thus each d in (9.2.1–2) should probably be replaced by a new *variable*, say l; otherwise $\Lambda(2,d,k)$, say, would be undefined, since (9.2.2) "defines" it to be $\Lambda\left(1,\, 7^{\cdot^{\cdot^{7^{d^d}}}}\Big\}yle3^d, k\right)$, which is not defined (even by (9.2.1)). Thus:

$$\Lambda(1,l,k) = k^{2^{(\frac{1}{2}l)^{2l}}} \quad \text{and} \quad \Lambda(\nu,l,k) = \Lambda\left(\nu-1,\, 7^{\cdot^{\cdot^{7^{l^l}}}}\Big\}3^l, k\right); \tag{9.2.1$'$–2$'$}$$

then use $\Lambda(n,d,k)$.

This Λ is one step beyond the "tower" functions, such as $2^{\cdot^{\cdot^{2}}}\Big\}ndk$ (which, like Kreisel's bound (9.1.2), are in $\mathcal{E}^4$ in the Grzegorczyk hierarchy—cf., e.g., Rose [1984]); i.e., $\Lambda \in \mathcal{E}^5 \setminus \mathcal{E}^4$. Few functions occurring in mathematical practice grow so quickly; on the other hand, the thesis does not seem to exclude the possibility that $k=1$ (e.g., for $K = R$), in which case the bound collapses, for all n and d, to 1! This is impossibly small, since, e.g., $1+X_1^2$ is not a perfect square in $R(X_1)$. Browsing through the 170-page proof (whose essentials are, otherwise, impeccable, as far as I can tell), I found the inequality $k \geq 1$ on p. 127, and $k \geq 2$ on p. 133; presumably the remedy is to *assume* that $k \geq 2$ (which includes the case $P(K) = 1$).

9.3 Friedman's bounds

In a manuscript ([1981], unpublished, but cited in Friedman, et al, [1983, p. 148]), Friedman considered positive semidefinite $f \in \mathbf{Z}[X]$ ($X = (X_1, \ldots, X_n)$) with coefficients whose absolute values are bounded by M. Combining Kreisel's ideas with some accelerated versions of real quantifier-elimination discovered in the seventies, he obtained (his Theorem 3.1) the representation $f = \sum_i r_i^2$ for some $r_i \in \mathbf{Q}(X)$ as in (2.4.1), with the bounds

$$2^{\cdot^{\cdot^{\cdot^{2^{8+6n^d \log^2 M}}}}}\Big\}10 \qquad \text{and} \qquad 2^{\cdot^{\cdot^{\cdot^{2^{7+6n^d \log^2 M}}}}}\Big\}9, \tag{9.3.1}$$

respectively, for the absolute values of the integer coefficients of the r_i, and for the number and degrees of the r_i. An advantage, mentioned by Friedman, is that these bounds, unlike Kreisel's and Daykin's, are *elementary* recursive in n, d, and M (i.e., they are in $\mathcal{E}^3$ in the Grzegorczyk hierarchy), and not merely primitive recursive, since the stacks of exponentials are of bounded height; moreover,

> "these bounds are of course very crude, and at every turn there are rich opportunities for improvements."

I could not find his definition of "log"; often in the complexity literature, $\log M$ means the number of binary digits in $M \geq 1$, so that, for example, $\log 1 = 1$. (If log were to mean "logarithm to base 2," say, then when $M = 1$, both bounds in (9.3.1) would become independent of n and d, which is impossible, since, e.g., $1 + X_1^2 + \cdots + X_n^2$ is not the sum of n squares in $\mathbf{R}(X)$, as Cassels showed in [1964], answering a question of Fine (cf. also Lam [1973, IX.2.4]).)

Friedman went on to prove a parametrized version (his Theorem 3.2) that is simpler than (5.8) or even (5.9)(b): Using the notation of (5.5), from n and d we can construct $p_i \in \mathbf{Z}(C)$ and $r_i \in \mathbf{Q}(C; X)$ such that, letting $f \in \mathbf{Z}[C; X]$ and P_{nd} be as above (5.8), $\forall (c; x) \in P_{nd} \times R^n$, $f(c; x) = \sum_i p_i(c) r_i(c; x)^2$; the absolute values of the integer coefficients of the p_i and r_i, and the number and degrees of same, are now bounded, respectively, by

$$2^{\cdot^{\cdot^{\cdot^{2^{8+n^d}}}}}\Big\}10 \qquad \text{and} \qquad 2^{\cdot^{\cdot^{\cdot^{2^{7+n^d}}}}}\Big\}9. \tag{9.3.2}$$

I see 3 problems with this second theorem.

(a) It does not assert that for $c \in P_{nd}$, each $p_i(c) \geq 0$; without this, it is not a solution to Hilbert's 17th problem.

(b) If we include the condition mentioned in (a), then the theorem itself becomes false on two counts:

(i) It asserts the affirmative answer to an old question of Kreisel, whether the disjunction of identities in (5.8) or (5.9)(b) can be contracted to a single identity ([1962], [1969, p. 102], [1974a], [1977a, p. 115–6], [1977b, footnote 1], and [1978]). For a counterexample, take $d \geq 4$: the set

$$\begin{aligned} P :&= \{ (a,b) \in R^2 \mid (X_1^2 + a)^2 + b \in R[X_1] \text{ is psd} \} \\ &= \{ (a,b) \mid b \geq 0 \ \vee \ (a \geq 0 \ \wedge \ b \geq -a^2) \} \end{aligned}$$

is not expressible as $\bigcap_i \{ (a,b) \mid p_i(a,b) \text{ is defined and } \geq 0 \}$, for any finite set of $p_i \in \mathbf{Q}(A,B)$ (or, for that matter, for any real analytic p_i when $R = \mathbf{R}$); rather, P is only a *union* of ≥ 2 such intersections, and thus ≥ 2 identities are needed in (5.8) or (5.9)(b); for details, cf. Delzell [1982a], [1982b], or [1994].

(ii) Friedman's second theorem seems also to assert the affirmative answer to another, even older question of Kreisel: given a psd $f \in \mathbf{R}[X]$, can we find *everywhere defined* $r_i \in \mathbf{R}(X)$ satisfying $f = \sum_i r_i^2$ as in (2.4.1)? Kreisel informed me that in a letter of 1956, E.G. Straus gave the first negative answer: there are psd f with "bad" points $x \in \mathbf{R}^n$ such that, no matter how the r_i in (2.4.1) are chosen, the denominator of some r_i must vanish at x. Specifically, he showed that if f is Hilbert's [1888] example of a psd homogeneous quaternary quartic (i.e., $(n,d) = (4,4)$) that is not a sum of squares in $\mathbf{R}[X]$, then f has a bad point at the origin. Examples simpler than Hilbert's were found by Motzkin (1967), R.M. Robinson (1973), and others.[51]

Whether we use (i) or (ii) to refute Friedman's second theorem (modified by (a)), any question of bounds is premature.

(c) Even if Friedman's second theorem is corrected as suggested in (a)–(b) above, the bounds themselves are still curious.

(i) If $n = 1$, then the bound on $\deg r_i$ given in (9.3.2) becomes independent of d; but this is impossible, considering $f(X_1) := X_1^d$ for large even d.

(ii) The bounds in (9.3.2) are smaller than those in (9.3.1) (unless $M = 1$ and 'log' means 'logarithm'); this is not necessarily a problem, but it is odd, since (9.3.1) applies to a single polynomial, while (9.3.2) attempts to apply to the general polynomial.

Thus the main problem with Friedman's second (parametrized) result is that it overlooks the fact that for $d \geq 4$, any sum-of-squares representation must vary *piecewise*-polynomially, at best. Once this piecewise character is recognized, however, it seems likely that Friedman's complexity analysis, *to the extent that it is correct for the non-parametrized case (9.3.1),* would extend straightforwardly to the parametrized version, yielding a similar bound. As to the essential correctness of Friedman's complexity analysis, perhaps someone will, some day, elaborate his argument sufficiently to settle the matter; certainly large parts of it, especially in the first part of his manuscript, are sound.

9.4 Pfister's bounds

We have already (in (9.1)) stated Pfister's celebrated 2^n-bound on the number of squares r_i^2 in (5.4.1), i.e., in the special case that the ground field K is real closed. Combining this with Cassels, Ellison, Pfister [1971] and Cassels [1964] yields $n + 2 \leq P(R(X)) \leq 2^n$ for all $n \geq 2$; nothing sharper is known on this (cf. Delzell [1987, p. 262] for details). To obtain his bound, Pfister studied a special class of quadratic forms, now called Pfister forms.[52] Changing R to $\mathbf{Q}$ entails the use of completely different techniques; we have $P(\mathbf{Q}(X_1)) = 5$ Pourchet [1971]

[51] Kreisel mentioned the existence of bad points in [1974a] and [1978] (but in [1977a, p. 116] he wrote as if the question were open). Not knowing this, Swan asked the same question in [1977, p. 228]; Choi and Lam [1977] replied by showing (or at least stating) that for $n \leq 2$, bad points never occur, while for every (n, d) with $n > 2$ and (even) $d \geq 4$ (except possibly $(3, 4)$), they do. They also gave the simplest-yet examples: $f = X_1^4X_2^2 + X_2^4X_3^2 + X_3^4X_1^2 - 3X_1^2X_2^2X_3^2$ and $f = X_4^4 + X_1^2X_2^2 + X_2^2X_3^2 + X_3^2X_1^2 - 4X_1X_2X_3X_4$. Similar ideas are in Brumfiel [1979, end of §8.5]. Meanwhile, *strictly* definite *homogeneous* $f \in \mathbf{R}[X]$ have no bad points other than the origin (reading between the lines of Habicht [1940] as we did in (5.2) above, or looking at (5.10) of Prestel [1975–84]); more generally, for psd f, the bad points lie among the zeros of f, a fact observed by Prestel and Knebusch in 1976 (Choi, Lam [1977]), and obvious from Stengle's theorem (5.14). In my thesis (1980) I showed that the codimension of the bad set is ≥ 3 (sharp).

[52] The latter continue to find applications, for example, in Scheiderer's solution [1989] of Bröcker's problem: A *basic open semialgebraic set S* is a set of the form $S = \cap_{i=1}^{r}\{\, x \in R^n \mid p_i(x) > 0 \,\}$, for some $r \geq 0$ and $p_i \in R[X]$; find the least $r := r(n)$ such that every such S can be written in this way. Scheiderer showed that $r(n) = n$.

and $6 \leq P(\mathbf{Q}(X_1, X_2)) \leq 8$ (Kato [1986], Appendix by J.-L. Colliot-Thélène); for $n \geq 3$ we do not know even whether $P(\mathbf{Q}(X)) < \infty$.

There is nothing underground about the bounds in the previous paragraph, so if that were the whole story, they would not belong in this section. But there is no published bound on the *degrees* of the r_i resulting from Pfister's transformations. In [1974], replying to Kreisel, Pfister showed (at least for $n > 3$) that if $r_i \in R(X) \quad (1 \leq i \leq l + 2^n)$ have degrees $\leq D$, then there are $s_j \in R(X) \quad (1 \leq j \leq 2^n)$ with degrees $\leq C(n)^{\frac{n^l-1}{n-1}} D^{n^l}$ such that $\sum r_i^2 = \sum s_j^2$; the constant $C(n)$ depends only on n, and could be determined explicitly; he speculated that it grows quickly with n.

L. Mahé [1990] extended Pfister's methods from Artin's theorem to variants of Stengle's *Positivstellensatz* (5.14), obtaining bounds (slightly more complicated than 2^n) on the number of squares in the left- and right-hand sides of (variants of) (5.14.1); as with Pfister's bound, K must be real closed, so that the nonnegative weights $p_I \in K$ can be absorbed into the squares.

9.5 Scowcroft's bounds

Scowcroft sketched a proof of Stengle's theorem in [1988, pp. 70–1] that is free of any constructively dubious principle. As he mentioned, the general recursive bounds for Stengle's theorem, obtained as in (5.5) and (5.11), are provably total in formal arithmetic (intuitionistic or classical, both having the same provably total recursive functions, as Kreisel observed in [1952] and [1958a]). He, too, mentioned that (weak) König's lemma suffices for the Artin-Schreier theory, but he did not attempt to exploit the conservation of **WKL** over **PRA** to get primitive recursiveness, as had been done (5.13) for Artin's theorem. (This is not to say that the latter isnecessary for constructive validity.)

9.6 Lombardi's bound

Lombardi got the following bound [1992] on the *degrees* of the summands in (5.14.1) (e.g., f^{2l+1}, $p_I G_I h_I^2 f$, etc.), and hence of the resulting r_i in (5.7.1)):

$$2^{\cdot^{\cdot^{2^{b+d(\lg d)+\lg\lg(1+s)}}}}\Big\}n+4,$$

where b is a constant; "lg" was not defined in the abbreviated, published exposition [1992]; but on p. 54 of the full-length version (the appendix to his Mémoire d'habilitation, 1990), he defined lg as the logarithm to base 2. As mentioned at (5.14.2), in the special case of Hilbert's 17th problem (5.7.1), $s = 0$, which ruins $\lg\lg(1+s)$; presumably, the remedy is to re-define lg to mean the number of binary digits. Lombardi left it to "the courageous reader" to use his method to figure out a bound on the number of squares; presumably it is similar. (Lombardi has informed me that he and M.-F. Roy have recently obtained elementary recursive bounds, consisting of a stack of approximately 5 exponentials; not all details have been written out, so for now, at least, their new bound belongs under the heading of §9.)

Lombardi did not apply logic to Stengle's original proof; so Kreisel might have described Lombardi's method as "ad hoc," from the point of view of unwinding (recall (1.9) above). Instead, Lombardi belongs to the Bishop school of Constructivism; for him the main purpose of finding a constructive proof [1988–91] of the *Positivstellensatz* (and the resulting primitive recursive bounds) was to establish its *validity.*

On pp. 325–6 he emphasized that his proof does not use the axiom of choice, implying, tacitly, that Stengle's use thereof was "essential." True, Stengle, like Artin, did appeal to (a variant of) the Artin-Schreier theorem ((2.1) and (5.1.1) above) that every formally real field K is orderable, and the original proof of this theorem did assume that—the underlying set of—K was well-ordered.

(a) In the first place, as Artin observed [1926, p. 110], at least for applications to theorems such as his (2.4.1) and Stengle's (5.14.1), we need apply the Artin-Schreier theorem only to fields of the form $K = \mathbf{Q}(b_1, \ldots, b_s)$, where the b_i are the coefficients of the given polynomial(s); such K can be recursively enumerated. (On the other hand, the Artin-Schreier method has another apparently non-constructive step: if a is algebraic over the formally real field K, then we are supposed to decide whether $K(a)$ is formally real or not; they did not say how, but this has nothing to do with the axiom of choice. Cf. also van der Waerden [1953, §71, Note following Theorem 7, p. 230].)

(b) Secondly, a little logical hygiene, courtesy of Kreisel, is salutary here: in [1956] (cf. also Krivine [1971, pp. 68, 97]) he observed that by relativising a proof to the constructible universe, the axiom of choice (and even the generalized continuum hypothesis and the axiom of constructibility) can be (*elementarily* recursively) eliminated from any proof of an *arithmetical* theorem,[53] a

[53] By Shoenfield [1961–66, p. 135], this result extends to Π^1_3-theorems; cf. Solovay [1990, pp. 18–9]. For the elimination of CH from proofs of theorems of second- and third-order arithmetic, cf. Platek [1969]; and even the *negation* of CH can be eliminated from proofs of Π^1_3-theorems of second-order arithmetic, by a result of

fact that has been used to sanitize proofs of results on the order of homotopy groups (by Serre) and on p-adic fields (by Ax and Kochen). And at least if $K = \mathbf{Q}$ (the most important example of a computable ordered field), Stengle's theorem is, indeed, arithmetic (in fact, it, like Artin's theorem, is Π_2, just as in (5.13)(c)). Thus, Stengle's theorem follows from ZF without further ado (cf. Delzell, González-Vega, Lombardi [1993, p. 63]).[54]

Corollary: It's not all that much harder to unwind a proof (say, of Stengle's theorem) with choice, than the corresponding proof without choice. This should be compared to the theme of Kreisel [1990b] and Kreisel, Macintyre [1982]: it's not much harder to unwind a proof (of a suitable theorem, e.g., Π_2) *with* the law of the excluded middle than without (and using only intuitionistic logic).

9.7 Summary of §9

No analysis (not even Pfister's or Mahé's, in (9.4)) of Artin's or Stengle's theorems has shown any prospect of yielding computational feasibility. Thus explicit bounds per se tend to fall between two stools: they are *not needed* for establishing the validity (constructive or otherwise) of Artin's or Stengle's theorems ((2.4.1), (5.14)), and, so far, they are *not enough* for—what mathematicians usually mean by—effective computation. Instead, a bound may be prized because it is elegant, memorable, and even applicable (e.g., Pfister's bound; cf. footnote 52).

10 Conclusion of Part I

10.1 Artin-Schreier theory largely constructive since 1955

In his review of A. Robinson's book, *Théorie métamathématique des Idéaux* (1955), Kreisel discussed what he called "the *fundamental principle*": for a first-order formula B and a set of first-order formulae A,

Solovay, mentioned in Moore [1990].

[54] While even ZF is not constructive, no single proof, such as Artin's or Stengle's, can use all the—infinitely many—axioms even of ZF (nor is ZF finitely axiomatizable: Montague [1961, 1966, p. 113]; sharpened by Kreisel, Levy [1968, p. 116]). So these theorems can obviously be proved in some *sub*system of ZF.

> "if $\neg B$ is consistent with every (B does not follow from any) finite subset of A then $\neg B$ is consistent with A" Kreisel [1956, p. 169].

After giving some examples of this principle, he wrote:

> "Robinson states that the methods used in the book are non-constructive; but it is not as bad as all that. Of course, if the set A (in the statement of the fundamental principle) is excessively ill-defined, the model he finishes with is ill-defined too; just as $(A \,\&\, B) \rightarrow (A \vee B)$ carries any ambiguity that may infect ill-defined propositions A and B. For reasonably well-defined classes A, a finitist version of the completeness theorem can be given by means of the ideas of Herbrand's aptly named *théorème fondamental.* (Of course, the algebraic terminology has to be tightened up too.) Herbrand's treatment is described very clearly in vol. 2 of Hilbert-Bernays' *Grundlagen der Mathematik.* The existence of such a treatment is the best reason for keeping to a non-constructive terminology (such as is employed in the present book); it means that we know how to 'read' the non-constructive terminology from a constructive point of view. It is reminiscent of the ϵ, δ definitions in analysis: once they are clearly understood one uses 'near', 'small', etc., quite freely and with confidence." Kreisel [1956, p. 173].

It seems to me that what is said here about ideal theory applies as well to Artin-Schreier theory, especially after Kreisel's sketches in the fifties.

In the whole Artin-Schreier theory (tacitly, when applied to *discrete* ordered fields), most of the techniques, such as Sturm's theorem, and its generalization by Tarski to quantifier-elimination, are already obviously finitary. At least, this was obvious to the experts of *Kreisel's* day—by now, so many new proofs of QE have been published that few modern readers look at Tarski's original proof; and, as mentioned in (4.4), some of these new proofs are less obviously constructive than Tarski's was. For another, more challenging example of a finitary result that was long known to insiders, there is Hollkott's [1941] finitary construction of the real closure of an arbitrary discrete ordered field, a thesis written under the direction of Artin, and later Zassenhaus[55]; it was never published, due to the war, but it was summarized in print

[55] I thank T. Sander in Dortmund for this information

by Krull [1942] and Zassenhaus [1970]. (Earlier constructions by Kronecker and Vandiver applied only to Archimedean ordered fields with a factorization algorithm—cf. Part II.)

There were, of course, some things that the experts did *not* know in those days:

1. Until 1955, it was not known whether Artin's theorem could be constructivized (again, tacitly, over *discrete* ordered fields).

2. Soon after the affirmative answer to (1) was found, the case of non-discrete ordered fields (such as the reals) became the next natural question (first raised in print in Kreisel [1962]); here the question can take different forms:

   (a) can the disjunctions in (5.8) or (5.9)(b) be contracted to a single identity (without introducing compound fractions—recall (6.38))? If not, then

   (b) can a continuously varying sum-of-squares representation be found?

   The answer to (a) is no (as mentioned in (9.3)(b)(i)); the answer to (b) is yes (cf. Delzell [1984] for real closed ground fields, and [1993] for arbitrary ordered ground fields; cf. also Delzell, González-Vega, Lombardi [1993], González-Vega, Lombardi [1993], and González-Vega, Lombardi [to appear]).

Finally, there was (only?) one result in Artin-Schreier theory that was shown to require a subtler hypothesis before constructivization is possible, even over discrete fields. That result is (2.1) (or (5.1.1)): Eršov gave [1977] a recursively presented formally real field with no recursive field ordering. This actually poses little problem "in practice," since his field has pathologies with which the K arising in Artin's main theorem (2.4), for example, are not afflicted. On the contrary, Eršov also proved that if a recursively presented (formally real) field has only finitely many orderings, then they are all recursive. (This was refined by Metakides and Nerode in [1979, pp. 314–5].)

Summary. In view of Kreisel's two sketches, and in view of what was known even earlier, the view that the bulk of Artin-Schreier theory still needs "epistemological rehabilitation" appears far-fetched. Indeed, few fields in mathematics have been more successfully constructivized than Artin-Schreier theory had been by 1955.

10.2 What was published, and what was known

Given the above situation, the facts that Kreisel, in the 1950's, (a) did not even mention such "issues" as whether QE is constructive over discrete ordered fields (4.4), or whether the real closure can be constructed finitarily (10.1), while, on the other hand, (b) he did deal with the then *challenging* questions, such as that actually asked by Artin about his own theorem, or whether intuitionistic logic could help for the case of non-discrete ordered fields, show, I think, that (c) he already knew as much as the experts of the day, and (d) he (like Artin) had a sense of *proportion* in these matters, i.e., as to what was *worth* mentioning.

For some (certainly not all) modern Constructivists, however, a sense of proportion is anathema: *all* old theorems must be rehashed by someone who has declared his allegiance to Bishop, before they can be cited in later publications. All exercises, however uninspired, must be worked out in detail, with credit claimed. All possible complexity bounds, however useless, must be estimated with the utmost precision. (Kreisel's commentary on the latter, on p. 35 of [1992], is just right—especially his hilarious footnote 7.)

Sometimes the new, politically correct expositions remain silent on the question whether there was anything seriously non-constructive about the original proofs (recall (4.4), or the end of §7). And even when this point is addressed, the diagnosis is often shallow (such as citing the "use" of the axiom of choice in a proof of an arithmetical theorem (9.6)(b)).

While Kreisel's sketches on Artin's theorem were sloppy by mathematical standards (footnotes 13, 23, 24, and 39; and subsections (5.6), (5.9), (7.2), and (9.1) above), and not presented in complete form, it is clear to me that he had worked out all the essential steps. Some evidence has already been presented here in Part I; more will come in Part II.

Moreover, Kreisel did more than constructivize individual theorems, such as Artin's: he showed how *general* methods from proof theory enable one to unwind *classes* of theorems. For example, Stengle's [1974] *Positivstellensatz* (5.14) belongs to roughly the same class as Artin's. When I needed the former in 1980 for the construction of a continuous version of Artin's theorem, I asked Kreisel whether Stengle's theorem could be proved constructively; he said yes, by the same general methods used in his unwinding of Artin's theorem. At the time, I did not understand those methods (not having read Hilbert, Bernays, for example), but I took his word for it that students of constructivity knew enough about Stengle's theorem—and I used it freely in [1984] (the other steps used in [1984] could be recognized as constructive even without Kreisel's insights of the 1950's).

So when Lombardi published a detailed constructive proof of Stengle's theorem in [1988–91]—not based on Scowcroft's 2-page, proof-theoretic sketch in [1988, pp. 70–1] (recall (9.5) above), and not consciously based on Kreisel's sketches; recall (1.4) and (9.6) above—it reminded me of Abel's and Jacobi's publications of material that Gauss had, to some degree, discovered 30 years earlier. Cox [1984] investigated Gauss' notebooks, sketchy publications, and letters, indicating that he had developed, around 1800, an enormous, rich theory of the arithmetic-geometric mean and theta functions:

> "Aside from the many other projects Gauss had to distract him, it is clear why he never finished this one: it was simply too big. Given his predilection for completeness, the resulting work would have been enormous. Gauss finally gave up trying in 1827 when the first works of Abel and Jacobi appeared. As he wrote in 1828, 'I shall most likely not soon prepare my investigations on transcendental functions which I have had for many years—since 1798—because I have many other matters which must be cleared up. Herr Abel has now, I see, anticipated me and relieved me of the burden in regard to one third of these matters, particularly since he carried out all developments with great concision and elegance' (see Gauss, *Werke,* X.1, p. 248)." Cox [1984, p. 328].

The parallel with Kreisel cannot be pushed too far. For one thing, Kreisel does not share the *aims* of most others who have worked on constructivity. For another, he doesn't seem to share Gauss' predilection for completeness. On the contrary, (a) Kreisel makes fun of "systematic" expositions, at least as written by others, and (b) it seems to me that the kind of lengthy, tedious calculations that typically arise in estimating bounds are simultaneously beneath and beyond Kreisel; he doesn't even *read,* much less write, such calculations. Thus I doubt that he had really worked out the details of the unwinding of Stengle's theorem, even privately; he was, probably, simply speaking off the cuff when he said that it could be unwound; as often happened, time proved him right.

References

Arnon, D.S., and Buchberger, B., eds.

[1988] *Algorithms in Real Alg. Geo.,* Ac. Press, 1988; *J. Sym. Comp.,* 5.

Arnon, D.S., and Mignotte, M.

[1988] On mechanical quantifier elimination for elementary algebra and geometry, in Arnon, Buchberger [1988, pp. 237–59].

Artin, E.

[1926] Über die Zerlegung definiter Funktionen in Quadrate, *Abh. math. Sem. Hamburg*, 5 (1926) 100–15; *Coll. Papers,* S. Lang, J.T. Tate, eds., Springer, 1965, pp. 273–88.

[1957] *Geometric Algebra,* Interscience, 1957.

Artin, E., and Schreier, O.

[1926] Algebraische Konstruktion reeller Körper, *Abh. math. Sem. Hamburg*, 5 (1926) 85–99; *Coll. Papers,* pp. 258–72.

[1927] Eine Kennzeichnung der reell abgeschlossenen Körper, *Abh. math. Sem. Hamburg*, 5 (1927) 225–31; *Coll. Papers,* pp. 289–95.

Basu, S., Pollack, R., and Roy, M.-F.

[199?] On the combinatorial and algebraic complexity of quantifier elimination, in preparation.

Benacerraf, P., and Putnam, H., eds.

[1964–83] *Philosophy of math.,* Prentice-Hall, 1964; Cambridge U., 1983.

Ben-Or, M., Kozen, D., and Reif, J.

[1986] The complexity of elementary algebra and geometry, *J. Computer and system sci.*, 32 (1986) 251–64.

Bochnak, J., Coste, M., and Roy, M.-F.

[1987] *Géométrie algébrique réelle*, Springer, 1987.

Bourbaki, N.

[1952–64] *Éléments de Mathématique*, Livre II: Algèbre; Chap. 6: Groupes et corps ord.; Chap. 7: Modules sur les anneaux princ., Fasc. XIV, Hermann, Vol. 1179 of Actualités Sci. Indust. (ASI), 1952, 1964.

[1954–60] *Éléments de Mathématique*, Livre I: Théorie des Ensembles; Chap. 1: Description de la math. formelle; Chap. 2: Théorie des ensembles; Fasc. XVII, Hermann, Vol. 1212 of ASI, 1954, 1960.

[1957–66] *Éléments de Mathématique*, Livre I: Théorie d. Ensem.; Chap. 4: Struct.; Fasc. XXII, Hermann, Vol. 1258, ASI, 1957, 1966.

Bridges, D., and Richman, F.

[1987] *Varieties of Constructive Mathematics,* Cambridge Univ., 1987.

Bröcker, L.

[1978] Über die Pythagoraszahl eines Körp., *Arc. Math.,*31 (1978) 133-6.

F. Browder, ed.

[1976] *Mathematical Developments Arising from Hilbert Problems,* Proc. Symp. Pure Math., 28, AMS, 1976.

Brown, J.

[1965] *Hilbert's Seventeenth Problem: A constructive proof, from the Herbrand Theorem*, Master's Thesis (supervised by J. Denton), McGill Univ., 1965, $2+43$ pages.

Brumfiel, G.W.

[1979] *Partially Ord. Rings & Semialg. Geom.,* Cambridge Univ., 1979.

Cajori, F.

[1908] Arithmetik, Algebra, Zahlentheorie; Chapter XX of *Geschichte der Mathematik,* 4, M. Cantor, ed., Teubner, 1908, pp. 37–198.

Cassels, J.W.S.

[1964] On the representation of rational functions as sums of squares, *Acta Arith.*, 9 (1964) 79–82.

Cassels, J.W.S., Ellison, W.J., and Pfister, A.

[1971] On sums of squares and on elliptic curves over function fields, *J. Number Theory*, 3 (1971) 125–49.

Choi, M.D., and Lam, T.-Y.

[1977] An old question of Hilbert, *Proc. Conf. Quadratic Forms—1976,* Queen's Papers Pure Appl. Math., 46 (1977) 385–405.

Choi, M.D., Knebusch, M., Lam, T.-Y., and Reznick, B.

[1982] Transversal zeros and positive semidefinite forms, *Géométrie Alg. Réelle et Formes Quad.,* Springer LNM, 959, 1982, pp. 273–98.

Coste, M., and Roy, M.-F.

[1979] Topologies for real algebraic geometry, *Topos theoretic methods in geom.,* Various Publ. Ser. 30, A. Kock, ed., Aarhus Univ., 1979.

[1988] Thom's lemma, the coding of real alg. numbers and the computation of the top. of semi-alg. sets, in Arnon, et al [1988, pp. 121–9].

Cox, D.A.

[1984] The arithmetic-geometric mean of Gauss, *L'Enseign. Math.*, 30 (1984) 275–330.

Daykin, D.E.

[1961] *Hilbert's 17th and Other Problems*, Ph.D. Thesis (examined by B.H. Neumann & R. Rado), Univ. Reading, 1961, $\mathrm{xi}+170$ pages.

Delzell, C.N.

[1982a] A finiteness theorem for open semi-alg. sets, with applications to Hilbert's 17th problem, *Ordered Fields and Real Alg. Geom.,* Dubois, Recio, eds., AMS Contemp. Math., 8, 1982, pp. 79–97.

[1982b] Case distinctions are necessary for representing polynomials as sums of squares, *Proc. Herbrand Symp. Logic Coll. 1981,* J. Stern, ed., North-Holland, 1982, pp. 87–103.

[1984] A continuous, constructive solution to Hilbert's 17th problem, *Invent. math.*, 76 (1984) 365–84.

[1987] Continuous Pythagoras numbers for rational quadratic forms, *J. Number Theory*, 26 (1987) 257–73.

[1993] Continuous, piecewise-polynomial functions which solve Hilbert's 17th problem, *J. reine angew. Math.*, 440 (1993) 157–73.

[1994] Nonexistence of analytically varying solutions to Hilbert's 17th problem, *Recent Advances in Real Alg. Geom. and Quad. Forms*, Jacob, et al, eds., AMS Contemp. Math., 155, 1994, pp. 107–17.

Delzell, C.N., González-Vega, L., and Lombardi, H.

[1993] A continuous and rational solution to Hilbert's 17th problem and several cases of the Positivstell., *Computational Alg. Geom.,* Eyssette, et al, eds., Birkhäuser, Progr. Math., 109, 1993, pp. 61–75.

Dieudonné, J.

[1946] Sur les corps ordonnables, *Boletim de Sociedade de Matemática de São Paulo*, 1 (1946) 69–75. *MR,* 9 (1948) 77.

van den Dries, L.

[1977] Artin-Schreier theory for commutative regular rings, *Ann. Math. Logic*, 12 (1977) 113–50.

[1982] Some applications of a model theoretic fact to (semi-)algebraic geometry, *Indag. Math.*, 44 (1982) 397–401.

[1988] Alfred Tarski's elimination theory for real closed fields, *J. Symb. Logic*, 53 (1988) 7–19.

Dubois, D.W.

[1967] Note on Artin's solution of Hilbert's 17th problem, *Bull. Amer. Math. Soc.*, 73 (1967) 540–1.

Dubois, D.W., and Efroymson, G.

[1970] Algebraic theory of real varieties. I. *Studies and Essays Presented to Yu-Why Chen on his Sixtieth Birthday*, 1970, pp. 107–35.

Erdös, P., Gillman, L., and Henriksen, M.

[1955] An isomorphism theorem for real-closed fields, *Ann. Math.*, 61 (1955) 542–54.

Eršov, J.L.

[1977] Theorie der Numerierung III, *Z. math. Logik Grundlagen Math.*, 23 (1977) 289–371.

Feferman, S.

[1974] Applications of many-sorted interpolation theorems, *Proc. Tarski Symp.* In: *Proc. Symp. Pure Math.,* 25, AMS, 1974, pp. 205–23.

[1995] Kreisel's "unwinding" program, this volume.

Ferrari, P.L.

[1987] A note on a proof of Hilbert's second ε-theorem, *J. Symb. Logic*, 52 (1987) 214–5.

[1989] The rank function and Hilbert's second ε-theorem, *Z. math. Logik u. Grundlagen d. Math.*, 35 (1989) 367–73.

Flannagan, T.B.

[1975] On an extension of Hilbert's second ε-theorem, *J. Symb. Logic*, 40 (1975) 393–7. (A correction is in Ferrari [1987], [1989].)

de Foncenex, D.

[1759] Reflexions sur les quantités imaginaires, *Miscellanea Taurinensia*, 1 (1759) 113–46.

Friedman, H.

[1979] On fragments of Peano arithmetic, *Mimeo*, May 1979, 7 pages.

[1981] On bounds for Hilbert's 17th problem I, *Mimeo*, 1981, 59 pages.

Friedman, H., Simpson, S.G., and Smith, R.L.

[1983] Countable algebra and set existence axioms, *Ann. Pure Appl. Logic,* 25 (1983) 141–81. Addendum, *APAL,* 28 (1985) 319–20.

Gauss, C.F.

[1799] Demonstratio Nova Theorematis Omnem Functionem Algebraicam . . . , Inaug. Diss., Helmstädt, 1799; *Werke*, 3, 1876, pp. 1–30. Cf. also Netto, *Die vier Gauss'schen Beweise . . . ,* W. Engelmann, 1890, 1904, 1913; Struik, *A Source Book in Mathematics, 1200–1800,* Harvard, 1969; van der Waerden, *A history of algebra . . . ,* Springer, 1985; *Encykl. math. Wiss. . . . ,* Vol. 1, Part 1, Teubner, 1900–1904, pp. 233–7; and *Encycl. Sci. Math.,* Tome I, Vol. 2, Fasc. 1, Teubner and Gauthier-Villars, 1907, pp. 189–205.

[1815] (Announcement of [1816]), *Göttingensche Gelehrte Anzeigen*, 204 (1815) 2017–9; *Werke*, 3 (1876) 105–6.

[1816] Demonstratio Nova Altera Theorematis Omnem Funct. Alg. . . . , *Comment. Soc. Regiae Sci. Gottingensis Recent.,* 3 (1816) 107–34; *Werke*, 3 (1876) 31–50. Cf. references in [1799]; D.E. Smith, *A Source Book in Math.,* 1, McGraw-Hill, 1929, Dover, 1959.

Gödel, K.

[1972] On an extension of finitary mathematics which has not yet been used (1972), in [1990, pp. 271–80]:

[1990] *Collected Works II*, S. Feferman, et al, eds., Oxford Univ., 1990.

González-Vega, L., and Lombardi, H.

[1993] A real Nullstellensatz and Positivstell. for the semi-polynomials over an ordered field, *J. Pure Appl. Algebra*, 90 (1993) 167–88.

[199?] Smooth parametrizations for several cases of the Positivstellensatz, *Math. Z.,* to appear.

Gorin, E.A.

[1961] Asymptotic properties of polynomials and algebraic functions of several variables, *Uspehi Mat. Nauk*, 16 (1961) 91–118; transl. *Russian Math. Surveys*, 16 (1961) 93–119.

Habicht, W.

[1940] Über die Zerlegung strikte definiter Formen in Quadrate, *Comm. Math. Helv.*, 12 (1940) 317–22; cf. Hardy, et al [1952–91]:

Hardy, G.H., Littlewood, J.E., and Pólya, G.

[1934–91] *Inequalities*, Cambridge Univ. Press, 1934, 1952, . . . , 1991.

Heegner, K.

[1952] Diophantische Analysis und Modulfunktionen, *Math. Z.*, 56 (1952) 227–53; cf. Mordell [1969, pp. 74–5].

van Heijenoort, J.

[1967] *From Frege to Gödel: A Source Book . . . ,* Harvard Univ., 1967.

Heintz, J., Recio, T., and Roy, M.-F.

[1991] Algorithms in real alg. geom. and applic. to computational geom., *DIMACS Discr. Math. Theor. Comp. Sci.,* 6 AMS (1991) 137–63.

Henkin, L.

[1960] Sums of squares, *Summaries of Talks . . . Summer Inst. Symb. Logic in 1957 at Cornell Univ.* Inst. Defense Analyses, Princeton, 1960, pp. 284–91. Reviewed by A. Robinson, *JSL,* 31 (1966) 128.

Herbrand, J.

[1929] *Recherches sur la théorie de la démonstration*, Thesis, Univ. Paris (written 1929, submitted 1930); cf. also W.D. Goldfarb, *Jacques Herbrand: Logical Writings*, Harvard, 1971, pp. 44–202.

Hilbert, D.

[1888] Über die Darstell. definiter Formen als Summe von Formenquad., *Math. Ann.*, 32 (1888) 342–50; *Ges. Abh.*, 2, 1933, pp. 154–61.

[1893a] Über die vollen Invariantensysteme, *Math. Ann.,* 42 (1893) 313–73; *Ges. Abh.,* 2, 1933, pp. 287–344. Cf. M. Ackerman, *Hilbert's Invariant Theory Papers*, Math Sci Press, 1978, pp. 225–301.

[1893b] Über ternäre definite Formen, *Acta. Math.*, 17 (1893) 169–97; *Ges. Abh.,* 2, 1933, pp. 345–66.

[1899] *Grundlagen der Geometrie.* Teubner, 1899 (transl. Open Court, 1902). (In footnote 38, we do *not* refer to later eds., e.g., 1971.)

[1900] Mathematische Probleme, *Göttinger Nachrichten* (1900) 253–97, and *Archiv der Math. u. Physik* (3rd ser.), 1 (1901) 44–53, 213–37; *Ges. Abh.,* 3, 1935, pp. 290–329. Transl. in *Bull. Amer. Math. Soc.,* 8 (1902) 437–79; and in Browder [1976, pp. 1–34].

[1923] Die logischen Grundlagen der Mathematik, *Math. Ann.,* 88 (1923) 151–65. *Ges. Abh.,* 3, 1935, pp. 178–91.

[1926] Über das Unendliche, *Math. Ann.,* 95 (1926) 161–90. Transl. in van Heijenoort [1967, pp. 367–92]. Transl. of first half ($\neq$ the half on CH) also in Benacerraf and Putnam [1964–83, pp. 183–201].

[1928] Die Grundlagen der Mathematik, *Abh. math. Sem. Hamburg,* 6 (1928) 65–85; *Grund. d. Math.,* Hamburger Math. Einzelschr., 5 Teubner, 1928, pp. 1–21; van Heijenoort [1967, pp. 465–79].

Hilbert, D., and Bernays, P.

[1934–70] *Grundlagen der Mathematik*, Vol. I (1934, 1968); Vol. II (1939, 1970), Grundl. Math. Wissens. ..., 40, 50, Springer. Kleene reviewed [1939] in *JSL,* 5 (1940) 16–20. Cf. also Leisenring [1969].

Hollkott, A.

[1941] *Finite Konstruktion geordneter algebraischer Erweiterungen von geordneten Grundkörpern,* Ph.D. Thesis, Univ. Hamburg, 1941, 65 pages. Review by Krull [1942]; summary in Zassenhaus [1970].

Hörmander, L.

[1955] On the theory of general partial differential operators, *Acta Math.,* 94 (1955) 161–248.

Isaacson, D.

[1967] A Constructive Solution of Hilbert's Seventeenth Problem by use of the Fundamental Theorem of Herbrand, Senior Undergraduate Thesis (supervised by Dreben), Harvard College (1967), 42 pages.

Jacobson, N.

[1964–75] *Lectures in Abstract Algebra, Vol. III: The Theory of Fields and Galois Theory;* D. van Nostrand (1964); Springer (1975).

[1974–89] *Basic Algebra*, 1 (1974, 1985); 2 (1980, 1989), W.H. Freeman.

Kato, K.

[1986] A Hasse principle for two-dimensional global fields (Appendix by J.-L. Colliot-Thélène), *J. reine angew. Math.,* 366 (1986) 142–83.

Kleene, S.C.

[1952] *Introduction to Metamathematics*, D. van Nostrand, 1952.

Kohlenbach, U.

[1992] Effective bounds from ineffective proofs in analysis: an application of functional interp. ..., *J. Symb. Logic,* 57 (1992) 1239–73.

Kreisel, G.

[1951–52] On the interpretation of non-finitist proofs—Parts I and II, *J. Symb. Logic,* 16 (1951) 241–67; and 17 (1952) 43–58 (erratum p. iv). Reviewed by Rosser in *JSL,* 18 (1953) 78–80.

[1956] Some uses of metamath., *British J. Phil. Sci.,* 7 (1956) 161–73.

[1957a] Hypotheses on algebraically closed extensions, *Bull. Amer. Math. Soc.,* 63 (Whole No. 647) (1957), 99.

[1957b] Hilbert's 17th problem, I and II. Same as [1957a]: 99–100.

[1958a] Mathematical significance of consistency proofs, *JSL,* 23 (1958) 155–82. Reviewed by A. Robinson, *JSL,* 31 (1966) 129; by Lorenzen, *MR,* 22 (1961) #6710; and by Schütte, *Zbl.,* 88 (1961) 15–6.

[1958b] Hilbert's programme, *Dialectica,* 12 (1958) 346–72; revised and expanded in Benacerraf and Putnam [1964–83].

[1958c] Wittgenstein's remarks on the foundations of mathematics, *British J. Phil. Sci.,* 9 (1958) 135–58.

[1960a] Sums of squares, *Summaries Talks ... Summer Inst. Symb. Logic, 1957, Cornell Univ.* Inst. Defense Analyses, Princeton, 1960, pp. 313–20. Reviewed by A. Robinson, *JSL,* 31 (1966) 128–9.

[1960b] Ordinal logics and the characterization of informal concepts of proof, *Proc. Internat. Cong. Mathematicians, 14–21 August 1958, Edinburgh*, J.A. Todd, ed., Cambridge, 1960, pp. 289–99.

[1962] Review of Goodstein, *MR,* 24A (1962) 336–7, #A1821.

[1965] Mathematical logic, *Lectures on Modern Mathematics*, Vol. III, T.L. Saaty, ed., John Wiley & Sons, 1965, pp. 95–195.

[1968] A survey of proof theory, *J. Symb. Logic,* 33 (1968) 321–88.

[1969] Two notes on the found.'s of set theory, *Dial.,* 23 (1969) 93–114.

[1974a] Letter to Mumford, May 21, 1974. *The Kreisel Papers*, Stanford University Archives.

[1974b] Reply to a letter from A. Pfister of October 1, 1974. *The Kreisel Papers*, Stanford University Archives.

[1977a] On the kind of data needed for a theory of proofs, in *Logic Colloq. 76*, R. Gandy, M. Hyland, eds., North Holland, 1977, pp. 111–28.

[1977b] Review of *L.E.J. Brouwer, Collected Works, Vol. I ...,* A. Heyting, ed.: *Bull. Amer. Math. Soc.,* 83 (1977) 86–93.

[1978] Review of Ershov, *Zbl.,* 374 (1978) #02027.

[1981] Neglected possibilities of processing assertions and proofs mechanically: Choice of problems and data, in *Univ.-Level Computer-Assisted Instruction at Stanford: 1968–1980,* P. Suppes, ed., Inst. Math. Studies Soc. Sci., Stanford Univ. 1981, pp. 131–147.

[1984] Frege's foundations & intuitionist. logic, *Monist,* 67 (1984) 72-91.
[1987] Proof theory: some personal recollections; Appendix, *Proof Theory*, by G. Takeuti, *2nd* ed., North-Holland, 1987, pp. 395–405.
[1990a] Logical aspects of computation: contributions and distractions, *Logic & Comp. Sci.,* Odifreddi, ed., Acad. Press, 1990, 205–78.
[1990b] Review of Troelstra and van Dalen [1988], *Bull. London Math. Soc.,* 22 (1990) 505–11.
[1992] On the idea(l) of log. closure, *Ann. Pu. Ap. Log.,* 56 (1992) 19-41.

Kreisel, G., and Krivine, J.L.
[1967–72] *Elements of Math. Logic (Model Theory)*: North-Holland, 1967, 1971 (English); Dunod, 1967 (French); Springer, 1972 (German).

Kreisel, G., and Levy, A.
[1968] Reflection princ. and their use for establishing the complexity of axiomatic syst., *Z. math. Logik Grund. Math.,* 14 (1968) 97–142.

Kreisel, G., and Macintyre, A.
[1982] Constructive logic versus algebraization I, *L.E.J. Brouwer Cent. Symp.*, Troelstra, et al, eds., North-Holland, 1982, pp. 217–60.

Krivine, J.L.
[1971] *Introduction to Axiomatic Set Theory*, D. Reidel, 1971.

Kronecker, L.
[1882] Grundzüge einer arith. Theorie der alg. Grössen, *Fest. Kummer's Fünfzigjährigem Doctor-Jubiläum 1881,* Reimer, 1882; *J. reine ang. Math.,* 92 (1882) 1–122; *Werke*, 2, 1897, 1968, pp. 237–387.

Krull, W.
[1942] Review of Hollkott [1941], *Zbl.,* 26 (1942) 198–9.

Ladrière, J.
[1951] Le théorème fondamental de Gentzen, *Revue phil. Louvain,* 49 (1951) 357–84. Reviewed by Fischer in *JSL,* 19 (1954) 129–30.

Lam, T.Y.
[1973] *The Algebraic Theory of Quadratic Forms*, Benjamin, 1973.
[1984] An intro. to real alg., *Rocky Mount. J. Math.,* 14 (1984) 767–814.

Lang, S.
[1953] The theory of real places, *Ann. Math.,* 57 (1953) 378–91.
[1965–93] *Algebra*, Addison-Wesley, 1965, 1971, 1993. (The false statement mentioned in (5.10) above was omitted from [1971–93].)

de Laplace, P.S.

[1812] Sur la rés. des éq. Théorème sur la forme de leurs racines imag., *Leçons Math. École Norm. 1795* (Cinq. Séance), in *J. École Poly.,* VII^{e} et VIII^{e} Cahiers (1812); *Œuvres,* 14, 1912, pp. 53–65.

Lazard, D.

[1988] Quantifier elimination: Optimal solution for two classical examples, in Arnon, Buchberger [1988, pp. 261–6].

Leisenring, A.C.

[1969] *Mathematical Logic and Hilbert's ε-theorems,* MacDonald Tech. & Sci.; Gordon & Breach, 1969. (Correction in Flannagan [1975].)

Littlewood, J.E.

[1948] Large numbers, *Math. Gazette,* 32 (1948) 163–71; *A Mathematician's Miscellany,* 1953; *Littlewood's Miscell.,* Bollobás, ed., 1968.

de Loera, J.A., and Santos, F.

[1995] An effective version of Pólya's theorem on positive definite forms, *J. Pure Appl. Algebra,* in press.

Łojasiewicz, S.

[1965] Ensembles semi-analytiques, I.H.E.S., prépublication (July 1965).

Lombardi, H.

[1988–91] Théorème des zéros réel effectif, et variantes, *Théorie des nom.,* Publ. Math. Besançon 1988–89, Fasc. 1, 27 pp. Théor. eff. des zéros réel . . . , Mém. d'habilit., 1990, 59 pp. Nullst. réel eff. et var., *CRAS,* Ser. I, 310 (1990), 635–40; Eff. real Null. and var., *Eff. Meth. Alg. Geom.,* Mora, et al, eds., Birkhäuser, 1991, pp. 263-88.

[1990] Une étude historique sur les problèmes d'effectivité en algèbre réelle, Memoire d'habilitation (1990).

[1992] Une borne sur les degrés pour le théorème des zéros réel effectif, *Real Algebraic Geometry,* Coste, et al, eds., Springer LNM, 1524, 1992, pp. 323–45; full-length version in Mém. d'habilit., 1990.

[1994] Relecture constructive de la théorie d'Artin-Schreier, preprint.

Lombardi, H., and Roy, M.F.

[1991] Constructive elementary theory of ordered fields, *Effec. Methods Alg. Geom.,* Mora, et al, eds., Birkhäuser, 1991, pp. 249–62; longer, French: *Théor. des nom.,* Publ. Math. Besançon 1990–91.

Łoś, J., Mostowski, A., and Rasiowa, H.

[1956] A proof of Herbrand's theorem, *J. Math. Pures et Appliqués,* 35 (1956) 19–25.

Mahé, L.

[1990] Level and Pythagoras number of some geometric rings, *Math. Z.,* 204 (1990) 615–29.

McKenna, K.

[1975] New facts about Hilbert's 17th problem, *Model Theory and Algebra ...,* D. Saracino, V.B. Weispfenning, eds., Springer Lect. Notes Math., 498, 1975, pp. 220–30. Cf. also Jacobson [1989].

Metakides, G., and Nerode, A.

[1979] Effective content of field theory, *An. Mat. Log.,* 17 (1979) 289-320.

Mints, G.E.

[1976] What can be done with PRA?, *Zap. Nauch. Sem. Leningradskogo ...,* 60 (1976) 93–102. Transl. *J. Sov. Math.,* 14 (1980) 1487–92.

Montague, R.

[1961–66] Fraenkel's addition to the axioms of Zermelo, *Essays Found. Math., Ded. to Fraenkel ...,* Y. Bar-Hillel, et al, eds. Jerusalem Acad. Press, 1961; Magnes Press, Hebrew U., 1966, pp. 91–114.

Mordell, L.J.

[1969] *Diophantine Equations,* Pure Appl. Math. 30, Acad. Press, 1969.

Moore, G.H.

[1990] Introductory note to *1947* and *1964*, in Gödel [1990, pp. 154–75].

Pfister, A.

[1967] Zur Darstellung definiter Funktionen als Summe von Quadraten, *Invent. math.,* 4 (1967) 229–37. English versions: Browder [1976], Lam [1973], Jacobson [1980, 1989].

[1974] Reply (October 1, 1974) to a letter from G. Kreisel. *The Kreisel Papers*, Stanford University Archives.

Platek, R.A.

[1969] Elimination of the continuum hypothesis, *JSL,* 34 (1969) 219–25.

Pourchet, Y.

[1971] Sur la représentation en somme de carrés des polynômes à une indet. sur un corps de nombres ..., *Acta Arith.,* 19 (1971) 89–104.

Prestel, A.

[1975–84] *Lectures on Formally Real Fields*, IMPA Lecture Notes, 22, Rio de Janeiro, 1975; Springer Lecture Notes in Math., 1093, 1984.

[1978] Remarks on the Pythagoras and Hasse number of real fields, *J. reine angew. Math.,* 303/304 (1978) 284–94.

Prieß-Crampe, S.

[1983] *Angeordnete Strukturen ...,* Ergeb. Math., 98, Springer, 1983.

Renegar, J.

[1991] Recent progress on the complexity of the decision problem for the reals, *DIMACS Ser. Discr. Math...,* 6, AMS, 1991, pp. 287–308.

Reznick, B.

[1995] Uniform denominators in Hilbert's seventeenth problem, *Math. Z.,* in press.

Robinson, A.

[1955] On ord. fields & def. forms, *Math. An.,* 130 (16 Dec. 1955) 257-71.

[1963] *Introduction to Model Theory and to the Metamathematics of Algebra,* North-Holland, 1963.

Rose, H.E.

[1984] *Subrecursion: functions & hierarchies,* Clarendon, Oxford, 1984.

Sacks, G.

[1972] *Saturated Model Theory,* Math. Lect. Notes Ser., Benjamin, 1972.

Scarpellini, B.

[1969] On the metamathematics of rings and integral domains, *Trans. Amer. Math. Soc.,* 138 (1969) 71–96.

Scheiderer, C.

[1989] Stability index of real varieties, *Invent. math.,* 97 (1989) 467–83.

von Schubert, F.T.

[1798] De inventione divisorum, *Nova acta scient. imp. Petropolitanae,* T. XI, ad annum 1793. Petropoli 1798, pp. 172–182.

Schütte, K.

[1960–77] *Beweistheorie,* Springer, Grundlehren math. Wissen..., 103, 1960; revised & transl. in Vol. 225, 1977.

Scowcroft, P.

[1988] A transfer theorem in constructive real algebra, *Ann. Pure and Appl. Logic,* 40 (1988) 29–87.

[1989] Some continuous Positivstellensätze, *J. Alg.,* 124 (1989) 521–32.

Seidenberg, A.

[1954] A new decision meth. for elem. alg., *Ann. Mat.,* 60 (1954) 365–74.

[1956] Some remarks on Hilbert's Nullstell., *Arch. Mat.,* 7 (1956) 235-40.

Serre, J.P.

[1949] Extensions de corps ordonnés, *CRAS,* 229 (1949) 576–7.

Shoenfield, J.R.

[1961–66] The problem of predicativity, *Essays on the Found. Math., Ded. A.A. Fraenkel...,* Bar-Hillel, et al, eds. Jerusalem Acad. Press, 1961; The Magnes Press, The Hebrew Univ., 1966, pp. 132–9.

[1967] *Mathematical Logic,* Addison-Wesley Series in Logic, 1967.

Sieg, W.

[1985] Fragments of arithmetic, *Ann. Pure Appl. Logic,* 28 (1985) 33–71.

Skewes, S.

[1933] On the diff. $\pi(x) - \mathrm{li}(x)$ (I), *J. Lond. Math. Soc.,* 8 (1933) 277–83.

[1955] On the difference $\pi(x) - \mathrm{li}(x)$ (II), *Proc. London Math. Soc.* (Third series), 5 (1955) 48–70.

Smale, S.

[1992] Theory of computation, *Mathematical Research Today and Tomorrow: Viewpoints of Seven Fields Medalists ...,* C. Casacuberta, et al, eds., Springer LNM, 1525, 1992, pp. 59–69.

Smirnov, V.A.

[1971] Élimination des termes ε dans la logique intuitionniste, *Logique et Méthodologie des Sciences en U.R.S.S.* Vol. 25, N° 98, Fasc. 4 (1971) of *Revue Internat. Phil.,* 512–9.

Solovay, R.M.

[1990] Introductory note to *1938, ..., 1940*, in Gödel [1990, pp. 1–25].

Springer, T.A.

[1952] Sur les formes quad. d'indice zero, *CRAS,* 234 (1952) 1517–9.

Stengle, G.

[1974] A Nullstellensatz and a Positivstellensatz in semialgebraic geometry, *Math. Ann.,* 207 (1974) 87–97.

Swan, R.G.

[1977] Topological extensions of projective modules, *Trans. Amer. Math. Soc.,* 230 (1977) 201–34.

Tait, W.W.

[1981] Finitism, *J. Phil.,* 78 (1981) 524–46; erratum, 618.

Tarski, A.

[1930–67] *A Decision Method for Elementary Algebra and Geometry*, Rand Corp., 1948; UC Press, Berkeley, 1951; announced in *Ann. Soc. Pol. Math.,* 9 (1930; published 1931) 206–7; and in *Fund. math.,* 17 (1931) 210–39. Page proofs of *The completeness of elementary algebra and geometry* were prepared in 1940 for *Actualités Sci. Indust.*, but not published until 1967, by Inst. B. Pascal.

[1954] Prime ideal theorems for set algebras and ordering principles. Preliminary report. *Bull. Amer. Math. Soc.,* 60 (1954) 391.

Thom, R.

[1965] Sur l'homologie des variétés algébriques réelles, in *Differential and Combinatorial Topology,* Cairns, ed., Princeton, 1965, pp. 255–65.

Troelstra, A.S., and van Dalen, D.

[1988] *Constructivism in Mathematics: An Introduction*, Vols. I and II, North-Holland, 1988. Reviewed by Kreisel in [1990b].

Vandiver, H.S.

[1936] On the ordering of real algebraic numbers by constructive methods, *Ann. Math.,* 37 (1936) 7–16.

Vaughan, R.C.

[1992] Foreword to the 1992 edition of A.E. Ingham's *The Distribution of Prime Numbers,* Cambridge Tracts Math. 30, 1932, 1990, 1992.

van der Waerden, B.L.

[1931–70] *Modern Algebra*, Vol. I and II; Springer, 1931, ..., 1967 (German); Frederick Ungar, 1949, 1953, 1970 (English).

Wessels, L.

[1977] Cut elimination in a Gentzen-style ε-calculus without identity, *Z. math. Logik u. Grundlagen d. Math.,* 23 (1977) 527–38.

Whiteley, W.

[1991] Invariant computations for analytic projective geometry, *J. Symb. Computation,* 11 (1991) 549–78.

Witt, E.

[1934] Zerlegung reeller algebraischer Funktionen in Quadrate. Schiefkörper ..., *J. reine angew. Math.,* 171 (1934) 4–11.

Zassenhaus, H.

[1970] A real root calculus, *Computational Problems in Abstract Algebra*, J. Leech, ed., Pergamon, 1970, pp. 383–92.

Kreisel's "Unwinding" Program

by Solomon Feferman

Through his own contributions (individual and collaborative) and his extraordinary personal influence, Georg Kreisel did perhaps more than anyone else to promote the development of proof theory and the metamathematics of constructivity in the last forty-odd years. My purpose here is to give some idea of just one aspect of Kreisel's contributions to these areas, namely that devoted to "unwinding" the constructive content of *prima-facie* non-constructive mathematical proofs.[1] This program was the subject of his first remarkable papers in the 1950's, and it has drawn his repeated attention ever since.

Anyone who is familiar with even a small part of Kreisel's writings knows that he inverts the usual ratio of technical work to discussion. He takes much greater pains to explain, at length, the significance of the work than to set it out in an organized step-by-step fashion. His attitude seems to be that if one has the right ideas, the details will look after themselves. And they did amazingly often (or, he could rely on more disciplined collaborators to look after them). However, in the specific area dealt with here, there are several important cases where the expected details are either problematic or simply missing. This can't help but affect my main aim here, which is to assess whether the

[1]The full range of Kreisel's contributions to proof theory and constructivity certainly deserves exposition and critical evaluation, but providing such would require a *much* more substantial investment of effort than that which was found possible here.

work on Kreisel's unwinding program, both his own and that of others, lives up to its claims. My overall conclusion is that while the general *theory* of unwinding launched by Kreisel is eminently successful, the supposed *applications* to date are few and far between; moreover, in some prominent cases, their status *even as applications* has to be put in question.

Kreisel's unwinding program was a reaction to Hilbert's consistency program. It aimed to substitute clear mathematical results for what were said to be vague, misplaced, crude foundational goals. But, as with his work on constructivity, Kreisel also sought to replace those by a more sophisticated stance about foundations, to be advanced by the technical results. In the end we must ask, as well, to what extent he was successful in doing so.

I believe that the general direction and character of Kreisel's contributions were influenced considerably by his early studies. Kreisel commenced university studies in mathematics at Trinity College, Cambridge in 1942, at the age of 19. He reports in several places (e.g. [1987, p. 395]) having read the second volume of Hilbert and Bernays' *Grundlagen der Mathematik* in that same year, so the level of his logical sophistication and direction of interests was already exceptional then.

In mathematics, Cambridge was noted for its predominant concentration on analysis and number theory, with such famous exemplars as Hardy, Littlewood, and Besicovitch. There were no logicians on the faculty (Russell was long gone), but some reinforcement of Kreisel's interests in logic and foundational matters was to be found in the lectures he attended on the philosophy of mathematics offered by his compatriot, the stormy petrel Ludwig Wittgenstein. According to Monk [1990, p. 498], in 1944 Wittgenstein declared Kreisel to be "the most able philosopher he had ever met who was also a mathematician."

During the war, Kreisel joined the British Admiralty in 1943 where he worked on problems of hydrodynamics and naval engineering until 1946. This experience with applied mathematics no doubt honed his skills with classical analysis but may also have encouraged a casual disregard for mathematical fastidiousness. Kreisel returned to Cambridge as a Research Fellow in 1946–48, receiving an M. A. in 1947. During this period he engaged in regular discussions with Wittgenstein on the philosophy of mathematics, but in later years he was extremely critical of Wittgenstein's ideas in this respect (cf. Monk [1990, pp. 499 and 642]). For the details of the further progress of Kreisel's career, cf. his Vita in this volume.

1 The Interpretation of non-Finitist Proofs

The title of this section is taken from that of the fundamental paper for Kreisel's program (K.P.) for unwinding non-constructive proofs; it appeared in *The Journal of Symbolic Logic* in two parts, Kreisel [1951], [1952]. This work modifies Hilbert's program (H.P.) in several important respects but initially retains some of the same language in a way that does not adequately convey the shifts of emphasis. Later, in Kreisel [1958], improved ways were used for describing what is to be accomplished, and thus [1958] supersedes [1951]and [1952] to a certain extent. However, to respect the historical progression of ideas, we shall not mix the two, but return to [1958] in §5 below.

For both H.P. and K.P. there is a basic syntactic distinction between free-variable (quantifier-free) formulas A_0 and those formulas A which contain bound variables. H.P. restricts attention to decidable A_0 with *free individual variables* $x, y, \ldots$ ranging over the natural numbers. Kreisel allows consideration of A_0 containing, in addition, *free function variables* $f, g, \ldots$; these are supposed to be decidable for each substitution instance by specific numerals and computable functions. A_0 is said to be *verifiable* if it is correct for each such substitution instance.

A necessary, but clearly not sufficient, condition on finitist proofs is that they consist of verifiable free variable formulas. Kreisel is careful to avoid claiming any characterization of finitist proofs ([1951, p. 242, ftn. 2]).[2] At any rate, proofs which contain formulas with bound variables (even just individual variables) are *non-finitist* on their face. The Hilbert program aimed to establish the consistency of formal systems for arithmetic, analysis, etc., by finitist reasoning. For systems containing a modicum of arithmetic, this would insure that free variable formulas $A_0(x_1, \ldots, x_n)$ established in such formalisms by non-finitist proofs (i.e. which contain some formulas with bound variables) are verifiable and, indeed, finitistically provable. However, even where H.P. succeeds, this tells us nothing about proofs of formulas A with bound variables. The "problem of non-finitist proofs" that Kreisel poses in [1951], [1952] is how to give "finitist sense" to such formulas.

To make this more precise, and to avoid the problem of saying just what is finitist (skirted as indicated above), Kreisel introduces the idea of an *interpretation*, which applies to systems Σ within an effectively specified formal language $\mathcal{L}(\Sigma)$; Σ itself need not be effectively given (for reasons to be explained below). An interpretation of Σ is taken to

[2]But in Kreisel [1960] he would later propose a formal characterization of the notion of finitist proof; cf. Kreisel [1987, p. 396] for one of his more recent assessments of that venture.

be an effective association with each formula A in $\mathcal{L}(\Sigma)$ of a sequence of free variable formulas $A_0^{(n)}$ such that:

(I_1) for each proof of A in Σ we can find an n such that $A_0^{(n)}$ is verifiable,

(I_2) for each proof of $\neg A$ in Σ and each n we can find a substitution instance which makes $A_0^{(n)}$ false.

There is also an obvious third condition which relates the interpretation of B to that of A when B is proved from A in Σ.

The best known example of an interpretation is provided by Herbrand's Theorem for the classical 1st order predicate calculus Σ. This associates with each formula A a sequence of formulas $A_0^{(n)}$ each of which is a finite disjunction of substitution instances of the quantifier-free matrix of A when A is taken in prenex form; details of the association will be recalled in the next section. It turns out that Herbrand's Theorem applies to extensions Σ of the predicate calculus by arbitrary verifiable formulas; these need not be effectively given.

Remark. The above notion of interpretation is too broad to be really useful as a general theoretical tool. As Kreisel points out, every usual formal system Σ with decidable proof predicate $\mathrm{Prf}_\Sigma(x, y)$ ("x is the number of a proof in Σ of the formula with number y") admits a *trivial interpretation* of each formula A by a single formula $A_0(x)$, namely $\neg\ \mathrm{Prf}_\Sigma(x, \ulcorner \neg A \urcorner)$. One would want the $A_0^{(n)}$ in an interpretation to have a closer contentual relation to A, and to be informative about that content. In this respect, the modified notion of interpretation introduced in Kreisel [1958] is an improvement; cf. §5 below. At any rate, (I_1), (I_2) serve for initial orientation to what is to be accomplished.

2 The No Counterexample Interpretation

The main result of Kreisel [1951], [1952] is to provide an informative recursive interpretation of classical 1st order arithmetic and its extension by verifiable free variable formulas. This is the so-called *no counterexample interpretation* (n.c.i.). Its form is illustrated by Π_3^0 formulas A, which in any case cover the specific applications to be discussed later. So let us consider

(1) $A \equiv \forall x \exists y \forall z R(x, y, z)$

with R quantifier-free. The n.c.i. is to be contrasted with the "naive" interpretation of A, which proceeds via its *Skolem normal form*, the 2nd order formula

$$(1)_{\text{S}} \quad \exists f \forall x, z R(x, f(x), z).$$

The naive interpretation seeks to find a computable f satisfying $(1)_{\text{S}}$ when (1) is proved. Simple recursion-theoretic examples show that, in general, this cannot be done when $\forall x \exists y \forall z R(x, y, z)$ is proved in arithmetic.[3]

In the classical predicate calculus, Skolem normal form is the appropriate one to consider for *satisfiability*; its dual, the *Herbrand normal form*, in this case

$$(1)_{\text{H}} \quad \forall x, f \exists y R(x, y, f(y)),$$

the more appropriate one for *validity* and hence (by completeness) for derivability. This is equivalent to

$$(2) \qquad \neg\exists x, f \forall y \neg R(x, y, f(y))$$

which, by Skolem form, is equivalent to

$$(3) \qquad \neg\exists x \forall y \exists z \neg R(x, y, z),$$

i.e. to A itself. A pair x, f such that $\forall y \neg R(x, y, f(y))$ would provide a *counterexample* to A; hence (2) and so, also $(1)_{\text{H}}$ can be read as asserting that *there is no counterexample to A*. The Herbrand form and its obvious generalization to arbitrary prenex A is the common formal starting point of both Herbrand's Theorem (H.T.) and Kreisel's n.c.i. for arithmetic.

Let us recall, briefly, some details of H.T. [4] For Σ the classical predicate calculus, extend the language $\mathcal{L}(\Sigma)$ by new free function variables $f, g, \ldots$ Then it is shown that A is derivable in Σ just in case the 1st-order version of its Herbrand form, i.e.

$$(4) \qquad \exists y R(x, y, f(y))$$

is derivable; moreover, that – according to H.T. – holds just in case there is a finite disjunction of substitution instances

$$(5) \qquad R(x, t_1, f(t_1)) \vee \cdots \vee R(x, t_k, f(t_k))$$

[3]The formula $\forall x \exists y \forall z [T(x, x, y) \vee \neg T(x, x, z)]$ using Kleene's T predicate provides one example; Kreisel gives an example from elementary analysis in Appendix I to [1952].

[4]Cf., e.g., Shoenfield [1967, pp. 52–55].

which is tautologous. Note that the t_i in (5) are built up by the function symbols in R together with f from the variable x and constants in R. These can be assumed to be ordered in such a way that when new variables $z_1, \ldots, z_k$ are substituted for $f(t_1), \ldots, f(t_k)$, resp., the $t_1, \ldots, t_k$ are transformed into terms $s_1, \ldots, s_k$, such that the free variables of $s_i (i = 1, \ldots, k)$ are contained in $\{x, z_1, \ldots, z_{i-1}\}$. That is, we have a disjunction

$$(6) \qquad R(x, s_1(x), z_1) \vee \cdots \vee R(x, s_k(x, z_1, \ldots, z_{k-1}), z_k)$$

which is tautologous when A is provable; moreover, A is derivable from each instance of (6) by Herbrand's direct rules for the predicate calculus. Note that the formulas (6) lie back in the original language $\mathcal{L}(\Sigma)$.

Granted these facts about H.T., Kreisel shows straightforwardly in [1951] that the sequence $A_0^{(n)}$ of all possible disjunctive substitution instances of the form (6) constitutes an interpretation (in his sense, as described in sec. 2) of the predicate calculus. He then goes on to give, as a second such interpretation, the n.c.i. for the predicate calculus, by returning to the form (4). Basically, this is to regard the choice of y satisfying $R(x, y, f(y))$ for arbitrary x, f as given by explicitly defined *functionals*

$$(7) \qquad y = F(f, x)^5$$

which are found through those disjunctions (5) that are tautologous. Namely, y may be taken to be $t_i (\equiv t_i(x, f))$ for the first i such that $R(x, t_i, f(t_i))$ holds. The n.c.i. for the predicate calculus simply associates with each A of the form (1) the sequence $A_0^{(n)}$ of all possible formulas of the form

$$(8) \qquad R(x, F(f, x), f(F(f, x)))$$

in which the F are specific explicit functionals built from terms of $\mathcal{L}(\Sigma)$ using definition by cases.[6] Kreisel sketches a proof of this in [1951] for arbitrary prenex A, using, instead of H.T., the Hilbert ε-calculus and results about the Hilbert substitution method (for it) from Hilbert and Bernays [1939].

The n.c.i. for arithmetic takes on a similar shape, only the description of the functionals required as well as the proof are more involved. Here Kreisel makes use of Ackermann [1940], which extended to arithmetic the substitution method for the ε-calculus, with termination

[5] More mnemonically, one might write $y = Y(f, x)$.

[6] That is, the (if ... then ... else ...) construction.

proved by effective transfinite induction up to Cantor's ordinal ϵ_0, the first fixed point of $\omega^\alpha = \alpha$.[7] What Kreisel's result shows is that if a prenex formula A is provable in arithmetic then its n.c.i. functionals can be defined using primitive recursive schemata and schemata for effective transfinite recursion up to any ordinal $\alpha < \epsilon_0$. In the case of a Π^0_3 formula A as in (1), these functionals may specifically be determined by

$$(9) \qquad F(f,x) = (\mu y)R(x,y,f(y))$$

for all (f,x). Since transfinite induction up to each ordinal $\alpha < \epsilon_0$ is provable in arithmetic, Kreisel's result gives a complete characterization of the n.c.i. functionals for that system.

Kreisel further observes – and stresses repeatedly – that for both the predicate calculus and arithmetic, the class of functionals needed for the n.c.i. is unaffected by the adjunction of purely universal (i.e. Π^0_1) axioms B. For if A follows from B, then $B \to A$ is provable in the system and hence so also is $\neg B \vee A$; passing to the prenex form for the latter and applying the n.c.i. to that gives the desired result. For example, if

$$A \equiv \forall x \exists y \forall z R(x,y,z) \qquad \text{and} \qquad B \equiv \forall u S(u),$$

with S quantifier-free, then $\neg B \vee A$ goes into

$$\forall x \exists u, y \forall z \left[\neg S(u) \vee R(x,y,z)\right],$$

and its Herbrand form is

$$\forall x, f \exists u, y \left[\neg S(u) \vee R(x,y,f(y))\right].$$

This leads to two functionals $F(f,x)$, $G(f,x)$, with

$$\neg S(G(f,x)) \vee R(x, F(f,x), f(F(f,x))).$$

Hence

$$R\left(x, F\left(f,x\right), f(F(f,x))\right)$$

holds if $\forall u S(u)$ holds, and indeed, if $S(m)$ holds for some one suitable m (depending on x and f). The functionals F, G lie in the same classes as described before for the predicate calculus, resp. arithmetic. Note that it is the addition of arbitrary (verifiable) Π^0_1 axioms which

[7] Such use of induction up to ϵ_0 had first been applied by Gentzen [1936] to obtain a consistency proof of arithmetic by means of a partial extension of his earlier cut-elimination procedure for the predicate calculus.

gives rise to the non-effective systems Σ allowed in Kreisel's notion of interpretation in §1 above.

The special case in all this of Π^0_2 formulas $\forall x \exists y R(x,y)$ provable in arithmetic (possibly extended by arbitrary Π^0_1 axioms) is of particular interest. Since the variable 'z' in (1) is missing, the functional F in (9) simply reduces to a function

$$(10) \qquad F(x) = (\mu y)R(x,y).$$

Then F can be defined by function schemata for primitive recursion and ordinal recursion up to ordinals $\alpha < \epsilon_0$. Since the provable Π^0_2 formulas of arithmetic correspond exactly to its provably recursive functions, one arrives at the following:

Theorem (Kreisel [1951], [1952]) *The provably recursive functions of arithmetic are exactly those which are ordinal recursive of order* $< \epsilon_0$. *Moreover, the same holds of any consistent extension of arithmetic by* Π^0_1 *axioms.*

3 Proofs of the n.c.i.

While the nature of the n.c.i. and results for it are clear enough, the proofs are quite another matter. Kreisel's own proof of the n.c.i. for arithmetic requires familiarity with the ins and outs of Hilbert's ε-calculus and with its application in Ackermann [1940]. For those few brought up on Hilbert and Bernays [1939] (as, e.g., Kreisel himself) this would not be an obstacle. But for the rest of us, the proof in Kreisel [1951], [1952] does not invite detailed study; it just looks like a thicket. I myself have never tried to wade through it, and don't know anyone who has.[8] The review by J. Barkley Rosser ([1953]) is extremely critical on a number of scores, to begin with that

> "... it is very difficult reading. In the first place, the subject matter is of considerable complexity. In the second place, there are many errors, too many to permit a complete listing ... Most are typographical ..."

[8] I asked my colleague, Grigori Mints, an expert on Hilbert's ε-substitution method, whether he had ever studied Kreisel's proof. He said that he had not because by the time he learned of the significance of the results, there were more understandable proofs available [see below]. After writing the above, I learned from Charles Parsons that back in the 50's he had managed to work his way through the proof in Kreisel [1951], [1952], though only by relying on Ackerman [1940] to help fill in the details.

Then,

> "[T]he word 'finitist' is itself used in an unfamiliar sense ... and furthermore this sense is never carefully explained ... Apparently 'finitist' means 'constructive' with the additional restriction that no bound variables shall occur."

As we shall see, Kreisel would deal with this objection in his [1958] paper.

Finally, in his review Rosser registered disappointment with the nature of the n.c.i. itself, pointing to provable Σ^0_2 sentences $\exists x \forall y R(x, y)$ in particular. For these, he said, one would wish to find a specific constant k such that $R(k, y)$ holds for all y, and that doing so

> "... would be of great value, since it would enable one to transform one of the existing non-constructive proofs of Siegel's theorem into a constructive proof, which is badly needed."

Instead, the n.c.i. only gives an $x = F(f)$ such that $R(F(f), f(F(f)))$ holds for all f. Presumably, Rosser was here referring to a theorem of C.L. Siegel guaranteeing that certain kinds of diophantine equations have at most a finite number of solutions. Kreisel would deal with this general objection some thirty years later in his paper [1982], as will be discussed in §6 below.

The first new proof of the n.c.i. for arithmetic was obtained via Gödel's functional (*"Dialectica"*) interpretation of intuitionistic arithmetic (Gödel [1958]), as was pointed out by Kreisel himself in his paper [1959]. This route breaks up the work into several more understandable steps:

(i) first the negative translation of classical into intuitionistic arithmetic,

(ii) then the application of Gödel's interpretation to the translations of prenex formulas,

(iii) and finally ordinal analysis of the resulting functionals.

The first two steps are quite direct, and show that the n.c.i. functionals for arithmetic, which lie at type level 2 in the type hierarchy, are generated by the schemata for Gödel's primitive recursive functionals of finite type.[9] The third step, which shows that every type 2 primitive

[9]Cf. Troelstra [1990, p. 225] and, for more details with a variant of Gödel's interpretation, Shoenfield [1967, pp. 223–227].

recursive functional is ordinal recursive of ordinal $< \epsilon_0$, is established in Schwichtenberg [1975], following earlier work of Tait [1967].

As an aside, the history of the development of Gödel's functional interpretation is of incidental interest here, as explained in Feferman [1993]. Gödel had arrived at this interpretation by 1941, when he lectured on it at Yale University, though he did not publish the work until 1958. In the 1930s, Gödel had explained several possible routes to obtain a constructive consistency proof for arithmetic going as little as possible beyond finitist methods. In a lecture he gave in 1938 for an informal seminar organized in Vienna by Edgar Zilsel, he explained a way of looking at Gentzen's consistency proof for arithmetic which is a clear anticipation of the n.c.i. Shorthand notes for this lecture were found in Gödel's *Nachlass* and transcribed as part of the Gödel editorial project a few years ago. These lecture notes appear in the third volume of Gödel's *Collected Works* ([1995]), together with an illuminating introductory note by Wilfried Sieg and Charles Parsons.

For those versed in Gentzen-style proof theory, the most direct and perhaps cleanest route to the n.c.i. for arithmetic make use of effective infinitary versions of cut-elimination. An excellent exposition of such is to be found in Schwichtenberg [1977, pp. 884–892], with application to the provably recursive functions of arithmetic and the n.c.i.[10]

4 Application to Littlewood's Theorem

The final part of Kreisel [1952, pp. 51–65], is supposed, in his words, to provide an application "... of the ideas of the [n.c.i.] to a theorem of analytic number theory whose interpretation has given trouble", namely the result of Littlewood in 1914 that the functions of integers

1. $\psi(x) - x$, and
2. $\pi(x) - \mathrm{li}(x)$,

change sign infinitely often. Here $\psi(x)$ is the log of the l. c. m. of numbers $\leq x$, $\pi(x)$ is the number of primes $\leq x$ and $\mathrm{li}(x)$ is the logarithmic integral $\int_0^x (1/\log u)du$ (treated as an improper integral at $u = 1$), which is asymptotic to $\pi(x)$. Littlewood's theorem for (1) is evidently in the form $\forall x \exists y R(x, y)$ with R (primitive) recursive; the same holds for (2) by use of suitable recursive approximations for $\mathrm{li}(x)$. Littlewood's result was surprising since as far as had been calculated,

[10]Cf. also Girard [1987, pp. 481–482].

$\pi(x) < \mathrm{li}(x)$; cf. Ingham [1932, p. 7][11], which shows representative values up to $x = 10$ billion.

Littlewood's proof of his theorem exposited in Ingham [1932, ch. V], is non-constructive: in both (1) and (2) existence of a $y > x$ is demonstrated for which a sign change occurs; a bound for y is given explicitly in case the Riemann Hypothesis (R.H.) is false, but a different argument for the existence of y is used if R.H. is true, and no explicit bound for y emerges in that case (see also Littlewood [1948]). As Ingham [1932, p. 7] put it,

> "Littlewood's theorem is a pure 'existence theorem' and we still know no numerical value of x for which $\pi(x) > \mathrm{li}(x)$."

That situation was shortly remedied by S. Skewes in 1933 who used *a different proof* of Littlewood's theorem for $\pi(x) - \mathrm{li}(x)$, assuming R.H., to put a bound of $10_3(34)$ on the first change of sign, where

$$10_1(n) = 10^n \quad \text{and} \quad 10_{i+1}(n) = 10^{10_i(n)};$$

later work by others from the 1960s on (beginning with Lehman [1966]) lowered this bound considerably (cf. Ingham [1990, p. ix]). What Kreisel claimed to do in [1952, pp. 51–52], was analyze *the original proof* of Littlewood's theorem to show how one could extract recursive bounds, using the idea of the n.c.i. applied to statements of the form $\forall u S(u) \to \forall x \exists y R(x, y)$, where $\forall u S(u)$ is a Π^0_1 equivalent of the R.H.[12] Informally, this was to revolve around the feature of the n.c.i. that $S(m)$ would only be needed for finitely many m; as Kreisel put it ([1952, p. 54]),

> "if the conclusion ... holds when the Riemann hypothesis is true, it should also hold when the Riemann hypothesis is nearly true: not *all* zeros [of the Riemann zeta function] need lie on $\sigma = 1/2$, but only those whose imaginary part lie below a certain bound ... and they need not lie *on* the line $\sigma = 1/2$, but near it."

The nature of this application is, however, obscure in a number of respects, despite the claim made for it in an introductory section to Kreisel [1951, p. 247]:

---

[11] Ingham's famous 1932 expository monograph on the distribution of prime numbers has been reprinted without change in 1990; that edition contains an additional foreword by R.C. Vaughan with supplementary up-to-date information and references.

[12] Curiously, it is not explicitly stated in Kreisel [1952] that the R.H. can be put in Π^0_1 form; this was only brought out in his 1958 paper discussed in §5 below.

> "Since [Littlewood's] proof was not developed in a formal system there can be no question of applying the results of the present work to it in a precise sense. But if one examines the official proof in [Ingham [1932]] (and the usual proofs of the standard theorems on complex variables used) ... it might fairly be said, I think, that it applies the principles of [the n.c.i.] in a straightforward manner without introducing 'new ideas' of proof."

First of all, even though the original proof was not presented in a formal system, one would want to know whether it can be formalized in arithmetic, at least in principle. Indeed, Kreisel says ([1952, p. 52]) that we could be sure of finding the required bounds if the proof were written out in one of the extensions of arithmetic that he considers. He says that he has

> "... discussed elsewhere [no reference given] how proofs in large parts of the theory of functions of a complex variable can be presented in [the system] Z_μ [of arithmetic]. Here we shall only give a method of constructing rational approximations to zeros of computable regular functions. This will enable us to deal with the present problem if we remember our familiar principle [about the n.c.i. for implications with Π^0_1 hypotheses]."

He then launches into a description of such methods which seem closely related to the work on constructive approximations to zeros of analytic functions from Kreisel [1952a]. The work on this purported application then goes on to sketch some modifications of the proof of Littlewood's theorem in Ingham [1932, ch. V], making use of those constructive methods of approximation.[13] However, no further indication is given as to whether or how any of this is to be formalized in arithmetic.[14] Nor is any statement made as to what precise conclusion one could draw if it were taken for granted that it could all be thus formalized, so that one could apply the n.c.i. for arithmetic to it. Thus its status as an application of the logical work is in question. Finally – setting aside the logical aspects (or lack thereof) – just considered as a piece of

[13]This presumes familiarity on the reader's part with the exposition loc. cit.

[14]Later work by a number of researchers, including H. Friedman, G. Mints, and S. Simpson, has shown that the bulk of classical analysis of the sort applied in Ingham [1932] can in fact be carried out in systems conservative over primitive recursive arithmetic (PRA); cf. Feferman [1988] for references. Further direct evidence for the formalization in PRA of complex analysis applied to number theory is provided by Cegielski [1992].

work in analytic number theory, no explicit bounds are extracted that one can point to, certainly none that would satisfy the mathematicians interested in such matters. Indeed, one is hard put to say exactly what the conclusion of this work is; of this, more in the next section.

5 Pushing the Program

Kreisel's program (K.P.) is formulated in his paper [1958, p. 155] in an improved way, without the misleading reference to finitism, as follows:

> "*To determine the constructive (recursive) content or the constructive equivalent of the non-constructive concepts and theorems used in mathematics*, particularly arithmetic and analysis."

This is meant to replace Hilbert's program (H.P.) calling for finitist or (more generally) constructive consistency proofs, which he says has two defects:

> "(1) Since the notion of constructive proof is vague, the whole formulation of the program is vague ... (2) The formulation does not cover too well the actual substance of the material contained in [Hilbert-Bernays [1939]]; e.g. the ε-theorems for the predicate calculus go far beyond establishing mere consistency ...".

Kreisel adds in (1) that though the problem of providing an exact formulation of what constitutes a finitist or constructive proof is an interesting one for the logician, H.P. is not attractive to the mathematician because of its vagueness. Regarding the use of 'constructive' in the statement of K.P., which might similarly be questioned, he says that this may be formulated in a precise way using his notion of *recursive interpretation* ([1958, p. 160]). That is an improved version of the notion of interpretation from [1951], as follows: A recursive interpretation of a system Σ in a fragment Σ_0 consists of two recursive functions $A \mapsto \langle A_0^{(n)} \rangle$ and $p \mapsto \pi(p)$ such that

(RI$_1$) if A is a formula of Σ then the $A_0^{(n)}$ are formulas of Σ_0, and if p is a proof of A in Σ then $\pi(p)$ is a proof of some $A_0^{(n)}$ in Σ_0,

(RI$_2$) A can be proved from each $A_0^{(n)}$ in Σ,

(RI$_3$) if $A \equiv \forall x \exists y R(x, y)$ then $A_0^{(n)}$ is $R(x, F_n(x))$ where the F_n are an enumeration of the provably recursive functions of Σ.

Once more, H.T. for the predicate calculus and the n.c.i. for the predicate calculus and for arithmetic provide examples of recursive interpretations. These are explained in a general way, though some details about the n.c.i. for the predicate calculus with equality are presented in the Appendix to [1958]. The rest of that paper is devoted to a description of examples – treated in other publications – of applications of recursive interpretations to results in mathematics and logic. On the mathematical side, two are in algebra, namely to Hilbert's 17th problem and Hilbert's *Nullstellensatz*, and one is in arithmetic, namely the application to $\pi(x) - \mathrm{l}i(x)$ discussed in the preceding section; I shall not go into the logical applications here.

Both the algebraic applications make use of H.T. or the ε-theorems with equality. In the case of Hilbert's 17th problem, this is used to extract uniform primitive recursive bounds on the number and degrees in the representation as sums of squares of rational functions, of a positive semi-definite polynomial over a real field – from Artin's result guaranteeing such a representation. (As he reported at the Cornell logic conference in 1957, Kreisel had also obtained the same sort of bounds by a direct analysis, without use of logic, of Artin's original proof.) This interesting application is discussed in detail in Delzell's contribution to the present volume, and will not be dealt with further here. The application to the *Nullstellensatz* appears to be rather straightforward, granted elimination theory for algebraically closed fields. (However, see Kreisel [1992, p. 33, ftn. 6] which refers to a correction of a "sloppy aside on proof-theoretic aspects of the *Nullstellensatz*" in his [1958] paper and also belittles his own work on that by comparison with more recent sharp results by algebraists.) For both theorems, a comparison is made in [1958, p. 167], with closely related results obtained by model-theoretic methods.

So we turn again to the earlier application to Littlewood's theorem; here there are several supplements to the sketch in [1952] discussed in the preceding section. First of all (cf. ftn. 12 above) it is now stated explicitly that R.H. is equivalent to a statement in Π^0_1 form, by an argument in outline based on the methods of Kreisel [1952a].[15] Secondly, the statement of Littlewood's theorem is shown to be in Π^0_2 form by use of recursive approximations to $\mathrm{li}(n)$. Finally, the way in which the n.c.i. for arithmetic is to be applied is explained more clearly. Once more, it is stated that proofs in analytic number theory, and in particular that of Littlewood's theorem, can be formalized in arithmetic

[15] Another proof that R.H. is Π^0_1, using only some particular classical results about the Riemann zeta function, is given in full detail in Davis, Matijasevič and Robinson [1976].

> "... to yield the bounds, at least *in principle*"

and that

> "if one examines proofs in analytic number theory *with a view* to a formalization in Z, one does not run into difficulties ... If one is interested in finding bounds, one will naturally formalize the given informal proof in as constructive a manner as possible." (Kreisel [1958, p. 171]).

But one can't expect to do this in a mechanical manner; Kreisel says that he does not conceive of mathematical logic as "the mathematics for morons."

Nevertheless, once more, no precise conclusion is stated as to exactly what can be drawn from application of the n.c.i. for arithmetic to Littlewood's theorem. This is all the more disappointing, as in the interim Skewes [1955] had published what Kreisel (op. cit. p. 170) calls an *ad hoc* solution which gives "the same bound as the method of [Kreisel [1952]]". Indeed, Skewes makes use of a weakening (H) of R.H.[16], under which he obtains an explicit bound x_1, for the first change of sign, as $x_1 = \exp\exp\exp(7.703)$ (or $e_3(7.703)$) and, under its negation (NH), an explicit bound x_2 as $10_4(3)$ and (slightly better) also as $e_4(7.705)$. Despite the close relationship, Skewes [1955] does not refer to Kreisel [1952], and Kreisel [1958] does not refer to Skewes' specific bounds.

It is only in two much later papers that Kreisel states explicit bounds on the first change of sign N of $\pi(x) - \mathrm{li}(x)$ resulting from his 1952 work. First, in [1968, p. 362] he writes that this gives roughly $\exp\exp\exp\exp 8$ (or $e_4(8)$) as a bound for N, while by comparison, Lehman [1966] gives $(1.65)10^{1165}$. He goes on to say that:

> "In view of these improvements, the present value of Kreisel [1952] ... consists not in the bounds themselves, but only in analyzing the *general nature* of these problems; it separates what bounds are got from quite general considerations and what improvements need special study. This type of analysis is a typical logical contribution ...".

Later, in [1977, p. 114], Kreisel says:

---

[16] Similar in character to that of Kreisel [1952]. As to the order of priority, it should be noted that, according to Littlewood [1948, p. 169], Skewes had obtained a numerical bound "free of hypotheses" in 1937 and that the method used was to be found in Skewes' Ph.D. thesis deposited in the Cambridge University Library; apparently only the bound was improved in Skewes [1955]. *N.B.* This footnote was omitted in Littlewood [1953].

> "[A]n essentially routine application of proof theory (ε-theorems or cut-elimination) applied to Littlewood's original proof, extracts a bound for N; cf. Kreisel [1952]. Of course this cannot be expected to give optimal bounds, for which further ideas are needed, and it doesn't. It gives a bound of about
>
> $$10^{10^{10^{34}}} \text{ compared to } (1.65)10^{1165}$$
>
> found in Lehman [1966]."

There's something fishy about all this, since the bound $10_3(34)$ shown here is the same as that obtained in Skewes [1933] (under R.H.) by his alternative proof of Littlewood's theorem, and is different from the bound claimed in Kreisel [1968]. Moreover, no mention is made in either paper of the still different bound in Skewes [1955] which Kreisel had stated in the above quote from [1958] to be the same as that obtained in his 1952 paper.

6 Applications to Finiteness Theorems

In his retrospective report [1987], Kreisel says that by the mid 1950s he had come to assume – "wrongly" – that the potential for striking mathematical applications of proof theory was low. In the following twenty years, in his technical work Kreisel concentrated on a number of other aspects of proof theory and constructivity, with influential side forays into other parts of logic, such as predicativity and generalized recursion theory. But he says (on p. 395) that he learnt in the late 1970s of many areas

> "where mathematicians wanted to unwind proofs of Π^0_2 theorems, but were not able to do so without logical guidance."

At the same time he began to formulate general results on unwinding Σ^0_2 theorems which led to applications to finiteness theorems in number theory; these are the matters to which the present section is devoted.

Recall from the description of H.T. in §2 above that if

(1) $\quad \exists x \forall y R(x, y)$

is proved in the predicate calculus then there are terms $s_i(y_1, \ldots, y_{i-1})$, $i = 1, \ldots, k$ such that

(2) $\quad R(s_1, y_1) \vee R(s_2(y_1), y_2) \vee \; \cdots \; \vee R(s_k(y_1, \ldots, y_{k-1}), y_k)$

is tautologous. Of special interest in number theory are statements asserting that a certain set $\{x|C(x)\}$ is finite, which we write in the form (1) as:

$$(3) \qquad \exists x \forall y (C(y) \to y < x).$$

Then the corresponding Herbrand disjunction may be rewritten as

$$(4) \qquad C(y_1) \wedge \cdots \wedge C(y_k) \to$$
$$y_1 < s_1 \ \vee \ y_2 < s_1(y_1) \ \vee \ \cdots \ \vee \ y_k < s_k(y_1, \ldots, y_{k-1}).$$

Kreisel observed that in such a case, if we have effective control of the s_i relative to their variables y_j $(j < i)$ then we can provide a bound on the *number* of y's such that $C(y)$ and – in some cases – on their *size.* For the latter, an obvious sufficient condition is that we have a number c such that $s_i(y_1, \ldots, y_{i-1}) \leq c$ for $i = 1, \ldots, k$, but we could rarely expect to meet this in practice. More interesting and potentially useful is when we have control over the *growth* of the s_i, leading to a bound on the number of y's such that $C(y)$, as illustrated in the following.

Theorem (Kreisel [1982]) *Suppose $C(y)$ is a number-theoretical predicate and that $s_i(y_1, \ldots, y_{i-1})$ $(i = 1, \ldots, k)$ are functions such that (4) holds for all natural numbers $y_1, \ldots, y_k$. If c_i $(i = 1, \ldots, k)$ are such that*

(i) $s_1 \leq c_1$,

(ii) $s_i(y_1, \ldots, y_{i-1}) \leq c_i + \max(y_1, \ldots, y_{i-1})$ *for* $1 < i \leq k$,

then $\Sigma_{i=1}^k c_i$ *is a bound for the number of* y*'s such that* $C(y)$.

The proof of this theorem is quite easy, taking only eight lines in Kreisel [1982, p. 41]. Another proof of the same length is given in Luckhardt [1989] along with some other results using variant growth conditions.

It should be noted that it is not required in Kreisel's theorem (or its variants) that the statement of finiteness be proved in the predicate calculus or any other given formal system, nor that the s_i are terms in a formal language, nor that C is decidable. Thus the only role of Herbrand's theorem here has been to suggest appropriate forms to consider for applications. This is exactly what one finds in the work of Luckhardt [1989] on "Herbrand analyses" of two proofs of the famous theorem of Roth [1955] (also known as the Thue-Siegel-Roth theorem) about diophantine approximations.[17] Roth's theorem states that if α is an irrational algebraic number then for any $\epsilon > 0$, there are at most a finite number of $q > 1$ such that for some (unique) integer p with $(p, q) = 1$,

$$(5) \qquad \left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{2+\epsilon}}.$$

The theorem guarantees that suitably related diophantine equations have only a finite number of solutions. For each specific α and $\epsilon = 1/n$ (say), Roth's theorem takes the logical form

$$(6) \qquad \exists m \forall q (C(q) \rightarrow q < m)$$

where

$$(7) \qquad C(q) \equiv q > 1 \wedge (\exists! p \in \mathbb{Z})[(p, q) = 1 \wedge |\alpha - pq^{-1}| < q^{-2-\epsilon}],$$

C can be shown to be decidable and hence (6) is in Σ^0_2 form, though as mentioned above, that information is not needed for the applications. At any rate (6) is a *prima facie* candidate for an Herbrand analysis, i.e. for the identification of suitable $s_i(q_1, \ldots, q_{i-1})$ $(1 \le i \le k)$ satisfying

$$(8) \qquad C(q_1) \wedge \ldots \wedge C(q_k) \rightarrow \\ q_1 < s_1 \ \vee \ q_2 < s_2(q_1) \ \vee \ \cdots \ \vee \ q_k < s_k(q_1, \ldots, q_{k-1})$$

for all $q_1, \ldots, q_k$. What Luckhardt does in his 1989 paper is to extract such s_i from two proofs of Roth's theorem, namely Roth's own in 1955 and that of Esnault and Viehweg [1984]. In both cases, by careful examination of the growth conditions, he is able to improve the bounds previously obtained on the number of q's with $C(q)$. In particular, Luckhardt obtains a bound from the second proof which is polynomial of low degree in n and $\log d$, where $n = 1/\epsilon$ and $d =$ degree α. It turns out that the same bounds were obtained independently by Bombieri and van der Poorten ([1988]) without any (explicit) Herbrand analysis.

7 Assessment of the Mathematical Applications

Besides the applications of proof theory discussed or indicated above, Kreisel mentions several areas, "... L-functions, Galois cohomology,

[17] The idea for this specific direction of application was already advanced by Kreisel in 1970; cf. Kreisel [1977, pp. 114–115].

ergodic theory, topological dynamics" that he says he learned about in the late seventies where (as quoted above) "mathematicians wanted to unwind proofs of Π^0_2-theorems, but were not able to do so without logical guidance". No references are given for these and I have not been able to chase them all down. There is a brief discussion in Kreisel [1990, pp. 247–248] of the question of obtaining a lower bound for $L(1)$, where L is the Dirichlet L-function. One has a standard non-constructive proof that $L(1) > 0$; Kreisel refers to a *modified proof* which can be unwound by hand "since it has been done." (For this, some indications are to be found in Kreisel [1981, p. 139, ftn. 2], and [1981a, pp. 150–152].)

The mentioned application to topological dynamics is presumably that made by Girard [1987] to extract bounds from two forms of the Furstenberg and Weiss [1978] proof (by those methods) of the famous combinatorial theorem of van der Waerden ([1927]) on arithmetic progressions in partitions of the natural numbers. That theorem asserts the existence for any p and k of an n such that if $\{0, \ldots, n-1\}$ is partitioned into k classes $C_1, \ldots, C_k$, at least one C_i contains an arithmetic progression of length p. Girard [1987, pp. 237–251 and 483–496] first applies cut-elimination to an "ad hoc" modified form of the Furstenberg-Weiss proof. He later applies the n.c.i. to the Furstenberg-Weiss proof closer to that originally given. The unwindings are shown to lead to bounds for $W(p, k)$, the least n as a function of p and k, in a sub-recursive hierarchy. In the case of the first, modified proof, this is at the level of the Ackermann non-primitive recursive function; that is the same bound as obtained by inspection from the original van der Waerden proof. In the case of the second proof, the bound for W extracted is at a somewhat higher level.

Discussion. What can we say now about these various claimed applications of proof theory to the unwinding of *prima-facie* non-constructive mathematical proofs? Here I would raise several questions:

1. Have there been enough applications so that one can speak of a definite direction of work, with a clear past and promising future?

2. In what sense are these really applications of proof theory?

3. Do the applications provide the kind of specific information sought by mathematicians?

As to 1, I think it is fair to say the number of applications is still disappointingly small. Moreover, the nature of the past applications is mixed (to the extent that one understands their "nature" at all), and the prospects for the future are far from clear. This is not an existing or emerging direction of work that one would propose for a thesis topic without a great deal of hesitation.

Concerning 2, here again the history is mixed. As we have seen, the unwinding of Littlewood's theorem, such as it is, is not really an application of the n.c.i., nor is the use of "Herbrand analysis" in the proofs of the Roth theorem an application of Herbrand's theorem. On the other hand, Kreisel's first treatment of Artin's solution of Hilbert's 17th problem did involve a genuine application of Herbrand's theorem or the ε-theorems[18], and Girard's treatments of the Furstenberg-Weiss proof are genuine applications of cut-elimination and the n.c.i. Of course, it is recognized that one does not have, in each of these cases, a matter of blindly formalizing existing proofs in some formal system and mechanically applying the transformations provided in principle by the relevant proof theory. Rather, these applications are "genuine" on their face because they apparently involve steps that correspond to those transformations in a significant way. On the other hand, there is no integral involvement of proof theory in the purported applications to the Littlewood and Roth theorems. The discussion in Luckhardt [1989, pp. 206–261], of the latter is pertinent. Luckhardt there makes the following points (among others):

(i) Herbrand's theorem (H.T.) is *not* necessary in principle ("*nicht prinzipiell nötig*")[19] in his applications;

(ii) knowledge of H.T. provides the concepts needed to interpret existing mathematical proofs as falling under a logical pattern; and

(iii) applications to specific cases are not simply obtained by substitution into a gross ("plump") scheme.[20]

With 'n.c.i.' or 'cut-elimination' or 'normalization' equally well substituted here for 'H.T.', I think (i)–(iii) pretty much speak for themselves in response to question 2 for most of the purported applications.

Finally, as to question 3, in the cases that one can make comparisons at all, the answer is: Yes ... but. There is no need to repeat the ambiguous outcome on the Littlewood theorem described at the conclusion of §5. And, as mentioned at the end of the preceding section, Luckhardt's improvement in the bound from the Esnault-Viehweg proof

[18]Cf. the discussion in Delzell's contribution to this volume).

[19]In fact, as I have stressed in §6, the use of H.T. was merely to suggest the *form* of the statements to be considered; the theorem itself is not used at all.

[20]See also Kreisel [1990, pp. 250–253], for further discussion along these lines.

was obtained independently by Bombieri and van der Poorten without any appeal to logic. Finally, in the case of the van der Waerden theorem, the main question that had interested combinatorists was whether the function W could be given a primitive recursive bound, i.e. essentially lower than the original Ackermann function bound. Here the striking result is that obtained by Shelah [1988] by a novel elementary combinatorial proof, with no use of logic, that gives a bound for W at level 4 in the "fast-growing" hierarchy.[21] For a typical appreciation by a combinatorist of that improvement, see the review Spencer [1990] of Shelah's paper.

As a postscript to the discussion of questions 1–3 above, one should mention the recent work of Kohlenbach [1993], [1993a], which concerns applications of a variant of Gödel's functional interpretation to results in Chebycheff approximation theory. The results in question fall under the general form of uniqueness theorems: for all $u \in U$, $v_1, v_2 \in V_u$,

$$G(u, v_1) = \inf\{G(u, v) : v \in V_u\} = G(u, v_2) \to v_1 = v_2$$

where U, V are complete separable metric spaces, V_u is compact in V and $G : U \times V \to \mathbb{R}$ is continuous. Kohlenbach's applications extract from classical proofs of results of the form above effective "moduli of uniqueness" Φ satisfying, for all $u \in U$, $v_1, v_2 \in V_u$, $n \in \mathbb{N}$:

$$\prod_{i=1}^{2} \big(G(u, v_i) - \inf\{G(u, v) : v \in V_u\} \leq 2^{-\Phi un}\big) \to d_V(v_1, v_2) \leq 2^{-n},$$

Note that this is an *a priori* estimate since Φ does not depend on v_1, v_2. There is no question that such unwindings are a genuine application of logic and, on the face of it, provide the kind of specific information of interest to mathematicians. Indeed, Kohlenbach's work apparently yields improvement of known estimates in connection with Chebycheff approximation. It is obviously premature to say *how* interesting this will be to the mathematicians in that field, and whether this direction of work augurs well for the future of applications of unwinding more generally.

8 K.P. versus H.P.

Kreisel's retrospective report ([1987]) begins:

---

[21] That is, at level f_4 where $f_1(x) = 2x$, $f_{i+1}(x) = f_i^{(x)}(1)$.

> "Like many others but particularly Gödel [1931] and Gentzen [1936] (on p. 564) who expressed their reservations discretely [sic] I was repelled by Hilbert's exaggerated claim for consistency as a sufficient condition for mathematical validity or some kind of existence."

Kreisel's assimilation of Gödel and Gentzen to his view in this respect is misleading, to say the least. Certainly Gentzen was totally committed to the Hilbert program, though he was more cautious than Hilbert about what we may claim to be achieved by a finitist consistency proof. And there is considerable evidence that Gödel took certain relativized forms of the Hilbert program (with consistency as its main aim) seriously all through his career. This is to be found both in his published work and in unpublished lectures whose texts have been found in Gödel's *Nachlass* and which appear in the third volume of his *Collected Works*.[22,23]

Be that as it may, Kreisel goes on to say, in the passage in question,

> "[B]ut unlike most others I was not only attracted by the logical wit of consistency proofs (which I learnt in 1942 from Hilbert-Bernays Vol. 2) but also by the so to speak philosophical question of making explicit the additional knowledge provided by those proofs (over and above consistency itself)."

He says that his answers took two forms:

> "(i) particular applications to mathematical proofs ... [and] (ii) general formal criteria such as functional interpretations to replace the incomparable condition of consistency; 'incomparable' because the aim of functional interpretations is meaningful without restriction on metamathematical methods."

In other words, the ("philosophical") aim was to substitute clear-cut mathematical results for inconclusive philosophical ones; the appeal of these moves (i) and (ii) is evident.

But despite the hoopla, the results in direction (i) – at least to date – have been largely disappointing. What about direction (ii)? Here I count the move a success, with such prime exemplars as the

[22] Cf. also Feferman [1993].

[23] Kreisel [1958a] and [1968] also formulated a relativized reductive form of H.P.; see Feferman [1988] for my own version which took that as its point of departure.

n.c.i. in Kreisel [1951] and Gödel's functional interpretation in his paper [1958] together with its many interesting extensions to systems of analysis and down to fragments of arithmetic (cf. Troelstra [1990] and Feferman [1993]), as well finally as the many results, using these methods and cut-elimination, classifying the provably recursive functions of various systems of arithmetic, analysis and set theory.[24] While these results may only (or primarily) be of interest to logicians, they are certainly presented in mathematically understandable terms independent of any normative – and hence (possibly) disputable – foundational doctrines. This direction of work is surely a success story and a going industry.[25] Apparently Kreisel has lost interest in it (perhaps for that very reason); for example, he is contemptuous of the "flashy precision" of ordinal analyses, which are one of the main technical tools in the characterization of the provably recursive functions of various theories. One can't fault him for losing interest in a subject that he helped launch years back; but, with or without his approval, this direction of work as a whole is sure to march on.

Finally, what about the foundational motivations for proof theory which (i) and (ii) were intended to replace? Here, I find Kreisel truly ambivalent. Over the course of his career he has certainly taken the philosophy of mathematics seriously and has devoted considerable thought and writing to it. And, while he railed constantly against the perpetuation of simple-minded traditional doctrines, substantial portions of his own work were motivated by more sophisticated foundational – as opposed to mathematical – concerns.[26] Indeed, in the period from the 1950's on when logic as technology was overtaking logic as a foundational tool, Kreisel led the way in promoting foundational concerns as the driving force in the pursuit of proof theory and constructivity, and those of us who followed him did so for exactly that reason. In more recent years, though, Kreisel has become increasingly dismissive of any attempt at systematic philosophical efforts, at least as practiced by everybody else, and perhaps even himself. I'm afraid that here, also, we will continue to be inspired by those original foundational aims while subjecting their development to his perennial insistence on critical examination — and re-examination, and re-examination. It is in that spirit that this piece is dedicated to Kreisel.

[24]Cf. e.g. Takeuti [1987] and its appendices by Pohlers and myself for a survey of various such results.

[25]The impressive results described in Rathjen [1994] give further evidence for this assessment.

[26]Kreisel's influential *Survey of Proof Theory* ([1968]) provides a prime example.

Acknowledgements. I wish to thank C. Delzell, A.B. Feferman, G. Mints, P. Odifreddi, and W. Sieg for their useful comments on a draft of this article.

References

Ackermann, W.

[1940] Zur Widerspruchsfreiheit der reinen Zahlentheorie, *Mathematische Annalen*, 117 (1940) 162–194.

Bombieri, E., and van der Poorten, A.J.

[1988] Some quantitative results related to Roth's theorem, *J. Austral. Math. Soc.*, Ser. A, 45 (1988) 233–248.

Browder, F.E. (ed.)

[1976] *Mathematical Developments arising from Hilbert Problems*, Proc. Sympos. Pure Maths. 28, Amer. Math. Soc., Providence, 1976

Cegielski, P.

[1992] Le Théorème de Dirichlet est Finitiste, *LITP* 92.40, Inst. Blaise Pascal, Univ. Paris VI & VII, 1992.

Davis, M., Matijasevič, Y., and Robinson, J.

[1976] Hilbert's tenth problem. Diophantine equations: positive aspects of a negative solution, in Browder [1976], pp. 323–378.

Delzell, C.N.

[1995] Kreisel's unwinding of Artin's proof, this volume.

[199?] Kreisel's unwinding of Artin's proof, Part II, in preparation.

Esnault, H., and Viehweg, E.

[1984] Dyson's lemma for polynomials in several variables (and the theorem of Roth), *Inventiones Mathematicae*, 78 (1984) 445–490.

Feferman, S.

[1987] Proof theory: A personal report, Appendix to Takeuti [1987], pp. 432–485.

[1988] Hilbert's program relativized: proof-theoretical and foundational reductions, *J. Symbolic Logic*, 53 (1988) 364–384.

[1993] Gödel's *Dialectica* interpretation and its two-way stretch, in *Computational Logic and Proof Theory*, LNCS 713, 1993, pp. 23–40.

Furstenberg, H., and Weiss, B.

[1978] Topological dynamics and combinatorial number theory, *J. d'Analyse Math*, 34 (1978) 61–85.

Gentzen, G.

[1936] Die Widerspruchsfreiheit der reinen Zahlentheorie, *Mathematische Annalen*, 112 (1936) 493–565.

Girard, J.Y.

[1987] *Proof Theory and Logical Complexity*, Vol. I, Bibliopolis, Naples, 1987.

Gödel, K.

[1931] Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I, *Monatsh. Math. Physik*, 38 (1931) 173–198. (Reprinted, with English translation, in Gödel [1986], pp. 144–195.)

[1958] Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes, *Dialectica*, 12 (1958) 280–287. (Reprinted, with English translation, in Gödel [1990], pp. 240–251.)

[1986] *Collected Works, Vol. I. Publications 1929–1936*, Oxford University Press, New York, 1986.

[1990] *Collected Works, Vol. II. Publications 1938–1974*, Oxford University Press, New York, 1990.

[1995] *Collected Works, Vol. III. Unpublished Essays and Lectures*, Oxford University Press, New York, 1995.

Hilbert, D., and Bernays, P.

[1939] *Grundlagen der Mathematik, Vol. II*, Springer, Berlin, 1939.

Ingham, A.E.

[1932] *The Distribution of Prime Numbers*, Cambridge Tracts in Mathematics 30, Cambridge University Press, Cambridge, 1932.

[1990] Reprinting of Ingham [1932], with a Foreword by R.C. Vaughan.

Kohlenbach, U.

[1993] Effective moduli from ineffective uniqueness proofs. An unwinding of de la Vallée Poussin's proof for Chebycheff approximations, *Annals Pure Appl. Logic*, 64 (1993) 27–94.

[1993a] New effective moduli of uniqueness and uniform a priori estimates for constants of strong unicity by logical analysis of known proofs in best approximation theory, *Numer. Funct. Anal. and Optimiz*, 14 (1993) 581–606.

Kreisel, G.

[1951] On the interpretation of non-finitist proofs, Part I, *J. Symbolic Logic*, 16 (1951) 241–267.

[1952] On the interpretation of non-finitist proofs, Part II, *J. Symbolic Logic*, 17 (1952) 43–58. Erratum, ibid, p. *iv*.

[1952a] Some elementary inequalities, *Konikl. Nederl. Akad. Wetenschappen*, A, 55 (1952) 334–338.

[1958] Mathematical significance of consistency proofs, *J. Symbolic Logic*, 23 (1958) 155–182.

[1958a] Hilbert's programme, *Dialectica*, 12 (1958) 346–372. (Revised version in *Philosophy of Mathematics: Selected Readings*, P. Benacerraf and H. Putnam eds., Prentice Hall, Englewood Cliffs, N. J., 1964, pp. 157–180.)

[1959] Interpretation of analysis by means of constructive functionals of finite type, in *Constructivity in Mathematics*, North-Holland, Amsterdam, 1959, pp. 101–128.

[1960] Ordinal logics and the characterization of informal concepts of proof, *Proc. Int'l. Cong. Mathematicians, 14-21 August 1958*, Cambridge Univ. Press, Cambridge, 1960, pp. 289–299.

[1968] A survey of proof theory, *J. Symbolic Logic*, 33 (1968) 321–388.

[1977] On the kind of data needed for a theory of proofs, in *Logic Colloquium 76*, North-Holland, Amsterdam, 1977, pp. 111–128.

[1981] Neglected possibilities of processing assertions and proofs mechanically: choice of problems and data, in *University-Level Computer-Assisted Instructions at Stanford: 1968–1980*, Instit. for Math. Studies in the Soc. Sciences, Stanford, 1981, pp. 131–147.

[1981a] Extraction of bounds: interpreting some tricks of the trade, *ibid.*, pp. 149–163.

[1982] Finiteness theorems in arithmetic: an application of Herbrand's theorem for Σ_2-formulas, in *Proceedings of the Herbrand Symposium*, North-Holland, Amsterdam, 1982, pp. 39–55.

[1987] Proof theory: Some personal recollections, in Takeuti [1987], pp. 395–405.

[1990] Logical aspects of computation: Contributions and distractions, in *Logic and Computer Science*, Odifreddi P. (ed.), Academic Press, San Diego, 1990.

[1992] On the idea(l) of logical closure, *Annals of Pure and Applied Logic*, 56 (1992) 19–41.

Lehman, R.S.

[1966] On the difference $\pi(x) - \mathrm{li}(x)$, *Acta Arithmetica*, 11 (1966) 397–410.

Littlewood, J.E.

[1948] Large numbers, *Mathematical Gazette*, 32 (1948) 163–171 (repr. in Littlewood [1953]).

[1953] *A Mathematician's Miscellany*, Methuen, London, 1953.

Luckhardt, H.

[1989] Herbrand-analysen zweier Beweise des Satzes von Roth: Polynomiale Anzahlschranken, *J. Symbolic Logic*, 54 (1989) 234–263.

Monk, R.

[1990] *Ludwig Wittgenstein: The duty of genius*, Penguin Books, London, 1990.

Pohlers, W.

[1987] Contributions of the Schütte school in Munich to proof theory, in Takeuti [1987], pp. 406–431.

Rathjen, M.

[1994] Admissible proof theory and beyond, *Proc. ICLMPS Uppsala*, 1991 (forthcoming).

Rosser, J.B.

[1953] Review of Kreisel [1951] and [1952], *J. Symbolic Logic*, 18 (1953) 78–79.

Roth, K.F.

[1955] Rational approximations of algebraic numbers, *Mathematika*, 2 (1955) 1–20, 168.

Schwichtenberg, H.

[1975] Elimination of higher type levels in definitions of primitive recursive functionals, in *Logic Colloquium '73*, North-Holland, Amsterdam, 1975, pp. 279–303.

[1977] Proof theory: Some applications of cut-elimination, in *Handbook of Mathematical Logic*, 1977, pp. 867–895.

Shelah, S.

[1988] Primitive recursive bounds for van der Waerden numbers, *J. Amer. Math. Soc.*, 1 (1988) 683–697.

Shoenfield, J.R.

[1967] *Mathematical Logic*, Addison-Wesley, Reading, 1967.

Spencer, J.

[1990] Review of Shelah [1988], *J. Symbolic Logic*, 55 (1990) 887–888.

Skewes, S.

[1933] On the difference $\pi(x) - li(x)$, *J. London Math. Soc.*, 8 (1933) 277–283.

[1955] On the difference $\pi(x) - li(x)$, *Proc. London Math. Soc.*, 5 (1955) 48–70.

Tait, W.W.

[1965] Infinitely long terms of transfinite type, in *Formal Systems and Recursive Functions*, North-Holland, Amsterdam, 1965, pp. 176–185.

Takeuti, G.

[1987] *Proof Theory*, 2nd edn., North-Holland, Amsterdam, 1987.

Troelstra, A.S.

[1990] Introductory note to *1958* and *1972*, in Gödel [1990], pp. 217–241.

van der Waerden, B.L.

[1927] Beweis einer Baudetschen Vermutung, *Nieuw Arch. Wisk*, 15 (1927) 212–216.

Some Proof Theory in the 1960's

by William A. Howard

To develop a constructive foundation of mathematics, two approaches suggest themselves.

1. Develop mathematics constructively (Brouwer, Bishop).

2. Pursue a modification of Hilbert's program such as the following. Assuming that we have a system $\mathcal{L}^*$ of constructive reasoning (not necessarily formalized), introduce classical theories $\mathcal{L}$ within which appropriate parts of ordinary mathematics can be formalized; then provide a reduction of $\mathcal{L}$ to $\mathcal{L}^*$ of the following sort. For suitable sentences A of $\mathcal{L}$, consider a constructive counterpart A^* and prove in $\mathcal{L}^*$:

   $$\text{if } p \text{ is a derivation, in } \mathcal{L}, \text{ of } A, \text{ then } A^*,$$

   where A^* may have parameters depending on p. For example, let $\mathcal{L}$ be classical first-order arithmetic. Let A be of the form $\forall x \exists y B(x, y)$, where $B(x, y)$ is quantifier-free. The following result is well-known. If p is a derivation, in $\mathcal{L}$, of A, then there is a function f, defined by transfinite recursion over a proper lower segment of ordinal notations less than ϵ_0, such that $\forall x B(x, f(x))$. All this is carried out in $\mathcal{L}^*$. The use of ordinals in $\mathcal{L}^*$ may be a temporary expedient, the intention being to justify this on more fundamental constructive grounds later. More generally, a part of the modified Hilbert program consists of the task of developing suitable systems $\mathcal{L}^*$ of constructive reasoning.

Although the formulation (2) is in the spirit of Hilbert and Gentzen (where, in the case of Hilbert, $\mathcal{L}^*$ consists of finitistic reasoning), explicit formulations of this sort were (I think) first emphasized by Kreisel.

The use of a program of this sort in order to *justify* classical mathematics has been severely criticized by Kreisel over the years. Kreisel's criticism is supported by the following phenomenon. As time has gone by, the constructive methods have become more and more abstract and complex. As an example of complexity: it is difficult to understand the ordering of the ordinal notations needed for the treatment of, say, Π^1_1 comprehension. Thus one feels that the constructive principles of $\mathcal{L}^*$ are no more "understandable" than the subsystem $\mathcal{L}$ of classical mathematics being treated. Kreisel has made repeated attempts to find new aims for proof theory. Aside from the question of the philosophical issues, these attempts have had a beneficial effect on the purely technical side of proof theory.

One view of proof theory is as follows. Proof theory is an instrument to be used in the exploration of constructive reasoning. The reductive program, above, challenges us to develop sufficiently strong constructive principles. Also, proof theory is one means of investigating the *connections* between constructive reasoning and classical mathematics.

1 Some Subsystems of Analysis

Schütte's *Beweistheorie* appeared in 1960. Schütte used a system of ramified type theory, of finite types and orders, and also a type-free system in which the propositional logic is restricted. Since these systems are not suitable for the formalization of ordinary mathematics, Schütte indicated how the mathematics can be changed so that it can be handled by these systems. This represents a radical departure from the sort of modified Hilbert program described above: a foundation of ordinary mathematics is not given; rather, the mathematics itself is changed. This situation was remedied by the papers of Kreisel [1962] and Spector [1962], in which various subsystems of analysis were introduced which turned out to be amenable to a constructive proof-theoretic treatment. The weakest of these subsystems center around the Σ^1_1 axiom of choice ($\Sigma^1_1 - AC$), the Δ^1_1 comprehension axiom ($\Delta^1_1 - CA$), and the corresponding rules (see below). These subsystems are adequate for the treatment of interesting parts of ordinary mathematics.

If an axiom has the form $P \to Q$, then by the corresponding rule of inference is meant: from P infer Q. Thus to $\Sigma^1_1 - AC$ and $\Delta^1_1 - CA$ correspond rules of inference; these will be denoted by $\Sigma^1_1 - RC$ (the Σ^1_1 rule of choice) and $\Delta^1_1 - CR$ (the Δ^1_1 comprehension rule), respectively.

The first proof-theoretic results about any of these axioms or rules appeared in Feferman [1964] – unless one counts the derivation of $\Sigma^1_1 - AC$ in Spector [1962, p. 26]. Feferman's emphasis was on autonomous progressions and "predicative provability". Hence his main results are stated in terms of a formal system IR whose proof-theoretic ordinal is the Feferman-Schütte ordinal Γ_0. But certain of his theorems provide a proof-theoretic treatment of the rules $\Delta^1_1 - CR$ and $\Sigma^1_1 - RC$ by means of smaller ordinals as follows. As in Feferman's paper, let H denote elementary analysis extended by $\Delta^1_1 - CR$. By 6.14, a proof p in H with (necessarily finite) length k can be relativized to a proof p^* of length $\omega^2 k$ in ramified analysis $RA(\omega^k)$ of rank (i.e., ramification order) ω^k. Hence Tait's ordinal bound for cut-elimination for $RA(\omega^k)$, [1968, pp. 222–224], yields a treatment of H by means of the ordinal $\phi(\omega, 0)$ where $\lambda x.\phi(b, x)$ is the b-th function in the Feferman-Schütte hierarchy of normal functions. Conversely, 6.19 – together with the fact that arithmetical transfinite recursion up to ω^{k+1} for any given $k < \omega$ can be derived in H – shows that transfinite induction up to $\phi(k, 0)$ can be proved in H. Presumably it is possible to extract a proof-theoretic treatment of $\Sigma^1_1 - RC$ by means of the ordinal of $RA(\omega^k)$ – ultimately $\phi(\omega, 0)$ – from Feferman's proof of 6.27.

Friedman [1967] showed that the proof-theoretic ordinal of $\Sigma^1_1 - AC$ is the same as that of ramified analysis $RA(\epsilon_0)$, consequently $\phi(\epsilon_0, 0)$.

Predicativity

The subsystems of analysis based on the axioms $\Sigma^1_1 - AC$, $\Delta^1_1 - CA$, or on the rules $\Sigma^1_1 - RC$, $\Delta^1_1 - CR$ arose from Kreisel's investigations of the idea of predicativity (see especially [1960a]). On the one hand, various aspects of predicativity had played an important role in the work of certain *mathematicians*, principally in the first decades of the twentieth century; e.g., Baire, Borel, and Lebesgue. On the other hand, *logicians* such as Wang, Lorenzen, and Schütte had proposed a foundation of mathematics using layered formalisms based on the idea of predicativity. The traditional idea of predicativity in mathematics is that of *predicative definitions*: An object (e.g., property, set, or function) is defined by reference to a previously defined object or class of objects. By this means, a new class of objects is obtained. This process might be repeated a finite number of times, or perhaps iterated into the transfinite – using ordinals to index the stages of the iteration. On the side of logic, Kreisel proposed a notion of *predicative proof*; in other words, proof based on predicative notions. This was connected with ideas about autonomous progressions.

2 Autonomous Progressions

The idea of an autonomous progression of formal systems was first presented in Kreisel [1960]. The following is a mature version given by Kreisel in [1970], where he emphasized the role of reflection.

First we consider Kreisel's informal notion of reflection. During the course of reasoning, a person may perform the step of recognizing that certain principles $\mathcal{P}$ (of proof, or, alternately, of definition) are valid – whatever might be meant by "valid" here. This act of recognition, called *reflection*, leads to an extension of the person's methods of proof.

A formal counterpart of this process is obtained as follows. If the person just mentioned is a logician, this person may decide to represent his (or her) principles $\mathcal{P}$ by formulating them in a formal system Σ. The process of reflection produces a new system Σ^* (an extension of Σ). Under suitable conditions, one may iterate this process into the transfinite, getting a hierarchy of formal systems Σ_b indexed by ordinals b. Which ordinals are to be used? Kreisel's answer is that one should consider the ordinals which are generated (starting with 0) by the following process. Having found a proof in Σ_b that c is an ordinal (where $c > b$), then introduce Σ_c. Presumably when c is obtained in this way, all smaller ordinals can be so obtained. The phrase "c is an ordinal" is shorthand for "c is a notation for an ordinal", and this is to be understood in terms of some characterization of the notion of ordinal. The choice of the characterization is itself a nontrivial issue. The ordinals just described are called *provable ordinals* (with respect to the notion of proof just considered). The hierarchy of formal systems obtained in this way is called an *autonomous progression*. In a variant of this, the indices b, c, denote linear or partial order relations. Now "c is an ordinal" is replaced by "c is well-founded" (according to some notion of well-foundedness).

Also, in [1970], Kreisel made a systematic study of the notion of transfinite iteration. He considered the close connection between this and reflection – a connection that becomes evident when we consider how we use reflection in building up ordinals in a concrete way.

Here are a couple of thoughts.

- Would it make sense to consider autonomous progressions with respect to other extension processes – not just reflection?

- I suspect that most rules of inference express, implicitly, some form of reflection.

By means of autonomous progressions based on reflection, Kreisel [1970, p. 496] proposed to capture the thinking of certain kinds of idealized logicians. Thus his aim was to establish the limits of what could be achieved by a logician who uses certain concepts and principles together with reflection on these:

> "... *we now ask ourselves: what is implicit in the given concepts* **together** *with the concept of reflection on these concepts?* (reflection of the kind analyzed above)".

Naturally, anyone examining Kreisel's theory of autonomous progressions must keep in mind the kind of reflection he analyzes.

A substantial part of the work being considered in the present essay can be viewed as follows. The purpose of the work is to investigate *systems of thought*. An intellectual difficulty in pursuing this subject is that we must be able to switch back and forth between two modes of thought:

1. To explore a system of thought, we think *within* the system.

2. We think *about* the system.

In (1) we perform an act of self-programming: we set up a system of rules and play within the rules. This is a characteristic human activity. It need not be conscious or voluntary. It may consist of the formation of habits. Social processes may play a crucial role. The system of thought will be governed by certain conceptions. As a part of (1) and (2) we may engage in systematic introspection, in an attempt to become aware of the conceptions which govern our thinking.

In his original use of an autonomous progression [1960], Kreisel's emphasis was on a certain notion of finitist proof but he indicated the possibility of using autonomous progressions to treat "the concept of predicative proof". In his review of Schütte's book [1960b, pp. 246–247], he proposed that ramified analysis (especially ramified analysis based on intuitionistic logic) be extended autonomously to ordinal ranks, saying that some procedure of this kind is necessary if the use of ramified analysis of ordinal ranks is to be regarded as predicative.

Feferman [1964, pp. 23–25] sketched a proof that for classical ramified analysis extended autonomously to ordinal ranks, the limit of the ordinals involved is Γ_0. Schütte, independently, proved that Γ_0 is an upper bound in the case of classical ramified type theory [1964] and that the ordinals less than Γ_0 are provable in the case of intuitionistic ramified analysis [1965].

In a seminar at Penn State in 1962 or 1963, Schütte said that Γ_0 is the limit of the predicatively provable ordinals (i. e., ordinals provable by predicative means). The following dialogue ensued.

> Howard: "The claim, that Γ_0 is the limit of the predicatively provable ordinals, seem to me to be dubious."
>
> Schütte: "It is Kreisel's opinion that the claim is correct."

Whether Kreisel himself thought that the claim had been established – either then or later – is another question. Indeed, there is a question as to what it would mean to say that the claim had been established. Some of the relevant papers will be considered below. In any case, in 1963 I felt that there were three difficulties with the claim.

Difficulty 1. What are the ideas and principles upon which the "predicative view" is supposed to be based?

Difficulty 2.a. In carrying out predicative reasoning by means of an autonomous progression of systems $RA(b)$ of ramified analysis, reflection principles are used at various points. But if the predicative person admits the validity of the appropriate reflection principles, then he (or she) can consider ramified analysis $RA(b_n)$ of rank b_n for an autonomous sequence of ordinals b_n with limit Γ_0. Reflecting on this, he is able to conclude "b_n is an ordinal, for every n" and hence "Γ_0 is an ordinal". In more detail: Reflection is involved in the following step. "If it can be proved in $RA(b)$ that c is an ordinal, then the use of $RA(c)$ is justified." Reflecting on this, he sees that the step is justified for general b and c; so he can consider the sequence of ordinals b_n just mentioned, and hence conclude that Γ_0 is an ordinal. Thus, by use of reflection, the person gets to Γ_0 and beyond. (In 1963 I probably would not have used the word *reflection* as systematically as I have here.)

By hindsight, the essence of Difficulty 2.a is:

Difficulty 2.b. What is to prevent the finitist or predicativist from making the following reflection for general b and c? "If it can be proved in Σ_b that c is an ordinal, then the use of Σ_c is justified."

Difficulty 3. (now stated in general for systems Σ_b):

1. Supposing that c is proved in Σ_b to be an ordinal, how does that justify the introduction of Σ_c?

In connection with (1) it is necessary to treat the following two items.

2. There are different ways of saying "c is an ordinal". The choice may be crucial.

3. Justification of a rule of inference of the form: If c has been proved to be an ordinal, then infer $\mathcal{A}(c)$.

The autonomous progression in [1960] for capturing finitism appears interesting but the details are sketchy. There is a lack of clarity in the treatment of well-foundedness since there are no function variables in the formalism. There are a few remarks about this in Kreisel [1968, p. 337, second paragraph] and [1970, Note 25, p. 511]. Other autonomous progressions are provided in Kreisel [1965, pp. 169–173]. These make a use of a predicate symbol O for ordinals in a restricted way; in particular, O is applied only to constants. The emphasis is on visualization (in a theoretical sense). Thus $O(b)$ means: b is a visualizable ordinal. The extent to which any of these systems deal with Difficulty 2.b and Difficulty 3 needs to be clarified. Kreisel [1970, Note 28, p. 512] is relevant here.

In [1970, pp. 499–501] Kreisel addresses Difficulties 2.b and 3 in the following way. His example consists of an autonomous progression of systems Σ_b for the analysis of finitism (he uses the notation P_b). First he considers an autonomous progression of systems P_b, where P_b expresses principles of definition by $\omega^{\omega b}$-recursion; then Σ_b is obtained from P_b by adding free variables for ω-sequences and a functional which expresses well-foundedness with respect to a standard system of order relations. Let $e(n)$ be defined by $e(1) = \omega$, and $e(k+1) = \omega^{e(n)}$ for positive integers k. For crucial b and c, a derivation, in Σ_b, of the statement that c is an ordinal, does not justify the introduction of Σ_c. To understand this, we should

> "recall the *long and tortuous process by which we convince ourselves that the formal derivations of each $P_{e(n)}$ provide proofs implicit in the notions treated*; a formal derivation in $P_{e(n)}$ in no way reflects this process ...".

To represent these mental processes, he introduces a progression of systems $\mathcal{P}_b$, where

> "a representation of a proof justifying Σ_b can be defined by b-recursion".

The proof just mentioned, or perhaps its representation, belongs to $\mathcal{P}_b$; it is not clear where the b-recursion takes place. In any case, Difficulty 3.1 is handled by introducing the systems $\mathcal{P}_b$. Finally, Difficulty 2.b is handled by the fact that

> "the whole sequence $\mathcal{P}_{e(n)}$ (for $n = 1, 2, \ldots$) cannot be *defined* in any particular $\mathcal{P}_{e(n)}$ ...".

Kreisel's discussion is sketchy. Is the progression of systems $\mathcal{P}_b$ an autonomous progression? Can Difficulties 2.b and 3 be handled by the systems $\mathcal{P}_b$ alone, or does the argument depend on an interaction between all three progressions?

It appears to me that in this discussion Kreisel is trying to capture the following phenomenon. When one works on a concrete level of thought with ordinals, one runs into increasing complexity, so that finally the details – at the level of concreteness on which one is working – become overwhelming. At that point, one gets control of the situation by going to a new level of abstraction (if one can think of the proper abstractions – which may require a new creative idea). Perhaps some insight into this situation can be gained by establishing a connection between the little progressions of arithmetical reasoning in Kreisel [1960], [1965], and [1970] and Girard's slow-growing hierarchy of number-theoretic functions (see Cichon and Wainer [1983]).

3 Mathematics vs. Logic

In order to understand how Kreisel arrived at the subsystems of analysis mentioned in §1, and also to gain a perspective on the idea of predicative provability, it is helpful to consider the program he formulates in [1968, pp. 323 and 360] and [1970, pp. 489–490]. In the light of various shortcomings of Hilbert's Program – some of them caused by genuine difficulties uncovered by Gödel's incompleteness theorems – Kreisel proposed the following program for the proof-theoretic investigation of the foundations of mathematics. The program has two parts. On the side of *mathematics*, suitable areas of ordinary mathematics should be isolated for formalization; in particular, suitable subsystems of analysis should be developed. On the side of *logic*, it is necessary to develop corresponding methods of proof to be used in the metamathematics; in other words, simple-minded constructivism – let alone crude finitism – is not enough. Furthermore, Kreisel felt that proof theory had been hampered by an ideological constriction of thought, and he proposed to remedy this by introducing methods not usually employed by proof theorists – in particular, model-theoretic considerations and the systematic use of recursion theory. Even if the ultimate goal requires the use of a constructive metamathematics, model-theoretic considerations may constitute a useful tool.

Kreisel arrived at the subsystems of analysis based on $\Sigma^1_1 - AC$, $\Delta^1_1 - CA$, or on the rules $\Sigma^1_1 - RC$, $\Delta^1_1 - CR$, by thinking model-theoretically about the hyperarithmetic hierarchy [1962]. This in turn was a part of his investigation of the idea of predicativity [1960a]. These two papers are concerned mainly with the side of his program dealing with the isolation of suitable areas of mathematics.

In the remark starting on p. 313 of [1962], Kreisel gives a convincing justification of $\Delta^1_1 - CR$ as a predicative principle of definition.

Feferman's paper [1964] is based on an interplay between autonomous progressions and predicative definitions. The theory IR in [1964] embodies the latter (in the light of Kreisel's justification of $\Delta^1_1{-}CR$ as a predicative principle of definition). In the axioms and rules of inference of IR there is no explicit use of reflection – since, in the axioms and in the premises and conclusions of the rules of inference, there is no explicit reference to provability or definability. Thus the theory IR is free from Difficulty 2.b. By contrast, the autonomous progression of theories H_a, based on a formalized ω-rule, employs reflection (as can be seen in the definition 5.15). Similarly for the autonomous progression of systems R_a of ramified analysis. Hence any attempt to capture the notion of predicative proof by employing these two autonomous progressions runs up against Difficulties 2.b and 3. This is not to deny the interest in Fefermans results about these progressions, including the fact that in both cases the limit of the ordinals is Γ_0.

In [1966] and [1974], Feferman introduces predicative versions of various of the usual axioms of set theory (some of these now being stated as rules). Somewhat less natural as principles of set theory are his rules of transfinite induction and recursion; it appears that they owe their origin to the system IR of Feferman [1964]. As in the case of IR, these set theories are free from Difficulty 2.b. My feeling is that IR and the set theories are not free from Difficulty 3.3.

In [1979] Feferman undertakes a systematic investigation of the questions: What are predicative principles; what is predicative reasoning? This is done through a certain system of second-order arithmetic based on two ground types: N, for natural numbers, and Π, for predicates. These generate function types and predicate types (both of level 1). The system is free from Difficulties 2 and 3. One would need some experience with it before being able to say just how natural it is – either as a system for logicians or a system for doing ordinary mathematics.

In any case, the paper makes a good contribution toward dealing with Difficulty 1.

Perhaps this is the best place to mention that it appears to me that Kreisel and Feferman, in various discussions of predicativity, have not made a sufficiently clear separation between the classical and constructive cases.

In the mid 1970s Harvey Friedman introduced a system ATR_0 of second-order arithmetic whose main axiom scheme says that if a linear ordering is well-founded, then arithmetic comprehension may be iterated along it. This expresses in a straightforward way the idea of the formation of hyperarithmetical properties. Is ATR_0 predicative? According to the discussions of Kreisel and Feferman – at least, the discussions that have seen – the answer would be "No", since a condition of well-foundedness is used as the premise of an axiom. On the other hand, as Friedman has shown, the ordinal of ATR_0 is Γ_0.

Studies by Friedman and Simpson have shown that various parts of ordinary mathematics can be developed in ATR_0; specifically, ATR_0 provides a good subsystem of analysis. (A good exposition, with further references, is provided by Friedman, McAloon, and Simpson [1982].) These studies – at least, in the early stages – provide a good example of the mathematical side of Kreisel's program mentioned at the beginning of the present section.

4 Is Γ_0 the Ordinal of Predicativity?

Kreisel and Feferman have given two arguments in support of the claim that Γ_0 is the limit of the ordinals "provable by predicative means".

1. Kreisel's argument by means of autonomous progressions. The most mature form of this argument is in [1970].

2. The empirical argument: Formal systems developed on the basis of various ways of looking at predicativity all turn out to have the proof-theoretic ordinal Γ_0.

As indicated in §2, Kreisel's argument by means of autonomous progressions is (in my opinion) inconclusive. On the other hand, in carrying out the argument, Kreisel examines a variety of fundamental issues; his discussion of these is of interest in its own right. The empirical argument has some plausibility. One is reminded of the empirical argument in support of Churchs Thesis; but there the argument is more convincing because the process of carrying out an action on the basis of instructions is familiar to us all. Predicativity – whatever it involves – does not involve anything anywhere near as basic.

Use of the phrase "predicative analysis" has become customary in referring to the relevant systems of Feferman and Schütte. This seems to be a reasonable usage. Presumably there will not be any breakdown of ideas which will necessitate the use of the phrase "what used to be called predicative analysis". At least, not very soon.

It appears that Feferman's idea of predicativity carries along with it the idea of countability ([1974], [1979, §4.2]). Admittedly, he is not very explicit about this. In any case, one can offer – with a certain degree of seriousness – the following support for the claim that Γ_0 is "the ordinal of predicativity". We note that $\Gamma_0 = \phi(\Omega, 0)$ in the Bachmann hierarchy, where Ω denotes the first uncountable ordinal. To get the ordinals less than Γ_0, one uses a system of normal functions indexed by ordinals of the first and second number classes. To get beyond Γ_0, one uses a system of normal functions indexed by (some) ordinals in the third number class – thus using uncountable ordinals, an impredicative notion.

5 Something Impredicative

As indicated in §1, Spector's paper [1962] played an important role in proof theory in the 1960s. Kreisel's influence on the paper can be seen from the fact that he collaborated with Spector in the early stages and also did a good deal of work on the paper in bringing it out after Spector's death. In [1987, pp. 159–160] Kreisel gives an interesting account of his role in getting Spector interested in the ideas that led to his paper.

To get the Gödel functional interpretation of classical analysis (second-order arithmetic with the full comprehension axiom) Spector used "bar recursion of finite type". Here "type" refers to the type of the nodes of the tree over which the recursion takes place. Admittedly bar recursion of type level greater than one has shown itself to be proof-theoretically intractable, but in the years after 1962 bar recursion of type zero (BR_0) showed itself to be amenable to a proof-theoretic treatment by means of the appropriate system of ordinal notations.

Let $H_2 + BI_0$ denote second-order Heyting arithmetic extended by adding Brouwer's Bar Theorem (BI_0) and Markov's Principle. Then BR_0 is just what is needed in order to carry out Gödel's functional interpretation of $H_2 + BI_0$. Let $Z_2 + (QF - BI_0)$ denote classical second-order arithmetic extended by adding BI_0 over trees given by quantifier-free formulas. Obviously the "negative translation" of classical logic into intuitionistic logic sends $Z_2 + (QF - BI_0)$ into $H_2 + (QF - BI_0)$. Hence, via Gödel's functional interpretation, a proof-theoretic treatment of BR_0 yields a proof-theoretic treatment of $Z_2 + (QF - BI_0)$. The system $Z_2 + (QF - BI_0)$ is intermediate in strength between "predicative" analysis and analysis based on Π^1_1 comprehension. It is not clear whether the full system $Z_2 + (QF - BI_0)$ is very interesting as a system within which to do ordinary mathematics, but Simpson [1982] has shown that the subsystems obtained by restricting the bar induction to Σ^1_1 and Π^1_1 formulas are of interest. In particular, by means of the latter he provides a significant improvement on the derivation of $\Sigma^1_1 - AC$ in Spector's paper.

In the work initiated by Spector's paper, we are dealing with the connections between intuitionistic mathematics, functionals of finite type, and classical analysis. This is a large subject matter. Kreisel has played a central role in its development. In accordance with the view of proof theory mentioned at the end of the Introduction, this subject matter is of particular interest as an exploration of constructive reasoning.

Kreisel's influence occurred not only through his papers by also through personal contacts and a truly vast amount of correspondence. There was a large-scale flow of ideas and information, with Kreisel at the center. A high point of this activity was a seminar, on subsystems of analysis, organized by Kreisel at Stanford in the summer of 1963. This gave rise to a mimeographed report, which became Volume I because there were so many ideas to be worked out or written down that Kreisel organized a second volume a year later.

My own contact with Kreisel has been mainly through a large number letters during the period from 1963 to 1973. Even when I was living next door to him during the summer of 1963, we wrote letters to each other.

An important topic introduced by Kreisel in the 1963 seminar at Stanford was the theory of generalized inductive definitions, in particular the theory ID_1 of noniterated positive inductive definitions – considered both intuitionistically and classically. Whereas BI_0 expresses the idea of transfinite induction on *well-founded* trees, ID_1 captures the idea of transfinite induction on *inductively generated* trees. The equivalence of ID_1 to BI_0 within a system of intuitionistic mathematics is proved in Kreisel and Troelstra [1970]. Inductive definitions iterated into the transfinite have played an important role in proof theory. A good reference for this, with an informative introduction by Feferman, is provided by Buchholz et al. [1981].

Conclusion. I hope that this exposition has provided an indication of the nature and importance of Kreisel's work and influence in proof theory.

References

Buchholz, W., Feferman, S., Pohlers, W., and Sieg, W.

[1981] Iterated Inductive Definitions and Subsystems of Analysis: Recent Proof-Theoretical Studies, *Lecture Notes in Mathematics* 897, Springer-Verlag, Berlin, 1981.

Cichon, E.A., and Wainer, S.S.

[1983] The slow-growing and the Grzegorczyk hierarchies, *The Journal of Symbolic Logic*, 48 (1983) 399–408.

Feferman, S.

[1964] Systems of predicative analysis, *The Journal of Symbolic Logic*, 29 (1964) 1–30.

[1966] Predicative provability in set theory, *Bulletin of the American Mathematical Society*, 72 (1966) 486–489.

[1974] Predicatively reducible systems of set theory, in T.J. Jech, ed., *Axiomatic Set Theory*, Amer. Math. Soc., 1974, pp. 11–32.

[1979] A more perspicuous formal system for predicativity, in K. Lorenz, ed., *Konstruktionen versus Positionen*, de Gruyter, Berlin, 1979, pp. 87-139.

Friedman, H.

[1967] *Subsystems of analysis and set theory*, Ph.D. thesis, Massachusetts Institute of Technology, 1967.

Friedman, H., McAloon, K., and Simpson, S.G.

[1982] A finite combinatorial principle which is equivalent to the 1-consistency of predicative analysis, in G. Metakides, ed., *Patras Logic Symposion*, North-Holland, Amsterdam, 1982, pp. 197–230.

Kreisel, G.

[1960] Ordinal logics and the characterization of informal concepts of proof, *Proceedings of the 1958 International Congress of Mathematicians*, Edinburgh, 1960, pp. 289–299.

[1960a] La Prédicativité, *Bulletin de la Société Mathématique de France*, 88 (1960) 371–391.

[1960b] Review of Schütte's *Beweistheorie*, *The Journal of Symbolic Logic*, 25 (1960) 243-249.

[1962] The axiom of choice and the class of hyperarithmetic functions, *Koninklijke Nederlandse Akademie van Wetenschappen*, 65 (1962) 307-319 (also in *Indagationes Mathematicae*, 24).

[1965] Mathematical logic, in T.L. Saaty, ed., *Lectures on Modern Mathematics*, vol. III, Wiley, New York, 1965, pp. 95–195.

[1968] A survey of proof theory, *The Journal of Symbolic Logic*, 33 (1968) 321-388.

[1970] Principles of proof and ordinals implicit in given concepts, in A. Kino, J. Myhill, and R. E. Vesley, eds., *Intuitionism and Proof Theory*, North-Holland, Amsterdam, 1970, pp. 489–516.

[1987] Gödel's excursions into intuitionistic logic, in P. Weingartner and L. Schmetterer, eds., *Gödel Remembered*, Bibliopolis, Naples, 1987, pp. 65–186.

Kreisel, G., and Troelstra, A.S.

[1970] Formal systems for some branches of intuitionistic analysis, *Annals of Mathematical Logic*, 1 (1970) 229–387.

Schütte, K.

[1960] *Beweistheorie*, Springer-Verlag, Berlin, 1960.

[1964] Eine Grenze für die Beweisbarkeit der transfiniten Induktion in der verzweigten Typenlogik, *Archiv f. math. Logik u. Grundlagenf.*, 7 (1964) 45–60.

[1965] Predicative well-orderings, in J. Crossley and M. Dummett, eds., *Formal Systems and Recursive Functions*, North-Holland, Amsterdam, 1965, pp. 280–303.

Simpson, S.G.

[1982] Σ^1_1 and Π^1_1 transfinite induction, in *Logic Colloquium '80* (Prague), North-Holland, Amsterdam, 1982, pp. 239-253.

Spector, C.

[1962] Provably recursive functionals of analysis: a consistency proof of analysis by an extension of principles formulated in current intuitionistic mathematics, in J.C.E. Dekker, ed., *Recursive Function Theory*, Proceedings of the Symposia in Pure Mathematics, vol. V, American Mathematical Society, Providence, 1962, pp. 1–27.

Tait, W.W.

[1968] Normal derivability in classical logic, in J. Barwise, ed., *The Syntax and Semantics of Infinitary Languages*, Lecture Notes in Mathematics, 72, Springer-Verlag, Berlin, 1968, pp. 204–236.

Bounds Extracted by Kreisel from Ineffective Proofs

by Hoarst Luckhardt

In [1987] Kreisel published a retrospective survey of his work. In this paper we concern ourselves with one aspect of Kreisel's work, namely, with the question of how one uses mathematical logic and in particular proof theory to make substantial contributions to mathematics. We discuss only logical and mathematical aspects, leaving aside epistemological questions.

1 A Theorem of Littlewood (1914)

Let $\pi(n)$ be the number of primes $\leq n$ and $\mathrm{li}(n)$ be the logarithmic integral. Up until the end of the nineteenth century no example of a sign change in $\pi(n) - \mathrm{li}(n)$ had been found by trial and error, and it even was conjectured that there was none. Littlewood proved that the sign of $\pi(n) - \mathrm{li}(n)$ changes infinitely often in fact. At that time his argument was regarded as a pure existence proof which deduced its Π^0_2 claim first from the Riemann hypothesis RH, and then from its negation $\neg RH$. A good (recursive) bound is of use in investigations of the distributions of primes.

In [1951], [1952] (see also [1958] and [1976]) Kreisel formulates the idea of the proof so that it consists of proofs of formulas in the first order predicate calculus

$$\begin{array}{rcl} \bigwedge z\, B_0(z) & \to & \bigwedge x \bigvee y A_0(x,y) \\ \bigvee z \neg B_0(z) & \to & \bigwedge x \bigvee y\, A_0(x,y) \end{array} \tag{1}$$

where A_0 and B_0 are predicates primitive recursive in their argument vectors. The Herbrand analysis (in any of several possible technical variants) yields terms r_i, s_j and t_k such that

$$\begin{array}{rcl} \bigwedge_{i=1}^{\ell} B_0(r_i(x)) & \to & \bigvee_{j=1}^{m} A_0(x, s_j(x)) \\ \neg B_0(z) & \to & \bigvee_{k=1}^{n} A_0(x, t_k(z,x)) \ . \end{array} \tag{2}$$

One can then eliminate B_0 and replace the Herbrand disjunction thus obtained by the weaker

$$\bigwedge x \ \bigvee y \le S(x) : A_0(x,y) \ .$$

Hence, contrary to the original view of the proof, we have discovered a recursive bound that was hidden there.

In [1955] Skewes gave the first ad hoc mathematical proof of the existence of a bound of the same order of magnitude (Kreisel [1958], p. 170); he showed that the first change of sign of $\pi(n) - \mathrm{li}(n)$ occurs for $n < 10^{10^{10^{1000}}}$. Lehman [1966] was able to reduce this bound to $1.65 \cdot 10^{1165}$ in 1966, and Te Riele [1987] improved this to $6.69 \cdot 10^{370}$ in 1984. New lower bounds for the magnitude of the variation and the number of sign changes can be found in articles by Pintz [1980], [1980a] and Kaczorowski [1984], [1985].

The equivalence of (1) and (2) is an example of a method of systematic elimination of quantifiers. Already in [1951] Kreisel (p. 247) emphasized that such methods are not primarily useful because of the possibility of mechanically applying them (for example, such applications are often impractical because of their complexity); rather, such logical results often bring critical points in mathematical arguments to our attention. They may, for example, suggest a basic form for the argument where one can find sensitive points which suggest how to proceed.

The transition from (1) to (2) is a *local* elimination which can be optimized by varying the proof. In [1958] some possibilities of this sort are suggested. A new observation is also given there: one does not only obtain a bound from the Herbrand disjunction, but also obtains, by "contraction", $R(x)$ and $S(x)$ such that $B_0(R(x)) \to A_0(x, S(x))$, and these functions R and S can be proved to be recursive in any conventional formal system $\mathcal{S}$ which has an $\mathbb{N}$–model without use of the

Π_1 premise $\bigwedge xB_0(x)$. This is also true for other elimination methods (which interpret Π_2^0 predicates as $\mathbb{N}$–functions) such as ε–elimination, cut–elimination, the no–counterexample–interpretation, functional interpretation and realizability (via $\neg\neg$–translation and Markov's principle). This has important *local* and *global* consequences.

If one examines a proof of $\bigwedge x \bigvee yA_0(x,y)$ *locally*, and takes the Π_1 lemmas in it to be the formula $\bigwedge zB_0(z)$, then, in order to construct the algorithm $S(x)$ of the end formula, one only needs the form of the lemmas and the fact that they are true. In other words, to extract an S–program one does not need to analyse the proofs of the Π_1 lemmas.

In order to formulate the global consequence we need the concept of "$\mathcal{S}$–provable recursive function". Let

$$\bigwedge \underline{x} \ \bigvee \underline{y} \ A_0(\underline{x},\underline{y}) \qquad (\underline{x} = x_1,\dots,x_m, \ \underline{y} = y_1,\dots,y_n)$$

be $\mathbb{N}$–true. Furthermore, let $\langle a_1,\dots,a_n\rangle$ be a coding with inverse functions $(.)_i, \ i=1,\dots,n$. Then

$$\bigwedge \underline{x} \ \bigvee \underline{y} \ A_0(\underline{x},\underline{y})$$

is first associated with the recursive function

$$\varphi(\underline{x}) = \mu z \cdot A_0(\underline{x},(z)_1,\dots,(z)_n)$$

and then with each y_i we associate

$$\Phi_i(\underline{x}) = (\varphi(\underline{x}))_i, \ i = 1,\dots,n.$$

These Φ_i are $\mathcal{S}$–provable recursive if and only if $\vdash_{\mathcal{S}} \ \bigwedge \underline{x} \ \bigvee \underline{y} \ A_0(\underline{x},\underline{y})$.

The global relationship between (1) and (2) is now as follows: The systems $\mathcal{S}$ and $\mathcal{S}$ + { all $\mathbb{N}$–true Π_1^0 formulas} have the same provable recursive functions. Kreisel stated already in [1958, p. 158] that independence results can be obtained from this fact as follows:

> The functions $\Phi_1,\dots,\Phi_n$ associated with $\bigwedge \underline{x} \ \bigvee \underline{y} \ A_0(\underline{x},\underline{y})$
> are not all $\mathcal{S}$–provable recursive $\iff$ $\bigwedge \underline{x} \ \bigvee \underline{y} \ A_0(\underline{x},\underline{y})$
> is unprovable in $\mathcal{S}$ + {all $\mathbb{N}$–true Π_1^0 formulas}.

The implication $\Longrightarrow$ was used 23 years later by Ketonen and Solovay [1981] for a refined proof of the independence of the modified Ramsey–function R' of Paris and Harrington from Peano–arithmetic PA together with all $\mathbb{N}$–true Π_1^0 formulas. They use a "logical" classification of the rate of growth of the PA–provable recursive functions given by Wainer which resulted from Kreisel's [1952] ordinal recursion of order

$< \epsilon_0$. Discussion of this topic can be found in [1977]. Bellin [1990], following a suggestion in this paper, developed with the use of proof analysis a parametric Ramsey theorem from which several known Ramsey theorems follow as special cases. (For further discussion of the role of "logical" parameters see §§5 and 6.)

2 Hilbert's Nullstellensatz (1893)

The geometric version of this theorem states that if the polynomials $f_1, \ldots, f_m \in k[x_1, \ldots, x_n]$ of degree $\leq d$ have no common zeros in an algebraically closed extension of k, then there exist polynomials $g_1, \ldots, g_m \in k[x_1, \cdots, x_n]$, such that $1 = f_1 g_1 + \cdots + f_m \cdot g_m$. From this the full Hilbert *Nullstellensatz* follows by means of Rabinowitsch's trick (1929).

By combining the proof–theoretic methods discussed above with algebraic ideas (contraction and elimination theory for the Skolem functions used), Kreisel proved in 1957 ([1957], [1958]) that the degrees of the polynomials g_i are *primitive recursively* bounded where the bound depends only on n and d and is independent of k. One can use model–theoretic methods to obtain a bound *recursive* in n and d (cf. Robinson [1966]).

Today we know (see Brownawell [1988]) that the induction proofs of Grete Hermann in 1926 already contained doubly exponential bounds on the degrees. The best (and nearly optimal) bound obtained so far is exponential; it was given by Brownawell in [1987].

3 Hilbert's 17th Problem (1900)

Pfister [1976] describes the history of this problem. The first contributions of mathematical logic to the problem were made by A. Robinson, Henkin, and Kreisel in the 1950s, who gave model–theoretic and proof–theoretic formulations of Artin's solution in a first–order language. In the process the result was generalized and the proof was varied and refined. Today we formulate Artin's solution of Hilbert's 17th problem as follows:

Let k be an ordered field and let R be a real–closed order extension of k. If $f \in k[x_1, \ldots, x_n]$ of degree d is positive semi–definite over R (i.e. $f \geq 0$ over R) then f can be represented as a non–negative weighted sum of squares

$$f(x_1, \dots, x_n) = \sum_{i=1}^{\lambda} p_i \cdot g_i(x_1, \dots, x_n)^2$$

where $p_i \in k,\ p_i \geq 0$ and $g_i \in k(x_1, \dots, x_n)$.

Kreisel has made two contributions ([1957a], [1960]) to this subject.

i) He adapted the proof-theoretic methods of §1 to this problem using a kind of Galois theory and thus gave a bound primitive recursive in n and d (independent of k) for λ and the degrees of the rational functions g_i. When one carries this method out exactly one obtains an exponential tower of height n whose uppermost term is $c \cdot d$, with $c > 0$. Using model–theoretic methods one obtains a recursive bound (cf. Robinson [1966]).

ii) He outlined the constructive content of Artin's proof.

For the case where k is already real–closed Pfister [1967] found the bound $\lambda \leq 2^n$ which is independent of d (however, the bound for the degrees of the g_i's is large). Henkin also showed that the p_i and the coefficients of the g_i can be piecewise–rationally constructed from the coefficients of f. This was the beginning of further "logical" developments. First Daykin [1960] improved Henkin's result by carrying out Kreisel's sketch ii) above. The outcome was a primitive recursive construction of finitely many polynomials $p_{ji}(\underline{c})$ and rational functions $g_{ji}(\underline{c}, \underline{x})$ such that

> $f = \sum_i p_{ji} \cdot g_{ji}^2$ for all j, and for every vector $\underline{c}$ of coefficients of f for which f is positive semi–definite over R {we write $\underline{c} \in P_{nd}$} $\bigvee_j \bigwedge_i \big[p_{ji}(\underline{c}) \geq 0 \wedge$ the denominator of $g_{ji}(\underline{c}, \underline{x})$ does not vanish identically in the variables x_k $\big]$

In this connection Kreisel posed the following questions in 1962:

a) Are the case distinctions with respect to j necessary (i.e. does one need a Herbrand disjunction of length > 1) ?

b) Is there a solution to Hilbert's 17th problem which is continuous in $\underline{c}$ on P_{nd} and continuous in the variables x_k on R^n (more exactly, for which there is a continuous extension)?

In [1978] Kreisel showed, with the aid of Stengle's Positivstellensatz of 1974, that continuity with respect to the variables x_k holds. After 1980 Delzell ([1981], [1981a], [1982], [1984]) gave nearly complete answers to a) and b):

a) For $d \leq 2$ case distinctions with respect to j are not necessary, but for $d > 2$ they are.

b) For $k = R$ there is a solution of Hilbert's 17th problem which is continuous in all variables.

4 L–Series

For all Dirichlet-characters χ mod k and all (real) t there has long existed an ineffective proof that $L(1+it,\chi) \neq 0$. In [1981], [1981a] and [1990] Kreisel considers the extraction of (lower) bounds for $L(1,\chi) > 0$ from such proofs for real characters χ which differ from the principal character χ_0. He shows that mathematically trivial changes in the proof simplify the unwinding of the proof considerably, clarify the dependence on the zeros of the ζ–function, and provably also have an effect on the bound itself. In [1981] and [1981a] this program has not yet been completed (see [1981a, p. 153]). In [1990, pp. 239, 247–248] Kreisel comes back to this problem with the comment that the analysis of the modified proof has now been carried out. He does not however give any details. It would be of interest, for example, to know explicitly how the result depends on k. Suppose the dependence were something like the one given in Siegel's ineffective theorem of 1936:

$$\bigwedge \varepsilon > 0 \bigvee \kappa_1(\varepsilon), \kappa_2(\varepsilon) > 0 \bigwedge \text{real}\, \chi \neq \chi_o \bmod k \bigwedge s \in \mathbb{C} \;:$$
$$|s-1| \leq \tfrac{\kappa_1(\varepsilon)}{k^{\varepsilon}} \quad \longrightarrow \quad |L(s,\chi)| > \tfrac{\kappa_2(\varepsilon)\cdot \log^2 4k}{2\cdot k^{\varepsilon}}\ .$$

This would have interesting consequences in analytical number theory.

Such interactions between the idea of the proof and the data in the proof occur also in other places as we see in a result of Girard from 1980 (covered in [1987, pp. 238–251, 484–496], and discussed in [1981a]). Girard analyses two proofs of Van der Waerden's theorem concerning arithmetic progressions, the proof of Fürstenberg and Weiss which uses dynamic systems and the axiom of choice AC, and a harmlessly simplified variant which does not use the axiom of choice. In the first case the analysis is done by cut–elimination and the no–counterexample interpretation, and in the second case only cut–elimination is applied. It extracts bounds which differ fundamentally from each other and from the bound which Shelah later found.

5 Instructive Examples

Productive proof analyses have to be discovered. A fruitful search requires experience in different fields and its sensitive combination. One can learn this from well–known examples, training oneself to recognize the important features of proofs, to note significances in the details and data of the analysis, and then to combine these into a new result. This learning from well–known examples plays an important role in Kreisel's work. We can only give a small sample here.

First, there are examples which one thinks one has long understood. In [1985] Kreisel compares the proofs of Euclid and Euler that there are infinitely many primes. It turns out that the bound given for the next prime which is implicit in Euler's proof is distinctly better than the explicit bound in Euclid's proof. An analytical example from [1985] is $a^n \to 0$ for $0 \leq a < 1$. One is surprised to see how simple Herbrand combinatorics can be used to give an elementary convergence module. Actually certain proofs should concentrate on proving a Herbrand disjunction from the very beginning.

Then there are also examples which one has had to work out himself. When one later regards the work again from a distance one can learn more from it. We collect several such observations from [1981], [1981a], [1989] and [1990].

The first observation is concerned with the particular meaning which proofs have in a first–order language and certain parts of it (for instance proofs where the Herbrand terms satisfy certain conditions on their rate of growth, see §6).

- The constructions in a proof are affected only by the form of its Π_1 lemmas and by the fact that these are true.
- Small changes in a proof can cause large changes in the data of the proof.
- The analysis of the proof must take both logical and mathematical aspects into account.

Now we turn to observations concerning the object and structure parameters of a proof.

1. First, there are the outer parameters of the given situation. Here one must identify which of these parameters are significant and, among these, which are dominant. Often geometric parameters are more important than arithmetic ones and after these come the outer formal and logical parameters.

2. Second, we have three kinds of inner parameters of a proof:

   a) hidden parameters – These parameters express the dependence of the proof on concrete magnitudes, for example, the dependence on the zeros of the ζ–function discussed in §4. The quality of the end result depends on how well these magnitudes can be established.

   b) the substructures that are sufficient for a solution or a partial solution.

   c) new *logical* object parameters (alias eigenvariables) which are needed for a sequence of quantifications leading to the end formula (on account of combinatorial considerations and the conditions on eigenvariables) – The Herbrand disjunction is a normal form for them.

These parameters were first consciously exploited *mathematically* in the analysis of the proofs of Roth's theorem (cf. §6).

6 Roth's Theorem (1955)

An algebraic irrational α has only finitely many very good rational approximations; i.e. for $\varepsilon > 0$ there are only finitely many $q \in \mathbb{N}$ such that

$$Rq \;\equiv\; q > 1 \;\wedge\; \bigvee! \, p \in \mathbb{Z} \;:\; (p,q) = 1 \;\wedge\; |\alpha - pq^{-1}| < q^{-2-\varepsilon} \;.$$

There were a number of earlier theorems of this sort with smaller exponents: Dirichlet 1842, Liouville 1844, Thue 1909, Siegel 1921, Schneider 1936, Dyson and Gelfond 1947, and Roth 1955. Following Dyson's method, Esnault and Viehweg gave a second proof in 1983. Both proofs are not effective; they give no bound either on the size or on the number of the denominators q. Davenport and Roth analysed the first proof in 1955 and in the process they found an exponential bound on the number of values for q which depends on d^2 and ε^{-2}.

While the examples in §§1–5 are basically concerned with Π_2 theorems, this theorem of Roth is, practically speaking, a Σ_2 theorem since R is decidable if ε is algebraic. Kreisel treats Roth's theorem in [1968], [1976], [1981b], [1989] and [1990]. Again his approach is based on Herbrand analysis. In [1981b] Kreisel gives arithmetic conditions on the rate of growth of Herbrand terms, and shows that then a bound on the number of values for q can be obtained. Unfortunately the Herbrand terms in both proofs of Roth's theorem grow much faster. The present author was able in [1989] to formulate stronger conditions on the rate of growth of Herbrand term which:

a) also give a bound on the number of values of q where the inner

   logical parameters mentioned in 5.2.c are now specifically *mathematically* exploited and which

b) apply to the Herbrand terms in both proofs of Roth's theorem.

Thus these proofs can be given a satisfactory logical interpretation. Thereby Davenport and Roth's bound b on the number of values was replaced by $\sqrt[4]{b}$, and from Esnault and Viehweg's proof we get, in fact, a polynomial bound on the number of values of q where the degrees of $\log d$ and ε^{-1} are small:

$$\#\{q : Rq\} < \frac{7}{3}\varepsilon^{-1} \log N_\alpha + 6 \cdot 10^3 \varepsilon^{-5} \log^2 d \cdot \log(50\varepsilon^{-2} \log d)$$

where $N_\alpha < \max(21 \log 2h(\alpha),\ 2\log(1+|\alpha|))$, and h is the logarithmic absolute homogeneous height.

In [1988] Bombieri and van der Poorten obtained essentially the same bound using a more or less ad hoc strategy of proof. Mueller and Schmidt gave lower bounds in [1989] by showing that the first summand in the above formula is best possible, and that there must be a further term involving $\log d$. One can find a discussion of the mathematical context of these results in Schmidt [1990, pp. 299–300] or [1991, p. 66]. No height bounds are known at this time. In Luckhardt [1989] a primitive recursive sequence of finite sets of natural numbers is given having only non–recursive bounds on the size of the elements while the number of elements is bounded recursively.

Eine persönliche Bemerkung. *Georg Kreisel habe ich wissenschaftlich auf Umwegen kennengelernt und weiß daher, daß es bei ihm kaum Kompromisse gibt. Die Konsequenzen daraus ließen ihn gelassen, so gelassen wie ein Unfall, den wir beide am 16. Oktober 1980 in Kronberg in meinem Auto (ohne physischen Schaden) überstanden.*

References

Bellin, G.

[1990] Ramsey interpreted: a parametric version of Ramsey's theorem, *Contemporary Math.*, 106 (1990) 17–37.

Bombieri, E., and van der Poorten, A.J.

[1988] Some quantitative results related to Roth's theorem, *J. Austral. Math. Soc. (Series A)*, 45 (1988) 233–248.

Brownawell, W.D.

[1987] Bounds for the degrees in the Nullstellensatz, *Annals Math.*, 126 (1987) 577–591.

[1988] Aspects of the Hilbert Nullstellensatz, in A. Baker (ed.), *New Advances in Transcendence Theory*, Cambridge University Press, 1988, pp. 90–101.

Daykin, D.E.

[1960] *Thesis*, Univ. of Reading, 1960 (unpublished).

Delzell, Ch.N.

[1981] Continuous sums of squares of forms, *Proc. L.E.J. Brouwer Centenary Symposium*, Noordwijkerhout, 1981, 65–75.

[1981a] Case distinctions are necessary for representing polynomials as sums of squares, *Proc. Herbrand Symposium*, 1981, pp. 87–103.

[1982] A finiteness theorem for open semi–algebraic sets, with applications to Hilbert's 17th problem, *Contemporary Math.*, 8 (1982) 79–97.

[1984] A continuous, constructive solution to Hilbert's 17th problem, *Invent. math.*, 76 (1984) 365–384.

Girard, J.Y.

[1987] *Proof theory and logical complexity*, Bibliopolis, Napoli, 1987.

Kaczorowski, J.

[1984] On sign-changes in the remainder-term of the prime-number formula I, *Acta Arith.*, 44 (1984) 365–377.

[1985] On sign-changes in the remainder-term of the prime-number formula II, *Acta Arith.*, 45 (1985) 65–74.

Ketonen, J., and Solovay, R.

[1981] Rapidly growing Ramsey functions, *Annals Math.*, 113 (1981) 267–314.

Kreisel, G.

[1951] On the interpretation of non-finitist proofs I, *J. Symbolic Logic*, 16 (1951) 241–267.

[1952] On the interpretation of non-finitist proofs II, *J. Symbolic Logic*, 17 (1952) 43–58.

[1957] Hypotheses on algebraically closed extensions, *Bull. Amer. Math. Soc.*, 63 (1957) 99.

[1957a] Hilbert's 17th problem, I and II, *Bull. Amer. Math. Soc.*, 63 (1957) 99–100.

[1958] Mathematical significance of consistency proofs, *J. Symbolic Logic*, 23 (1958) 155–182.

[1960] Sums of squares, *Summaries of talks presented at the Summer Institute for Symbolic Logic, Cornell University, 1957*, Communications Research Division, Institute for Defense Analyses, Princeton, N.J., 1960, pp. 313–320.

[1968] Hilbert's programme and the search for automatic proof procedures, *Symposium on automatic demonstration, Versailles 1968*, Springer Lecture Notes in Mathematics 125, pp. 128–146.

[1976] On the kind of data needed for a theory of proofs, *Logic Colloquium '76*, pp. 111–128.

[1977] From foundations to science: justifying and unwinding proofs. - Set theory, *Foundations of mathematics*, Proc. Symp. Belgrade, 1977, pp. 63–72.

[1978] Review of Ershov, *Zentralblatt*, 374 (1978) 02027.

[1981] Neglected possibilities of processing assertions and proofs mechanically: choice of problems and data, in P. Suppes (ed.), *University-level computer-assisted instruction at Stanford 1968–1980*, Standford University, 1981, pp. 131–147.

[1981a] Extraction of bounds: interpreting some tricks of the trade, *ibid.*, pp. 149–163.

[1981b] Finiteness theorems in arithmetic: an application of Herbrand's theorem for Σ_2–formulas, *Proc. Herbrand Symposium*, Marseille, 1981, pp. 39–55.

[1985] Proof theory and the synthesis of programs: potential and limitations, *Springer Lecture Notes in Computer Science*, 203 (1985) 136–150.

[1987] Proof theory: some personal recollections, in G. Takeuti, *Proof theory*, Studies in Logic, vol. 81, North Holland, 1987, 2nd ed., pp. 395–405.

[1989] Review of Luckhardt, *Zentralblatt*, 669 (1989) 03024.

[1990] Logical aspects of computation: contributions and distractions, in P. Odifreddi (ed.), *Logic and Computer Science*, Academic Press, London, 1990, pp. 205–278.

Lehman, R.Sh.

[1966] On the difference $\pi(x) - li(x)$, *Acta Arith.*, 11 (1966) 397–410.

Luckhardt, H.

[1989] Herbrand-Analysen zweier Beweise des Satzes von Roth: polynomiale Anzahlschranken, *J. Symbolic Logic*, 54 (1989) 234–263.

Mueller, J., and Schmidt, W.M.

[1989] On the number of good rational approximations to algebraic numbers, *Proc. Amer. Math. Soc.*, 106 (1989) 859–866.

Pfister, A.

[1967] Zur Darstellung definiter Funktionen als Summe von Quadraten, *Invent. math.*, 4 (1967) 229–237.

[1976] Hilbert's seventeenth problem and related problems on definite forms, *Proc. Symposia Pure Math.*, 28 (1976) 483–489.

Pintz, P.

[1980] On the remainder term of the prime number formula I, *Acta Arith.*, 36 (1980) 341–365.

[1980a] On the remainder term of the prime number formula II, *Acta Arith.*, 37 (1980) 209–220.

te Riele, H.J.J.

[1987] On the sign of the difference $\pi(x)-li(x)$, *Math. Comp.*, 48 (1987) 323–328.

Robinson, A.

[1966] Reviews of Kreisel [1958] and [1960], *J. Symbolic Logic*, 31 (1966) 128–129.

Schmidt, W.M.

[1990] Diophantische Approximationen und diophantische Gleichungen, *Jahresbericht der Deutschen Mathematiker-Vereinigung*, Jubiläumstagung, 1990, pp. 297–326.

[1991] *Diophantine Approximations and diophantine equations*, Springer Lecture Note in Mathematics 1467, 1991.

Skewes, S.

[1933] On the difference $\pi(x)-li(x)$ I, II, *J. London Math. Soc.*, 8 (1933) 277–283 and *Proc. London Math. Soc.*, 5 (1955) 48–70.

Completeness for Intuitionistic Logic

by David M$\underline{\text{c}}$Carty

> When one provides an axiomatic foundation for logic, as, for example, is done in *Principia Mathematica*, the question arises whether the axioms initially adopted are "complete," that is, whether they actually suffice for the formal deduction of every correct proposition of logic.
>
> Gödel [1930]

The general area is metamathematics; the neighborhood is completeness for Heyting's predicate logic **HPL** and four of Kreisel's articles on that subject. First, in 1957, Gödel proved that weak forms of completeness are not (plausibly) to be had within a strictly intuitionistic metamathematics. He showed that completeness implies forms of Markov's Principle, which is refused by most constructivists. Kreisel [1962] reports upon a refinement of Gödel's proof.

Second, negative formulae are logical compounds of negated atoms made without the aid of $\vee$ and $\exists$. As Gödel and others had shown in the thirties, these are formulae on which **HPL** agrees in derivability with its conventional counterpart. One might say that these represent statements which are "conventional" or "constructively boring," formulae on which intuitionism cannot show its stuff. Motivation for constructive reasoning is often fostered on the grounds that it - and not conventional reasoning - better captures the "real" mathematical meanings of positive statements containing $\vee$ and $\exists$. Kreisel [1958a] contains an intuitionistic proof that **HPL** is weakly complete for negative formulae. It may disappoint that **HPL** be proved complete first for "boringly conventional" brands of statement outside its realm of concern.

Third, Kreisel [1958b] displays the connection between topological completeness theorems such as Tarski's and standard completeness questions for **HPL**. Forging the connection is an idea of Brouwer. A Brouwerian choice sequence is an "incomplete object" whose course of values may not be fully determinate, though subject to possible further requirements. A limiting case, which might be thought degenerate *a priori*, is a sequence on which no (first-order) requirement is permitted. Brouwer imagined this brand of "lawless" sequence, mentioned it in a 1924 letter to Heyting but did not exploit it in any early publication. It is surprizing that the first (strikingly nondegenerate) use of lawless sequences was to a study – mathematical logic – Brouwer himself discouraged. Tarski and Beth had already obtained completeness theorems for **HPL**, exploiting conventional logic over topological interpretations. Kreisel proved that, granting lawless sequences, completeness over topological models implies standard completeness.

Finally, Kreisel [1970] gives a proof of incompatibility between completeness of **HPL** and Church's Thesis.

These theorems and their surrounds underscore peculiarities. First, it is peculiar that the concepts 'disappoint,' 'discourage,' 'surprize,' put to such good use in the critique of serious graphic and musical art, get refused by serious philosophers in the critique of our greatest, mathematical art – and in logic, among its artier regions. Second, on such "natural" foundational views as Kreisel's informal rigour Kreisel [1967], it would be peculiar to uncover the fact that tough spadework in foundations is more agonistic than journalistic. In formalizing and regularizing, logicians are not always attempting faithful, formal record of extant, elusive, informal concepts. Instead, they seem to struggle - with their peers and with themselves - to articulate constraints on future mathematics and to see constraints enforced. Given informal rigour, it would be peculiar if it turned out to be the political history of the struggle for constraint that grants mathematical concepts their intrinsic characters.

On the picture of foundations Kreisel [1967] draws, that is, on the picture of informal rigour, at any stage in a living mathematics, logicians are seen to be struggling towards regularizations, even formalizations, of concepts and notions Kreisel believes already to subsist informally. Contours of the original concepts are, mysteriously, supposed to guide and, eventually, to justify future efforts with the concept. In the best cases, the happy results of struggle will consist in definitive proofs that our formal concepts agree with their informal originals. Logicians are thought able often (perhaps even always) to peer into the past for certification of their foundational efforts. Current theories, current utterances are certified - to the extent that they are - when they accord with previously subsisting – but unseen – informal concepts or determinate facets of those concepts.

This accord must possess two further properties. First, whether a fixed statement or formulation accords with a concept or not is a matter of a liaison with the past. Hence, it lies outside personal control. It is not subject to the will. Second, the intuitive concept Kreisel imagines to presubsist its mathematical expression should be independent of its formalization, if the certification it is intended to offer is to be noncircular. Hence, it cannot rely for its certification upon details of the formalization itself. As we shall see, there is some question whether Kreisel's research into intuitionistic completeness satisfies these conditions.

On an opposing view, mathematicians of the foundations are not struggling towards formalizations which are successful in so far as they accurately report upon the dispositions of informal concepts subsisting prior to formalization, concepts which, though on the tips of their minds, were only articulable later. Foundational researchers do not give tentative renderings today of something that will someday banish from foundations the sense of tentativeness. Mathematical formulations and proofs are not always, or even normally, the records of informal notions. They are, rather, constraints imposed upon future approaches to the notions in question, notions emerging very slowly and haltingly. Perhaps a mathematical concept is nothing more than a series of constraints laid down over time. By their natures, these constraints are conditions over which selected individuals can often exercise willful control. Constraints are subject to individual and corporate designs.

The following is neither comprehensive survey nor elementary introduction. Important materials pertinent to completeness, materials which appear (perhaps too prominently) in Dummett [1977], in Troelstra [1975] and in Troelstra and van Dalen [1988], are omitted. Among these are the researches of Veldman and de Swart into "exploding" Beth and Kripke models, which contribute to inquiries set in motion by Kreisel and Gödel. Closely akin is Friedman's elegant constructive proof of completeness for minimal logic, set out in Troelstra and van Dalen [1988, p. 685ff]. Also omitted are replies to a question put, in effect, in Kreisel [1965, p. 147]: "Is there a completeness theorem, even for intuitionistic propositional logic, without lawless sequences?" Negative answers are available from McCarty [1991].

Preliminaries

It is standard to separate weak and strong concepts of intuitionistic completeness for single formulae.

Definition. (Strong Completeness) *For any formula ϕ, if ϕ is valid, then ϕ is a theorem of* **HPL**.

Definition. (Weak Completeness) *For any formula ϕ, if ϕ is not a theorem of* **HPL***, then ϕ fails to be valid.*

Note. Only with special permission does weak completeness entail strong. They are coextensive under the influence of a form of Markov's Principle.

Markov's Principle or **MP** asserts that, if $M(n)$ is a primitive recursive numerical predicate, then $\neg\neg\exists n.M(n)$ implies $\exists n.M(n)$:

Definition. Markov's Principle (MP) *is the schema*

$$\neg\neg\exists n M(n) \to \exists n M(n)$$

for primitive recursive predicates $M(n)$.

Definition. Generalized Markov's Principle (MP$\boldsymbol{\alpha}$) *is the schema*

$$\forall\alpha\neg\neg\exists n M(\alpha, n) \to \forall\alpha\exists n M(\alpha, n)$$

for primitive recursive predicates $M(\alpha, n)$ with α ranging over binary-valued natural number functions.

Definition. Weak Generalized Markov's Principle (WMP$\boldsymbol{\alpha}$) *is the schema*

$$\forall\alpha\neg\neg\exists n M(\alpha, n) \to \neg\neg\forall\alpha\exists n M(\alpha, n)$$

for primitive recursive $M(\alpha, n)$ as above.

The latter two versions of **MP** take into consideration the possibility that M carry parameters for higher-order constructs.

Definition. Markov's Principle for decidable predicates (MPD) *is the claim that, for $P(n)$ a decidable predicate of numbers,*

$$\textit{if} \quad \neg\neg\exists n P(n), \quad \textit{then} \quad \exists n P(n).$$

Many sources reserve the expression 'Markov's Principle' for **MPD**. For the first version of **MP** just displayed, the term 'primitive recursive Markov's Principle' is in common use. Smorynski [1973, p. 360ff] contains a brief survey of several forms of Markov's Principle and their behaviors in **HA**. McCarty [1988b] examines equivalents of **MPD** within intuitionistic set theory and second-order arithmetic. Specialists maintain that **MP** or **MPD** is a test for "non-intuitionistic" in that, if Φ proves **MP** intuitionistically, then Φ is not itself intuitionistically correct. Usually, the efficacy of this test is not confirmed by a proof. Instead, it may be claimed that versions of Markov's Principle are inconsistent with the "constructive meanings of the connectives and quantifiers." An expression of doubt about this claim appears below. Suffice it to note that none of the above formulations of Markov's Principle are derivable in (appropriately natural extensions of) **HA**. See Kreisel [1958c].

Many intuitionistic formal systems, while not deriving **MP**, are yet closed under the correlative rule **MR**:

Definition. (**Markov's Rule (MR)**) *A formal system S is closed under Markov's Rule provided that, for primitive recursive $M(n)$, whenever $S \vdash \neg\neg\exists n M(n)$, then $S \vdash \exists n M(n)$.*

Any intuitionistic system admitting the Friedman-Dragalin translation Friedman [1978] is closed under **MR**.

In intuitionistic contexts, one has the following

Definition. Church's Thesis (CT) *asserts that if $\forall x \exists y \phi(x, y)$, there exists an index e such that $\{e\}$ is total and uniformizes ϕ.*

Let **WCT** be the consequence of **CT** asserting that every decidable set of natural numbers is not not recursive. **CT** is consistent with an array of "intuitionistic" formal systems, including the ZF set theory **IZF** and set theory plus Brouwer's Continuity Theorem. (For axioms, see Troelstra and van Dalen [1988, p. 624].) **CT** entails that every decidable property of natural numbers is recursive. **CT** is commonly employed as a test for "nonconstructiveness" as follows: if **CT** implies that statement ϕ fails, then (it is argued) ϕ is not provable constructively.

As mentioned, a negative formula of pure logic is one composed entirely of negated atoms combined without the use of $\vee$ or $\exists$:

Definition. (Negative Formula) *A first-order formula is negative when it contains neither $\vee$ nor $\exists$ and when, in it, every atomic subformula appears prefixed by a negation. For a formula of elementary arithmetic to be negative it suffices that it contain neither $\vee$ nor $\exists$.*

Negative translation theorems, as in Gödel [1933] and Kolmogorov [1926], reveal that negative formulae express properties shared with conventional systems. Let '**CPL**' stand for conventional first-order logic and '**HA**' for Heyting's intuitionistic arithmetic. Then, in primitive recursive arithmetic,

Theorem. *For negative ϕ, $\mathbf{CPL} \vdash \phi$ if and only if $\mathbf{HPL} \vdash \phi$.*

Theorem. *For negative ϕ of elementary arithmetic, $\mathbf{PA} \vdash \phi$ if and only if $\mathbf{HA} \vdash \phi$.*

Gödel's Intuitionistic Incompleteness

Gödel's opinion (quoted above) notwithstanding, once *Principia Mathematica* had popularized an axiomatic foundation for conventional logic, the completeness question did not so much *arise* as painfully emerge. In view of, say, the Frege-Hilbert correspondence, it would be more accurate to admit that the question was *forced* into the light. Completeness would not have arisen of its own accord with Frege, Russell or Whitehead. Close onto *fifty years* elapsed between Frege's formulation of quantifier logic and the statement of the question in Hilbert and Ackermann. Had the problem arisen "naturally," it would have done so for Skolem and, in Skolem [1922], he may well have solved it.

More surprizing than misconstruing the past of *conventional* completeness is anticipating the future of its *intuitionistic* analogue. Gödel's dissertation asked after an intuitionistic completeness proof for intuitionistic logic. This would seem a case of true arising, seemingly *ex nihilo* and without provision of axiomatic foundations for the relevant logic. Formal intuitionistic logic did not at the time exist. Heyting's formalization of intuitionistic logic and arithmetic Heyting [1930a, b and c], main lines of which are now considered definitive, would not appear in print until 1930. Kolmogorov [1925] contains an incomplete system for predicate logic and raises a completeness question for a fragment of its propositional part. But this article, written in Russian, was unknown to Gödel in 1929.

Further statements Gödel then made about the intuitionistic completeness question proved, in light of experience, uncannily accurate:

> In conclusion, let me make a remark about the *means of proof* used in what follows. Concerning them, no restriction whatsoever has been made. In particular, essential use is made of the principle of the excluded third for infinite collections (the nondenumerable infinite, however, is not used in the main proof). It might perhaps appear that this would invalidate the entire completeness proof. For what is to be proved can, after all, be viewed as a kind of decidability (every expression of the restricted functional calculus either can be recognized as valid through finitely many inferences or its validity can be refuted by a counterexample). ... From the intuitionistic point of view, the entire problem would be a different one, because already the meaning of the statement 'A system of relations *satisfies* a logical expression' (that is, the sentence obtained through substitution is true) would be a fundamentally different one. For we would then have to require that the existential assertions occurring in the expression be constructively proved. It is clear, moreover, that an intuitionistic completeness proof (with the alternative: provable or refutable by counterexamples) could be carried out only through the solution of the decision problem for mathematical logic. Gödel [1929, pp. 63-65]

In retrospect, Gödel's prognosis was surprizing on all major counts:

1. The inference forms which are actually instantiated in his original completeness proof are irreducibly conventional.

2. There will be no intuitionistic proof of any theorem

3. roughly as general as the theorem formulated by Gödel, on pain of solving problems now known recursively unsolvable.

4. The completeness problem(s) of intuitionistic mathematics are, as Gödel had it, "fundamentally different." And this is due in no small part to uncertainty surrounding "the meaning of the statement 'A system of relations *satisfies* a logical expression.' "

Ad 1. In his dissertation, Gödel [1929] had succeeded in proving that, for denumerable first-order Γ,

either Γ is satisfiable or Γ is formally refutable.

Note. One reports that Gödel had unquestionably proved this even though Tarski's *Wahrheitsbegriff* would not be published until 1933 and an identifiable model *theory*, including an unrestricted definition of satisfaction, would not arise until after World War II.

Let ϕ be any first-order formula underivable from Γ. Given conventional logic, $\Gamma \cup \{\neg\phi\}$ is not formally refutable. By Gödel's dissertation, $\Gamma \cup \{\neg\phi\}$ is satisfiable. Therefore, Γ fails to entail ϕ semantically and

$$\Gamma \nvdash \phi \text{ implies that } \Gamma \not\models \phi.$$

Contraposition yields the statement known to undergraduates:

$$\Gamma \models \phi \text{ implies that } \Gamma \vdash \phi.$$

The indebtedness of completeness proofs to conventional logic is in plain sight. Even the first step,

from "$\Gamma \nvdash \phi$" to "$\Gamma \cup \{\neg\phi\}$ is not refutable,"

is nonconstructive for propositional logic. The final inference of the series, an instance of contraposition, is also out of constructive bounds.

Ad 2. Each member of this trio of completeness statements

- **A**. Either Γ is satisfiable or it is refutable.
- **B**. If Γ is not refutable, then it is satisfiable.
- **C**. If $\Gamma \models \phi$ then $\Gamma \vdash \phi$.

is intuitionistically unprovable, even assuming the intuitionistic set theory **IZF**. Were either of the first two intuitionistically provable, unsolvable problems would be solvable, as Gödel opined.

Were **A** true for single formulae then, intuitionistically,

> for arbitrary ϕ, either $\neg\phi$ is provable in conventional first-order logic or it is not.

Church's Thesis enforces the intuitionistically inadmissible but formally consistent (with **IZF**) demand that every decidable numerical property be recursive. So, were Church's Thesis true and **A**, theoremhood in first-order logic would be decidable.

Next, assume that Γ in **B** is **PA**, conventional first-order arithmetic. By the negative translation, **PA** is not refutable unless Heyting's intuitionistic arithmetic **HA** is. So, if **B** were intuitionistically true, then **PA** would be satisfiable. But Church's Thesis implies that the satisfiability of **PA** violates the First *In*completeness Theorem. Let $\Im$ be a model of **PA**. Since **PA** obeys the law of the excluded third,

$$\text{for all sentences } \phi \text{ of } \mathbf{PA}\text{'s language, either } \Im \models \phi \text{ or } \Im \not\models \phi.$$

The theory of $\Im$ is, therefore, a complete decidable extension of **PA**. By Church's Thesis, the theory of $\Im$ is recursive. Given that **IZF** is consistent with Church's Thesis and proves the incompleteness of arithmetic, none of the principles embodied in **IZF** proves **B**.

In addition, **B** entails a restricted but unacceptable form of the law of the excluded third, without help from Church's Thesis. Let S be a set of natural numbers. Consider a denumerably infinite collection of propositional atoms $\{p_n : n \in \omega\}$. Let T_S be

$$\{p_n : n \in S\} \cup \{\neg p_n : n \notin S\} \cup \mathbf{CL}.$$

CL is a set of axioms for conventional propositional logic. T_S is plainly consistent. But, just as plainly, if T_S had a model $\Im$ then, for all $n \in \omega$,

$$\text{either } \Im \models p_n \text{ or } \Im \not\models p_n$$

Hence, S satisfies the *principle of testability*:

$$\forall n \in \omega(\neg\neg n \in S \vee \neg n \in S).$$

So, if either **B** or **A** were true, testability would hold.

Even a weakened rendering of **C** does not stand the test of consistency with Church's Thesis. Consider this consequence of **C**:

$$\text{if } \Gamma \nvdash \phi \text{ then } \Gamma \not\models \phi.$$

Take $\perp$ to be a standard formal contradiction. As just proved on the assumption of Church's Thesis, $\mathbf{PA} \models \perp$. Hence, **C** is not provable in **IZF**, since **PA** is irrefutable if **HA** is. An argument similar to the above, but with more careful treatment of negation, shows that this consequence of **C** entails

$$\forall S \neg\neg\forall n \in \omega(\neg\neg n \in S \vee \neg n \in S),$$

which is intuitionistically unacceptable.

Ad 3. In his original proofs of completeness, Gödel did not rely upon any noncircular mathematical definition or concept of the truth of an arbitrary formula in a structure. Nor did he call for an analysis of the respective contributions of connectives and quantifiers to the satisfaction of a formula or to the truth of a sentence. Gödel's 1929 notion of satisfaction was a piecemeal business: each formula of the first-order language determines its individual truth conditions by substitution. Let R^A be a subset of, f^A a unary function on and a a member of the nonempty domain A. This A Gödel would (in 1929) have termed not "domain" but '*Denkbereich*' – 'domain of thought.' Formula Rfx is satisfied by the system

$$< R^A, f^A, a >$$

just in case "it yields a proposition that is true (in the domain in question) when it is substituted in the expression" Rfx. Gödel [1930, p. 69] Overlooking insult to use and mention, Gödel's intention is clear: Rfx is satisfied by the given system over *Denkbereich* A just in case the sentence $R^*f^*(a^*)$ is true, wherein the item denoted using the asterisk names the respective element of the system. This notion of satisfaction leaves the general concept "is true" unexplicated. It relies on nothing more exacting than a native ken of how to symbolize and unsymbolize statements.

In the case of intuitionism, little more is currently available. For a logic **L** to be complete, it must be the case that every intuitively correct **L**-formulae is **L**-provable. It is nowise obvious, however, that there is a general concept of what it means intuitionistically for a predicate formula to be intuitively correct. There has been little general agreement among intuitionists over the range of the quantifier in the expression 'given any structure $\Im$, ϕ is satisfied by $\Im$.' Nor is there a working concept of interpretational structure operating throughout intuitionistic thought.

These facts, albeit untoward, are in keeping with Brouwer's constraints upon intuitionism. Among the most notorious was that intuitionistic mathematics remain, relative to philosophy and metaphysics, an absolutely free creation. In consequence, the Brouwerian intuitionist may live under an obligation not to establish, once and for all, what it is to be an intuitionistic *Denkbereich*. Were he or she to set *a priori* limits on what counts as a legitimate domain of intuitionistic thought, intuitionists may no longer be able to refuse the efforts of philosophical logicians to exert semantic authority over them.

If we understand such expressions as '$R^*f^*(a^*)$ is true' in Gödel to express an intuitionistic condition, to refer implicitly to the existence of a construction which would guarantee the truth of the sentence, then Gödel's notion expresses a constraint on truth to which Kreisel appealed in his first 1958 article on the completeness of intuitionistic predicate logic. In the opening section of Kreisel [1958a, p. 317], one surprizingly entitled *The intended interpretation*, the author provided this definition of completeness and employed Gödel's retail conception of satisfaction. I quote:

> Suppose the predicate symbols of the formula Φ are P_1, $\ldots$, P_k. Then 'Φ' should be provable in the calculus (viz., that of Heyting) if and only if, for all species of individuals and all predicates $P_1^*, \ldots, P_k^*$ defined on such a species, the proposition $\Phi(P_1^*, \ldots, P_k^*)$ holds intuitionistically.

Kreisel, in the same article [1958, p. 318], asserted that the notion of truth or "holding" at work here is an intuitionistic one; he writes:

> Naturally, we mean here intuitionistic truth, i.e. constructive provability.

The author of the article then attempted to reassure his readership that the vagueness of the notion is not injurious by writing, "It is to be expected that, if [the completeness statement] is true, *it can be sharpened by a weakening of the hypothesis* in such a way that the dubious totalities [of all species and properties] are eliminated." Kreisel [1958a, p. 318] (There is, however, no indication of means to deal with such ambiguities in the [real] situation in which completeness fails.)

One should ask whether Kreisel's importation of Gödel's satisfaction into intuitionism competently represents any intuitionistic concept of *absolutely general* validity. If one makes explicit the implicit reference to constructions in the definition of validity just given, then, according to Kreisel's importation, Heyting's predicate calculus is intuitionistically complete on condition that

> 'Φ' is provable in the calculus if and only if, for all species of individuals and all predicates $P_1^*, \ldots, P_k^*$ defined on such a species, there is a constructive proof of the proposition $\Phi(P_1^*, \ldots, P_k^*)$ obtained by substitution.

Since the clause asserting the existence of the proof lies within the scope of the universal quantifiers over species and predicates, this concept attaches to a formula provided that, for each system of species and properties, a relevant reason for the formula to be true as interpreted is constructible. There is no guarantee, apart perhaps from an axiom of uniformity on species, that these reasons be wholly uniform; they may depend for their identity crucially upon idiosyncracies of the particular species under consideration.

This is no clear relative of validity in the conventional case. When a formula is classically valid, the completeness theorem assures us that a single line of reasoning, spelt out by a derivation in a first-order system, suffices to cover all the possible instantiations of the formula. Nor does it seem closely related to the content of Brouwer's utterances on the topic. When Brouwer does refer to (what is in effect) the validity of a form, he does seem to ask that its truth be certified by a constructive proof of validity which applies uniformly and schematically across all possible instantiations. (The reader should consider, as an example, the proof on page 12 of Brouwer [1981] that

$$\neg\neg\neg p \leftrightarrow \neg p$$

is a valid form.) Clear formulation of Brouwer's idea would require a straightforward uniformization of the Gödel-Kreisel proposal.

Note. In defining satisfaction, the present writing does not rely upon a formula-by-formula grasp of what it is to symbolize (or unsymbolize). In effect, another suggestion of Kreisel is adopted: "to introduce a so-called semantic definition of intuitionistic truth along the lines of Tarski's definition of truth," only keeping in mind "that the logical constants in the definition would have to be interpreted intuitionistically." Kreisel [1958a, p. 318] One might say that a constraint is here being assayed: that the *intuitionistische Weltauffassung* will eventually grow to incorporate an intuitionistic model theory resting upon set theory. Whether this constraint is binding remains to be seen.

Completeness and Church's Thesis

Nothing like conventional proofs of completeness yielding **A** through **C** above will be forthcoming in the name of Brouwer's intuitionism. The counterexamples to Gödel's general completeness theorems earlier proferred call upon infinite collections Γ of formulae, for **PA** is provably inequivalent to any finite extension of **HA**. Therefore, the question of completeness for finitely many formulae remains unanswered in the foregoing. Thanks to results – soon to be described – of Gödel and Kreisel, specialists commonly insist that the road to finitary forms of intuitionistic completeness is permanently blocked. This is despite the facts, first, that there is now no generally accepted mathematical representation for intuitionistic truth, satisfaction or validity. (There seems insufficient agreement among intuitionists on these topics to aver that their efforts are guided by common concepts of truth and validity.)

Secondly, certain principles upon which the presumptive roadblocking relies are yet unproved and, perhaps, unprovable. What looks to be the initial roadblock – to the constructive provability of completeness – involves Church's Thesis.

Kreisel proved Kreisel [1958a, pp. 322–323] constructively that completeness holds for **HPL** with respect to negative formulae:

Theorem. *Constructively,* **HPL** *is weakly complete for all negative formulae.*

Proof. Assume that formulae are pure – without function and individual signs. Let ϕ be a negative formulae and let Φ be the standard arithmetization of the statement that ϕ is unprovable from **HPL**. From the first negative translation theorem above, Φ is equivalent in **PA** to the statement that ϕ is underivable in **CPL**. The Hilbert-Bernays form of the *conventional* completeness theorem for first-order logic Smorynski [1977, p. 860] shows that there exist suitable Δ_2 predicates of arithmetic such that, if ϕ^* is the interpretation of ϕ supplied by these predicates, then

$$\mathbf{PA} \vdash (\Phi \rightarrow \neg\phi^*).$$

Φ can be assumed negative. Conventional **PA** underwrites the ability to insist that the Δ_2 predicates be negative, too. In that case, ϕ^* itself is negative and the negative translation theorem yields also that

$$\mathbf{HA} \vdash (\Phi \rightarrow \neg\phi^*).$$

HA is certainly sound. Therefore, if ϕ is unprovable, it is not valid. □

Note. Intuitionistic nontheorems such as $\phi \vee \neg\phi$ cannot have conventional countermodels. So, more extensive completeness results for **HPL**, ones embracing formulae well outside the "conventional zone" of negative formulae, would call for unconventional intuitionistic assumptions. Kreisel [1958a, p. 326] commented:

> "The proofs of weak completeness for the fragments of the predicate calculus considered above are special in the following respect: *they can be formalized in a subsystem of classical mathematics,* namely, Heyting's arithmetic. It is clear that there is no completeness proof even for the propositional calculus in such a system."

Here is a straightforward and obvious proof of Kreisel's result on completeness and **CT** from M$\underline{\text{c}}$Carty [1991, p. 334]. Afterwards comes a sketch of the original proof and later refinements.

Theorem. *Weak completeness is inconsistent with* **CT**.

Proof. The idea is to modify a proof earlier described, that using **PA** together with **CT**. Let **Q** be a finite set of **HA** axioms sufficient to insure the **HPL**-provability of the simple properties of Kleene's **T** predicate and the finite number of instances of Gödel's Fixed-Point Theorem required below. Let **Test** be the sentence expressing the testability of the halting predicate:

$$\forall x \forall y(\neg\exists z \mathbf{T}(x,y,z) \vee \neg\neg\exists z \mathbf{T}(x,y,z))$$

As usual, $\exists z\mathbf{T}(x,y,z)$ means that the partial recursive function with index x converges on input y.

HA suffices to prove consistency of **Q**+**Test**:

$$\mathbf{Q} + \mathbf{Test} \ \not\vdash \bot\ .$$

If weak completeness were here to hold, then $\neg$(**Q**+**Test**) should fail to be valid. But this is not the case: on the assumption of **CT**, $\neg$(**Q**+**Test**) turns out to be logically true.

Consider the prospect that $\Im \models \mathbf{Q} + \mathbf{Test}$. Then $\Im$ obeys the law of excluded third for such formulae as $\neg\exists z\mathbf{T}(n,m,z)$. Associated with these are formulae one might call *instances*:

Definition. (**Instances**) *An instance of* $\neg\exists z\mathbf{T}(n,m,z)$ *– or simply an instance – is any formula* χ *constructed by applying the standard fixed-point calculation to* $\neg\exists z\mathbf{T}(n,x,z)$.

It follows that $\mathbf{Q} + \mathbf{Test}$ deduces (intuitionistically) that

$$\chi \leftrightarrow \neg\exists z\mathbf{T}(n,\ulcorner\chi\urcorner,z).$$

All instances are either true or false in $\Im$. **CT** entails that there is a partial recursive function $\{e\}$ such that, for any instance χ, $\{e\}$ outputs 0 if $\Im \models \chi$ and is undefined otherwise. Given the representability and functionality properties of the **T** predicate formally provable from **Q**, each instance χ is such that

$$\Im \models \exists z\mathbf{T}(e,\ulcorner\chi\urcorner,z) \text{ if and only if } \Im \models \chi.$$

Now, fixed-points in **Q**, set to work on the formula $\neg\exists z\mathbf{T}(e,x,z)$, produces a special instance χ such that

$$\Im \models \chi \text{ if and only if } \Im \models \neg\exists z \mathbf{T}(e, \ulcorner\chi\urcorner, z).$$

$\neg$(**Q**+**Test**) is, therefore, both intuitionistically valid and **HPL**-unprovable. □

Note. Since the result's statement is negated – **HPL** is *not* complete – an anticlassical principle weaker that **CT** would have sufficed.

Kreisel first remarked *en passant* at Kreisel [1962, p. 140] that **CT** blocks weak completeness for single formulae. A proof sketch for the following more general theorem appeared later, in Kreisel [1970].
Theorem. **CT** – *in the presence of other significant assumptions – implies that intuitionistic validity is not recursively enumerable.*

Proof: Proofs of the Gödel-Kreisel theorems Kreisel [1962] show how to extract from any primitive recursive binary tree T a formula ϕ_T such that, intuitionistically,

$$\text{all constructive paths through } T \text{ are noninfinite iff } \phi_T \text{ is valid.}$$

The nontrivial assumptions that enter into the proof of the displayed biconditional include **DC**, the axiom of dependent choice. With **CT**, the displayed statement entails

$$\text{all recursive paths through } T \text{ are noninfinite iff } \phi_T \text{ is valid.}$$

By applying a negative translation to theorems of classical recursion theory proved by Jockusch, one can show that there is a recursive function p such that, for all e,

$$p(e) \notin W_e \text{ iff every recursive path through } T(p(e)) \text{ is noninfinite,}$$

where W_e is the eth r.e. set and $T(n)$ is the nth primitive recursive binary tree. It follows intuitionistically from the facts displayed that, for any e,

$$\phi_{T(p(e))} \text{ is valid iff } p(e) \notin W_e.$$

Now, if the set of intuitionistic validities is r.e., so is

$$\{\phi_{T(p(e))} : \phi_{T(p(e))} \text{ is valid}\}.$$

Let the index of this set be j. Then,

$$\phi_{T(p(j))} \text{ is valid iff } p(j) \in W_j.$$

On the other hand,

$$\phi_{T(p(j))} \text{ is valid iff } p(j) \notin W_j. \quad \square$$

This proof underwent improvement in Leivant [1976], where its mathematical and recursion-theoretic overheads were reduced. Leivant's proof also afforded the intelligence that

$$\{\phi : \phi \text{ is intuitionistically valid}\}$$

is inconsistent with a weak form of **CT** asserting that every decidable number-theoretic predicate is not not r.e.

Although the notion of intuitionistic validity remains only loosely circumscribed, Kreisel [1970, p. 126] posed the following question as "an analogue to Skolem-Löwenheim":

> **Problem 4:** *Is there a definition (in the theory of species of natural numbers or even in arithmetic) of the species of constructively valid formulae in the language of predicate calculus (without having to decide Church's Thesis)?*

Here is a set of theorems from M$\underline{\text{c}}$Carty [1988a] which reply to the arithmetic part of this question. Here, "Γ is categorical" denotes the same situation it denotes in conventional metamathematics: that absolutely all models of Γ are isomorphic. In any extension of **HA** containing the means necessary to speak of arbitrary structures and containing an axiom of nonchoice, one can prove that

Theorem. *Given* **WCT** *plus* **MPD**, *first-order intuitionistic arithmetic* **HA** *is categorical.*

As did the proofs of Kreisel and Leivant, this one also makes essential use of the behavior of inseparable r.e. sets in conventional mathematics. In fact, the argument for this theorem is a constructivization of Tennenbaum [1959]. A more delicate version of the same argument shows that there is a sentence Θ, drawn from among theorems of **HA**, such that,

Theorem. *Assuming* **WCT** *and* **MPD**, Θ *is categorical.*

It follows with very little work that

Corollary. *Assuming* **WCT** *and* **MPD**, *intuitionistic validity is not arithmetically definable.*

It would be worth asking, "What can we say about intuitionistic validity without **MPD**?" For an answer, this definition is apposite:

Definition. (**Θ-stable**) *Formula ϕ of arithmetic is Θ-stable just in case Θ proves all ϕ's universalized subformulae to be invariant with respect to double negation. In other words, if ϕ is Θ-stable, for all subformulae of ϕ of the form $\forall x\psi$,*

$$\Theta \vdash (\psi \leftrightarrow \neg\neg\psi).$$

Then we can prove that,

Theorem. *Assuming* **WCT**, *intuitionistic validity is not definable by any Θ-stable formula.*

A fortiori, intuitionistic validity is not r.e. if **WCT** holds. Nor will it be definable by any negative formula or formula devoid of universal quantifiers.

Note. In the proofs of this section, the impact of even perfectly simple facts of classical recursion theory on the metamathematics of intuitionistic systems is apparent. Yet explicit applications, even borrowings, of recursion-theoretic notions remain in intuitionistic mathematics relatively rare. This is surprizing if one recalls the number and variety of conventional theorems about such entities of intuitionistic interest as Dedekind-finite sets – which logicians examined under the title 'isols' – ordered sets and decidable or semidecidable models already obtained by recursive mathematicians. There is here no recommendation that extant theorems about recursive structures be carted slavishly over into intuitionistic mathematics. Mindless efforts would unlikely be successful and would, most certainly, prove miserably dull.

Completeness and Markov's Principle

The relations between **MP** and completeness discovered by Gödel and refined and published by Kreisel [1962, p. 140] may be the most significant facts about **MP** – and about completeness – now known. The facts are these:

1. Strong completeness implies $\mathbf{MP}\boldsymbol{\alpha}$.

2. Weak completeness implies $\mathbf{WMP}\boldsymbol{\alpha}$.

3. Strong completeness for negative formulae implies **MP**.

The last is said to demonstrate that Kreisel's theorem on the weak completeness of the negative fragment Kreisel [1958a] cannot, within the strictures of pure intuitionism, be upgraded to strong completeness.

Again, exposition of the original Gödel-Kreisel argument is preceded by a simpler alternative exploiting the fact that many intuitionistic systems are closed under **MR**.

Theorem. *Strong completeness implies* **MP**.

Proof. Let $M(n)$ be primitive recursive and let $\mathbf{Q}$ be a natural finite set of **HA** axioms and recursion equations sufficient for

$$\mathbf{N} \models M(n) \text{ iff } \mathbf{Q} \vdash M(n).$$

to hold for all n. Here, $\mathbf{N}$ is the standard model. It follows immediately that

$$\mathbf{N} \models \exists x M(x) \text{ iff } \mathbf{Q} \vdash \exists x M(x).$$

$\mathbf{Q}$ can be presumed to deduce the Friedman-Dragalin translation of any of its own formula. So, $\mathbf{Q}$ is closed under **MR**. Now, assume that $\mathbf{N} \models \neg\neg\exists n M(n)$ and that strong completeness holds for formulae of the language of $\mathbf{Q}$. Let $\Im$ be any model of $\mathbf{Q}$. Since $\mathbf{N} \models \neg\neg\exists n M(n)$, it follows from the line displayed above that $\neg\neg(\mathbf{Q} \vdash \exists n M(n))$. Soundness of **HPL** yields that $\Im \models \neg\neg\exists n M(n)$. Altogether, then, this shows that

$$\mathbf{Q} \models \neg\neg\exists n M(n).$$

From the assumption of completeness, one sees that $\mathbf{Q} \vdash \neg\neg\exists n M(n)$. Since $\mathbf{Q}$ is closed under **MR**,

$$\mathbf{Q} \vdash \exists n M(n).$$

Finally, the soundness theorem enters in once more to produce

$$\mathbf{N} \models \exists n M(n).$$

Therefore, **MP** follows from strong completeness. □

Similar arguments, but utilizing extensions of $\mathbf{Q}$, lead to parallel proofs of the more general Gödel-Kreisel results concerning $\mathbf{MP}\boldsymbol{\alpha}$. Moreover, one can adjust the arguments so that, rather than assuming completeness for $\mathbf{Q}$'s language – a language containing functions symbols in addition to predicates – one can get by with completeness for a purely predicate language.

In the Gödel-Kreisel proofs Kreisel [1962], instead of applications of **MR**, one finds use of **DC**, a principle of dependent choice. Also at work is a form of Herbrand's Theorem for **HPL** allowing the reduction of an **HPL**-derivation of a negated prenex formula to the propositional derivation of a negation of a conjunction of suitably chosen instances.

Theorem. *Strong completeness implies* $\mathbf{MP}\boldsymbol{\alpha}$*, and Weak completeness implies* $\mathbf{WMP}\boldsymbol{\alpha}$*.*

Proof. At its roughest, the idea is to use completeness to prove $\mathbf{MP}\boldsymbol{\alpha}$ by interpolating the correctness of a first-order formula between the assumption of the antecedent and the conclusion of the consequent of Markov's Principle. One begins with$M(\alpha, n)$ and shows how (primitive recursively) to associate with it a distinctive formula ϕ_M whose validity "expresses" the assertion $\forall\alpha\exists n M(\alpha, n)$ in the sense that conditions **A** and **B** hold:

- **A** If $\forall\alpha\neg\neg\exists n M(\alpha, n)$ then $\models \phi_M$ and
- **B** If $\vdash \phi_M$ then $\forall\alpha\exists n M(\alpha, n)$.

The burden of proof rests on demonstrating **A** and **B**:

Lemma A. *If* $\forall\alpha\neg\neg\exists n M(\alpha, n)$ *then* $\models \phi_M$*.*

Proof. Consider the formula ψ_M which is

$$S_M \to T_M.$$

S_M is to represent the "operating system" of the primitive recursive "machine" $M(\alpha, n)$. T_M expresses the claim that, on α, the run has been successful: that there is an n such that machine $M(\alpha, n)$ halts on n. Less metaphorically, S_M is the lengthy but predictable formula which can be written in a language of pure predicate logic and sets out the primitive recursive defining equations for $M(\alpha, n)$. S_M also requires that zero (in the form of a "zero predicate") and the successor relation be well-behaved. T_M is the existential formula asserting that the defining conditions for $M(\alpha, n)$ are satisfied for some n.

Take ϕ_M to be $\neg\neg\psi_M$, equivalently,

$$\neg(S_M \wedge \neg T_M).$$

To prove the lemma for this ϕ_M, assume that $\forall\alpha\neg\neg\exists n M(\alpha, n)$ or $\forall\alpha\neg\forall n\neg M(\alpha, n)$, entertain the prospect that $\neg\psi_M$ be satisfied and show that this supposition leads to a contradiction. So, assume that $\Im \models (S_M \wedge \neg T_M)$. Then, $\Im \models S_M$. Since S_M was constructed to assert that zero and successor are well-behaved, one can use **DC** to extract from $\Im$ a substructure $\wp$ which is an isomorph of the natural number structure with zero and successor. One then proves by induction, using the truth of S_M in $\Im$, that $M(\alpha, n)$ holds in $\mathbf{N}$ just in case its defining condition holds in $\wp$. From the other assumption, that $\Im \not\models T_M$, one can extract from $\wp$ a binary function α such that

$$\forall n \neg M(\alpha, n),$$

contradicting the initial insistence that $\forall\alpha\neg\forall n\neg M(\alpha, n)$ holds. Therefore, the arbitrarily selected structure $\Im$ must actually satisfy $\neg\neg\psi_M$ and this verifies that ϕ_M is valid, provided $\forall\alpha\neg\neg\exists n M(\alpha, n)$. □

Lemma B. *If* $\vdash \phi_M$ *then* $\forall\alpha\exists n M(\alpha, n)$.

Proof. Start from the assumption that ϕ_M, above constructed, is derivable in **HPL**. Let α be any binary number-theoretic function. It is easy to see, using the standard (intuitionistic) algorithms for manipulating quantifiers, that $\neg\psi_M$ or $(S_M \wedge \neg T_M)$ is provably equivalent to a prenex formula Θ. If ϕ_M is a theorem, then so is $\neg\Theta$, and $\neg\Theta$ is in a form suitable for the application of Herbrand's Theorem. From this, one obtains a quantifier-free formula provable in propositional logic. The latter formula will, therefore, hold universally over the standard structure $\mathbf{N}$ and that fact is seen to reduce to a disjunction of simple formulae equivalent to

$$M(\alpha, n_0) \vee M(\alpha, n_1) \vee \ldots \vee M(\alpha, n_k)$$

for some numbers $n_0, n_1, \ldots n_k$. $\exists n M(\alpha, n)$ follows directly. □

Note. The vivid talk of "machines" and "halting" was not entirely heuristic. Those familiar with Büchi's proof of the undecidability of multiadic predicate logic Büchi [1962] via explicit formal representations of the workings of machines will readily mark the general similarity between his constructions and the preceding.

Who's afraid of Markov's Principle?

The Gödel-Kreisel theorems concerning completeness and Markov's Principle are generally interpreted so as to bar any prospect for a strictly intuitionistic proof of completeness. After presenting a conventional proof of completeness with respect to Beth models – but before describing the Gödel-Kreisel theorems – Troelstra and van Dalen [1988, p. 694] write:

> The results of the preceding section might lead us to believe that completeness for full [**HPL**] for the notion of intuitionistic validity is within reach. We shall show that, nevertheless, we cannot expect to achieve this.

The idea behind this interpretation of the theorems seems to be that, if an intuitionist could prove the completeness of **HPL**, then he or she would be required to accept **MP**. But no true intuitionist can accept **MP**, so completeness will remain unprovable. This idea seems unexceptionable, even though there may well be no convincing argument that, if an intuitionist is to remain faithful to the basic premises of the subject, then he or she is obliged to refuse **MP**. This could only come as a surprize if one demanded that advances in the foundations of mathematics look for their justification wholly to the past rather than to the future, that they can only be adequate if they accord with the outlines of the intuitive concepts which earlier (possibly secret) cogitations have inked in and if the accord is securable with strict argument. **MP**'s failure is not a theorem of some intuitionistic semantical theory – currently under development.

Rather, it would appear that the unproved assertion of its informal intuitionistic unprovability acts as a constraint which has been set down on the course of intuitionism to come. For this point of view to be convincing, attempted intuitionistic disproofs of **MP** must be unconvincing. One especially wants to query seemingly well-grounded assertions to the effect that **MP** stands in conflict with the *meanings* of the intuitionistic logical signs. Such a query may take off by pointing out that it does not seem encumbent upon us to maintain that the logical signs of the intuitionist differ in meaning from those of the ordinary mathematician. After all, $\models$ in an intuitionistic metatheory commutes as readily – and as happily – with the logical signs as it does in the usual case. Besides, it is not always clear that the best course of action, in the face of intuitionistic mathematics, would be to contend that the presumptive divergence of intuitionism from its conventional brother is due fundamentally to matters of meaning. The intuitionists differ in their mathematics and that seems (significant) enough. Third, despite the efforts of logicians since Bolzano, no conception of meaning has been enunciated which is sufficiently rich that it contains, if only implicitly, all the truths of even so everyday a field as elementary arithmetic. Hence, it seems unlikely that we could reduce all of the intuitionist's "deviance" in mathematics to a circumscribable array of “deviant” meanings. Fourth, on the difference in meaning account, there are problems about the very understanding of what intuitionists are up to. If we assume divergence in meaning, it seems difficult to interpret such familiar claims as “Brouwer attacked standard mathematics by charging the failure of the law of the excluded third.” If the intuitionist’s assertions differ in meaning from ones classically expressed, it may well be that what the intuitionist refuses in denying the validity of $\phi \vee \neg\phi$ is not what one usually asserts by demanding it. Perhaps this worry about meaning is part of the message Kreisel [1964, footnote 3] meant to convey in a remark on Hilbert:

> Considering that the intended meaning of the intuitionistic disjunction is different from that of ordinary disjunction, the rejection of *tertium non datur* is much more like depriving non-commutative algebra of the rule $ab = ba$ than a boxer of the use of his fists.

Moreover, even if talk about divergence in meaning were the best route to an explanation of the apparent peculiarities of intuitionistic mathematics, it would not follow that the contested meanings can, even in simple cases, be boiled down to a disagreement over the meanings of the *logical* signs. Were intuitionism’s disagreements about mathematics fundamentally semantical, it would not follow that every principle of intuitionistic mathematics was to be called up to judgment before the court of pure logic. One does not pass judgments on the adequacy of principles of everyday mathematics on the basis of their congruence with the meanings of the classical logical signs (in, say, truth tables). Could one not coherently and, at times, correctly write down, using the meanings of the conventional signs, “There are more than two truth-values,” even though this is clearly in conflict with the meanings of the signs?

Lastly, talk of reducing intuitionism’s differences to disagreements over meaning is a kind of talk about meaning that makes some philosophical logicians feel good inside. It is just the sort of talk which they are equipped to handle. If that talk were true and those divergent meanings could be philosophically illuminated, then philosophical logicians would exert a measure of control over the intuitionists. However, this manner of control is at odds with the *basic* idea of intuitionism: that the certification for any piece of mathematics should be neither a bit of logic nor a bit of metaphysics. In the hands of philosophers, it seems that what begins as casual talk of alternative meanings turns out, very quickly, to be both logic and metaphysics – more often the latter.

Even if we grant that intuitionistic mathematics can be explicated by reference to characteristically intuitionistic meanings of the logical signs, anyone who maintains that **MP** is thereby seen to be unacceptable has some explaining to do. It is well known that **MP** is *Dialectica*-interpretable and, hence, that it is consistent with such theories as $\mathbf{HA}^{\omega}$, intuitionistic arithmetic in all finite types, plus the Axiom of Choice. **MP** is also demonstrably consistent with **IZF** plus Dependent Choice plus Brouwer's Theorem. Therefore, if some aspect of the mere *explanation* of the meanings of the logical operators stands at odds with **MP**, whatever that part is, its mathematical powers would exceed the proof-theoretic reach of all of intuitionistic set theory. **MP** is inconsistent with theories of lawless sequences and with Brouwer's ideas on the creative subject. Even so, is it true that, by refusing **MP** on the basis of meanings alone, some part of the meanings of the connectives and quantifiers requires acceptance of such matters as lawless sequences?

For an influential example of the attempt to undermine the acceptability of **MP** by attending to the meanings of the logical signs, consider Professor Dummett's arguments [1977, pp. 246 ff]. Concerning **MP** in the form

$$\neg\neg\exists n.P(n) \to \exists n.P(n)$$

where $P(n)$ is decidable, he wrote:

> the intuitionistic statement that $\neg\neg\exists n.P(n)$, or $\neg\forall n\neg P(n)$, does not express the conventional proposition that there exists an n such that $P(n)$, or that it will not happen that we check each n in turn, and find, in every case, that $\neg P(n)$. The intuitionistic statement merely expresses that we shall never be able to prove that $\forall n\neg P(n)$; i.e. that, for however large a number m we may have verified that $\forall n \leq m\neg P(m)$, the possibility will remain open that we may find an $n > m$ for which $P(n)$; and, from this proposition, $\exists n.P(n)$ does not follow, even in its usual sense.

If one avoids interpretations on Beth or Kripke models and understands the "possibility" to which Dummett adverts as double negation, then his final assertion here is equivalent to the claim that $\exists n.P(n)$ does not follow logically from its double negation, even in the presence of the assumption that $P(n)$ is decidable. This is uncontestable; **MP** is no principle of intuitionistic logic. What remains highly contestable is the implicit claim that, if **MP** be true at all, it must be true as a matter of logic (plus principles of meaning). It remains open to suggest that **MP** is a mathematically correct statement governing the behavior of constructions which guarantee the intuitionistic truth of such statements as "$P(n)$ is decidable." If it be possible that, for decidable $P(n)$, there exist an m such that $P(m)$ then it seems reasonable to insist that this possibility be revealed by working the construction recipe that, for each n, decides the truth of $P(n)$. Here, **MP** says that, by working the recipe to attempt to enumerate the truths $\neg P(0)$, $\neg P(1)$, $\ldots$, mathematicians will eventually discover $P(m)$ for some m. That is *one* way in which the "possibility" mentioned by Dummett may get manifested.

Note. Read as a strict justification for **MP**, the argument just given is not wholly circular. It may well reduce the truth of **MP** for decidable $P(n)$ to that of **MP** for primitive recursive $P(n)$.

Lawless Sequences and Topological Models

Even in a purely intuitionistic metatheory, one can readily see that **HPL** is sound with respect to reinterpretations of the connectives and quantifiers over topological spaces τ. The route to these reinterpretations is familiar to anyone who knows how to read the connectives of ordinary logic as correlative operations over Boolean algebras. In this section, think of τ as the usual topology on Cantor Space 2^ω, as generated by the clopen basis of "finite sequence sets":

$$\{\{\alpha : \alpha \in s\} : s \text{ is a finite binary sequence}\}.$$

Let D be any inhabited domain. As is standard, one stipulates that there is a map v such that, for each predicate R of arity n and each n-length sequence d of members of D, $v(R(d)) \in \tau$. (This formulation is to cover the case $n = 0$.) Intuitionistically or conventionally, v can be extended uniquely to accept as inputs all the sentences in the language, which is assumed to be pure. The language may even be expanded with (names for) elements of D. 'I' below stands for the interior operator on τ and '$\Rightarrow$' stands for the operation which, for any sets A and B, outputs $A \Rightarrow B = \{x : \text{ if } x \in A \text{ then } x \in B\}$. The conditions for one such schema of extension are as follows.

1. $v(\bot) = \emptyset$
2. $v(\phi \wedge \psi) = v(\phi) \cap v(\psi)$
3. $v(\phi \vee \psi) = v(\phi) \cup v(\psi)$
4. $v(\phi \rightarrow \psi) = I[v(\phi) \Rightarrow v(\psi)]$

5. $v(\exists x\phi) = \bigcup_{d\in D} v(\phi(d))$
6. $v(\forall x\phi) = I[\bigcap_{d\in D} v(\phi(d))]$

Then ϕ is said to be *valid in* τ (or $\models_\tau \phi$) whenever $\gamma \in v(\phi)$ holds for all binary sequences γ.

No special assumptions are required for an intuitionistic proof of a soundness theorem:

Theorem. *If* $\vdash \phi$ *then* $\models_\tau \phi$*.*

In various cases, this theorem opens a direct route to intuitionistically correct independence results. For example, consider $p \vee \neg p$ and continue to think of τ as the standard topology on Cantor space. Set

$$v(p) = \{\alpha : \exists n.\alpha(n) = 0\}.$$

Then $v(\neg p) = \emptyset$ and we see that

$$v(p \vee \neg p) \neq 2^\omega.$$

It follows that

$$\not\models p \vee \neg p.$$

With the aid of conventional metatheory, Tarski [1938] showed that this propositional counterexample procedure is general: intuitionistic propositional logic is complete with respect to the seemingly nonstandard notion of validity $\models_\tau$, even if τ is restricted to Cantor space. Dana Scott later saw – and recorded in Scott [1957] – how to enhance the constructivity of Tarski's result. What remains unclear is the relation between completeness for topological models and the general completeness questions Gödel posed originally. One wants to know how, if at all, information about interpretations over topological spaces can yield constructive information about ordinary relational structures conceived intuitionistically.

In Kreisel [1958], the required connection was forged by means of the concept of lawless sequence. The behavior of a sentence ϕ over a sequence α lawless with respect to τ turns out to be equivalent to the behavior of ϕ over an ordinary structure $\Im^\alpha$. As motivation for this connection, take another look at the clauses for topological interpretation v just displayed. Please note that, except for $\rightarrow$ and $\forall$, the clauses defining $v(\phi)$ are precisely those one would write for defining an ordinary truth assignment. Apart from these two signs, the function v commutes with the logical operations, viewed as set functions. The noncomplying clauses, those for $\rightarrow$ and $\forall$, are noncomplying only in the intervention of the interior operator I. If there were legitimate means, therefore, to suppress the effect of I in these two cases, topological models would be nothing more than set-theoretic reformulations of ordinary structures. Kreisel realized, in effect, that one can legitimately suppress the effect of taking interiors if, in thinking of the underlying Cantor space, focus is restricted from arbitrary binary sequences to certain special sequences α. These α's would have to be such that, for any subset S of 2^ω which might appear in the crucial clauses defining topological models,

$$\alpha \in S \Leftrightarrow \alpha \in I(S).$$

In short, such α are never "pinned down" by being forced to run through the boundary of a set S. Put in the present heuristic terms, what Kreisel saw was this: for these conditions to obtain, it suffices to pick out those binary sequences which are *lawless* in 2^ω and to permit ourselves the luxury of specifying structures with parameters ranging over lawless sequences.

For the present, assume that the class of lawless sequences in Cantor space is just the collection of those sequences obeying the principles of *density* and *open data*. If α is a sequence in Cantor space, then $\alpha \mid n$ is its initial segment of length n. s ranges over finite binary sequences. 'α' and 'β' are reserved for use as variables over lawless sequences.

First, density tells us that there are enough lawless sequences to go around in 2^ω, i.e.,

$$\forall s \exists \alpha . \alpha \in s.$$

Second, open data spells out the *lawlessness* in lawless sequences. Here, $\phi(\alpha)$ is to be a predicate of lawless sequences in which α is the only free lawless parameter. Then open data is

$$\forall \alpha [\phi(\alpha) \rightarrow \exists s \forall \beta \in s \phi(\beta)].$$

(The need for a restriction on predicates $\phi(\alpha)$ will be clear to anyone who attempts to apply open data to such predicates as $\alpha = \beta$.) Prosaically put, open data prohibits one from implicating a lawless α in the truth of a "nonlawless" claim $\phi(\alpha)$ about it without also implicating, by that very act, the truth of ϕ for a whole neighborhood s of lawless sequences $\beta \in 2^\omega$ surrounding α.

An alternative way to retell the story of open data is by speaking of kinds of interiors. Let S be any set of binary sequences specifiable without recourse to lawless parameters. Say that α is in the *lawless interior* of S whenever α is contained in a basic neighborhood all of whose **lawless** members belong to S as well. Writing 'LI' for the operation of taking lawless interiors with respect to τ, one can express open data as follows.

$$\forall\alpha[\alpha \in S \to \alpha \in LI(S)].$$

If one can show that – as far as lawless sequences and sets in the image of v are concerned – lawless interiors coincide with interiors, then every topological model will give rise to an array of ordinary structures indexed by $\alpha \in v(\phi)$. That they so coincide is provable from the *Fan Theorem* for lawless sequences:

$$\forall\alpha\exists n.\psi(\alpha \mid n) \to \exists m\forall\alpha\exists n \leq m.\psi(\alpha \mid n).$$

Here, ψ is a predicate of finite binary sequences. The idea behind the Fan Theorem is that, if every lawless path α through 2^ω strikes the barrier set up by ψ at some stage of α's expansion, then there will be a single stage by which time any α will have struck ψ. The Fan Theorem is, then, an intuitionistic "positivization" of König's Lemma – but restricted here to lawless sequences.

Lemma. *For open sets S specifiable without lawless parameters,*

$$\forall\beta \in \alpha \mid n.\beta \in S \text{ only if } \alpha \mid n \subseteq S.$$

Proof. A straightforward application of the Fan Theorem and the density condition. □

Now the fundamental result:

Theorem. *For any lawless α, the assignment $\alpha \in v(\phi)$ to sentences ϕ is a truth assignment.*

Proof. Remembering that open data allows the move, for lawless α, from α's membership in a set to α's membership in its lawless interior, one sees that it is enough to prove two claims: first, that

$$\alpha \in LI[v(\phi) \Rightarrow v(\psi)]$$

only if

$$\alpha \in I[v(\phi) \Rightarrow v(\psi)]$$

and, second, that

$$\alpha \in LI[v(\forall x\phi)]$$

only if

$$\alpha \in I[v(\forall x\phi)].$$

Simple arguments show that both these follow from the lemma. □

Fix a lawless sequence α and consider the structure $\Im^\alpha$ determined by the truth assignment $\alpha \in v(\phi)$. So, for sentences ϕ, let

Definition. $\Im^\alpha \models \phi$ *if and only if* $\alpha \in v(\phi)$.

The theorem assures us that this definition is coherent. The following results are more or less immediate.

Corollary 1 $\models_\tau \phi$ *if and only if, for all lawless* α, $\Im^\alpha \models \phi$.

Corollary 2 *Strong completeness for topological interpretations over* τ *implies strong completeness in the usual sense.*

Corollary 3 *Weak completeness for topological interpretations over* τ *implies weak completeness in the usual sense.*

Proof. Proof of first corollary requires the last lemma. The other corollaries are simple consequences of the first. □

Lawless sequences and informal rigour

The expression 'lawless sequence' is itself a misnomer, a *nom de mathematique*. Unless the collection of "lawless" sequences satisfies some thoroughly restrictive laws, such as open data, it would be useless. But this is merely a terminological quibble. What is worth investigating is the status of lawlessness within informal rigour. Analyses of lawlessness are proffered, e.g., Troelstra and van Dalen [1988, p. 833], as paradigms of what Kreisel (is thought to have) meant by 'informal rigour.' This is especially heroic, given that lawless sequences are dark horse candidates for such a paradigmatic role.

On the extant analyses, there were and are no intuitive concepts on which one could lay noetic hands, concepts subsisting **prior** to the construction of any formal theory of lawless sequences and to which that theory is to be responsible in a noncircular fashion. These would be concepts sufficient to justify density and open data – without begging the question – but not exhausted in that justification. Consider the sample analysis which Troelstra and van Dalen recommend for the preformal predecessor of the axiomatic concept of lawless sequence: the image of successive throws of a die. Their contention is that, from the informal concept of successive throws of a binary die, one can extract (in the sense of 'informally demonstrate') the principles of density and open data. From the outset, certain modifications in the primal concept are required. Some are inconsequential; others are not. First, one must suppose that the die is binary, specifically, that the die faces are not adorned with numerals for one through six but numerals for either one or two. Second, the binary die must be idealized: its properties must be impossibly refined relative to those of a terrestrial die. It must, for example, be perfectly fair and not liable, even after millions of throws, to favor one possible outcome over another. Also, it must be constructed of materials which resist wear, even unto the end of time. But, the more the conceptual tailoring, the more hollow rings any insistence that the notion of 'physically unattainable binary die' remains intuitive - and remains capable of justifying open data without circularity.

Worse, there is nothing in the "throw of the die" model – so far – corresponding to the density restriction. There is nothing that will guarantee *a priori* that any finite binary series chosen will be one which is realized in a series of throws of the special, idealized die. The longer the finite series, the less likely it is that such guarantee can be issued. The supporter of the "die throws" analysis cannot respond that it is not actual throws of the die to be considered but merely possible throws and, if the latter is the target of analysis, it is plain that any finite series of 0's and 1's represents a possible array of throws. But, once one allows *all* possible series of throws, the concept of lawless sequence gets left behind. It would not, it seems, be out of the question for an infinite sequence of possible outcomes of a fair die to be coextensive with a recursive sequence and, hence, demonstrably nonlawless.

At a price, one can always modify the original idea again. Perhaps density can be guaranteed by thinking of an infinite collection of perfect dice, each of which is "loaded" for a period of time. Let $lh(s)$ be the length of finite binary sequence s. Imagine that, for each s, there is a "perfectly loaded" die D_s such that, for the first $lh(s)$ throws, D_s is weighted to mimic the components of s in its outcomes and that, after this time, the weighting is removed and D_s acts as a perfect randomizer. In this way, given initial segments are determined during the loaded sequences of throws. But this extensive modification only hands us over a concept at odds with the principle of open data. Any two sequences of throws which appear in the same neighborhood of Cantor space may well – and probably were – determined by perfectly loaded dice D_s and D_t such that s and t have different lengths. This is a violation of open data.

One might persist in this vein, slapping ever more restrictions onto the original intuitive notion so that, eventually, some ultimate modification comes to respect both density and open density. Yet, the more you travel this way, the farther behind you leave any recognizable intuitive concept which might arguably serve as pre-formal backing to the notion of lawless sequence. So thoroughly modified a concept of throws of a die is so abstracted from our everyday experience that the deduction of density and open data from it could hardly be less than circular. For the only real knowledge of such abstracted notions would, most likely, have been provided by a grasp of and conviction in density and open data in the first place, perhaps aided and abetted by the mathematical laws of probability. This route does not seem to bring forward an intuitive conception so much as a concoction, made up specially to conform to a preëxistent ken of the axioms for lawlessness.

More success may be found in turning things around and attempting to extract a quasi-intuitive conception from the formal results. Kripke, in applying what people later called 'Kripke models' Kripke [1965], gave a metaphysical gloss on his formal ideas which enforced, across the philosophical community at least, constraints first set down by Tarski's thought on topological models. Instead of speaking of nodes in a tree or points of an abstract partial order, Kripke referred to elements of his models as "states of information," as stages within potential courses of inquiry. Hence, he thought of (some of) the branches running through the model structures as following epistemically possible courses of mathematical investigation. He wrote:

> Now, in general, in a model structure (**G**, **K**, R), we interpret **G** as the present "evidential situation." If **H** is any situation, we say **H** R **H'** if, as far as we know, at the time **H**, we may later get enough information to advance to **H'**. (Kripke [1965, p. 99])

Following Kreisel, Kripke also explained Kripke [1965, pp. 100–101] how countermodels for intuitionistic formulae constructed on his partially-ordered structures might be construed as assertions concerning lawless sequences. Each Kripke model is, after all, equivalent to some topological model built with 'truth-values' drawn from τ. One might be better advised, then, to reverse the "from the past into the future" order of justification suggested by informal rigour and allow lawless sequences to be given their best and most intuitive understanding in terms of the metamathematics, in this case, in terms of mathematically and epistemically possible avenues of investigation.

Concluding Comments

It seems difficult, in light of the sort of research into intuitionistic completeness inaugurated in 1957 and 1958, to credit an idea underlying informal rigor: that it is by staring into the conceptual past that we find license for the foundational work of today and the future. According to a contrary view, we must often peer in the opposite direction, for it is the future outcome of the constraints (among which are "intuitive concepts" which follow upon formal work) which we set down today that will prove or disprove their mathematical value. On that view, Kreisel's insistence on "standard interpretations" of intuitionistic validity Kreisel [1958a] notwithstanding, it would be most accurate to say that no such interpretation of intuitionistic validity existed prior to his – and others' – efforts to constrain it. Second, what is arguably the central result: that completeness – even for pure intuitionistic predicate logic – implies Markov's Principle, operates as a constraint on research into completeness. In the absence of a convincing argument to the conclusion that **MP** is intuitionistically incorrect, one is left – as far as mathematics is concerned – with little more than a *determination* that **MP** be counted as incorrect and, hence, that full completeness cannot be proved. And this determination will prove itself – or be overthrown – in times to come. Third, the position is similar with regard to lawlessness: there seems to have been no intuitive concept sufficient to underwrite the axioms behind Kreisel's proof that topological completeness implies completeness. That proof itself, however, may point a future way to attractively "informal" images of lawless sequence.

Lastly, Gödel's seemingly supernatural ability to predict in his thesis the fate of research into intuitionistic completeness is no surprize, no miracle, if, as Wittgenstein once suggested, grammatical forms of mathematical "predictions" are taken not as indicative, future tense, but as imperatives or optatives. For you often exercise control over the contraints which you yourself lay down for the future. With sufficient power, you might even shape them by shaping their interpretations, just as you may shape the way your own commands are carried out.

Acknowledgement. The author wishes to thank Piergiorgio Odifreddi for his helpful comments and suggestions.

Bibliography

Beeson, M.

[1985] *Foundations of Constructive Mathematics*, Springer-Verlag, New York, 1985, xxiii+466 pp.

Brouwer, L.E.J.

[1981] *Brouwer's Cambridge Lectures on Intuitionism*, D. van Dalen (ed.), Cambridge University Press, Cambridge, UK, 1981, xii+109 pp.

Büchi, J.R.

[1961] Turing machines and the Entscheidungsproblem, *Mathematische Annalen*, 148 (1962) 201–213.

van Dalen, D.

[1973] Lectures on intuitionism, in *Cambridge Summer School in Mathematical Logic*, A. Mathias and H. Rogers (eds.), Springer-Verlag, New York, 1973, pp. 1–94.

Dummett, M.

[1977] *Elements of intuitionism*, Oxford Logic Guides, Clarendon Press, Oxford, 1977, xii+467 pp.

Friedman, H.

[1977] The intuitionistic completeness of intuitionistic logic under Tarskian semantics, *Xeroxed typescript*, The State University of New York at Buffalo, March 1977, 10pp.

[1978] Classically and intuitionistically provably recursive functions, in *Higher Set Theory*, G. Müller and D.S. Scott (eds.), Springer-Verlag, New York, pp. 21–27.

Gödel, K.

[1929] On the completeness of the calculus of logic, in *Collected Works. Volume I. Publications 1929–1936*, S. Feferman et al. (eds.), Oxford University Press, New York, 1986, pp. 61–101.

[1930] On the completeness of the calculus of logic, *Collected Works. Volume I. Publications 1929–1936*, S. Feferman et al. (eds.), Oxford University Press, New York, 1986, p. 125.

[1933] On intuitionistic arithmetic and number theory, *Collected Works. Volume I. Publications 1929–1936*, S. Feferman et al. (eds.), Oxford University Press, New York, 1986, pp. 286–295.

Heyting, A.

[1930a] Die formalen Regeln der intuitionistischen Logik, *Sitzung. Preuss. Akad. Wissensch., physik. math. Kl.*, 1930, pp. 42–56.

[1930b] Die formalen Regeln der intuitionistischen Mathematik II, *ibidem*, pp. 57–71.

[1930c] Die formalen Regeln der intuitionistischen Mathematik III, *ibidem*, 1930, pp. 158–169.

Kolmogorov, A.N.

[1926] On the principle of the excluded middle, in *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*, J. van Heijenoort (ed.), Harvard University Press, Cambridge, MS., 1977, pp. 414–437.

Kreisel, G.

[1958a] Elementary completeness properties of intuitionistic logic with a note on negations of prenex formulae, *The Journal of Symbolic Logic*, 23 (1958) 317–330.

[1958b] A remark on free choice sequences and the topological completeness proofs, *The Journal of Symbolic Logic*, 23 (1958) 369–388.

[1958c] The nonderivability of $\neg(x)A(x) \to (\exists x)\neg A(x)$, A primitive recursive, in intuitionistic formal systems (abstract), *The Journal of Symbolic Logic*, 23 (1958) 456–457.

[1962] On weak completeness of intuitionistic predicate logic, *The Journal of Symbolic Logic*, 27 (1962) 139–158.

[1964] Hilbert's Programme, in *Philosophy of Mathematics. Selected Readings*, P. Benacerraf and H. Putnam (eds.), Prentice-Hall, Englewood Cliffs, NJ., 1964, pp. 157–183.

[1965] Mathematical logic, in *Lectures in Modern Mathematics III*, T.L. Saaty (ed.), Wiley and Sons, New York, 1965, pp. 95–195.

[1967] Informal rigor and completeness proofs, in *Problems in the Philosophy of Mathematics*, I. Lakatos (ed.), North-Holland, Amsterdam, 1967, pp. 138–171.

[1970] Church's Thesis: a kind of reducibility axiom for constructive mathematics, in *Intuitionism and Proof Theory*, A. Kino, J. Myhill and R.E. Vesley (eds.), North-Holland, Amsterdam, 1970, pp. 121-150.

Kripke, S.A.

[1965] Semantical analysis of intuitionistic logic, *Formal Systems and Recursive Functions*, J. Crossley and M. Dummett (eds.), North-Holland, Amsterdam, 1965, pp. 92–130.

Leivant, D.

[1976] Failure of completeness properties of intuitionistic predicate logic for constructive models, *Annales Scientifiques de l'Université de Clermont. Series Mathematiques*, 13 (1976) 93–107.

[1985] Intuitionistic formal systems, *Harvey Friedman's Research on the Foundations of Mathematics*, L. Harrington et al. (eds.), Elsevier Science Publishers, New York, 1985, pp. 231–255.

M<u>c</u>Carty, D.C.

[1986] Realizability and recursive set theory, *Annals of Pure and Applied Logic*, 32 (1986) 153–183.

[1988a] Constructive validity is nonarithmetic, *The Journal of Symbolic Logic*, 53 (1988) 1036–1041.

[1988b] Markov's Principle, isols and Dedekind-finite sets, *The Journal of Symbolic Logic*, 53 (1988) 1042–1069.

[1991] Incompleteness in intuitionistic metamathematics, *The Notre Dame Journal of Formal Logic*, 32 (1991) 323–358.

Scott, D.S.

[1957] Completeness proofs for the intuitionistic sentential calculus, *Summaries of Talks presented at the Summer Institute of Mathematical Logic in 1957 at Cornell*, mimeographed, 1957, pp. 231–241.

Skolem, T.

[1922] Some remarks on axiomatized set theory, in *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*, J. van Heijenoort (ed.), Harvard University Press, Cambridge, MS, 1977, pp. 290–301.

Smorynski, C.

[1973] Applications of Kripke models, in *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis*, A.S. Troelstra (ed.), Spring-Verlag, New York, 1973, pp. 324–391.

[1977] The incompleteness theorems, in *Handbook of Mathematical Logic*, J. Barwise (ed.), North-Holland, New York, 1977, pp. 821–865.

de Swart, H.

[1976] Another intuitionistic completeness proof, *The Journal of Symbolic Logic*, 41 (1976) 644–662.

Tarski, A.

[1938] Der Aussagenkalkül und die Topologie, *Fundamenta mathematicae*, 31 (1938) 103–134.

Tennenbaum, S.

[1959] Non-archimedean models of arithmetic, *Notices of the American Mathematical Society*, 6 (1959) 270.

Troelstra, A.S.

[1975] Completeness and validity for intuitionistic predicate logic, *Report 76-05. Department of Mathematics. University of Amsterdam*, February 1976, 28 pp.

Troelstra, A.S., and D. van Dalen

[1988] *Constructivism in Mathematics. An Introduction*, vol. II, North-Holland, New York, xvii + 345–879 + LII pp.

Veldman, W.

[1976] An intuitionistic completeness theorem for intuitionistic predicate logic, *The Journal of Symbolic Logic*, 41 (1976) 159–166.

Density and Choice for Total Continuous Functionals

by Helmut Schwichtenberg

In his seminal paper *Interpretation of analysis by means of constructive functionals of finite types*, in the volume *Constructivity in Mathematics* edited by A. Heyting in 1959, Georg Kreisel has studied partial and total continuous higher order functionals. The paper states two important theorems on these notions:

- the density theorem, which says that any finite functional can be extended to a total one, and
- the choice principle for total continuous functionals, which says that whenever for any total x there is a total y such that Rxy, where the relation R is given by a total Boolean-valued functional, then this dependency can be realized by a total functional f (which can be found uniformly in R).

Both of these theorems also have effective versions, in a natural sense of the word.

Kreisel's paper and the one by Kleene ([1959]) on *Countable functionals* in the same volume, which independently introduced equivalent concepts, have given rise to a lot of activity aimed at understanding and developing the mathematical notions involved. First of all there is the unpublished work of Tait in the so-called *Stanford Report*, which is based directly on Kreisel's work. Another early contribution of essential impact was Platek's thesis [1966], who had started his work with Kreisel and later also worked with Dana Scott. Particularly Scott, in a number of attempts ([1970], [1982]), contributed a lot to the subject. His work influenced Yuri Ershov to come up with an elaborated theory, with a somewhat topological flavour ([1974a], [1975], [1977]). Many other people have also made essential contributions, among them Feferman, Gandy, Hyland, Normann and Berger, mostly again in a rather general context. In particular Normann's work [1993] can be viewed as establishing the density theorem in a more general setting.

My aim in this paper is to give complete proofs of the density theorem and the choice principle for total continuous functionals in the natural and concrete context of the partial continuous functionals (Ershov [1977]), essentially by specializing more general treatments in the literature. The proofs obtained are relatively short and hopefully perspicious, and may contribute to redirect attention to the fundamental questions Kreisel originally was interested in.

Obviously this work owes much to other sources. In particular I have made use of work by Scott [1982] (whose notion of an information system is taken as a basis to introduce domains), Roscoe [1987], Larsen and Winskel [1984], and Berger [1990].

The paper is organized as follows. Section 1 treats information systems, and in Section 2 it is shown that the partial orders defined by them are exactly the (Scott) domains with countable basis. Section 3 gives a characterization of the continuous functions between domains, in terms of approximable mappings. In Section 4 cartesian products and function spaces of domains and information systems are introduced. In Section 5 the partial and total continuous functionals are defined. Section 6 finally contains the proofs of the two theorems above; it will be clear that the same proofs also yield effective versions of these theorems.

1 Information Systems

The basic idea of information systems is to provide an axiomatic setting for to describe approximations of abstract objects (like functions or functionals) by concrete, finite ones. We do not attempt to analyse the notion of 'concreteness' or finiteness here, but rather take an arbitrary countable set A of 'data objects' or 'tokens' as a basic notion to be explained axiomatically. In order to use such data objects to build approximations of abstract objects, we certainly need a notion of 'consistency', which determines when the elements of a finite set of data objects are consistent with each other. At our present level of generality there clearly is no way to define this notion concretely, hence it again must be described axiomatically. Finally we need an 'entailment relation' between consistent sets X of data objects and single data objects a, which expresses the fact that the information contained in X supersedes the information in a. The following axioms on these notions have been given by Scott [1982].

Definition 1.1 *An information system is defined to be a structure* $(A, \mathrm{Con}, \vdash)$ *where A is a countable set (the tokens),* Con *is a nonempty set of finite subsets of A (the consistent sets), and* $\vdash$ *is a subset of* $\mathrm{Con} \times A$ *(the entailment relation) which satisfy:*

(i) $X \subseteq Y \in \mathrm{Con}$ *implies* $X \in \mathrm{Con}$;

(ii) $a \in A$ *implies* $\{a\} \in \mathrm{Con}$;

(iii) $X \vdash a$ *implies* $X \cup \{a\} \in \mathrm{Con}$;

(iv) $X \in \mathrm{Con}$ *and* $a \in X$ *implies* $X \vdash a$;

(v) $(X, Y \in \mathrm{Con}$ *and* $(\forall b \in Y) X \vdash b$ *and* $Y \vdash c)$ *implies* $X \vdash c$.

Any countable set A can be turned into an information system $\mathbf{A}$ by letting the set of data objects be A, $\mathrm{Con} = \{\emptyset\} \cup \{\{a\} : a \in A\}$ and $X \vdash a \iff a \in X$.

Definition 1.2 *The elements or ideals of an information system*

$$\mathbf{A} = (A, \mathrm{Con}, \vdash)$$

are defined to be those subsets z of A which satisfy:

(i) $X \subseteq^{\mathrm{fin}} z$ *implies* $X \in \mathrm{Con}$ *(z is consistent);*

(ii) $X \subseteq^{\mathrm{fin}} z$ *and* $X \vdash a$ *implies* $a \in z$ *(z is deductively closed).*

The set of all elements of $\mathbf{A}$ is written $|\mathbf{A}|$. It clearly follows that the intersection of any number of ideals is an ideal again.

Lemma 1.3 *Suppose* $\mathbf{A} = (A, \mathrm{Con}, \vdash)$ *is an information system. Then*

(i) $\emptyset \in \mathrm{Con}$;

(ii) $(X \in \mathrm{Con}$ *and* $Y \subseteq X$ *and* $Y \vdash a)$ *implies* $X \vdash a$.

Proof. (i) follows because Con is by assumption nonempty, and axiom (i).

(ii) holds by axiom (v) since, by axiom (iv), $X \vdash b$ for each $b \in Y$. □

Suppose $\mathbf{A} = (A, \mathrm{Con}, \vdash)$ is an information system and $B \subseteq A$. We define $\overline{B}$, the *deductive closure* of B, by

$$\overline{B} = \{a \in A : X \vdash a \text{ for some } X \subseteq^{\text{fin}} B\}.$$

$\overline{B}$ is the set of tokens deducible from B. Note that by part (ii) of the previous Lemma, if $X \in \mathrm{Con}$, then $\overline{X} = \{a \in A : X \vdash a\}$.

Lemma 1.4 *Suppose* $\mathbf{A} = (A, \mathrm{Con}, \vdash)$ *is an information system,* $X \in \mathrm{Con}$ *and* Y *a finite subset of* A*. Then*

(i) if $X \vdash a$ *for every* $a \in Y$*, then* $Y \in \mathrm{Con}$*;*

(ii) $\overline{X} \in |\mathbf{A}|$*;*

(iii) $\overline{\emptyset} \subseteq z$ *for every* $z \in |\mathbf{A}|$*.*

Proof. (i) We prove $X \cup Y \in \mathrm{Con}$ by induction on the size of Y; the result then follows by axiom (i). In case $Y = \emptyset$ there is nothing to prove, so suppose $Y = Y' \cup \{a\}$ and $X \cup Y' \in \mathrm{Con}$. Since $X \vdash a$ by assumption, we have $X \cup Y' \vdash a$ by Lemma 1.3(ii) and hence $X \cup Y \in \mathrm{Con}$ by axiom (iii).

(ii) If Y is a finite subset of $\overline{X}$, then $Y \in \mathrm{Con}$ by (i); hence $\overline{X}$ is consistent. If Y is a finite subset of $\overline{X}$ and $Y \vdash a$, then $X \vdash a$ by the definition of $\overline{X}$ and axiom (v); hence $\overline{X}$ is deductively closed.

(iii) Let $z \in |\mathbf{A}|$ and $a \in \overline{\emptyset}$, i.e. $\emptyset \vdash a$. Since z is deductively closed, it follows that $a \in z$. □

In the light of (i) it is convenient informally to allow '$\vdash$' to be used as a relation between Con and Con. So from now on '$X \vdash Y$' will mean '$X \vdash a$ for each $a \in Y$'.

Lemma 1.5 *If* $\mathbf{A} = (A, \mathrm{Con}, \vdash)$ *is an information system and* $U \subseteq A$ *is consistent, then* $\overline{U} \in |\mathbf{A}|$*.*

Proof. We first show that $\overline{U}$ is consistent. If $Y = \{a_1, a_2, \dots, a_n\}$ is any finite subset of $\overline{U}$, then for each $i \in \{1, 2, \dots, n\}$ there exists a finite subset Z_i of U such that $Z_i \vdash a$. But $Z = Z_1 \cup \dots \cup Z_n$ is a finite subset of U, and so is in Con. Now $Z \vdash a_i$ for each i by Lemma 1.3(ii). Lemma 1.4(i) then tells us that $Y \in \mathrm{Con}$.

To show deductive closure, suppose that Y is a finite subset of $\overline{U}$ and that $Y \vdash b$. Exactly as above we can find a finite subset Z of U such that $Z \vdash Y$. But then $Z \vdash b$ by axiom (v), so $b \in \overline{U}$ as required. $\square$

2 Complete Partial Orders and Domains

Let $(I, \leq)$ be a partial order, i.e. $\leq$ is a reflexive and transitive subset of $I \times I$ satisfying $i \leq j \leq i \implies i = j$ (antisymmetry). We say that I is *directed* if it is nonempty and for any $i, j \in I$ there is a $k \in I$ such that $i \leq k$ and $j \leq k$.

A partial order $(D, \leq)$ having a least element $\perp$ is said to be *complete* (and we say that D is a *complete partial order*, abbreviated to *cpo*) if every directed subset $M \subseteq D$ has a least upper bound $\bigsqcup M$.

A point x of a cpo D is said to be *compact* or *finite* if, for every directed collection $M \subseteq D$ such that $x \leq \bigsqcup M$, there is a $y \in M$ such that $x \leq y$. Let $\mathbf{B}_D$ denote the collection of compact elements of D; $\mathbf{B}_D$ is called the *basis* of D.

The cpo D is *algebraic* if, for every $x \in D$, the set $M = \{x_0 \in \mathbf{B}_D : x_0 \leq x\}$ is directed and $x = \bigsqcup M$. A cpo D is *bounded complete* or *consistently complete* if every bounded subset of D (or equivalently every bounded finite subset of D) has a least upper bound.

We call bounded complete algebraic cpo's *Scott domains* or just *domains*.

With any information system $\mathbf{A} = (A, \mathrm{Con}, \vdash)$ we can associate the partial order $(|\mathbf{A}|, \subseteq)$. Our aim in this section is to show that the partial orders obtained in this way are exactly the Scott domains with countable basis.

Theorem 2.1 *Suppose* $\mathbf{A} = (A, \mathrm{Con}, \vdash)$ *is an information system. Then* $(|\mathbf{A}|, \subseteq)$ *is a domain with the countable basis* $\{\overline{X} : X \in \mathrm{Con}\}$.

Proof. $(|\mathbf{A}|, \subseteq)$ clearly is a partial order. It has $\overline{\emptyset}$ as its least element by Lemma 1.4(iii). To prove completeness, let $M \subseteq |\mathbf{A}|$ be directed. Then $\bigcup M$ is consistent and deductively closed and hence is the least upper bound of M.

We now show that, for any $z \in |\mathbf{A}|$, that z is compact if and only if $z = \overline{X}$ for some $X \in \mathrm{Con}$. First assume that z is compact. Clearly $M := \{\overline{X} : X \subseteq^{\mathrm{fin}} z\}$ is directed and $z = \bigcup M$. Since by assumption z is compact, there is a finite subset X of z such that $z \subseteq \overline{X}$, hence $z = \overline{X}$. Conversely, let $X = \{a_1, a_2, \ldots, a_n\} \in \mathrm{Con}$ and $M \subseteq |\mathbf{A}|$ directed such that $\overline{X} \subseteq \bigcup M$. Then $a_i \in z_i$ for some $z_i \in M$. Since

M is directed, there is $z^* \in M$ such that $z_i \subseteq z^*$, hence $X \subseteq z^*$ and therefore $\overline{X} \subseteq z^*$. This means that $\overline{X}$ is compact.

The algebraicity of $(|\mathbf{A}|, \subseteq)$ now is obvious since, for every $z \in |\mathbf{A}|$, $\{\overline{X} : X \in \mathrm{Con} \text{ and } \overline{X} \subseteq z\}$ clearly is directed and has z as its least upper bound.

Finally every bounded subset $M \subseteq |\mathbf{A}|$ has $\overline{\bigcup M}$ as its least upper bound, and hence $(|\mathbf{A}|, \subseteq)$ is bounded complete. $\Box$

We now prove that any domain with countable basis can in fact be represented in this way, as the domain of an appropriate information system associated with it. To obtain this result we need a Lemma saying that in a cpo least upper bounds of finite sets of compact elements are compact, when they exist.

Lemma 2.2 *Suppose* $(D, \leq)$ *is a cpo,* $X \subseteq^{\mathrm{fin}} \mathbf{B}_D$ *and* $\bigsqcup X$ *exists. Then* $\bigsqcup X \in \mathbf{B}_D$*.*

Proof. Suppose $M \subseteq D$ is directed and $\bigsqcup X \leq \bigsqcup M$. Let $X = \{e_1, e_2, \dots, e_n\}$. Since $e_i \leq \bigsqcup M$ and $e_i \in \mathbf{B}_D$ there is a $x_i \in M$ with $e_i \leq x_i$. Since M is directed, there is a $x^* \in M$ such that $x_i \leq x^*$ for every $i \in \{1, 2, \dots, n\}$. Hence $\bigsqcup X \leq x^*$. $\Box$

Definition 2.3 *Let D be a domain with countable basis $\mathbf{B}_D$. Define* $ID = (\mathbf{B}_D, \mathrm{Con}, \vdash)$*, where* Con *and* $\vdash$ *are defined by*

$$\begin{aligned} X \in \mathrm{Con} &\iff X \subseteq^{\mathrm{fin}} \mathbf{B}_D \text{ and } X \text{ is bounded in } D \\ X \vdash e &\iff X \in \mathrm{Con} \text{ and } e \leq \bigsqcup X. \end{aligned}$$

Let D and E be partial orders. A function $f\colon D \to E$ is *monotone* if

$$d \leq d' \implies f(d) \leq f(d'),$$

and *orderpreserving* if

$$d \leq d' \iff f(d) \leq f(d').$$

A bijective orderpreserving function $f\colon D \to E$ is called an *isomorphism*. Now let D and E be cpo's. f is *continuous* if it is monotone and if for any directed set $M \subseteq D$

$$f(\bigsqcup M) = \bigsqcup \{f(d) : d \in M\}.$$

Note that from the monotonicity of f the directedness of

$$\{f(d) : d \in M\}$$

follows, and hence the supremum on the right hand side exists in E. Note also that any surjective orderpreserving function $f: D \to E$ must be injective, hence an isomorphism and therefore also continuous.

It is easy to see that the concatenation of two continuous functions is continuous again.

Lemma 2.4 *Suppose D, E and F are cpo's. Let $f: D \to E$ and $g: E \to F$ be continuous functions. Then $g \circ f: D \to F$ is continuous.*

Proof. Clearly $g \circ f$ is monotone. Furthermore we have that

$$\bigsqcup_{d \in M} g(f(d)) = g(\bigsqcup_{d \in M} f(d)) = g(f(\bigsqcup M)). \quad \square$$

Theorem 2.5 *Let D be a domain with countable basis. Then ID is an information system with domain of elements $|ID|$ isomorphic to D. The isomorphism pair is*

$$\varphi: D \to |ID| \quad \textit{given by} \quad \varphi(d) = \{e \in \mathbf{B}_D : e \leq d\},$$
$$\psi: |ID| \to D \quad \textit{given by} \quad \psi(z) = \bigsqcup z.$$

Proof. It is easy to see that ID is an information system.

φ is well-defined: We have to show that $\varphi(d) \in |ID|$. Let us first prove that $\varphi(d)$ is consistent. So let $e_1, e_2, \ldots, e_n \leq d$. We have to show that $\{e_1, e_2, \ldots, e_n\} \in \mathrm{Con}$. But this is clear, since $\{e_1, e_2, \ldots, e_n\}$ is bounded by d. We now prove that $\varphi(d)$ is deductively closed. Choose again $e_1, e_2, \ldots, e_n \leq d$ and let $e \leq \bigsqcup\{e_1, e_2, \ldots, e_n\}$. We have to show $e \in \varphi(d)$, i.e. $e \leq d$. But this follows from $e_1, e_2, \ldots, e_n \leq d$.

ψ is well-defined: Let $z \in |ID|$. By Theorem 2.1 it suffices to show that z is directed. So let $X \subseteq^{\text{fin}} z$. Then $X \subseteq \mathbf{B}_D$ and $X \in \mathrm{Con}$, hence X is bounded in D, hence $\bigsqcup X$ exists. Now by Lemma 2.2 $\bigsqcup X \in \mathbf{B}_D$. Therefore $\bigsqcup X \in z$ by the deductive closure of z.

$\psi \circ \varphi = \mathrm{id}_D$: Since D is algebraic, we have that $\varphi(d) = \{e \in \mathbf{B}_D : e \leq d\}$ is directed and $d = \bigsqcup \varphi(d)$.

$\varphi \circ \psi = \mathrm{id}_{|ID|}$: We have to show that $\varphi(\bigsqcup z) = z$ for every $z \in |ID|$. To prove $\supseteq$, let $z \subseteq \mathbf{B}_D$ be consistent and deductively closed. Recall that $\varphi(\bigsqcup z) = \{e \in \mathbf{B}_D : e \leq \bigsqcup z\}$. Now let $e \in z$. Then $e \leq \bigsqcup z$, hence $e \in \varphi(\bigsqcup z)$. To prove $\subseteq$, let $z \in |ID|$ and $e \in \varphi(\bigsqcup z)$, i.e. $e \leq \bigsqcup z$. We must show $e \in z$. Now since z is directed, from $e \leq \bigsqcup z$ and $e \in \mathbf{B}_D$ we can conclude $e \leq e'$ for some $e' \in z$. Since z is deductively closed and $e \leq \bigsqcup\{e'\}$, it follows that $e \in z$.

φ, ψ clearly are monotone. Hence φ and ψ are continuous and therefore form an isomorphism pair. $\square$

3 Mappings on Information Systems

Let $\mathbf{A} = (A, \mathrm{Con}_A, \vdash_A)$ and $\mathbf{B} = (B, \mathrm{Con}_B, \vdash_B)$ be information systems. We want to study 'information respecting' mappings from $\mathbf{A}$ into $\mathbf{B}$. Such a mapping is given by a relation r between Con_A and B, where Xrb intuitively means that whenever we are given the information $X \in \mathrm{Con}_A$ on the argument, then we know that at least the data object b belongs to the value.

Definition 3.1 *Let* $\mathbf{A} = (A, \mathrm{Con}_A, \vdash_A)$ *and* $\mathbf{B} = (B, \mathrm{Con}_B, \vdash_B)$ *be information systems. A relation* $r \subseteq \mathrm{Con}_A \times B$ *is an approximable mapping if it satisfies*

(i) Xrb_i *for all* $i \in \{1, 2, \dots, n\}$ *implies* $\{b_1, b_2, \dots, b_n\} \in \mathrm{Con}_B$*;*

(ii) Xrb_i *for all* $i \in \{1, 2, \dots, n\}$ *and* $\{b_1, b_2, \dots, b_n\} \vdash_B b$ *implies* Xrb*;*

(iii) $X \vdash_A X'$ *and* $X'rb$ *implies* Xrb*.*

We write $r\colon \mathbf{A} \to \mathbf{B}$ to mean that r is an approximable mapping from $\mathbf{A}$ to $\mathbf{B}$.

Theorem 3.2 *Let* $\mathbf{A} = (A, \mathrm{Con}_A, \vdash_A)$ *and* $\mathbf{B} = (B, \mathrm{Con}_B, \vdash_B)$ *be information systems. With any approximable mapping* $s\colon \mathbf{A} \to \mathbf{B}$ *we can associate a continuous function* $|s|\colon |\mathbf{A}| \to |\mathbf{B}|$ *by*

$$|s|(z) := \{b \in B : Xsb \textit{ for some } X \subseteq^{\mathrm{fin}} z\}.$$

Conversely, with any continuous function $f\colon |\mathbf{A}| \to |\mathbf{B}|$ *we can associate an approximable mapping* $r_f\colon \mathbf{A} \to \mathbf{B}$ *by*

$$Xr_fb :\iff b \in f(\overline{X}).$$

These assignments are inverse to each other, i.e. $f = |r_f|$ *and* $s = r_{|s|}$*.*

Proof. We first show that $|s|$ is well-defined. So let $z \in |\mathbf{A}|$.

$|s|(z)$ is consistent. Let $b_1, b_2, \dots, b_n \in |s|(z)$. Then there are $X_1, X_2, \dots, X_n \subseteq^{\mathrm{fin}} z$ such that X_isb_i. Hence

$$X := X_1 \cup X_2 \cup \dots \cup X_n \subseteq^{\mathrm{fin}} z$$

and Xsb_i by Definition 3.1(iii). Now from Definition 3.1(i) we can conclude that $\{b_1, b_2, \dots, b_n\} \in \mathrm{Con}_B$.

$|s|(z)$ is deductively closed. Let

$$b_1, b_2, \ldots, b_n \in |s|(z) \qquad \text{and} \qquad \{b_1, b_2, \ldots, b_n\} \vdash_B b.$$

We must show $b \in |s|(z)$. As before we find $X \subseteq^{\text{fin}} z$ such that Xsb_i. Now from Definition 3.1(ii) we can conclude Xsb and hence $b \in |s|(z)$.

The monotonicity of $|s|$ clearly follows from the definition. To see that $|s|: |\mathbf{A}| \to |\mathbf{B}|$ is continuous, let $M \subseteq |\mathbf{A}|$ be directed. Then we have

$$\begin{array}{rcll} |s|(\bigcup M) & = & \{b \in B : (\exists X \subseteq^{\text{fin}} \bigcup M) X s b\} & \\ & = & \{b \in B : (\exists z \in M)(\exists X \subseteq^{\text{fin}} z) X s b\} & (M \text{ is directed}) \\ & = & \bigcup_{z \in M} \{b \in B : (\exists X \subseteq^{\text{fin}} z) X s b\} & \\ & = & \bigcup_{z \in M} |s|(z). & \end{array}$$

Now let $f: |\mathbf{A}| \to |\mathbf{B}|$ be continuous. It is easy to verify that r_f is indeed an approximable mapping. Furthermore

$$\begin{array}{rcll} f(z) & = & f(\bigcup_{X \subseteq^{\text{fin}} z} \overline{X}) & (|\mathbf{A}| \text{ is algebraic}) \\ & = & \bigcup_{X \subseteq^{\text{fin}} z} f(\overline{X}) & (f \text{ is continuous}) \\ & = & \{b \in B : (\exists X \subseteq^{\text{fin}} z) b \in f(\overline{X})\} & \\ & = & \{b \in B : (\exists X \subseteq^{\text{fin}} z) X r_f b\} & \\ & = & |r_f|(z). & \end{array}$$

Finally, for any approximable mapping $s: \mathbf{A} \to \mathbf{B}$ we have

$$\begin{array}{rcll} Xsb & \iff & (\exists Y \subseteq^{\text{fin}} \overline{X}) Y s b & \text{by 3.1(iii)} \\ & \iff & b \in |s|(\overline{X}) & \\ & \iff & X r_{|s|} b. \quad \square & \end{array}$$

4 Products and Function Spaces

We discuss some canonical ways to construct new cpo's from given ones:

- the cartesian product $D \times E$, consisting of all pairs (d, e) with $d \in D$ and $e \in E$ under the component-wise order, and
- the function space $D \to E$, consisting of all continuous functions from D to E under the pointwise order.

We introduce these two constructs, and define operations on information systems corresponding to them.

The *cartesian product* of two cpo's $(D, \leq_D)$ and $(E, \leq_E)$ is defined to be $(D \times E, \leq)$ with the *component-wise order* $\leq$, i.e.

$$(d, e) \leq (d', e') \iff d \leq_D d' \text{ and } e \leq_E e'.$$

We want to show that $(D \times E, \leq)$ is a cpo again. For the proof we need to know that the supremum of a directed subset $M \subseteq D \times E$ operates component-wise. Denoting the left and right projection functions by π_{left} and π_{right} we have

Lemma 4.1 *Suppose D and E are cpo's. Let $M \subseteq D \times E$ be directed. Then $\pi_{\text{left}}(M)$ and $\pi_{\text{right}}(M)$ are directed too, and*

$$\bigsqcup M = (\bigsqcup \pi_{\text{left}}(M), \bigsqcup \pi_{\text{right}}(M)).$$

Proof. We first show that e.g. $\pi_{\text{left}}(M)$ is directed. So let $d, d' \in \pi_{\text{left}}(M)$. Then there are $e, e' \in E$ such that $(d, e), (d', e') \in M$. Since M is directed, we have $(d, e), (d', e') \leq (d'', e'')$ for some $(d'', e'') \in M$. Therefore $d, d' \leq_D d'' \in \pi_{\text{left}}(M)$.

Hence $(\bigsqcup \pi_{\text{left}}(M), \bigsqcup \pi_{\text{right}}(M))$ exists. Clearly this pair is an upper bound of M. It is also a least upper bound. To see this suppose that (d', e') is another upper bound, i.e. $(d, e) \leq (d', e')$ for all $(d, e) \in M$. But then $d \leq_D d'$ for all $d \in \pi_{\text{left}}(M)$, i.e. $\bigsqcup \pi_{\text{left}}(M) \leq_D d'$. $\square$

Corollary 4.2 *Suppose D and E are cpo's. Then the set $D \times E$ equipped with the component-wise order is a cpo again.*

The continuity of π_{left} and π_{right} also follows from Lemma 4.1.

We now define a corresponding operation on information systems, assigning to any $\mathbf{A} = (A, \text{Con}_A, \vdash_A)$ and $\mathbf{B} = (B, \text{Con}_B, \vdash_B)$ an

$$\mathbf{A} \times \mathbf{B} = (C, \text{Con}, \vdash) \text{ such that } |\mathbf{A} \times \mathbf{B}| \cong |\mathbf{A}| \times |\mathbf{B}|.$$

Without loss of generality we may assume that A and B are disjoint. But then any pair (x, y) of elements $x \in |\mathbf{A}|$ and $y \in |\mathbf{B}|$ can be approximated in each component separately. Hence we choose as data objects in $\mathbf{A} \times \mathbf{B}$ simply the union $C := A \cup B$. Consistency and entailment then is inherited in the obvious way from $\mathbf{A}$ and $\mathbf{B}$: $X \in \text{Con} \iff X \cap A \in \text{Con}_A$ and $X \cap B \in \text{Con}_A$, and $X \vdash c \iff (c \in A \implies X \cap A \vdash_A c)$ and $(c \in B \implies X \cap B \vdash_B c)$. The elements of $\mathbf{A} \times \mathbf{B}$ then are exactly the unions of the elements of $\mathbf{A}$ and of $\mathbf{B}$: $|\mathbf{A} \times \mathbf{B}| = \{x \cup y : x \in |\mathbf{A}| \text{ and } y \in |\mathbf{B}|\}$.

Lemma 4.3 *If $\mathbf{A} = (A, \text{Con}_A, \vdash_A)$ and $\mathbf{B} = (B, \text{Con}_B, \vdash_B)$ are information systems with $A \cap B = \emptyset$, then $\mathbf{A} \times \mathbf{B}$ defined as above is an information system, and*

$$|\mathbf{A} \times \mathbf{B}| = \{x \cup y : x \in |\mathbf{A}| \textit{ and } y \in |\mathbf{B}|\} \cong |\mathbf{A}| \times |\mathbf{B}|.$$

Corollary 4.4 *Suppose D and E are domains with countable bases. Then the set $D \times E$ equipped with the component-wise order is again a domain with countable basis.*

Proof. $D \times E \cong |ID| \times |IE| \cong |ID \times IE|$. $\square$

The *function space* of two cpo's $(D, \leq_D)$ and $(E, \leq_E)$ is defined to be $(D \to E, \leq)$, where $D \to E$ is the set of all continuous functions from D to E, and the order $\leq$ is defined pointwise, i.e.

$$f \leq g \iff \forall d\, f(d) \leq g(d).$$

We want to show that $(D \to E, \leq)$ is a cpo again. This will follow from the observation that the supremum of a directed subset of $D \to E$ is formed pointwise. To prove this we will make use of the fact that the order of forming least upper bounds is irrelevant provided all relevant suprema exist. More precisely we have

Lemma 4.5 *Suppose D is a cpo and $d_{ij} \in D$ for all $i \in I$ and $j \in J$. Assume that*

(i) $\bigsqcup_{i \in I} d_{ij}$ *exists for any $j \in J$,*

(ii) $\bigsqcup_{j \in J} d_{ij}$ *exists for any $i \in I$, and*

(iii) $\bigsqcup_{i \in I} \bigsqcup_{j \in J} d_{ij}$ *exists.*

Then also $\bigsqcup_{j \in J} \bigsqcup_{i \in I} d_{ij}$ exists, and we have

$$\bigsqcup_{i \in I} \bigsqcup_{j \in J} d_{ij} = \bigsqcup_{j \in J} \bigsqcup_{i \in I} d_{ij}.$$

Proof. We first prove that $\bigsqcup_{i \in I} \bigsqcup_{j \in J} d_{ij}$ is an upper bound of all $\bigsqcup_{i \in I} d_{ij}$. But this is clear, since $d_{ij} \leq \bigsqcup_{j \in J} d_{ij}$ for all $i \in I$ and hence $\bigsqcup_{i \in I} d_{ij} \leq \bigsqcup_{i \in I} \bigsqcup_{j \in J} d_{ij}$.

It remains to prove that $\bigsqcup_{i \in I} \bigsqcup_{j \in J} d_{ij}$ is the least upper bound of all $\bigsqcup_{i \in I} d_{ij}$. So assume that d is some upper bound. We must show $\bigsqcup_{i \in I} \bigsqcup_{j \in J} d_{ij} \leq d$. But this is clear, since $d_{ij} \leq \bigsqcup_{i \in I} d_{ij} \leq d$ for any $i \in I$ and $j \in J$ by assumption. $\square$

Lemma 4.6 *Suppose D and E are cpo's. Let $\{f_i : i \in I\} \subseteq D \to E$ be directed. Then*

$$f(d) := \bigsqcup_{i \in I} f_i(d)$$

is a well-defined continuous function from D to E, and f is the least upper bound of $\{f_i : i \in I\}$.

Proof. Let $d \in D$. Then $\{f_i(d) : i \in I\}$ is directed since $\{f_i : i \in I\}$ is, hence $\bigsqcup_{i \in I} f_i(d)$ exists, i.e. f is well-defined. f is continuous, for if $\{d_j : j \in J\}$ is directed, we have

$$f(\bigsqcup_{j \in J} d_j) = \bigsqcup_{i \in I} \bigsqcup_{j \in J} f_i(d_j) = \bigsqcup_{j \in J} \bigsqcup_{i \in I} f_i(d_j) = \bigsqcup_{j \in J} f(d_j),$$

where the middle equality follows from Lemma 4.5.

It remains to be shown that f is the l.u.b. of $\{f_i : i \in I\}$. Clearly f is an upper bound, since $f_i(d) \leq \bigsqcup_{i \in I} f_i(d) = f(d)$ for any $d \in D$. It is also the least upper bound, for if g is any other upper bound, we have $f_i(d) \leq g(d)$ and hence $f(d) = \bigsqcup_{i \in I} f_i(d) \leq g(d)$ for any $d \in D$. $\square$

Corollary 4.7 *$D \to E$ is a cpo if D and E are.*

We now define a corresponding operation on information systems, assigning to any $\mathbf{A} = (A, \mathrm{Con}_A, \vdash_A)$ and $\mathbf{B} = (B, \mathrm{Con}_B, \vdash_B)$ an

$$\mathbf{A} \to \mathbf{B} = (C, \mathrm{Con}, \vdash) \text{ such that } |\mathbf{A} \to \mathbf{B}| \cong |\mathbf{A}| \to |\mathbf{B}|.$$

We first have to decide on the set of data objects. The simplest finite piece of information on a continuous function from $|\mathbf{A}|$ to $|\mathbf{B}|$ is to say that if we are given the information $X \in \mathrm{Con}_A$ on the argument, then this suffices to know at least the information $b \in B$ on the value. Hence we let $C := \mathrm{Con}_A \times B$. Now when a set $\{(X_1, b_1), \ldots, (X_n, b_n)\}$ of such data objects is to be called consistent? Clearly just in the case when for any subset $I \subseteq \{1, 2, \ldots, n\}$ such that the left hand sides $\bigcup\{X_i : i \in I\}$ are consistent, i.e. may be viewed as approximations to a common argument, then the corresponding right hand sides $\{b_i : i \in I\}$ should be consistent too, i.e. it should be possible to view them as approximations to a single value. Therefore Con is defined to be the set of all $\{(X_1, b_1), \ldots, (X_n, b_n)\} \subseteq C$ such that

$$\forall I \subseteq \{1, \ldots, n\}. \bigcup_{i \in I} X_i \in \mathrm{Con}_A \implies \{b_i : i \in I\} \in \mathrm{Con}_B.$$

For the definition of the entailment relation $\vdash$ it is helpful to first define the notion of an *application* of $W := \{(X_1, b_1), \ldots, (X_n, b_n)\} \in \mathrm{Con}$ to $X \in \mathrm{Con}_A$:

$$\{(X_1, b_1), \ldots, (X_n, b_n)\}X := \{b_i : X \vdash_A X_i\}.$$

From the definition of Con we know that this set is in Con_B. Clearly application is *monotone in the second argument*, in the sense that

$$X \vdash_A X' \implies WX \vdash_B WX'.$$

In fact we even have $WX' \subseteq WX$ if $X \vdash_A X'$. Now define $W \vdash (X, b)$ by

$$W \vdash (X, b) :\iff WX \vdash_B b.$$

Lemma 4.8 *If* $\mathbf{A}$ *and* $\mathbf{B}$ *are information systems, then so is* $\mathbf{A} \to \mathbf{B}$ *defined as above.*

Proof. Let $\mathbf{A} = (A, \mathrm{Con}_A, \vdash_A)$ and $\mathbf{B} = (B, \mathrm{Con}_B, \vdash_B)$. The axioms 1.1(i), (ii) and (iv) are clearly satisfied.

For (iii), suppose

$$\{(X_1, b_1), \dots, (X_n, b_n)\} \vdash (X, b),$$

i.e. $\{b_j : X \vdash_A X_j\} \vdash_B b$. We have to show that

$$\{(X_1, b_1), \dots, (X_n, b_n), (X, b)\} \in \mathrm{Con}.$$

So let $I \subseteq \{1, \dots, n\}$ and suppose $X \cup \bigcup_{i \in I} X_i \in \mathrm{Con}_A$. We must show that $b \cup \{b_i : i \in I\} \in \mathrm{Con}_B$. Let $J \subseteq \{1, \dots, n\}$ consist of those j with $X \vdash_A X_j$. Then also $X \cup \bigcup_{i \in I} X_i \cup \bigcup_{j \in J} X_j \in \mathrm{Con}_A$. Since

$$\bigcup_{i \in I} X_i \cup \bigcup_{j \in J} X_j \in \mathrm{Con}_A,$$

from the consistency of $\{(X_1, b_1), \dots, (X_n, b_n)\}$ we can conclude that $\{b_i : i \in I\} \cup \{b_j : j \in J\} \in \mathrm{Con}_B$. But $\{b_j : j \in J\} \vdash_B b$ by assumption. Hence

$$\{b_i : i \in I\} \cup \{b_j : j \in J\} \cup \{b\} \in \mathrm{Con}_B.$$

For (v), suppose

$$W \vdash \{(X_1, b_1), \dots, (X_n, b_n)\} \vdash (X, b).$$

We have to show that $WX \vdash_B b$. Note that from $X \vdash_A X_i$ we can conclude $WX \vdash_B WX_i$ by the monotonicity of application in the second argument. Hence

$$WX \vdash_B \bigcup \{WX_i : X \vdash_A X_i\} \vdash_B \{b_i : X \vdash_A X_i\} \vdash_B b. \quad \square$$

Note that with the above definitions of the entailment relation $\vdash$ in $\mathbf{A} \to \mathbf{B}$ application is also *monotone in the first argument*, i.e.

$$W \vdash W' \implies WX \vdash_B W'X.$$

To see this let $W' = \{(X_1, b_1), \dots, (X_n, b_n)\}$ and observe that

$$WX \vdash_B \bigcup\{WX_i : X \vdash_A X_i\} \vdash_B \{b_i : X \vdash_A X_i\} = W'X.$$

Lemma 4.9 *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be information systems. Then the elements of* $\mathbf{A} \to \mathbf{B}$ *are exactly the approximable mappings from* $\mathbf{A}$ *to* $\mathbf{B}$.

Proof. Let $\mathbf{A} = (A, \mathrm{Con}_A, \vdash_A)$ and $\mathbf{B} = (B, \mathrm{Con}_B, \vdash_B)$. Suppose $r \in |\mathbf{A} \to \mathbf{B}|$. Then $r \subseteq \mathrm{Con}_A \times B$ is consistent and deductively closed. We have to show that r satisfies the axioms 3.1(i)–(iii) of approximable mappings.

(i) Let Xrb_i for $i \in \{1, \dots, n\}$. We must show $\{b_1, \dots, b_n\} \in \mathrm{Con}_B$. But this clearly follows from the consistency of r.

(ii) Let Xrb_i for $i \in \{1, \dots, n\}$ and $\{b_1, \dots, b_n\} \vdash_B b$. We must show that Xrb. But by the definition of $\vdash$ in $\mathbf{A} \to \mathbf{B}$ we have $\{(X, b_1), \dots, (X, b_n)\} \vdash (X, b)$, hence Xrb by the deductive closure of r.

(iii) Let $X \vdash_A X'$ and $X'rb$. We want Xrb. But $\{(X', b)\} \vdash (X, b)$ since $\{(X', b)\}X = \{b\}$ (which follows from $X \vdash_A X'$), hence again Xrb by the deductive closure of r.

For the other direction suppose that $r \colon \mathbf{A} \to \mathbf{B}$ is an approximable mapping. We must show that $r \in |\mathbf{A} \to \mathbf{B}|$.

1. Consistency of r.

   Suppose $X_i r b_i$ for all $i \in \{1, \dots, n\}$ and

   $$X := \bigcup\{X_i : i \in I\} \in \mathrm{Con}_A$$

   for some $I \subseteq \{1, \dots, n\}$. We must show that

   $$\{b_i : i \in I\} \in \mathrm{Con}_B.$$

   Now from $X_i r b_i$ and $X \vdash_A X_i$ we obtain Xrb_i by 3.1(iii) for any $i \in I$, and hence $\{b_i : i \in I\} \in \mathrm{Con}_B$ by 3.1(i).

2. Deductive closure of r.

   Suppose $X_i r b_i$ for all $i \in \{1, \ldots, n\}$ and

   $$W := \{(X_1, b_1), \ldots, (X_n, b_n)\} \vdash (X, b).$$

   We must show Xrb. By the definition of $\vdash$ for $\mathbf{A} \to \mathbf{B}$ we have $WX \vdash_B b$, which is $\{b_i : X \vdash_A X_i\} \vdash_B b$. Further by assumption $X_i r b_i$ we know $X r b_i$ by 3.1(iii) for all i with $X \vdash_A X_i$. Hence Xrb by 3.1(ii). $\square$

Theorem 4.10 *Let* $\mathbf{A} = (A, \mathrm{Con}_A, \vdash_A)$ *and* $\mathbf{B} = (B, \mathrm{Con}_B, \vdash_B)$ *be information systems. Then*

$$|\mathbf{A} \to \mathbf{B}| \cong |\mathbf{A}| \to |\mathbf{B}|.$$

The isomorphism pair is

$$\begin{array}{lll} \varphi\colon |\mathbf{A} \to \mathbf{B}| \to (|\mathbf{A}| \to |\mathbf{B}|) & \textit{given by} & \varphi(s) = |s|, \\ \psi\colon (|\mathbf{A}| \to |\mathbf{B}|) \to |\mathbf{A} \to \mathbf{B}| & \textit{given by} & \psi(f) = r_f. \end{array}$$

Proof. $\psi \circ \varphi = \mathrm{id}_{|\mathbf{A} \to \mathbf{B}|}$ and $\varphi \circ \psi = \mathrm{id}_{|\mathbf{A}| \to |\mathbf{B}|}$ has been shown in Theorem 3.2.

- φ is monotone.

  Suppose $s, s' \in |\mathbf{A} \to \mathbf{B}|$ and $s \subseteq s'$. We must show that $|s| \leq |s'|$ in the pointwise order of the continuous functions in $|\mathbf{A}| \to |\mathbf{B}|$. So let $z \in |\mathbf{A}|$. Then we have

  $$\begin{array}{rcl} |s|(z) & = & \{b \in B : Xsb \text{ for some } X \subseteq^{\mathrm{fin}} z\} \\ & \subseteq & \{b \in B : Xs'b \text{ for some } X \subseteq^{\mathrm{fin}} z\} \\ & = & |s'|(z). \end{array}$$

- ψ is monotone. Suppose $f, f'\colon |\mathbf{A}| \to |\mathbf{B}|$ are continuous and $f(z) \subseteq f'(z)$ for all $z \in |\mathbf{A}|$. Then for any $X \in \mathrm{Con}_A$ and $b \in B$ we have

  $$\begin{array}{rcl} X r_f b & \Longleftrightarrow & b \in f(\overline{X}) \\ & \Longrightarrow & b \in f'(\overline{X}) \\ & \Longleftrightarrow & X r_{f'} b. \quad \square \end{array}$$

Corollary 4.1 *Suppose D and E are domains with countable bases. Then the set $D \to E$ of all continuous functions from D to E equipped* *with the pointwise order is again a domain with countable basis.*

Proof. $D \to E \cong |ID| \to |IE| \cong |ID \to IE|$. $\square$

The cpo's together with the continuous functions between them form a cartesian closed category. We only need here the continuity of application, hence we restrict ourselves to a proof of this fact. We first prove the well-known universal property of cartesian products. Another way to put this is to say that a function into a product is continuous if its component functions are continuous. Note that the converse follows from Lemma 2.4 and the continuity of the projections.

Lemma 4.12 *Suppose F, D and E are cpo's. Let $f\colon F \to D$ and $g\colon F \to E$ be continuous functions. Then there is a unique continuous function $h\colon F \to D \times E$ such that $\pi_{\mathrm{left}} \circ h = f$ and $\pi_{\mathrm{right}} \circ h = g$. ($h$ is denoted by $f \times g$.)*

Proof. Suppose h is such a function. Then for any $x \in F$

$$h(x) = (f(x), g(x)),$$

so h is uniquely determined. For existence, define h as in the equations above. Clearly h is monotone. It remains to prove that h is continuous. So let $M \subseteq F$ be directed. Then we have

$$\begin{aligned} h(\bigsqcup M) &= (f(\bigsqcup M), g(\bigsqcup M)) \\ &= (\bigsqcup f(M), \bigsqcup g(M)) \\ &= (\bigsqcup \pi_{\mathrm{left}}(h(M)), \bigsqcup \pi_{\mathrm{right}}(h(M))) \\ &= \bigsqcup h(M) \qquad \text{by Lemma 4.1.} \quad \square \end{aligned}$$

The next thing to do is to show that the sections ι_1^e and ι_2^d of the identity function on $D \times E$ are continuous. By these we mean the functions $\iota_1^e\colon D \to D \times E$ given by $\iota_1^e(d) = (d, e)$ for any fixed $e \in E$, and $\iota_2^d\colon E \to D \times E$ given by $\iota_2^d(e) = (d, e)$ for any fixed $d \in D$. Clearly e.g. ι_1^e is monotone. To see that it is continuous, let $M \subseteq D$ be directed. Then clearly $\{(d, e) : d \in M\}$ is directed too, and $\bigsqcup_{d \in M}(d, e) = (\bigsqcup M, e)$.

Now we can show that a function from a product $D \times E$ is continuous if its sections are continuous, or as one might say if it is continuous in each component separately. Note that the converse follows from the continuity of the functions ι_1^e and ι_2^d.

Lemma 4.13 *Suppose D, E and F are cpo's. Then a function*

$$f\colon D \times E \to F$$

is continuous if and only if it is continuous in each component separately, i.e. if all sections $f_1^e : D \to F$ *given by* $f_1^e(d) = f(d,e)$ *and* $f_2^d : E \to F$ *given by* $f_2^d(e) = f(d,e)$ *are continuous.*

Proof. If f is continuous, then its sections $f_1^e = f \circ \iota_1^e$ and $f_2^d = f \circ \iota_2^d$ are continuous by Lemma 2.4 and the continuity of the functions ι_1^e and ι_2^d.

Conversely, we first show that f is monotone. So let $(d,e) \leq (d',e')$. Then

$$f(d,e) = f_1^e(d) \leq f_1^e(d') = f(d',e) = f_2^{d'}(e) \leq f_2^{d'}(e') = f(d',e').$$

Now assume that $M \subseteq D \times E$ is directed. We must show

$$f(\bigsqcup M) = \bigsqcup f(M).$$

Clearly $f(\bigsqcup M)$ is an upper bound of $f(M)$. Now suppose u is some upper bound of $f(M)$. We must show $f(\bigsqcup M) \leq u$. Now by Lemma 4.1 $\bigsqcup \pi_{\mathrm{left}}(M)$ and $\bigsqcup \pi_{\mathrm{right}}(M) =: e^*$ exist and we have

$$\begin{array}{rcll} f(\bigsqcup M) & = & f(\bigsqcup \pi_{\mathrm{left}}(M), \bigsqcup \pi_{\mathrm{right}}(M)) & \\ & = & \bigsqcup_{d \in \pi_{\mathrm{left}}(M)} f(d, \bigsqcup \pi_{\mathrm{right}}(M)) & (f_1^{e^*} \text{ is continuous}) \\ & = & \bigsqcup_{d \in \pi_{\mathrm{left}}(M)} \bigsqcup_{e \in \pi_{\mathrm{right}}(M)} f(d,e) & (f_2^d \text{ is continuous}). \end{array}$$

Let $d \in \pi_{\mathrm{left}}(M)$ and $e \in \pi_{\mathrm{right}}(M)$. Then $(d,e') \in M$ for some $e' \in E$ and $(d',e) \in M$ for some $d' \in D$. Since M is directed, there is some $(d'',e'') \in M$ with $d,d' \leq d''$ and $e,e' \leq e''$. Hence from the monotonicity of f we get $f(d,e) \leq f(d'',e'') \leq u$. Since this holds for any $e \in \pi_{\mathrm{right}}(M)$, we have $\bigsqcup_{e \in \pi_{\mathrm{right}}(M)} f(d,e) \leq u$, and since this again holds for any $d \in \pi_{\mathrm{left}}(M)$, we get

$$\bigsqcup_{d \in \pi_{\mathrm{left}}(M)} \bigsqcup_{e \in \pi_{\mathrm{right}}(M)} f(d,e) \leq u. \quad \square$$

Now we can show that application is continuous.

Lemma 4.14 *Let D and E be cpo's. Then* apply: $(D \to E) \times D \to E$ *given by* $\mathrm{apply}(f,d) = f(d)$ *is continuous.*

Proof. By Lemma 4.13 it suffices to prove that apply is continuous in each argument separately. For the second argument this is trivial, and monotonicity in the first argument follows from the pointwise definition of $\leq$ in $D \to E$. So assume $\{f_i : i \in I\} \subseteq D \to E$ is directed. Then we have

$$\begin{array}{rcll} \mathrm{apply}(\bigsqcup_{i\in I} f_i, d) & = & (\bigsqcup_{i\in I} f_i)(d) & \\ & = & \bigsqcup_{i\in I} f_i(d) & \text{by Lemma 4.6} \\ & = & \bigsqcup_{i\in I} \mathrm{apply}(f_i, d). & \square \end{array}$$

5 Partial Continuous Functionals

We now construct the partial continuous functionals. First we define for any simple type ϱ an information system $\mathbf{C}_\varrho = (C_\varrho, \mathrm{Con}_\varrho, \vdash_\varrho)$ by letting

$$\mathbf{C}_{\mathrm{nat}} := \mathbf{N},$$

i.e. the information system built from the set of natural numbers as described after Definition 1.1, and

$$\begin{array}{rcl} \mathbf{C}_{\varrho\to\sigma} & := & \mathbf{C}_\varrho \to \mathbf{C}_\sigma, \\ \mathbf{C}_{\varrho\times\sigma} & := & \mathbf{C}'_\varrho \times \mathbf{C}''_\sigma, \end{array}$$

where $\mathbf{C}'_\varrho$, $\mathbf{C}''_\sigma$ are variants of $\mathbf{C}_\varrho$, $\mathbf{C}_\sigma$ with disjoint sets of data objects (e.g. $C'_\varrho = \{(0, a) : a \in C_\varrho\}$ and $C''_\sigma = \{(1, a) : a \in C_\sigma\}$).

The information systems $\mathbf{C}_\varrho$ have a rather pleasant property, which amounts to the possibility to locate inconsistencies in two-element sets of data objects; this property has been called coherent by Plotkin [1978, p. 210]. It is defined as follows.

An information system $\mathbf{A} = (A, \mathrm{Con}, \vdash)$ is called *coherent* if for any set $X \subseteq^{\mathrm{fin}} A$ we have

$$X \in \mathrm{Con} \iff (\forall a, b \in X)\{a, b\} \in \mathrm{Con}.$$

Clearly $\mathbf{C}_{\mathrm{nat}} = \mathbf{N}$ is coherent, and furthermore we have

Lemma 5.1 *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be information systems.*

(i) If $\mathbf{B}$ *is coherent, then so is* $\mathbf{A} \to \mathbf{B}$.

(ii) If $\mathbf{A}$ *and* $\mathbf{B}$ *are coherent, then so is* $\mathbf{A} \times \mathbf{B}$.

Proof. Let $\mathbf{A} = (A, \mathrm{Con}_A, \vdash_A)$ and $\mathbf{B} = (B, \mathrm{Con}_B, \vdash_B)$.

(i) Let $\{(X_1, b_1), \ldots, (X_n, b_n)\} \subseteq \mathrm{Con}_A \times B$ and assume

$$\forall i, j.1 \leq i < j \leq n \to \{(X_i, b_i), (X_j, b_j)\} \in \mathrm{Con}. \qquad (*)$$

We have to show that $\{(X_1, b_1), \dots, (X_n, b_n)\} \in \mathrm{Con}$. So assume $I \subseteq \{1, \dots, n\}$ and $\bigcup_{i \in I} X_i \in \mathrm{Con}_A$. We show $\{b_i : i \in I\} \in \mathrm{Con}_B$. Now since $\mathbf{B}$ is coherent by assumption, it suffices to show that $\{b_i, b_j\} \in \mathrm{Con}_B$ for all $i, j \in I$. So let $i, j \in I$. By assumption we have $X_i \cup X_j \in \mathrm{Con}_A$ and hence by $(*)$ and the definition of Con also $\{b_i, b_j\} \in \mathrm{Con}_B$.

(ii) We may assume that A and B are disjoint. Let $a_1, \dots, a_n \in A$ and $b_1, \dots, b_m \in B$ and assume that any two–element subset of $\{a_1, \dots, a_n, b_1, \dots, b_m\}$ is consistent. Then $\{a_1, \dots, a_n\} \in \mathrm{Con}_A$ and $\{b_1, \dots, b_m\} \in \mathrm{Con}_B$ since $\mathbf{A}$ and $\mathbf{B}$ are coherent. Therefore $\{a_1, \dots, a_n, b_1, \dots, b_m\} \in \mathrm{Con}$ by definition of Con. □

Corollary 5.2 *The information systems $\mathbf{C}_\varrho$ are all coherent.*

We now let $\mathrm{Cont}_\varrho := |\mathbf{C}_\varrho|$ be the domain of elements of $|\mathbf{C}_\varrho|$; these elements are called the *partial continuous functionals* of type ϱ. The partial continuous functionals are partial in the sense that the object undefined, i.e. $\perp := \overline{\emptyset}$, is a possible value. We now can introduce easily the total continuous functionals, i.e. those which for total arguments only produce total values; this idea clearly can be turned into a valid inductive definition. We will show that the total functionals are dense in the sense that any finite partial continuous functional (i.e. any $\overline{X}$ with $X \in \mathrm{Con}_\varrho$) can be extended to a total functional.

The set G_ϱ of the *total* functionals of type ϱ is the subset of Cont_ϱ defined by

$$\begin{array}{lcl} G_{\mathrm{nat}} & := & \{z \in \mathrm{Cont}_{\mathrm{nat}} : z \neq \emptyset\} = \{\{n\} : n \in \mathbf{N}\}, \\ G_{\varrho \to \sigma} & := & \{z \in \mathrm{Cont}_{\varrho \to \sigma} : \forall x \in G_\varrho . zx \in G_\sigma\}, \\ G_{\varrho \times \sigma} & = & \{z \in \mathrm{Cont}_{\varrho \times \sigma} : z \cap C_\varrho \in G_\varrho \text{ and } z \cap C_\sigma \in G_\sigma\}. \end{array}$$

Here we have written zx instead of the more correct $|z|(x)$ (cf. Theorem 4.10). We will continue to do so below.

We now define a relation $\sim_\varrho$ on G_ϱ. For $z, z' \in G_{\mathrm{nat}}$ let

$$z \sim_{\mathrm{nat}} z' \iff z = z'.$$

For $z, z' \in G_{\varrho \to \sigma}$ let

$$z \sim_{\varrho \to \sigma} z' \iff \forall x \in G_\varrho . zx \sim_\sigma z'x.$$

For $z, z' \in G_{\varrho \times \sigma}$ let

$$z \sim_{\varrho \times \sigma} z' \iff \pi_{\mathrm{left}} z \sim_\varrho \pi_{\mathrm{left}} z' \text{ and } \pi_{\mathrm{right}} z \sim_\sigma \pi_{\mathrm{right}} z'.$$

Clearly $\sim_\varrho$ is an equivalence relation.

We obviously want to know that $\sim_\varrho$ is compatible with application. The only nontrivial part of this argument is to show that

$$x \sim_\varrho y \implies zx \sim_\sigma zy.$$

This has first been proved by Ershov [1975], as an application of the density theorem. However, a simpler proof has later been found by Longo and Moggi [1984], and we present that one. First we need some Lemmata. The first one notes a rather obvious fact on total functionals.

Lemma 5.3 *If $z \in G_\varrho$, $z' \in \mathrm{Cont}_\varrho$ and $z \subseteq z'$, then $z' \in G_\varrho$.*

Proof. By induction on ϱ.

- For the ground type nat the claim is obvious.
- $\varrho \to \sigma$

  Assume $z \in G_{\varrho\to\sigma}$ and $z \subseteq z'$. We must show $z' \in G_{\varrho\to\sigma}$. So let $x \in G_\varrho$. We have to show $z'x \in G_\sigma$. But $z'x \supseteq zx \in G_\sigma$, so the claim follows by induction hypothesis.
- $\varrho \times \sigma$

  The claim follows easily from the monotonicity of the projection functions. □

Lemma 5.4 *For any $z_1, z_2 \in \mathrm{Cont}_{\varrho\to\sigma}$ and $x \in \mathrm{Cont}_\varrho$ we have*

$$(z_1 \cap z_2)x = z_1x \cap z_2x.$$

Proof. By the definition of $|r|$ in Theorem 3.2 we have

$$\begin{array}{rcl} |z_1 \cap z_2|x & = & \{b \in C_\sigma : (\exists X \subseteq^{\text{fin}} x)(X, b) \in z_1 \cap z_2\} \\ & = & \{b \in C_\sigma : (\exists X_1 \subseteq^{\text{fin}} x)(X_1, b) \in z_1\} \cap \\ & & \{b \in C_\sigma : (\exists X_2 \subseteq^{\text{fin}} x)(X_1, b) \in z_2\} \\ & = & |z_1|x \cap |z_2|x. \end{array}$$

The part $\subseteq$ of the middle equality is obvious. For $\supseteq$, let $X_i \subseteq^{\text{fin}} x$ with $(X_i, b) \in z_i$ be given. Choose $X = X_1 \cup X_2$. Then clearly $(X, b) \in z_i$ (as $\{(X_i, b)\} \vdash (X, b)$ and z_i is deductively closed). □

Lemma 5.5 *For any $z, z' \in G_\varrho$ we have $z \sim_\varrho z' \iff z \cap z' \in G_\varrho$.*

Proof. By induction on ϱ.

- For the ground type nat the claim is obvious.
- $\varrho \to \sigma$

$$\begin{array}{rcll} z \sim_{\varrho\to\sigma} z' & \iff & \forall x \in G_\varrho . zx \sim_\sigma z'x & \\ & \iff & \forall x \in G_\varrho . zx \cap z'x \in G_\sigma & \text{by ind. hyp.} \\ & \iff & \forall x \in G_\varrho . (z \cap z')x \in G_\sigma & \text{by Lemma 5.4} \\ & \iff & z \cap z' \in G_{\varrho\to\sigma}. & \end{array}$$

- $\varrho \times \sigma$

$$\begin{array}{rcl} z \sim_{\varrho\times\sigma} z' & \iff & \pi_{\text{left}} z \sim_\varrho \pi_{\text{left}} z' \quad \text{and} \quad \pi_{\text{right}} z \sim_\sigma \pi_{\text{right}} z' \\ & \iff & (\pi_{\text{left}} z) \cap (\pi_{\text{left}} z') \in G_\varrho \text{ and} \\ & & (\pi_{\text{right}} z) \cap (\pi_{\text{right}} z') \in G_\sigma \\ & \iff & \pi_{\text{left}}(z \cap z') \in G_\varrho \quad \text{and} \quad \pi_{\text{right}}(z \cap z') \in G_\sigma \\ & \iff & z \cap z' \in G_{\varrho\times\sigma}. \quad \Box \end{array}$$

Theorem 5.6 *For any $x, y \in G_\varrho$ and $z \in G_{\varrho\to\sigma}$ we have*

$$x \sim_\varrho y \implies zx \sim_\sigma zy.$$

Proof. Since $x \sim_\varrho y$ we have $x \cap y \in G_\varrho$ by Lemma 5.5. Now $zx, zy \supseteq z(x \cap y)$ by Theorem 3.2, and hence $zx \cap zy \in G_\sigma$ by Lemma 5.3. But this implies $zx \sim_\sigma zy$ again by Lemma 5.5. $\Box$

We now prove the density theorem, which says that any finite partial continuous functional (i.e. any $\overline{X}$ with $X \in \text{Con}_\varrho$) can be extended to a total functional. This result is essentially due to Kreisel [1959]. The proof we give here has been obtained by specialization of a proof for a more general density theorem (for arbitrary f_0–spaces) given by Berger in [1990], [1993] (this proof also appears in Stoltenberg Hansen [1994]). Berger's proof in turn has been obtained by abstracting some general features from a proof of a density theorem given by Ershov[1974a], [1975].

Theorem 5.7 (Density) *For any $X \in \text{Con}_\varrho$ we can find an $x \in G_\varrho$ such that $X \subseteq x$.*

Proof. Call a type ϱ *dense* if

$$\forall X \in \text{Con}_\varrho \exists x \in G_\varrho . X \subseteq x.$$

Furthermore call a type ϱ *separating* if

$$\forall X_1, X_2 \in \text{Con}_\varrho (X_1 \cup X_2 \notin \text{Con}_\varrho \to \exists \vec{z} \in G . \emptyset \neq \overline{X_1}\vec{z} \neq \overline{X_2}\vec{z} \neq \emptyset).$$

Here $\vec{z} \in G$ means that $\vec{z}$ is a sequence of total z_i or of 0 or 1 such that $X_j\vec{z}$ is of type nat.

We prove by simultaneous induction on ϱ that any type ϱ is dense and separating. This extension of our claim is helpful to get the induction through.

- For the ground type nat both claims are obvious.
- $\varrho \to \sigma$ is separating

  This will follow from the inductive hypotheses that ϱ is dense and σ is separating. So let $W, W' \in \mathrm{Con}_{\varrho\to\sigma}$ such that

  $$W \cup W' \notin \mathrm{Con}_{\varrho\to\sigma}.$$

  Since $\mathbf{C}_\varrho$ is coherent by Corollary 5.2 there are $(X, a) \in W$ and $(X', a') \in W'$ such that $X \cup X' \in \mathrm{Con}_\varrho$ but $\{a, a'\} \notin \mathrm{Con}_\sigma$. Since ϱ is dense we have a $z \in G_\varrho$ such that $X \cup X' \subseteq z$. Hence

  $$a \in \overline{W}z \quad \text{and} \quad a' \in \overline{W'}z.$$

  Now since σ is separating there are $\vec{z} \in G$ such that

  $$\emptyset \neq \overline{\{a\}}\vec{z} \neq \overline{\{a'\}}\vec{z} \neq \emptyset,$$

  hence also

  $$\emptyset \neq \overline{W}z\vec{z} \neq \overline{W'}z\vec{z} \neq \emptyset.$$

- $\varrho \to \sigma$ is dense

  This will follow from the inductive hypotheses that ϱ is separating and σ is dense. So let

  $$W = \{(X_i, a_i) : i \in I\} \in \mathrm{Con}_{\varrho\to\sigma}.$$

  Consider i, j such that $\{a_i, a_j\} \notin \mathrm{Con}_\sigma$. Then $X_i \cup X_j \notin \mathrm{Con}_\varrho$. Since ϱ is separating, there are $\vec{z}_{ij} \in G$ and k_{ij}, l_{ij} such that

  $$\{k_{ij}\} = \overline{X_i}\vec{z}_{ij} \neq \overline{X_j}\vec{z}_{ij} = \{l_{ij}\}.$$

  We clearly may assume that $\vec{z}_{ij} = \vec{z}_{ji}$ and $(k_{ij}, l_{ij}) = (l_{ji}, k_{ji})$.

  Now define for any $X \in \mathrm{Con}_\varrho$ a set I_X of indices $i \in I$ such that 'X behaves as X_i with respect to the $\vec{z}_{ij}$'. More precisely, let

  $$I_X := \{i \in I : \forall j. \{a_i, a_j\} \notin \mathrm{Con}_\sigma \to \overline{X}\vec{z}_{ij} = \{k_{ij}\}\}.$$

We first show that

$$\{a_i : i \in I_X\} \in \mathrm{Con}_\sigma. \tag{1}$$

Since $\mathbf{C}_\sigma$ is coherent it suffices to show that $\{a_i, a_j\} \in \mathrm{Con}_\sigma$ for all $i, j \in I_X$. So let $i, j \in I_X$ and assume $\{a_i, a_j\} \notin \mathrm{Con}_\sigma$. Then we have

$$\overline{X}\vec{z}_{ij} = \{k_{ij}\} \quad \text{as } i \in I_X$$

and

$$\overline{X}\vec{z}_{ji} = \{k_{ji}\} \quad \text{as } j \in I_X,$$

and because of $\vec{z}_{ij} = \vec{z}_{ji}$ and $k_{ij} \neq l_{ij} = k_{ji}$ we could conclude that $\overline{X}\vec{z}_{ij}$ would be inconsistent. This contradiction proves $\{a_i, a_j\} \in \mathrm{Con}_\sigma$ and hence (1).

Since (1) holds and σ is dense by induction hypothesis, we can find a $y_{I_X} \in G_\sigma$ such that $a_i \in y_{I_X}$ for all $i \in I_X$. Now define a relation $r \subseteq \mathrm{Con}_\varrho \times C_\sigma$ by

$$Xra :\Longleftrightarrow \begin{cases} a \in y_{I_X}, & \text{if } \overline{X} \text{ is defined on all } \vec{z}_{ij}; \\ \{a_i : i \in I_X\} \vdash_\sigma a, & \text{otherwise,} \end{cases}$$

where 'x is defined on $\vec{z}$' means $x\vec{z} \neq \emptyset$. We will show that $r \in G_{\varrho\to\sigma}$ and $W \subseteq r$.

For $W \subseteq r$ we have to show $X_i r a_i$ for all $i \in I$. But this holds, since clearly $i \in I_{X_i}$ and hence $\{a_j : j \in I_{X_i}\} \vdash_\sigma a_i$ and also $a_i \in y_{I_{X_i}}$.

We now show that r is an approximable mapping; by Lemma 4.9 this means $r \in \mathrm{Cont}_{\varrho\to\sigma}$. To prove this we have to verify conditions (i) to (iii) from Definition 3.1.

(i) $Xrb_1, \ldots, Xrb_n \implies \{b_1, \ldots, b_n\} \in \mathrm{Con}_\sigma$.
If $\overline{X}$ is defined on all $\vec{z}_{ij}$ the claim follows from the consistency of y_{I_X}. If not, the claim follows from the properties of $\vdash_\sigma$.

(ii) $Xrb_1, \ldots, Xrb_n, \{b_1, \ldots, b_n\} \vdash_\sigma a \implies Xra$.
If $\overline{X}$ is defined on all $\vec{z}_{ij}$ the claim follows from the deductive closure of y_{I_X}. If not, the claim follows from the properties of $\vdash_\sigma$.

(iii) $X \vdash_\varrho X', X'ra \implies Xra$.
Case 1. $\overline{X'}$ is defined on all $\vec{z}_{ij}$. Then also $\overline{X}$ is defined on all $\vec{z}_{ij}$. From $X'ra$ we get $a \in y_{I_{X'}}$. We have to show that

$a \in y_{I_X}$. Now since X and X' are defined on all $\vec{z}_{ij}$ and $X \vdash_\varrho X'$, they must have the same values on the $\vec{z}_{ij}$, hence $I_{X'} = I_X$ and therefore $y_{I_{X'}} = y_{I_X}$.
Case 2. Otherwise. We get $\{a_i : i \in I_{X'}\} \vdash_\sigma a$ from $X'ra$. Now from $X \vdash_\varrho X'$ we can conclude $I_{X'} \subseteq I_X$, by the definition of I_X. Hence $\{a_i : i \in I_X\} \vdash_\sigma a$, and also $a \in y_{I_X}$ (since $a_i \in y_{I_X}$ for all $i \in I_X$, and y_{I_X} is deductively closed). Therefore Xra.

It remains to show that $r \in G_{\varrho\to\sigma}$. So let $x \in G_\varrho$. We must show

$$|r|(x) = \{a \in C_\sigma : \exists X \subseteq^{\text{fin}} x.Xra\} \in G_\sigma.$$

Now for any $i, j \in I$ we have $x\vec{z}_{ij} \neq \emptyset$, hence there is some $X_{ij} \subseteq^{\text{fin}} x$ such that $\overline{X_{ij}}\vec{z}_{ij} \neq \emptyset$. Let $X \subseteq^{\text{fin}} x$ be the union of all the X_{ij}. Then clearly $\overline{X}$ is defined on all $\vec{z}_{ij}$ and hence Xra for all $a \in y_{I_X}$. Therefore $y_{I_X} \subseteq |r|(x)$ and hence $|r|(x) \in G_\sigma$ by Lemma 5.3.

- $\varrho \times \sigma$ is dense and separating

  This follows easily from the induction hypotheses that ϱ and σ are dense and separating. □

As an application of the Density Theorem we prove a choice principle for total continuous functionals. This result was first proved by Kreisel [1959]; our proof essentially follows Berger [1993].

Theorem 5.8 (Choice Principle for Total Continuous Functionals) *We can construct an object*

$$r \in \text{Cont}_{(\varrho\to\sigma\to\text{nat})\to\varrho\to\sigma}$$

such that for any $F \in G_{\varrho\to\sigma\to\text{nat}}$ satisfying

$$(\forall x \in G_\varrho)(\exists y \in G_\sigma)Fxy = \{0\}$$

we have $rF \in G_{\varrho\to\sigma}$ and

$$(\forall x \in G_\varrho)Fx(rFx) = \{0\}.$$

Proof. Let $Y_0, Y_1, Y_2, \ldots$ be an enumeration of Con_σ. By the Density Theorem 5.7 we can find for any Y_n a $y_n \in G_\sigma$ such that $Y_n \subseteq y_n$. Define for any $W \in \text{Con}_{\varrho\to\sigma\to\text{nat}}$, $X \in \text{Con}_\varrho$ and $a \in C_\sigma$

$$\begin{aligned} Wr(X,a) \quad \iff \quad & (\exists m, \vec{k}.(\forall i < m)\overline{W}\,\overline{X}y_i = \{k_i + 1\} \wedge \\ & \overline{W}\,\overline{X}y_m = \{0\} \wedge a \in y_m) \ \vee \ \emptyset \vdash_\sigma a. \end{aligned}$$

We first show that r is an approximable mapping. To prove this we have to verify conditions (i) to (iii) from Definition 3.1.

(i) If

$$Wr(X_1, a_1), \ldots, Wr(X_n, a_n)$$

then

$$\{(X_1, a_1), \ldots, (X_n, a_n)\} \in \mathrm{Con}_{\varrho\to\sigma}.$$

Assume the premise and $X := \bigcup_{i\in I} X_i \in \mathrm{Con}_\varrho$. We must show that $\{a_i : i \in I\} \in \mathrm{Con}_\sigma$. If $\emptyset \vdash_\sigma a_i$ for all $i \in I$ we are done. Otherwise for all $i \in I$ such that $\emptyset \not\vdash_\sigma a_i$ the numbers m_i in the definition of $Wr(X_i, a_i)$ are all the same, $= m$ say. Hence $\{a_i : i \in I\} \subseteq y_m$, and the claim follows from the consistency of y_m.

(ii) If

$$Wr(X_1, a_1), ..., Wr(X_n, a_n), \{(X_1, a_1), ..., (X_n, a_n)\} \vdash_{\varrho\to\sigma} (X, a)$$

then

$$Wr(X, a).$$

Assume the premise and $I := \{i : X \vdash_\varrho X_i\}$. We know that $\{a_i : i \in I\} \vdash_\sigma a$. If $\emptyset \vdash_\sigma a_i$ for all $i \in I$ we are done. Otherwise for all $i \in I$ such that $\emptyset \not\vdash_\sigma a_i$ the numbers m_i in the definition of $Wr(X_i, a_i)$ are all the same, $= m$ say. Hence $\{a_i : i \in I\} \subseteq y_m$, and the claim follows from the deductive closure of y_m.

(iii) If

$$W \vdash_{\varrho\to\sigma\to\mathrm{nat}} W', W'r(X, a)$$

then

$$Wr(X, a).$$

This clearly follows from the definition of r; the m from $W'r(X, a)$ can be used for $Wr(X, a)$.

We finally show that for all $F \in G_{\varrho\to\sigma\to\mathrm{nat}}$ satisfying

$$(\forall x \in G_\varrho)(\exists y \in G_\sigma) Fxy = \{0\}$$

and for all $x \in G_\varrho$ we have $rFx \in G_\sigma$ and $Fx(rFx) = \{0\}$. So let F and x with these properties be given. By assumption there is a $y \in G_\sigma$ such that $Fxy = \{0\}$. Hence by the definition of application there is a $Y_n \in \mathrm{Con}_\sigma$ such that $Fx\overline{Y_n} = \{0\}$. Since $Y_n \subseteq y_n$ we also have

$Fxy_n = \{0\}$. Clearly we may assume here that n is minimal with this property, i.e. that

$$Fxy_0 = \{k_0 + 1\}, \ldots, Fxy_{n-1} = \{k_{n-1} + 1\}.$$

We show that $rFx \supseteq y_n$ or more precisely $||r|(F)|(x) \supseteq y_n$; this suffices by Lemma 5.3. Recall that

$$|r|(F) = \{(X, a) \in \mathrm{Con}_\varrho \times C_\sigma : (\exists W \subseteq^{\mathrm{fin}} F) Wr(X, a)\}$$

and

$$\begin{aligned} ||r|(F)|(x) &= \{a \in C_\sigma : (\exists X \subseteq^{\mathrm{fin}} x)(X, a) \in |r|(F)\} \\ &= \{a \in C_\sigma : (\exists X \subseteq^{\mathrm{fin}} x)(\exists W \subseteq^{\mathrm{fin}} F) Wr(X, a)\}. \end{aligned}$$

Now let $a \in y_n$. By the choice of n we then get $X \subseteq^{\mathrm{fin}} x$ and $W \subseteq^{\mathrm{fin}} F$ such that

$$(\forall i < n) \overline{W}\, \overline{X} y_i = \{k_i + 1\} \qquad \text{and} \qquad \overline{W}\, \overline{X} y_n = \{0\}.$$

Therefore $Wr(X, a)$ and hence $a \in ||r|(F)|(x)$. $\square$

We finally comment of the notion of effectivity in this context. In particular we want to define what it means for a partial continuous functional of an arbitrary simple type ϱ to be computable.

An information system $(A, \mathrm{Con}, \vdash)$ is called *effective* if the (countable) set A, the set Con and the relation $\vdash$ are all decidable. Since decidability always means decidability relative to a given set we must assume here that the data objects are taken from a fixed given set, e.g. the set of natural numbers or the set of all strings over some finite alphabet. It is easy to see that all the operations on information systems introduced in Section 4 preserve effectivity.

In an effective information system it makes sense to talk about a recursively enumerable (r.e.) ideal, since any ideal is a set of data objects. The *computable* elements of an effective information system are defined to be its r.e. ideals. Clearly application does not carry us out of the realm of computable elements, i.e. if $r \in |\mathbf{A} \to \mathbf{B}|$ and $x \in |\mathbf{B}|$ are computable, then so is rx (or more precisely $|r|(x)$). Also we can compute an r.e. index of $|r|x$ from those of r and x.

Let us now look somewhat closer at our particular information systems $\mathbf{C}_\varrho$, whose elements are the partial continuous functionals. A *computable* functional of type ϱ is defined to be a computable element of $\mathbf{C}_\varrho$. The set of all computable functionals of type ϱ is denoted by

Comp_ϱ.

Theorem 5.9 (Effective Density) *For any* $X \in \mathrm{Con}_\varrho$ *we can find effectively an* $x \in G_\varrho \cap \mathrm{Comp}_\varrho$ *such that* $X \subseteq x$.

Proof. By inspection of the proof of the Density Theorem 5.7. To see that r (in the proof that $\varrho \to \sigma$ is dense) is effective one needs that

- $\{a_i : i \in I_X\} \vdash_\sigma a$ implies $a \in y_{I_X}$ for all X and a, since by definition $a_i \in y_{I_X}$ for all $i \in I_X$. Hence

$$\begin{aligned} Xra \quad \Longleftrightarrow \quad & \{a_i : i \in I_X\} \vdash_\sigma a \text{ or} \\ & (a \in y_{I_X} \text{ and } \overline{X} \text{ is defined on all } \vec{z_{ij}}). \end{aligned}$$

- If $\overline{X}$ is defined on all $\vec{z_{ij}}$, one can actually compute I_X (and not only an enumeration procedure for I_X). $\square$

Theorem 5.10 (Effective Choice Principle) *There is an*

$$r \in \mathrm{Comp}_{(\varrho\to\sigma\to\mathrm{nat})\to\varrho\to\sigma}$$

such that for any $F \in G_{\varrho\to\sigma\to\mathrm{nat}}$ *satisfying*

$$(\forall x \in G_\varrho)(\exists y \in G_\sigma)Fxy = \{0\}$$

we have $rF \in G_{\varrho\to\sigma}$ *and*

$$(\forall x \in G_\varrho)Fx(rFx) = \{0\}.$$

Proof. Immediate from the proof of the choice principle for total continuous functionals in 5.8. $\square$

References

Berger, U.

[1990] *Totale Objekte und Mengen in der Bereichstheorie*, Ph.D. Thesis, Mathematisches Institut der Universität München, 1990.

[1993] Total sets and objects in domain theory, *Annals of Pure and Applied Logic*, 60 (1993) 91–117.

Ershov, Y.L.

[1974] The model G of the theory BR, *Soviet Math. Doklady*, 15 (1974) 1158–1160.

[1975] Theorie der Numerierungen II, *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 21 (1975) 473–584.

[1977] Model C of partial continuous functionals, in R. Gandy and M. Hyland (Eds.), *Logic Colloquium 1976*, North Holland, Amsterdam, 1977, pp. 455–467.

Kleene, S.C.

[1959] Countable functionals, in A. Heyting (Ed.), *Constructivity in Mathematics*, North Holland, Amsterdam, 1959, pp. 81–100.

Kreisel, G.

[1959] Interpretation of analysis by means of constructive functionals of finite types, in A. Heyting (Ed.), *Constructivity in Mathematics*, North Holland, Amsterdam, 1959, pp. 101–128.

Longo, G., and Moggi, E.

[1984] The hereditary partial effective functionals and recursion theory in higher types, *Journal of Symbolic Logic*, 49 (1984) 1319–1332.

Larsen, K.G., and Winskel, G.

[1984] Using information systems to solve recursive domain equations effectively, in *Proceedings of the Conference on Abstract Datatypes, Sophia-Antipolis, France*, Lecture Notes in Computer Science, vol. 173, Springer, Berlin, 1984, pp. 109–129.

Normann, D.

[1993] Closing the gap between the continuous functionals and recursion in ${}^3E$, submitted to the *Proceedings of the Sacks conference*, MIT, Cambridge, 1993.

Platek, R.A.

[1966] *Foundations of recursion theory*, Ph.D. Thesis, Department of Mathematics, Stanford University, 1966.

Plotkin, G.D.

[1977] LCF considered as a programming language, *Theoretical Computer Science*, 5 (1977) 223–255.

[1978] $\mathbf{T}^\omega$ as a universal domain, *Journal of Computer and System Sciences*, 17 (1978) 209–236.

Roscoe, A.W.

[1987] *Notes on domain theory*, 1987, unpublished notes, Oxford University, 1987.

Schwichtenberg, H.

[1991] Primitive recursion on the partial continuous functionals, in M. Broy (Ed.), *Informatik und Mathematik*, Springer, Berlin, 1991, pp. 251–269.

Scott, D.

[1970] *Outline of a mathematical theory of computation*, Technical Monograph PRG–2, Oxford University Computing Laboratory, 1970.

[1982] Domains for denotational semantics, in E. Nielsen and E. Schmidt (Eds.), *Automata, Languages and Programming*, Lecture Notes in Computer Science, vol. 140, pages 577–613, Springer, Berlin, 1982.

Stoltenberg-Hansen, V., Griffor, E., and Lindström, I.

[1994] *Mathematical Theory of Domains*, Cambridge University Press, Cambridge, 1994.

KREISEL'S PHILOSOPHY

Mathematical Logic: What Has it Done for the Philosophy of Mathematics?

by Carlo Cellucci

1 Prologue

The aim of this paper is not to provide any systematic reconstruction of Kreisel's views but only to discuss some claims concerning the relationship between mathematical logic and the philosophy of mathematics that repeatedly occur in his writings. Although I do not know to what extent they are representative of his present position, they correspond to widespread views of the logical community and so seem worth discussing anyhow. Such claims will be used as reference to make some remarks about the present state of relations between mathematical logic and the philosophy of mathematics.

Kreisel's views greatly influenced me in the Sixties and the Seventies. His critical remarks on the foundational programs taught me that one could and should have an approach to the subject of mathematical logic less dogmatic, corporative and even thoughtless than the one the logical community sometimes used to have. This is even more true today when the professionalization of mathematical logic generates a flood of results but few new ideas and the lack of ideas leads to the sheer byzantinism of most current production in mathematical logic.

In the past few years, however, I have come to the conclusion that Kreisel's criticism has not been radical enough: his main worry seems to have been to preserve as much as possible – to save the savable – of the tradition of mathematical logic. His critical remarks have focused on the defects of the foundational schools, thus drawing attention away from the intrinsic defects of mathematical logic itself instead of stressing the need of putting the subject on new grounds replacing it by a more adequate approach. This seems absolutely necessary, not only to rectify misconceptions about the nature of mathematics, such as those spelled out in Crowe [1988], for which mathematical logic has great responsibility, but also to meet the challenges that logic has to face as a result of the development of entirely new subjects such as computer science and artificial intelligence. To such challenges mathematical logic has been so far hopelessly unequal, contrary to McCarthy's [1963, p. 69] hope that the relationship between these new subjects and mathematical logic could be as fruitful as the one between analysis and physics.

In this paper I will try to explain the reasons of my disagreement with Kreisel, as much as it can be done in the limited space allowed to me.

2 Mathematical Logic and the Philosophy of Mathematics

Most of Kreisel's work in the Sixties and the Seventies can be seen as an attempt to answer the question occurring in the title. In particular, a paper bearing the same title [1967] contains claims on the role of mathematical logic in the philosophy of mathematics that seem widespread in the logical community, or at least, as Kreisel would put it, in the silent majority.

Kreisel's remarks in [1967] concerning the contribution of mathematical logic to the philosophy of mathematics seem somewhat ambiguous. While stressing the defects of mathematical logic, he does not go so far as to consider replacing it by a substantially different, more comprehensive paradigm: his dependence on the tradition of mathematical logic is too strong to allow it. At most he would like to convert the machinery originally developed by mathematical logic to further the grand aims of the foundational schools into new machinery intended to serve more limited aims in mathematical practice, such as "unwinding" proofs [1977]. Behind such program there seems to be the same old illusion that has deceived so many mathematical logicians, from

Frege [1967, pp. 6–7] and Hilbert [1967, pp. 383–384] to Gödel [1990, p. 140]: in the latter's words, the hope that mathematical logic would contribute to theoretical mathematics to the same extent as the decimal system of numbers has contributed to numerical computations.

Such a hope has been constantly ridiculed by mathematicians like Dieudonné [1988, pp. 21–22] who point out that mathematical logic has had no serious influence on mathematical practice; that if by any chance all work done in mathematical logic since 1925 disappeared, this would make no difference to the working mathematician; that, as a mathematical discipline, mathematical logic is marginal with respect to the rest of mathematics. Even logicians like Longo [1991, p. 120] agree that mathematicians could hardly see any direct relevance to their work of set theory or model theory, just to mention the most mathematical areas of mathematical logic, let alone of proof theory or recursion theory. While any algebraist is aware of the relevance of geometry or even analysis to his own discipline, as well as any analyst is aware of the relevance of other disciplines to his own, and so on so forth with all possible combinations, this is not so for mathematical logic.

There are two alternative views concerning mathematical logic: either, as suggested, e.g., by Brouwer [1975] and recently by Barwise [1988], it is considered simply as a part of mathematics, or, as maintained by the founding fathers of mathematical logic, from Frege to Hilbert, it is conceived as the logic of mathematics. In this respect Kreisel's position [1967, p. 201] seems somewhat eclectic. On the one hand, he considers mathematical logic as a part of mathematics because model theory belongs to set theory and the theory of formal systems belongs to combinatorial arithmetic. On the other hand, perhaps primarily, he views it as the logic of mathematics insofar as it is meant to express in what terms mathematical experience is to be analysed. In the latter capacity mathematical logic is intended to be a tool in the philosophy of mathematics.

"Philosophy of mathematics" is meant by Kreisel not in the broad sense of the philosophical tradition but only in the strict and rather peculiar sense of the foundational schools: philosophy of mathematics consists of the foundations of mathematics. As a matter of fact, in [1967] the discussion about the philosophy of mathematics is confined to the views of the foundational schools. Such a narrow approach is related to his view, expressed e.g. in [1984, p. 82], that philosophy deals with notions that may serve only at an early stage, when we know too little about the phenomenon involved in order to ask sensible specific questions, and that preoccupation with these notions may draw attention away from genuinely rewarding questions. Contrary to Kreisel's outlook, the central role of philosophy of mathematics in the philosophical tradition and the limited scope of the foundational schools has been recently reasserted, e.g., by Prawitz [1993, pp. 87–88, 96–97].

The narrowness of Kreisel's approach appears from the fact that in the philosophical tradition several questions about the nature of mathematics have been discussed that cannot be reduced to those of the foundational schools. A macroscopic example is provided by the problem of developing a logic of mathematical discovery, which either was rejected as psychological by Frege or was trivialized by Hilbert replacing it by the *Entscheidungsproblem*.

Kreisel's conception of the philosophy of mathematics wholly determines the role he assigns to mathematical logic in [1967]. For example, since he assumes that mathematical logic is intended to be a tool in the philosophy of mathematics, just like the theory of partial differential equations is a tool in what used to be called natural philosophy [1967, p. 201], he considers the mathematical apparatus of mathematical logic as neutral instead of viewing it as strictly dependent on the aims of the foundational schools. Such is the case of formalization, whose importance as a tool is stressed in [1967, pp. 201–202]. Now, far from being neutral, the notion of formal system is strictly dependent on the foundational aims of Frege and Hilbert and would be useless, or at least inadequate, for other aims. Evidence for this is provided by automated theorem proving and logic programming, where the use of formal systems of mathematical logic has generally been very disappointing: in order to make them more efficient, compensating the defects of formal systems, it has been necessary to add several so-called "extra-logical" features.

3 Successes and Failures of Mathematical Logic

Also successes and failures of mathematical logic are assessed by Kreisel with respect to the foundational schools. Successes, in his view [1967, pp. 202–203], include the notions of mechanical process and logical inference.

Now, the former is useful to state precisely the formalist position, in particular, as stressed by Gödel [1986a, p. 369], the concept of formal system, but is of little use in establishing a theory of real computation, e.g. concurrent computation. For example, the notion of mechanical process is inadequate to dealing with the computations performed by operating systems, database servers etc. In such computations the system continually accepts requests from its environment and reacts to such requests changing state (e.g. storing new data in a database or deleting previously available data), which affects the reactions to future requests. In general these computations are not intended to terminate at any time: the behaviour of the system is best explained in terms of conceptually infinite sequences of state transitions.

On the other hand the notion of logical inference, being analysed in terms of mechanical rules in accordance with the formalist position, is of little use in developing a theory of real reasoning, in particular of the kind of reasoning actually used in mathematical practice that has little to do with the reasoning used in mathematical textbooks: at best a "rational reconstruction" intended only for pedagogical use. Indeed it has been argued by Etchemendy [1990, pp. 156–159] that the notion of logical inference is inadequate even to the more limited aim of analyzing the intuitive notion of logical consequence.

In Kreisel's view [1967, p. 203], failures of mathematical logic primarily include its incapability of building up on its own main results, such as Gödel's completeness and incompleteness theorems. Now, again such results essentially concern the notion of formal system. Kreisel does not seem to see the main failure of mathematical logic in its incapability of developing a realistic theory of computation and reasoning. Even his critical remarks about the analyses provided by mathematical logic are not sharp enough to determine a definite break with the toy projects of most mathematical logic.

It is true that Kreisel [1967, p. 203] stresses that, since Gödel's incompleteness theorem entails that the known axioms of arithmetic or set theory leave many problems formally undecided, the process that led to the known axioms is relevant to the progress of mathematics and should have been taken more seriously after Gödel's discovery, while on the contrary most mathematical logicians took it less seriously. However his analysis of such a process does not go beyond the slogan of the so-called "informal rigour": slogan, because informal rigour has little to do with the actual process that led to the known axioms. For example, even in the case of one of Kreisel's [1967a, pp. 143–145] favourite examples of informal rigour - the alleged derivation of Zermelo's axioms from the description of the cumulative type structure - the actual process must have been quite different if Zermelo formulated the cumulative type structure only in [1930], twenty two years after stating his axioms. Kreisel [1970, p. 502] even claims that, in formulating the cumulative type structure, Zermelo was helped very little by the many pages of technical papers that had appeared since 1908. This seems doubtful in view of the fact that a definite anticipation of the cumulative type structure can be found, e.g. in Mirimanov [1917].

Moreover Kreisel's notion of informal rigour seems to involve identifying the process that led to the known axioms simply with the process by which rules and definitions are obtained, analyzing given intuitive notions and putting down their properties [1967a, p. 138]. This neglects the fact that what is most important to the progress of mathematics is not so much how rules and definitions are obtained from intuitive notions as how the relevant intuitive notions are chosen among several possible alternatives.

4 The Basic Assumption of Mathematical Logic

As I have argued at length in [1993], all work in mathematical logic and the foundations of mathematics depends on the following basic assumption:

> The method of mathematics is the axiomatic method and the object of mathematical logic is the study of the axiomatic method.

Such an assumption is also made by Kreisel. Indeed, on the one hand, in [1988, p. 166] he claims that, after the discovery in the last century of the so-called abstract reasoning, which involves expressing in axiomatic terms what we feel to be essential about an argument, we have never looked back. On the other hand, in Kreisel and Krivine [1967, p. v] he explicitly identifies the object of mathematical logic with the study of the axiomatic method. In his view, model theory is intended to present the principles of the axiomatic method in set theoretic terms while proof theory is intended to present them in combinatorial terms.

Clearly the above assumption affects the whole business of the role of mathematical logic in the philosophy of mathematics. For, if the method of mathematics is to be identified with the axiomatic method, then several questions concerning the nature of mathematics are simply left out. On the one hand, all problems concerning mathematical change are neglected. As stressed by Kitcher [1983, p. 149], the latter include the question why mathematicians propound different statements at different times; why they abandon certain forms of language for other ones; why certain mathematical problems wax and wane in importance; why standards and styles of proof are modified. On the other hand, all logical questions that would be significant from the viewpoint of an approach to the method of mathematics different from the axiomatic method, such as the already mentioned problem of a logic of mathematical discovery, are excluded from the domain of mathematical logic as being extra-logical. All mathematical logicians have to say about the mathematical method and the logic of mathematical discovery seems to be expressed by Braithwaite's [1953, p. 352] statement, apparently not meant as a joke, that in the axiomatic method we start from the beginning and go on to the end, both logically and epistemologically.

Indeed, all the above questions are neglected or played down in Kreisel's writings. Thus in [1984, pp. 82–83] all he has to say about Lakatos [1976] is that Lakatos has a relativistic concept of proof and mixes up the fact that the literature contained some errors and cases where the original proof was adequate for the old questions but not for the new questions. In [1988, p. 167] he claims that the process of discovering a new axiom to prove some given result is often less demanding than discovering a proof of it from given axioms and moreover has a similar flavour, which patently conflicts with mathematical experience: just think of the frustrated attempts to settle Cantor's continuum hypothesis. Moreover, whereas the process of discovering proofs from given axioms can be mechanized in an absolutely trivial way by means of the so-called British Museum algorithm (Newell, Shaw and Simon [1983, p. 55]), the process of discovering new axioms cannot be so mechanized.

While, contrary to Wittgenstein's [1978, V.24] claim, it would be unrealistic to consider the basic assumption of mathematical logic as having had a pernicious influence on mathematics, it seems fair to say that it has had a pernicious influence on our picture of both mathematics and logic. For, it has led to the one-sided "logical" view of the nature of mathematics, sharply criticized e.g. by Poincaré [1952], Wittgenstein [Wi] or Lakatos [1976], which caused the regrettable neglect of several important questions concerning the nature of mathematics.

Of course alternative approaches to both the mathematical method and the object of logic are possible, involving different ways of conceiving logic, philosophy of mathematics and their mutual relations. Having recently outlined one such approach in [1993a], [1995] I will not make any new proposal here, confining myself to briefly pointing out some of the difficulties arising from the basic assumption of mathematical logic. Indeed, such an assumption is crucial in the present context because it is the critical issue in assessing Kreisel's [1967] claims concerning the contribution of mathematical logic to the philosophy of mathematics.

5 The "Grand Theory" Version

While one tends to consider the axiomatic method as a single method, as a matter of fact it has several forms. An extreme one is the "grand theory" version, according to which a single theory is to be used to develop the whole of mathematics. Such a view has been implemented through the universal systems developed by Frege, Russell and Zermelo at the turn of the century.

Through such systems, Frege [1984, p. 235] wanted to fix the fundamental axioms upon which the whole of mathematics rests; Russell (Whitehead and Russell [1910, p. 2]) wanted to attain the complete enumeration of all ideas and steps of reasoning employed in mathematics; Zermelo [1967, p. 200] wanted to develop the logical foundations of all of arithmetic and analysis. Their basic aim was reductive: to show that all mathematical notions and principles could be reduced to the primitive notions and principles of the universal system. Indeed, their reductionism was of a very strong kind: as Quine [1960, pp. 265–266] points out, there is an important sense in which physicalism may be said to be less clearly reductive than the grand theory version.

The assumption behind the universal systems of Frege, Russell and Zermelo is that there is just one domain, a fixed and all-embracing universe that comprehends everything about which there can be any discourse. As stressed by van Heijenoort [1985, p. 79], the universes of Frege, Russell or Zermelo should not be mistaken for the universes of set-theoretical semantics considered in current mathematical logic. While in the latter one considers several alternative universes, each universe including all we want to consider at a certain stage in a certain context, the universe of the grand theory version is not just "one" universe: it is "the" universe, consisting of all that there is.

This entails that, strictly speaking, in the grand theory version no metasystematic question can be raised about the universal system: raising such a question would involve getting outside the system, which would be impossible because the latter is the universal system of all mathematics. This peculiar character of the grand theory version is stressed by Wittgenstein [1961, 5.61] when he states that the limits of the world are also the limits of logic that for him consists of Russell's universal system. From this viewpoint one cannot say in logic that something is in the world whereas something else is not; otherwise logic would have to get outside the limits of the world and only in that way one would be able to consider such limits from the other side also.

In particular the rejection of metasystematic questions includes completeness. Completeness in the metasystematic sense is replaced by completeness in an empirical sense, i.e. the question whether the modes of reasoning provided by the axioms of the universal system can exhaust the intuitive modes of reasoning actually used by mathematicians. Thus for Frege [1984, p. 235], in order to test whether a list of axioms is complete, we have to try and derive from them all the proofs of the branch of learning to which they relate using only purely logical laws. For Russell (Whitehead and Russell [1910, p. v]) the justification for the axioms of the universal system must be inductive, i.e. it must lie in the fact that the theory in question enables us to derive ordinary mathematics. For Zermelo [1930, p. 200] the justification for the axioms of set theory must lie in the fact that they enable us to derive the results of set theory as it is historically given, i.e. as created by Cantor and Dedekind.

Notwithstanding its rejection of any metasystematic question, it seems fair to say that the grand theory version has been conclusively refuted by Gödel's incompleteness theorem, which entails that each axiom system for type theory or set theory is inadequate. Indeed, in his original paper Gödel [1986, p. 145] explicitly refers to the universal systems of Russell, Zermelo and their variants. In spite of its spectacular failure, however, the grand theory version seems to have preserved its attraction, even on Gödel. This, for instance, appears from his suggestion [1990a, p. 151] that the contemplation of a formalism gives rise to new axioms which are as evident and justified as those with which one started; that this process of extension can be iterated into the transfinite; that all these steps can be collected together in some non-constructive way; that for such a concept of demonstrability some completeness theorem can hold. Gödel [1990, pp. 140–141] even goes so far as expressing the hope that Leibniz's project of a *characteristica universalis* might, after all, be realized. Such a hope must be seen against the background of his ideal of logic as a universal science (Gödel, [1990, p. 119]), i.e. a science prior to all others, containing the ideas and principles underlying all sciences. Apart from Gödel's utopian dreams, however, the claims of the grand theory version appear unrealistic and at present do not seem to have many supporters.

Quite independently of its refutation by Gödel's incompleteness theorem, the grand theory version has some intrinsic defects. I will only mention a couple of them.

First of all, proposing a single theory to develop the whole of mathematics is purely ideological and does not help in practice. E.g., even when number theory is reduced to some universal system, we do not change our manner of doing it. Of course, we are aware that there is a sense in which our numerical calculations and number-theoretical proofs could be translated into the universal system, but doing number theory is still different from making calculations and proving number-theoretic results in that system: e.g., it would be ridiculous to make calculations using the encoding of natural numbers in terms of the primitive notions of the universal system.

Moreover, there is a good reason for proving number-theoretic results in number-theory instead of deriving them from the axioms of the universal system through such an encoding: proofs are significantly less complex, not in the rough sense that they contain less symbols, but in the more relevant sense that they are more easily grasped.

Secondly, the grand theory version does not explain why, among all possible results of the universal system, we choose only those that are actually used in current mathematical practice, or why we find certain results more interesting than others. Actually, in the grand theory version only the axioms of the universal system matter. All theorems are implicitly contained in them because they are ultimately obtained by repeated applications of fairly elementary logical rules, so elementary that most of them were already known to Chrysippus. Indeed, as already mentioned, there is an absolutely trivial mechanical procedure that, given sufficient time and space, will generate all theorems using such rules.

6 The "Big Theory" Version

A more moderate form of the axiomatic method is the "big theory" version, according to which a single theory will be used to develop not the whole of mathematics but only the whole of applicable mathematics, i.e. that part of everyday mathematics that finds its application in the other sciences. A typical implementation of such a version is provided by the systems for reverse mathematics (Simpson [1985]), i.e. subsystems of second order arithmetic conservative over primitive recursive arithmetic with respect to purely universal sentences: their intended use is to develop significant portions of analysis and algebra.

According to supporters of reverse mathematics, such as Simpson [1988, p. 362], the main advantage of such systems depends on the fact that, while the first few levels of the cumulative type structure bear some resemblance to external reality, the rest are a huge extrapolation based on a crude model of abstract thought processes. Now, it seems far-fetched to maintain that the first few levels of the cumulative type structure bear any resemblance to external reality: e.g., Hilbert [1967, p. 376] would have rejected such a view on the ground that the infinite is not to be found anywhere in reality. Plausibility apart, the big theory version appears just an attempt to save the savable of the grand theory version, stopping the leaks produced by Gödel's incompleteness theorem, and as such inherits all of its intrinsic defects.

Simpson [1988, pp. 359–360] claims that encoding mathematical objects – such as real numbers, continuous functions, complete separable metric spaces, etc. – as sets of integers in an admittedly arbitrary way does not vitiate the assumption that the systems for reverse mathematics reflect mathematical practice. This claim shows to what extent his approach is purely ideological and does not take into account the problems involved in actually developing mathematical practice within a single system. Such problems are not merely practical but affect the plausibility of the whole picture of mathematics provided by the big theory version.

7 The "Little Theory" Version

A further version of the axiomatic method is the "little theory" version, according to which a number of axiomatic theories will be used in the course of developing mathematics. The motivation commonly provided for this version is that it allows to separate the role of different axiom systems and to determine the scope of the conclusions derivable from each axiom system. Such a view was held by Hilbert in [1902, Introduction], was implemented by him in [1902, §§15, 17, 24–26, 28–29] and has been recently proposed in Farmer, Guttman and Thayer [1992] as useful to provide a more effective approach to mechanized reasoning, in particular to automated theorem proving.

The central notion of this version of the axiomatic method is that of interpretation of a theory $\mathcal{T}$ into another one $\mathcal{T}'$. As Kreisel [1967, pp. 264–265] points out, an interpretation does no more than connect a strictly limited use of notions, described by the axioms of $\mathcal{T}$, to another limited use of the same or other notions, described by $\mathcal{T}'$. If the interpretation satisfies some obvious conditions, this is enough 1) to establish the consistency of $\mathcal{T}$ in terms of the consistency of $\mathcal{T}'$; 2) to reuse theorems of $\mathcal{T}$ in $\mathcal{T}'$; and 3) to select a subdomain of an arbitrary model of $\mathcal{T}'$ together with distinguished values that determine a model of $\mathcal{T}$.

In Farmer, Guttman and Thayer [1992] the above features are said to represent the main advantage of the little theory approach over other approaches. In particular, the importance of 2) is especially stressed [1992, p. 272]: e.g. a theorem of the theory of groups will translate into a theorem, say, of the theory of fields by observing that multiplication over the nonzero elements has the structure of a group. Interpretations allow to transfer a theorem from the theory in which it was originally proved to other theories. Thus, trough the notion of interpretation of a theory into another one, mathematical reasoning is not concentrated in a single system but is distributed over a network of theories linked by interpretations. From this viewpoint, 1) and 3) are also useful [1992, pp. 576–577]. For, 1) guarantees that a theory $\mathcal{T}$ is consistent if $\mathcal{T}'$ is a well-established theory trusted to be consistent. On the other hand, if one has a conception of what sort of structures ought to satisfy $\mathcal{T}$, 3) allows to isolate a structure satisfying $\mathcal{T}$ within any structure satisfying $\mathcal{T}'$. If such structures cannot be made to look as one expected, then one has reason to suppose that the formulation of $\mathcal{T}$ could be wrong.

Such claims about the virtues of the little theory version seem hardly convincing. According to this view of the mathematical method, mathematicians first state definitions and axioms for a given theory, then prove theorems in that theory by purely logical rules and only afterwards transfer results from one theory to another via interpretations. This picture seems a parody of mathematical practice. As Longo [1991, pp. 119–120] points out, instead of first stating definitions and axioms, mathematicians first hint results either by wild analogies or by deep connections between different areas. A variety of languages are used in the very same proof to establish new connections and propose new ideas. Really deep proofs rarely fit into a single axiomatic framework or use a single language. Only later on an incomplete and only partly axiomatic framework is established to clean up deductions. And when a complete, ad hoc, axiomatic framework and a single language is established for a certain theorem, then the result is already too old and has little to say about current research.

8 Mathematical Logicians and the Axiomatic Method

It is psychologically interesting, although somewhat shocking, to see that, whenever the defects of the axiomatic method are pointed out to them, mathematical logicians lose their self-control.

This is understandable in view of the fact that Hilbert [1935, p. 152] hailed the axiomatic method as the decisive weapon used by Zermelo to rescue mathematics from the precarious situation in which it had fallen because of the attacks of Kronecker and Poincar who, under the pretext of paradoxes, denied set theory the right to exist. Indeed Hilbert [1935a,

p. 160] intended to use the axiomatic method to restore mathematics to its old reputation of unquestionable truth. Moreover in [1935, p. 156] he claimed that whatever can be the object of scientific thought falls under the axiomatic method and hence under mathematics, and that, in the sign of the axiomatic method, mathematics is called to play a leading role in science. In [1935a, p. 161] he even went so far as saying that to proceed axiomatically is nothing but to think consciously. In his view [1967a, p. 475] the rules of the axiomatic systems of mathematical logic express the technique of our thinking and are the rules according to which our thinking actually proceeds.

Hilbert [1935a, pp. 159–160] was obsessed by the idea that mathematics should be defended from the insidious attacks of Kronecker, Brouwer and Weyl. As it appears from [1935b, p. 387], he was also troubled by contemporary philosophers who speculated on the intrinsic limitations of our knowledge, to whom he opposed his motto: "We must know, we will know". He viewed the axiomatic method as the main defense weapon against such enemies. Contemporary mathematical logicians, far from having been made wiser by Gödel's incompleteness results, are subject to similar obsessions and often use equally emphatic tones.

A striking example is provided by Simpson [1988, pp. 357–358] who laments that the validity of mathematics is under siege because of the attacks of people like Wigner, Kline and even Russell; that unfortunately none of the existing foundational schools – logicism, formalism, intuitionism, platonism – offers any real defense against such attacks; that the latter are part of a general assault against reason; that mathematicians and logicians should feel the duty to get on with the task of defending their discipline. Simpson [1988, p. 362] also stresses that the need to defend the integrity of mathematics, far from having abated, has been made more urgent than ever by Gödel's incompleteness theorem; that the latter unfortunately supplied heavy artillery for all would-be assailants of mathematics and indeed is quoted by Kline with monotonous repetition and devastating effect; that the assault on mathematics rages as never before.

It is somewhat disconcerting to see Simpson in the Eighties using the same unrestrained tones as Hilbert in the Twenties.

9 Kreisel and the Axiomatic Method

Such emphatic tones are not to be found in Kreisel because they are alien to his style. Style apart, however, his attachment to the axiomatic method does not seems to be less passionate than Hilbert's or Simpson's. While criticizing the grand theory version, Kreisel is all out to defend a more moderate version of the axiomatic method, as it appears from the following two examples.

First, Kreisel [1973, pp. 602–603] sharply criticizes the empirical approach to completeness envisaged by the grand theory version. While maintaining that the defects of the latter most talked about can in fact be fairly easily corrected, he claims that the real difficulties of the grand theory version are closely connected with its empirical approach to completeness. In his view, instead of trying to analyze the concept of logical validity and to show that the rules of the system generate all logically valid formulas, the grand theory version derives a lot of such formulas. Instead of looking for some global features of ordinary mathematical concepts, for example, that any proposition about some specific objects such as the natural numbers is either true or is false, and instead of comparing these features with the properties of the universal system, it derives a lot of arithmetic propositions from the universal system. Contrary to Kreisel's view, it can be argued (Cellucci [1992]) that the use of an empirical approach to completeness, instead of being a defect, shows that, at least originally, the grand theory version had an open view of logic which, in the light of Gödel's incompleteness theorem, was much wiser and, at least potentially, more fruitful than Hilbert's subsequent closed view.

To a grand theory version providing no metasystematic theory of the universal system, no criticism from without, Kreisel [1973, p. 603] opposes a more moderate, although somewhat indeterminate, version of the axiomatic method whose main successes are Gödel's completeness and incompleteness theorems. As he states in [1988, pp. 162–163], while the grand theory version follows the tradition of natural history which is content with a compact description of data that happen to catch our attention, such a more moderate version, essentially originating from Hilbert, e.g. in [1935], pays attention to global mathematical properties of axiomatic systems such as completeness or incompleteness. Kreisel points out that this side of Hilbert's view of the axiomatic method got lost in Hilbert's later emphasis, for example, on real and ideal elements. Furthermore, he acknowledges that completeness would not be generally plausible if thought processes were the main object of study. In his opinion, however, this does not affect the plausibility of the basic idea.

Moreover Kreisel [1973, p. 604] acknowledges that Gödel's incompleteness theorem raises problems that he would have hoped to avoid. Specifically, as long as all true propositions are derivable from the axioms of the universal system, there is no question of determining whether that axiomatization is correct. But, in his view, this does not discredit the idea of a correct axiomatization: it only means that the choice of the axiomatization will be sensitive to its details, and such a sensitivity provides a rational means for finding a, or even the, correct axiomatization.

Kreisel [1988, p. 164] even tries to defuse the difficulty caused to the axiomatic method by Gödel's incompleteness theorem by playing down its sensational character. He claims that the completeness and the incompleteness theorem cannot both be sensational: if completeness was a sensation, then incompleteness is not. As evidence for such claim he mentions the fact that, two years before Gödel's incompleteness result, the number theorists Siegel and Weil discounted the possibility of a complete axiomatic system for diophantine equations. His aim seems to be to suggest that incompleteness is a predictable, even expected feature of axiomatic systems and hence does not affect the plausibility of the axiomatic method, but this seems to contradict his critical remarks on the empirical approach to completeness of the grand theory version.

The second example is provided by Kreisel's [1978, p. 86] patent resentment at Bourbaki's [1962, pp. 45–46] implicit criticism that, if all axiomatizations were categorical like Peano's second order axioms for number theory, then charging the axiomatic method with sterility would be fully justified. He interprets this criticism as a reaction against the preoccupation of the foundational schools with analysing informal notions. He acknowledges that some of such analyses have been of little use: e.g. one could hardly maintain that Gauss's *Disquisitiones arithmeticae* would have been better if Gauss had started with Peano's axioms. He also acknowledges that, in practice, the informal notions we start with often turn out to be unmanageable or unrewarding and that it is simply better to axiomatize which properties of such notions have been used for some striking conclusion.

Kreisel, however, considers the charge of sterility against categorical second order axioms a shallow one because in his view, by neglecting, say, Peano's categorical second order axioms, one loses an explanation of the choice of the first order axioms for number theory, which are obtained from Peano's axioms by considering induction not for arbitrary sets of integers but only for sets defined by first order formulas in the language of arithmetic. Here Kreisel seems to forget his own remark [1967a, p. 148] that the first order axioms resulting from Peano's axioms yield a weak decidable system – the theory of successor arithmetic – inadequate for formulating current informal number theory. Other axioms must be added, such as the recursion equations for addition and multiplication, whose choice requires additional considerations.

Moreover Kreisel seems to suggests that, while the passage to first order axioms raises problems, categorical second order axioms are totally unproblematic. Indeed in [1978, p. 81] he claims that the latter provide a precise and convincing analysis of the corresponding notions. Such a claim seems to neglect the basic distinction between characterizing a structure and axiomatizing it. While the aim of characterizing a structure is to distinguish that structure from other structures, the aim of axiomatizing it is to organize the truths of that structure as a deductive system. Now, by Gödel's [1986] original version of the incompleteness theorem for Peano's categorical second order system, a categorical characterization of a given structure need not provide a complete axiomatization of it. Moreover, for each categorical characterization that happens to be incomplete, there are infinitely many better categorical characterizations: better, in the sense of allowing to derive additional theorems not derivable from the first characterization. Thus no single categorical characterization can ultimately be adequate.

It might be thought that any categorical characterization of a given structure would allow to derive all "normal" truths of that structure, as distinguished from the "pathological" ones such as Gödel sentences. However, as pointed out by Corcoran [1980, p. 204], there are categorical characterizations of a given structure from which some of the most elementary normal truths of that structure are not derivable: for example, there are categorical characterizations of the set of all integers from which one cannot derive that zero is not a successor or that addition is commutative. This shows that there exists a sharp difference between characterizing a structure and axiomatizing it, hence categoricity cannot be a hallmark of good axiomatization.

Such a difference is perhaps connected with Kant's [1968, pp. 93–96] distinction between defining a mathematical concept, e.g. a circle, and comprehending its properties: we can define many mathematical concepts without fully comprehending their properties. This also helps to focus Feferman's [1981, p. 320] claim that the concept of natural number is a crystal-clear mathematical concept and indeed that, if anything is a candidate for being such, this is it. Now, the concept of natural number is not clear enough to allow a best characterization of it.

In any case, independently of the plausibility of Kreisel's claims about completeness and incompleteness or about categorical second order axiomatizations, the very fact that he makes such claims provides clear evidence that he is entirely committed to supporting the view that the method of mathematics is the axiomatic method, a view which seems no longer plausible after Gödel's incompleteness result.

10 Incidentals and Essentials

The axiomatic method, especially in the extreme form of the grand theory version, presents mathematics as a profoundly uniform subject based on an allegedly unique method. On the contrary, as Wittgenstein [1978, III.46] forcefully stresses, mathematics is a motley of techniques of proof which often are loosely related.

A typical mistake of the supporters of the axiomatic method is to consider what was only a trend of certain parts of mathematics in the late nineteenth and early twentieth centuries as a permanent feature of mathematics, and indeed as its essence. Just think of Hilbert's already mentioned claims that whatever can be the object of scientific thought falls under the axiomatic method; that to proceed axiomatically is nothing but to think consciously; that the formal rules of mathematical logic express the technique of our thinking and are the rules according to which it actually proceeds. Or just think of Kreisel's claim that after the discovery in the last century of abstract reasoning, which involves expressing in axiomatic terms what we feel to be essential about an argument, we have never looked back. On the contrary, as stressed e.g. in Rota [1991, p. 166], the method of mathematics has changed many times in the past and it would be rash to assume that it will not change again.

The basic presupposition for any reasonable philosophy of mathematics ought to be that there is no feature of mathematics that can be considered as permanent, each stage of development of mathematics having its own features. E.g., the Appel-Haken computer based proof of the Four Colour Problem is perhaps more closely related to empirical pre-euclidean proofs than to late nineteenth and early twentieth centuries axiomatic proofs. As Wittgenstein [1978, V.52] points out, even centuries ago a philosophy of mathematics was possible, a philosophy of what mathematics was then. Indeed centuries ago there were philosophies of mathematics incomparably more articulated than those produced by the foundational schools. For example Plato and Aristotle developed views of mathematical knowledge where the latter is seen as a part of scientific knowledge, forming an integrated system with it, contrary to the foundational schools where theoretical mathematics is treated as a secluded subject, essentially unrelated to the rest of scientific knowledge. That a genuine philosophy of mathematics should integrate mathematics within a general theory of scientific knowledge and a complete account of rational inference is stressed, e.g., by Kitcher [1983, p. 227].

It is not accidental that the most passionate supporters of the view that mathematics is a profoundly uniform subject based on the axiomatic method should be found among mathematical logicians rather than among mathematicians. While for the latter the axiomatic method is just a tool, for mathematical logicians it is the essence of their subject: without the axiomatic method there would be no mathematical logic. Thus for mathematical logicians to defend the axiomatic method is to defend their very right to be on the map.

The axiomatic method provides an extremely simplified view of mathematical knowledge. From the viewpoint of mathematical logic this has the advantage that the method can be easily approached in terms of its straightforward notions, such as set-theoretical consequence, formal system or mechanical process. On the other hand, such notions are patently inadequate to dealing with the variegated features of past and contemporary mathematics. On account of this it appears amusing that Kreisel should claim that philosophy deals only with immature notions and that preoccupation with them draws attention away from genuinely rewarding questions. As a matter of fact just the opposite is true. As already mentioned, both Plato and Aristotle developed a philosophy of the mathematics of their time incomparably more articulated than the limited view of contemporary mathematics provided by mathematical logic. In particular Aristotle – the first theorizer of the axiomatic method – was wise enough not to claim that the latter is the method of mathematics: he only claimed that it is a pedagogical method. Contrary to his modern followers, he did not confuse the presentation of mathematics with the process of mathematical discovery. It is somewhat ironical that the misconceptions of mathematical logicians about the nature of mathematics should be rectified by historians of mathematics like Crowe [1988] rather than by professional philosophers of mathematics.

A necessary non-sufficient condition for the success of mathematical logic would be to show that the whole of mathematics can be constructed as a single axiomatic system. In view of the fact that, because of Gödel's incompleteness theorem, this task cannot be carried out by the grand theory version, the big theory version undertakes to carry out it not with respect to all of mathematics but only with respect to scientifically applicable mathematics, that is, with respect to that part of everyday mathematics that finds its applications in the other sciences. But the price that must be paid for such an attempt is that it involves a fairly heavy encoding machinery which is both arbitrary and burdensome. Moreover, the machinery is very artificial, as it should be expected in view of the fact that it is not intended to help the working mathematician but only to achieve a purely ideological aim, totally unrelated to mathematical practice. Similar reservations apply to the little theories version, where the encoding machinery of the big theory version is replaced by the equally artificial and limited notion of interpretation of a theory into another one.

In any case, showing that the whole of mathematics could be constructed as a single axiomatic system would not have been sufficient for mathematical logic to be successful. As I have already mentioned, there would have remained, on the one hand, all questions concerning mathematical change and, on the other hand, all logical questions, such as that of a logic of mathematical discovery, that are excluded from the domain of mathematical logic but are essential to the understanding of mathematical experience. As Wittgenstein [1978, V.25] once said, while mathematical logicians believe that, by neglecting such questions, they keep mum only about incidentals, what they actually do is to confine themselves to the verbal expression of mathematical results, which is a mere shadow, and to keep mum about essentials.

11 Epilogue

The most important result of mathematical logic – Gödel's incompleteness theorem – suggests that the axiomatic method cannot be the method of mathematics, that the basic concepts of mathematical logic are inadequate and, generally, that the basic assumption on which the entire mathematical logic depends is untenable. Ten years after Gödel's discovery Post [1965, p. 345] expressed his continuing amazement that current views on the nature of mathematics were affected by it only to the point of seeing the need of many formal systems instead of a universal one. On the contrary, he considered inevitable that Gödel's result would lead to a reversal of the entire axiomatic trend of the late nineteenth and early twentieth centuries that would make the latter only a phase of mathematical thinking. Post [1965, footnote 12] also anticipated that, with the bubble of mathematical logic as a universal logic machine finally burst, mathematical logic would become the indisputable means for revealing and developing its own limitations, and that it would therefore deserve to be called "mathematics become self-conscious".

Sixty years after Gödel's discovery I must still express my continuing amazement that the situation has not much changed. While mathematical logic has continued to reveal its own limitations, it can hardly be considered as mathematics become self-conscious. Mathematical logicians still seem to hold that Gödel's incompleteness theorem affects them only to the point of seeing the need of many formal systems instead of a universal one, and surely are not willing to concede that it should lead to a reversal of the entire axiomatic trend of the late nineteenth and early twentieth centuries.

The logical conclusion to be drawn from Gödel's result should be:

1. negatively, that one must give up the basic assumption of mathematical logic that the method of mathematics is the axiomatic method and that the object of logic is the study of the axiomatic method;

2. positively, that a more sophisticated view of the mathematical method and a more sophisticated approach to logic is required.

But mathematical logicians are most illogical people: even when they agree that the axiomatic method provides a distorted picture of mathematical practice, they go on as before as if they had forgotten their own criticisms or the latter did not concern them. On account of past experience it would be daring to trust in their resipiscence. But, if the latter will not promptly intervene, it is only too easy to foresee that they will be doomed to the fate traditional (Aristotelian) logic eventually had to meet: increasing irrelevance followed by inevitable extinction.

References

Barwise, K.J.

[1988] Review of the Ω-bibliography of mathematical logic, *Bulletin of the American Mathematical Society*, 19 (1988) 524–528.

Bourbaki, N.

[1962] L'architecture des mathématiques, in F. Le Lionnais (ed.), *Les grands courants de la pensée mathématique*, Blanchard, Paris, 1962, pp. 35–47.

Braithwaite, R.B.

[1953] *Scientific explanation*, Cambridge University Press, Cambridge, 1953.

Brouwer, L.E.J.

[1975] The effect of intuitionism on the classical algebra of logic, in *Collected works*, vol. I, ed. A. Heyting, North-Holland Publ. Co, Amsterdam, 1975, pp. 551–554, 611.

Cellucci, C.

[1992] Gödel's incompleteness theorem and the philosophy of open systems, in D. Miéville (ed.), *Kurt Gödel*, CRS, Travaux de logique 7, Université de Neuchâtel, 1992, pp. 103–127.

[1993] Gli scopi della logica matematica, in *Peano e i fondamenti della matematica*, Accademia Nazionale di Scienze Lettere e Arti, Modena 1993, pp. 73–138.

[1993a] From closed to open systems, in J. Czermak (ed.), *Philosophy of Mathematics: Proceedings of the 15th International Wittgenstein Symposium*, Hölder-Pichler-Tempsky, Wien, 1993, pp. 206–220.

[1995] The growth of mathematical knowledge: an open world view, to appear in *Proceedings of the conference Growth of Mathematical Knowledge*, Penn State, 1995.

Corcoran, J.

[1980] Categoricity, *History and Philosophy of Logic*, 1 (1980) 187–207.

Crowe, M.J.

[1988] Ten misconceptions about mathematics and its history, in W. Aspray and P. Kitcher (eds.), *History and philosophy of modern mathematics*, Minnesota Studies in the Philosophy of Science, vol. XI, University of Minnesota Press, Minneapolis, 1988, pp. 260–277.

Dieudonné, J.

[1981] Logica e matematica nel 1980, in P. Rossi (ed.), *La nuova ragione*, Il Mulino, Bologna, 1981, pp. 15–25.

Etchemendy, J.

[1990] *The concept of logical consequence*, Harvard University Press, Cambridge, Mass., 1990.

Farmer, W.M., Guttman J.D., and Thayer, F.J.

[1992] Little theories, in D. Kapur (ed.), *Automated deduction*, CADE-11, Springer-Verlag, Berlin, 1992, pp. 567–581.

Feferman, S.

[1981] The logic of mathematical discovery vs. the logical structure of mathematics, in P.D. Asquith and I. Hacking (eds.), *PSA 1978*, vol. II, Philosophy of Science Association, East Lansing, MI 1981, pp. 309–327.

Frege, G.

[1967] Begriffsschrift, a formula language, modeled upon that of arithmetic, for pure thought, in J. van Heijenoort (ed.), *From Frege to Gödel. A source book in mathematical logic, 1879-1931*, Harvard University Press, Cambridge, Mass., 1967, pp. 5–82.

[1984] On Mr. Peano's conceptual notation and my own, in B. McGuinness (ed.), *Collected papers on mathematics, logic, and philosophy*, Blackwell, Oxford, 1984, pp. 234–248.

Gödel, K.

[1986] On formally undecidable propositions of Principia Mathematica and related systems I, in S. Feferman et al. (eds.), *Collected works*, vol. I, Oxford University Press, Oxford, 1986, pp. 144–195.

[1986a] Gödel, On undecidable propositions of formal mathematical systems, in S. Feferman et al. (eds.), *Collected works*, vol. I, Oxford University Press, Oxford, 1986, pp. 346–371.

[1990] Russell's mathematical logic, in S. Feferman et al. (eds.), *Collected works*, vol. II, Oxford University Press, Oxford, 1990, pp. 119–141.

[1990a] Remarks before the Princeton bicentennial conference on problems in mathematics, in S. Feferman et al. (eds.), *Collected works*, vol. II, Oxford University Press, Oxford, 1990, pp. 150–153.

Hilbert, D.

[1902] *The foundations of geometry*, Open Court, Chicago, 1902.

[1935] Axiomatisches Denken, in *Gesammelte Abhandlungen*, vol. III, Springer-Verlag, Berlin, 1935, pp. 146–156.

[1935a] Neubegründung der Mathematik. Erste Mitteilung, in *Gesammelte Abhandlungen*, vol. III, Springer-Verlag, Berlin, 1935, pp. 157–177.

[1935b] Naturerkennen und Logik, in *Gesammelte Abhandlungen*, vol. III, Springer-Verlag, Berlin, 1935, pp. 378–387.

[1967] On the infinite, in J. van Heijenoort (ed.), *From Frege to Gödel. A source book in mathematical logic, 1879-1931*, Harvard University Press, Cambridge, Mass., 1967, pp. 369–392.

[1967a] The foundations of mathematics, in J. van Heijenoort (ed.), *From Frege to Gödel. A source book in mathematical logic, 1879-1931*, Harvard University Press, Cambridge, Mass., 1967, pp. 464–479.

Kant, I.

[1968] Der einzig mögliche Beweisgrund zu einer Demonstration des Daseins Gottes, in *Kants Werke. Akademie Textausgabe*, vol. II, de Gruyter, Berlin, 1968, pp. 63–164.

Kitcher, P.

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

Kreisel, G.

[1951] On the interpretation of non-finitist proofs, part I, *The Journal of Symbolic Logic*, 16 (1951) 241–267

[1952] On the interpretation of non-finitist proofs, part II, *The Journal of Symbolic Logic*, 17 (1952) 43–58.

[1967] Mathematical logic: what has it done for the philosophy of mathematics?, in R. Schoenman (ed.), *Bertrand Russell, Philosopher of the Century*, Allen & Unwin, London, 1967, pp. 201–272.

[1967a] Informal rigour and completeness proofs, in I. Lakatos (ed.), *Problems in the philosophy of mathematics*, North-Holland, Amsterdam, 1967, pp. 138–171.

[1970] Principles of proof and ordinals implicit in given concepts, in J. Myhill et al. (eds.), *Intuitionism and proof theory*, North-Holland, Amsterdam, 1970, pp. 489–516.

[1973] Bertrand Arthur William Russell, Earl Russell 1872–1970, *Biographical Memoirs of Fellows of The Royal Society*, 19 (1973) 583–620.

[1977] From foundations to science: justifying and unwinding proofs, *Recueil des travaux de l'Institut Mathématique de Belgrade*, 2 (1977) 63–72.

[1978] Wittgenstein's lectures on the foundations of mathematics, Cambridge 1939, *Bulletin of the American Mathematical Society*, 84 (1978) 79–90.

[1984] Frege's foundations and intuitionistic logic, *The Monist*, 67 (1984) 72–91.

[1988] Review of Gödel's Collected Works, Volume I, *Notre Dame Journal of Formal Logic*, 29 (1988) 160–181.

Kreisel, G., and Krivine, J.L.

[1967] *Elements of mathematical logic*, North Holland, Amsterdam, 1967.

Lakatos, I.

[1976] *Proofs and refutations*, ed. J. Worrall and E. Zahar, Cambridge University Press, Cambridge, 1976.

Longo, G.

[1991] Notes on the foundations of mathematics and computer science, in G. Corsi and G. Sambin (eds.), *Nuovi problemi della logica e della filosofia della scienza*, vol. II, CLUEB, Bologna, 1991, pp. 117–127.

McCarthy, J.

[1963] A basis for a mathematical theory of computation, in P. Braffort and D. Hirschberg (eds.), *Computer programming and formal systems*, North Holland, Amsterdam, 1963.

Mirimanoff, D.

[1917] Les antinomies de Russell et de Burali-Forti et le probleme fondamental de la theorie des ensembles, *L'Enseignement Mathématique*, 19 (1917) 37–52.

Newell, A., Shaw, J.C., and Simon, H.A.

[1983] Empirical explorations with the logic theory machine: A case study in heuristics, in J. Siekmann and G. Wrightson (eds.), *Automation of reasoning*, vol. I, Springer-Verlag, Berlin, 1983, pp. 49–73.

Poincaré, H.

[1952] *Science and method*, Dover, New York, 1952.

Post, E.

[1965] Absolutely unsolvable problems and relatively undecidable propositions: account of an anticipation, in M. Davis (ed.), *The undecidable*, Raven Press, Hewlett, N.Y., 1965, pp. 340–433.

Prawitz, D.

[1993] Remarks on Hilbert's program for the foundation of mathematics, in G. Corsi et al. (eds.), *Bridging the gap: philosophy, mathematics, and physics*, Boston studies in the philosophy of science, vol. 140, Kluwer, Dordrecht, 1993, pp. 87–98.

Quine, W.V.

[1950] *Methods of logic*, Holt, Reinehart and Winston, New York, 1950.

[1960] *Word and object*, The MIT Press, Cambridge, Mass., 1960.

Rota, G.C.

[1991] The pernicious influence of mathematics upon philosophy, *Synthese*, 88 (1991) 165–178.

Simpson, S.G.

[1985] Reverse mathematics, in A. Nerode and R. Shore (eds.), *Recursion theory*, Proc. Symp. Pure Maths. 42, American Mathematical Society, Providence, R.I., 1985, pp. 461–471.

[1988] Partial realizations of Hilbert's program, *The Journal of Symbolic Logic*, 53 (1988) 349–363.

van Heijenoort, J.

[1985] Absolutism and relativism in logic, in *Selected essays*, Bibliopolis, Naples, 1985, pp. 75–83.

Whitehead, A.N., and Russell, B.

[1910] *Principia Mathematica*, vol. I, Cambridge University Press, Cambridge, 1910.

Wittgenstein, L.

[1961] *Tractatus Logico-philosophicus*, Routledge & Kegan Paul, London, 1961.

[1978] *Bemerkungen über die Grundlagen der Mathematik*, ed. G.H. von Wright et al., 2nd ed., Blackwell, Oxford, 1978.

Zermelo, E.

[1930] Über Grenzzahlen und Mengenbereiche: Neue Untersuchungen über die Grundlagen der Mengenlehre, *Fundamenta mathematicae*, 14 (1930) 29–47.

[1967] Investigations in the foundations of set theory I, in J. van Heijenoort (ed.), *From Frege to Gödel. A source book in mathematical logic, 1879-1931*, Harvard University Press, Cambridge, Mass., 1967, pp. 200–215.

Kreisel's Church

by Piergiorgio Odifreddi

> "Church's thesis has, within logic, a similar function to dogmas and doctrines within the Church. The faithful get excited at the cost of being ridiculous to outsiders."
>
> G.K. (24.XII.92)

In *Classical Recursion Theory* (Odifreddi [1989]) I dedicated 20 pages to a discussion of Church's Thesis. It was Kreisel (K. from now on) who alerted me at the subtleties of the subject, and at the insufficient treatments available in print. Suffering of a seemingly widespread illness, I understood only part of what he said or wrote; and even of what (I thought) I understood, I made use for my own purposes.[1] However, his name in boldface in a number of places in that discussion may

[1] As a related example of how things may come to be used in a (purposely) distorted way, I quoted on p. 2 of my book a parallel that K. had made in §4 of [1985], between Euclid's *Book X* (which classified irrationals by means of a notion of degree) and the theory of Turing degrees.

His purpose was to draw attention to the fact that the shift from the classification of *Book X* to the measures of irrationality in modern number theory (based on diophantine approximations) required an absolute level of imagination, while nothing in (classical) Recursion Theory "approaches the philosophical detachment from the original set-up that was so essential for progress in the parallel from number theory".

My purpose was to claim that Recursion Theory is part of classical mathematics, and I thus presented the parallel as an item of the subject's pedigree, claiming that "degrees were used for the purpose of a classification of reals already in Euclid's *Book X*".

K. has whipped me more than once for "squandering a cute quotation".

have created an illusion that I was reporting on his views: making him uncomfortable, and others confused.

I intend to attempt such a report here, letting as much as possible K.'s original words and formulations speak for themselves; in the hope not of satisfying him (an obviously impossible task), but of dispelling such a confusion.

Since K.'s range of interests (even in the limited area of concern to this article) is quite wide, the reader is advised to browse among the various topics, looking for ones of his or her own interest.

1 General Remarks

Observations about general aspects of Church's Thesis are scattered in K.'s papers and reviews, especially in:

- *Analysis of the general concept of computation* (1.(c) and 4.(c) of [1971]),
- *Principal distinctions* (II.(a) of [1972]),
- *Church's Thesis and the ideal of Informal Rigour* ([1987]).

The following unstructured selection somehow reflects the occasional character of those observations.

1.1 Informal rigour

K. states in [1987] that Church's Thesis is a candidate for informal rigour, "a venerable [2000-year-old] ideal in the broad tradition of analysing precisely common notions or, as one sometimes says, notions implicit in common reasoning [at a given time]". He repeatedly warns us, at the end of [1987], about the obvious risks involved in any such enterprise; specifically:

- First, "there is no end in sight to the possibilities of coherent and imaginative analysis in the tradition of informal rigour". What is in doubt is the adequacy of the common notions to be analysed, not only to the phenomena for which they are intended, but even to our practical knowledge of them.
- Second, "even if the contributions [to informal rigour in general, and Church's Thesis in particular] were more central than they are, the market would be limited by the background knowledge

> needed for more than an illusion of understanding. It is a hallmark of philosophical questions that they present themselves to those of us who do not have such knowledge (and even as not requiring any)".

It is then of no surprise to know that K. sees work about Church's Thesis, and more generally informal rigour, as an answer to the question: How to talk in the face of ignorance? Specifically, an answer taught by philosophy to those not satisfied with the easiest answer: stay silent.

Be that as it may, the common notion to be analysed here is of course effective computability, and work around it is a candidate not only for the *pursuit* of the ideal of informal rigour, but also for the *examination* of the pursuit itself. In other words, not only to show that informal rigour can be *achieved*, but also to discover if and how it can be *used* when achieved: and "it's a sight more difficult to find any use for (the truth of) such a thesis than to decide its truth".

§2 of [1987] reminds us that there are two opposite principles in the foundational literature: on the one hand, "the words 'essence of computation' are a directive to look for *one* variant or, equivalently, to what is common to all [of them]"; on the other hand, "the familiar homily 'it all depends' (on situation, purposes, etc.) suggests the need for an endless array of variants". K.'s position, supported by experience in mathematics with relatively few so-called basic structures, is to look for "*relatively few* variants [that] could be adequate for *relatively many* situations".

1.2 Variants of Church's Thesis

K. proposed in 2.7 of [1965] to consider the variants of Church's Thesis obtained by specifying 'effectively calculable' as: mechanically, constructively, humanly, and physically realizable.

As noted in 2.(c).(iii).(β) of [1966], "at the time [of Church], one would have been prepared to regard *effective, intuitionistic, constructive, mechanical, formal* as equivalent when applied to rules! After all, less than ten years before Church formulated his thesis, von Neumann and Herbrand took it for granted that *finitist* and *intuitionistic* had the same meaning! But, perhaps, it would be historically more correct not to call it Church's Thesis; for, once alerted to the difference between intuitionistic and mechanical rules, he would surely have formulated the thesis for the latter".

[1972] however points out that "it seems safe to say that the *sensational* aura around references in popular philosophy to Turing's analysis or to Church's Thesis reflect the – conscious or unconscious – assumption that the humanly effective, not only the mechanically effective definitions are in question".

In any case, 2.715 of [1965] states that explicit support for Church's Thesis exists only in the case of mechanically computable functions. Precisely, it "consists above all in the analysis of machine-like behavior and in a number of closure conditions, for example diagonalization", for which K. refers to the discussion in Kleene [1952].

1.3 Equivalent characterizations of recursiveness

In 2.715 of [1965] one finds the first statement of a point that K. will often repeat,[2] asserting that support for Church's Thesis is *not* to be found in empirical evidence such as the equivalence of different characterizations of recursiveness: "what excludes the case of a *systematic* error?". For comparison, he quotes "the overwhelming empirical support for: if an arithmetical identity is provable at all, it is provable in classical first order arithmetic; they all overlook the principle involved in, for example, consistency proofs".

In 1.(c).(i) of [1971] K. stresses the fact that "the mathematics has here much the same role as in the natural sciences: to state rival hypotheses and to help one deduce from them a consequence, an *experimentum crucis* which distinguishes between them; one will try to avoid artefacts and systematic error. Equivalence results do not play a special role, simply because one good reason is better than 20 bad ones, which may be all equivalent because of systematic error". A note adds that "the familiar emphasis on stability or equivalence results is not rooted in some kind of 'common sense', but in a positivistic philosophy of research which rejects the objectivity of [informal rigour]. An equivalence result allows one to act in accordance with this doctrine without formally adopting it: the result allows one to evade the issue (for the time being)".

In II.(a).(v) of [1972] the mantra is chanted again: "equivalence of different notions such as definability in λ-calculus or by Post rules is often said to provide evidence: *evidence for what?* Such equivalences may indeed provide evidence of *some* interest. But they cannot provide evidence for equivalence to a notion which is *not* among those considered! And if the intended notion is explicitly *included* among

[2]Such repetitions were obviously needed, and should be contrasted with the repetitions of the argument of equivalence in practically every textbook of Recursion Theory, till this day.

the notions considered then there is no need for equivalence proofs - on the principle that one good reason is better than 20 bad ones".

Finally, in 1.2.1 of [1990a]: "the conventional 'evidence' neglects safeguards against systematic errors, and thus the axiom of experimental science (derived from experience, not only doctrine) that the most insidious errors are not at all random, but systematic".

A shift of emphasis

§3 of [1987], while recalling "the curious 'evidence' provided by equivalence between various definitions, as if not every notion had many definitions", introduces a new twist: "those totally absorbed in pursuing such equivalences do not ask whether the schemes are all equally sensible or equally silly. Less obviously, they do not ask whether the details that are left out in the matching are significant; enough to make one scheme practically superior, at least occasionally".

As adumbrated in the last quotation, in recent years K.'s criticism has taken a positive side too, well expressed in 3.(b).(ii) of [1987a]: "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. ... It is an object of research to *discover* which description suits particular problems, even though it may well be that other descriptions tend to force themselves on us".

This shift of emphasis (from what equivalent descriptions have in common to their different potentials) is repeated in 1.2.1 of [1990a]: "those different schemes are not viewed as different, so-called informal analyses of a familiar notion, but are simply the logical aspects of different (mechanical) processors".

1.4 Church's Superthesis

Mere equivalence of characterizations hides a neglected aspect, for which K. introduced in 4.(c).(i) of [1971] a special name. He called Church's *Superthesis* a stronger version of Church's Thesis, in which one not only claims that certain mathematical tasks are *equivalent* to recursive ones, but rather that each such task is *equal* to some program for an idealized computer: "to each mechanical rule or algorithm is assigned a more or less specific programme, modulo trivial conversions, which can be seen to define the same computation *process* as the rule".

K. has been attentive to positive evidence for the superthesis. In particular:

- In [1972] he notes that Turing's analysis of the notion of computability does establish a version of the superthesis, for the notion of mechanically computable function.

- In [1972a] and §4 of [1987] he reports on work by Barendregt,[3] establishing the superthesis for: reduction of terms in the λ-calculus, execution of programs by Turing machines, and evaluations according to the computation diagrams for partial recursive functions given by Kleene. Thus "not only the classes of functions defined by the different familiar schemes are equal, but the definitions themselves match so as to preserve computation steps".

§5 of [1987] stresses the value of the emphasis on the superthesis in another direction. "Common sense says: If you want to find out about things, for example, processes, don't hide them in black boxes! Try to look at them. Specifically, in connection with a refutation of Church's Thesis, don't rely on the off-chance of some process being grossly non-mechanical; so much that not even its effects that strike the eye, the so-called output, can be computed mechanically from the input".

1.5 Turing's analysis

§3 of [1987] recalls that among the equivalent characterizations of recursiveness, the one in terms of Turing machines has a particular intrinsic value: "Turing's description of computations, by the rules of his universal machine, is so vivid that it would establish a common notion together with its elementary properties even if it were not present before. ... This constituted essential progress for informal rigour, and is not changed by the many defects of the notion", some of which are at issue here.

As noted in II.(a).(i) of [1972], the distinction between mechanically and humanly computable functions was clearly presupposed in Turing's attempt [1936] to establish that "a machine can reproduce all steps that a human computer can perform".[4]

[3] In the supplementary Part II of his unpublished dissertation, briefly discussed on p. 43 of the second edition of his book [1981].

[4] In II.(b).(ii) of [1972] K. notes how Turing's introduction of progressions of formal systems (called by him 'ordinal logics') may be taken as showing that he did not take his own claim too seriously.

However, Gödel [1972] noticed a problem in the details of Turing's assumptions about distinguishable states of mind (which, however, does not invalidate his analysis of *mechanical* instructions). Precisely, Turing proposed a compactness argument to establish that the number of such states is finite, but according to Gödel he disregarded the fact that mind develops, and thus that such number (though finite at any given moment) may tend to infinity. Gödel's remark is accepted by K. in II.(a).(i) of [1972], Note 4.(c).(ii) of [1987a], and Appendix I of [1990].

In II.(a) of [1972] K. finds that "an even more important error in Turing's argument consists in a kind of *petitio principii* assuming that the basic relations between (finite) codes of mental states must themselves be mechanical". While "in the case of (Turing) machines whose states are finite spatio-temporal configurations it is quite clear how to code states by natural numbers, ... coding (mental) states of the human computer is a much more delicate matter". In particular, "even if we assume a coding by *finite* configurations, ... what is the *arithmetic* character of the relation (between codes) induced by meaningful relations between the mental states considered"?

In a word, the problem here is that "the human computations are more 'complicated' or, better, more abstract than the objects on which they operate (our thoughts may be more complicated than the objects thought about)". In contrast, "the mechanical computations and their arguments are on a par".

According to Note 4.(c).(iii) of [1987a], Gödel's view on K.'s objection that the coding operators may be non-recursive, "was that we know so little about the details that only very simple assumptions can be convincing. But here, in contrast to his reaction to other, apparently comparable cases, ... he rejected the thought that we may know too little for *anything* convincing".

1.6 Perfect fluids vs. perfect computers

In §3 of [1987] K. reminds us that in the 19th century "not only geometric notions [such as area] were analysed with informal rigour, but also those belonging to the aptly named subject of rational mechanics, with notions of uneven scientific value, including the notoriously imperfect notion of perfect liquid".

Experience with such a notion was considered in §2 of [1985] as an object lesson, not to be forgotten in the context of perfect computers: "progress was made by shifts of emphasis away from the original context. The two dimensional motion of such liquids provides a valid description of - not merely, as is sometimes said, a metaphor for - the notion of function of complex variable. The latter is firmly established in mathematics, even used in parts of mathematical physics, but just not primarily in successful hydrodynamics".

The exceptionality of the idea of perfect fluid is spelled out by K. in unpublished notes of 1989:

- "First, by and large it is a very imperfect idealization compared to, say, celestial mechanics of the planets. There is no area of familiar experience of fluids where the neglected aspects - viscosity, compressibility, turbulence, etc. - are absent to a comparable degree as, say, friction and air resistance in outer space.

- Secondly, the mathematical properties of that idealization belong to function theory, one of the jewels of mathematics. Specifically, the potentials and stream lines of these ideal motions are simply - given by - the real parts of functions of one complex variable.

  The same theorem is both one of the most useful mathematical tools and, applied to the idealization, one of its severest limitations. It is Cauchy's Theorem (on the vanishing of integrals round closed curves), which implies, for such ideal flow, that a stream does not exert any drag of any cylinder".

In the words of §3 of [1987], the example of perfect fluids shows that, when a common notion has been shown to have a mathematical equivalent by means of informal rigour, "there is the possibility of *discovering* other areas, in pure mathematics or its applications, where the mathematical equivalent is suited to describing the facts".

Tacitly, in the case of recursiveness 'other areas' means ones not directly connected to computability.

A first success: Higman's theorem

§3 of [1987] states that, "as is (or should be) well known, the prototype of such a discovery is Higman's answer to the question: Which finitely generated groups can be embedded in finitely presented groups? It is given in recursion-theoretic terms,[5] and is a model of evidence for the use of a notion to tell us what we want to know about (groups)".

§6 adds that "a mere corollary to [Higman's] positive answer is a finitely presented group with unsolvable word problem; in other words, something of concern to Church's Thesis. So Higman's answer shifts attention away from the latter. The answer, in terms of recursiveness, is

[5] Namely: exactly the recursively presented ones, i.e. those for which the set of words equal to 1 is recursively enumerable, are embeddable.

tested by its contribution to the demands of group theory; not primarily by the validity of Church's Thesis in any of its versions".

1.7 Infinitistic character of recursiveness

K. has noted in recent years that an obstacle to simple-minded applications of recursion theory to problems such as those discussed below (especially in Sections 3 and 4) lies in the infinitistic character of the notion of recursiveness, which makes all finite sequences of numbers automatically recursive.

This was adumbrated in §6 of [1987], where he noted that "it is generally assumed that there [is no experimental consequence of the existence of irrational numbers], and it seems very plausible that there is no single measurement that could be interpreted to establish irrationality; or rationality, for that matter. For the record, I am not persuaded that (ir)rationality results have no experimental implications at all. ... Be that as it may, problems of similar flavour come up with the two demarcations, between rational/irrational and computable/non-computable".

The suggestion becomes explicit in Appendix I.3 of [1990]: "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 [of being recursive] 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)".

The point is reiterated in [1992]: "It is well known that *infinitesimal* properties like irrationality or (repeated) differentiability have no place in so-called phenomenological interpretations, that is those that strike the naked eye. Now, computability in the logical sense is quite coarsely *infinitistic*: every finite sequence of (hereditarily finite) data is computable in that sense. This does not exclude a physically suitable interpretation, for example, by reference to some appropriate microtheory, but this matter is demanding".

For direction and comparison, K. refers in Appendix I.3 of [1990] to experience with partial differential equations, where "conditions on solutions being once or twice differentiable are, often demonstrably, mathematically significant; most simply, for admitting or excluding a particular P.D.E. as (even) a candidate for a theory of (the aspects of) the phenomena considered"; for example, as noted in [1982], "the most visible features of many phenomena obeying the wave equation, such as caustics (images) in optics, occur only with weak solutions", i.e. their second derivative exists but is not continuous. "But, again, every observational set of data is consistent with those conditions and also with their negation".

2 Constructive Mathematics

Church's Thesis for constructive mathematics was discussed by K. in:

- *Church's Thesis* (2.7 of [1965]),
- *Church's Thesis: a kind of reducibility axiom for constructive mathematics* ([1970]),
- *Church's Thesis for effective definitions of number theoretic functions* (Part II of [1972]),
- *Laws of thought: this side of the pale* (§5 of [1987]).

In its simplest setting, it amounts to saying that "every constructive number theoretic function has an equivalent definition by means of a certain kind of computation procedure".[6]

The reason to consider constructive mathematics is recalled in §5 of [1987]: "originally Church's Thesis was intended and understood in the sense [that] effectiveness for the ideal mathematician was meant. The recursive undecidability results were advertized under the slogan: what mathematicians cannot do. ... In view of how little is known about the outer limits of mathematical imagination, Church's Thesis in its original sense is simply beyond the pale. If anything remotely like it is to be pursued, some shift of emphasis is required, ... and the intuitionistic variant presents itself at least as a candidate. ... The link with the common notion in question is the meaning of intuitionistic logic as originally explained by Brouwer and Heyting: in terms of mental constructions (of the ideal mathematician)".

[6]Extensions to functions of higher types are also considered in [1965], with the role of the recursive functions variously played by Kreisel's and Kleene's continuous functionals, Gödel's primitive recursive functionals, and Kleene's recursive functionals. A different extension, to partial functions, is considered in [1972]. We don't discuss these extensions here.

2.1 Mechanical and constructive rules

K. notes in §1 of [1970] that there is an issue here, since constructive and mechanical are not equivalent: "it is almost banal that we understand non-mechanical rules; on the contrary too detailed, that is 'too' mechanical rules only confuse the human computer".

This is sharpened in §7 of [1987]: "everyday experience of creative and mechanical thinking shows that the former is simply more congenial to us, less prone to errors, and accordingly more reliable; but also (perhaps disappointingly) the latter can be more efficient. Thus, a modern computer sums $\sum_{1 \leq n \leq 100} n$ more quickly - not more reliably - by routine addition than Gauss did at the age of 6 by use of a bright idea. (Computers do *mechanical* work more reliably than people.)"

In 2.35 of [1965], elaborated in §1 and §2 of [1970], K. goes as far as proposing the following as a specific example of a constructive but apparently non-mechanical function.[7] Given a constructively valid formal system $\mathcal{F}$ for arithmetic, constructively enumerate its proofs, and associate to n:

- 0 if either the conclusion of the n-th proof is not an existential assertion, or it is but the proof does not provide an explicit witness for it;

- $m+1$ if the conclusion of the n-th proof is an existential assertion, and the proof provides m as an explicit witness for it.

Transformation of this (obviously constructive) function into an equivalent mechanical one encounters a number of obstacles: $\mathcal{F}$ does not necessarily have the so-called *Constructive* $\exists$*-Rule* (if an existential assertion is provable, then so is some of its numerical instances); even if it does, $\mathcal{F}$ does not necessarily admit recursive procedures that associate numerical witnesses to provable existential assertions; even if it does, the problem still remains of knowing whether one of such procedures (which are not all necessarily equivalent, in the sense of providing the same witnesses for the same provable existential assertions, unless the system has the so-called $\exists$*-Stability*[8]) is equivalent to the function

[7] Note 9 of [1970] criticizes Kalmar [1959], an inconclusive paper whose title had attracted some attention, and that claimed to contain an example with similar properties.

[8] Warning: the term 'stability' in this context does not fit the usual meaning of stability w.r.t. (small) changes of data, and is instead applied to changes in interpretation.

This kind of stability is made insignificant, from a proof-theoretical point of view, by the kind of instability (this time in the usual meaning of the word) discovered above, and if so which one.[9]

The main characteristic of this example is isolated in II.(a) of [1972] as "the passage between a formal derivation ... and the corresponding mental act, namely the proof expressed by the derivation". In particular, because of this reference to mental acts the definition above is not even meaningful from a set-theoretical standpoint!

The example was shown in 1.(b) of [1971a] to be mechanically computable (for a large class of formal systems, including Heyting's Arithmetic) by use of normalization techniques. But the mere fact that relatively advanced work was needed to answer the question establishes that the latter was genuinely problematic.

2.2 Formal versions of Church's Thesis

In 2.72 of [1965] K. proposes two formal versions of Church's Thesis, for systems for constructive mathematics:

CT1 $\forall f \exists e \forall x \exists z [T_1(e,x,z) \wedge f(x) = U(z)]$

CT2 $\forall x \exists y R(x,y) \ \rightarrow \ \exists e \forall x \exists z [T_1(e,x,z) \wedge R(x, U(z))]$.

The former, suitable for second-order systems with functional variables (for lawlike functions), says directly (via the Normal Form Theorem) that every function is recursive. The latter, suitable also for first-order systems, expresses (via the axiom of choice, extracting a function from a $\forall\exists$ form) the fact that every function is recursive.

Naturally, there is no question of *provability* of CT1 or CT2 in usual systems for constructive mathematics, since in them the corresponding classical systems are interpretable, and in the latter CT1 and CT2 are false.

2.3 Consistency of Church's Thesis

2.723 of [1965] states that both CT1 and CT2 are *consistent* with the systems for constructive mathematics considered there (intended for treatments of free choice sequences, generalized inductive definitions, and bar recursion), as well as with those in Kleene [1952]. This consistency result is extended in Kreisel and Troelstra [1970] to the theory of species of natural numbers (an intuitionistic analogue of classical

by Girard's in proofs of a theorem by Van der Waerden, and reported in his book [1986].

[9] An example of a system with both the Constructive $\exists$-Rule and $\exists$-Stability is Heyting's Arithmetic.

analysis with the comprehension axiom and the axiom of dependent choices), and hence to a number of theories (including the ones just quoted) that can be modelled in it.

As Note 9 of [1970] states, "the main purpose of consistency results is to help avoid fruitless lines of research, since our principal interest is the refutation of Church's Thesis": "consistency results exclude even a 'weak' refutation in the sense of showing the absurdity of a proof, not only a 'strong' one in the sense of exhibiting a counterexample" (e.g. of the kind considered in 2.1 above).

In the opposite direction, II.(c).(ii).(β) of [1972] notices that an inconsistency of CT1 or CT2 "would only mean the absurdity of assuming the existence of a proof, and it would not establish a counterexample". In other words, inconsistency would only show that Church's Thesis cannot be proved by methods in the system considered, but it would fall short of providing an example of a constructive function that is not recursive.

2.4 Validity of Church's Thesis

The question of *validity* of CT1 and CT2 is posed in 2.75 of [1965], with the remark that "there is no reason why the question should not be decidable [in the negative] by means of evident axioms about constructive functions", whose discovery is described as "one of the really important open problems" (and, in 2.(c).(iii).(β) of [1966], "one of the more feasible problems at the present time").

Obviously, one is not thinking here of axioms *stated* in the language of constructive mathematics, but justified by explicitly non-constructive, or otherwise arbitrary interpretations. Here are three examples:

- Spector's *bar recursive functionals*.

  2.(c).(iii).(β) of [1966] states, and 2.b.(iii) of [1971a] proves, that they are inconsistent with Church's Thesis. But, "because of excessive extensionality conditions imposed on them, the contradiction is of little interest".

- Various notions of *choice or lawless sequence* α.

  As noted in 4.b of [1970], all known such notions naturally satisfy the negation of Church's Thesis: "for dice α (or lawless sequences) you don't expect to prove that successive values of α will follow a recursive, or for that matter, any law".

- Brouwer's *thinking subject*.

  This is an analysis of mathematics into ω stages, and states that every set of natural numbers is constructively enumerable (over the natural numbers). The application one has in mind here is to the set of all (numbers coding) constructive proofs, and hence to the possibility of enumerating such proofs constructively in an ω-ordering.

  As noted in 4.c of [1970], "the thinking (freely creating) subject will not convince himself that his (mathematical) behaviour is subject to any law".

  The assumption of the thinking subject is actually provably inconsistent with Church's Thesis.[10]

2.5 Church's Thesis as a reducibility axiom

K. points out in 2.74 of [1965] how the consistency results quoted above show that Church's Thesis plays a somewhat similar role in intuitionistic mathematics as Gödel's constructible sets in set theory:[11] "not only is consistent with the known axioms, but it can also be used to show the *formal* character of interesting open questions".

This is quoted not as a mere possibility in principle, but with explicit examples: in particular, the result that "the rules of intuitionistic predicate logic cannot be proved complete [w.r.t. the intended semantics] by any method consistent with Church's Thesis" (a result sketched in [1962] and 2.741 of [1965], and fully proved in Technical Note I of [1970][12]). In particular, as noted in §3 of [1970], this shows that "the notion of constructive validity of first- order formulas depends on problematic properties of the basic notion of constructive function

[10] The idea of the proof, due to Kripke and reported in Note 10 of [1970], is the following. If Church's Thesis holds, every constructively enumerable set is recursively enumerable. But in constructive mathematics one can show that there is a set which is not recursively enumerable (for example, the usual complement of the Halting Problem). One thus has a counterexample to the thinking subject assumption.

[11] Incidentally, as noted in II.(a).(i) of [1972], the constructible sets were proposed by Gödel as an analysis of humanly effective definitions (and the letter 'L' stood for 'lawlike'). Later Gödel expanded the analysis to the notion of ordinal-definable sets.

[12] The idea of the proof is the following. In [1962] K. had proved that constructive completeness of predicate logic implies (actually, is equivalent to) a constructive version of König's Lemma. Consider an infinite recursive (hence, constructive) tree with no infinite recursive branch: if König's Lemma holds constructively, such a tree has an infinite constructive branch, which cannot be recursive. One thus has a counterexample to Church's Thesis.

(like second-order validity, but unlike first-order validity in the classical case)".

The parallel with set theory is explored in [1970] and supplemented in (d) of [1971b], where K. compares:

- on the one hand: the abstract notion of set, its basic properties described by the Zermelo-Fraenkel axioms, Gödel's constructible model, the assumption $V = L$, and a non-axiomatizability result for infinitary predicate calculus following from it;

- on the other hand: the abstract notion of constructive arithmetical function, its basic properties described by Heyting's axioms, Kleene's realizability model, the assumption of Church's Thesis, and the non-axiomatizability result for intuitionistic predicate calculus following from it, and quoted just above.

In this context, axioms refuting Church's Thesis would play a role similar to set-theoretical axioms (such as the existence of measurable cardinals) contradicting $V = L$.

2.6 Church's Rule

For systems for which (consistency of) Church's Thesis is not known to hold, or it actually fails, one can restrict attention to (consistency of) closure under Church's Rule, i.e. the assertion that if the premise of CT2 is provable then so is the conclusion.

Technical Note II of [1970] warns that Church's Rule is genuinely problematic: even if CT2 is valid, provability of the premise implies only validity of the conclusion, not necessarily its provability (because of incompleteness of usual systems).

As noted in 2.7231 of [1965], the first *consistency* result of closure under Church's Rule was obtained by Kleene for his system in [1952] (a result strengthened in 2.723 of [1965] to consistency of Church's Thesis).

Closure under Church's Rule of the theory of species of natural numbers without choice was proved in Technical Note II of [1970]. The result was extended to the theory with choice in Kreisel and Troelstra [1970] (a result supplemented there by consistency of Church's Thesis), while the proof was simplified in 2.a.(ii) of [1971a], and (b) of [1971b].

In general, II.(c) of [1972] showed that every sound formal system satisfying the Constructive ∃-Rule (quoted in 2.1 above) is closed under

Church's Rule.[13] This implies that for a *refutation* of Church's Rule one can only look at systems that either are not formal or do not satisfy the Constructive $\exists$-Rule. III.3 of [1974] points out that the insistence on considering formal systems satisfying the Constructive $\exists$-Rule was a systematic error that precluded the possibility of disproving closure under Church's Rule. This is balanced by §5 of [1987], where it is noted that, however, "there are no rewarding candidates of systems in sight that can be established with informal rigour to hold for the constructions of the ideal mathematician, but do not have both the two properties [of being formal and satisfying the constructive $\exists$-Rule]".[14]

The assessment of results about Church's Rule is a delicate matter, discussed in II.(c).(ii).(α) of [1972]: *closure* under the rule refers to provability in the system, and thus if true is significant only for systems complete for constructive mathematics (for which however Church's Thesis would hold), and if false is merely a symptom of incompleteness; on the other hand, *inconsistency* would instead disprove Church's Thesis.

3 Theories of Mathematical Reasoning

The possibility of a theory of mathematical reasoning was touched upon by K. in:

- *Mechanistic theories of reasoning* (§4 of [1966]),
- *Genetic theories of effective definitions* (II.(b) of [1972]).

As a general point, the end of §4 of [1966] states that "the use of technically advanced machinery in analysing reasoning is encouraging; after all, Aristotle thought about reasoning; one would like to see clearly what one has that he did not have! (It is no comfort to know that over 2000 years have passed since his time unless one sees just *how* one has used the experience of these 2000 years.)"

[13] The proof is the following. If $\forall x \exists y R(x, y)$ is provable, let n be given: then $\exists y R(n, y)$ is provable, and by the Constructive $\exists$-Rule so is $R(n, m)$ for some m. But the system is formal, and by enumerating its theorems one can find (one such) m. This defines a recursive function f such that, for every n, $R(n, f(n))$ is provable.

[14] Specially concocted intuitionistic formal systems not satisfying the Constructive $\exists$-Rule do exist by the incompleteness theorems: an explicit example is given in II.(c) of [1972].

3.1 Individual reasoning

In discussing formalist rules of reasoning, K. notes that (in the terminology introduced in 1.4) what is at stake here is the superthesis for (mathematical) reasoning. According to §5 of [1987], "the pioneers, in particular Frege, had of course a lot to say about the distinction between thought processes and their results. He called the latter 'objective' thoughts, ... [and] saw a principal use of his objective analysis (ignoring subjective processes) in the greater security it gave to common reasoning".

However, Note 31 of [1965] remarks that "the conviction, probably, is not merely that such rules happen to generate the provable statements in a particular domain of mathematics, but that ... this is really all that goes on". K. proposes a parallel with the early days of chemistry, where "one did not merely mean that the particular integral ratios in chemical reactions happen to be *formally* explained by an atomic hypothesis, but that there were such things as atoms". And notices that the attraction of formalism "derives at least partly from this: long before electronic computers were thought of, one could see more or less how behavior according to such formal rules could be realized by a *mechanism*, that is an old fashioned mechanism in the sense of a Turing machine".

This is reiterated in 2.(a).(i) of [1966]: "Probably the major attraction of formalization was that it suggested the possibility of a mechanistic theory of human reasoning, in particular, that [mathematical] propositions not only can be decided by means of formal rules, but that something like repeating application of such rules is all that goes on even if we consciously think of reasoning differently; more precisely, that the higher nervous system consists of a mechanism whose behaviour is given by the formal rules, as an electronic computer is a mechanism whose physical behaviour realizes certain mechanical laws (the 'instructions' which it is given)".

4.(a).(i) of [1966] notices however that "it is remarkable how little work was done on this even in areas, such as predicate logic, where the set of valid statements is recursively enumerable. The least one would have to do is to show that there is something mechanical about the *actual* choice of proofs, not only about the set of results".

The point is taken up again in 4.(c).(i) of [1971], where the notion of superthesis is applied to "Frege's empirical analysis of logical validity in terms of his formal rules; the superthesis would then correspond to an assignment of specific deductions, modulo trivial conversions, to intuitive logical proofs. Here ... the theorem proved does not determine the process, that is the proof (*a fortiori*, not the formal description of the process); in fact, not even in propositional logic: thus we have at least two obviously different proofs of the theorem $(p \wedge \neg p) \rightarrow (p \rightarrow p)$, one using $p \rightarrow p$ and $q \rightarrow (r \rightarrow q)$ with $q = p \rightarrow p$ and $r = p \wedge \neg p$, the other using $(p \wedge \neg p) \rightarrow s$ with $s = p \rightarrow p$".

K. notes there that what is now called the Curry-Howard isomorphism (between derivations in intuitionistic calculi and terms in corresponding typed λ-calculi) provides an example of work in the direction of the superthesis in the sense just discussed.

3.2 Collective reasoning

The cooperative phenomenon (in the language of statistical mechanics) of the mathematical community, and its behavior with respect to arithmetic problems, is considered in Note 29 of [1965], and in 4.(a).(ii) of [1966]: "This behavior seems asymptotically stable. We certainly have no *better* theory at present than this: a statement will be accepted if true". Now this theory is certainly not recursive and hence not mechanistic, "but the whole issue is whether reasoning is mechanistic, and so it is a *petitio principii* to require that only mechanistic theories of reasoning are admitted".

In II.(b).(i) of [1972] a shift from truth to provability is made: the possibility of "an all-encompassing formal system $\mathcal{F}$ for the whole of mathematics (or even the part dealing with number theoretic predicates)" is considered, and it is noted that such a formal system would establish Church's Thesis for humanly realizable functions.[15] Equivalently, any example of a humanly realizable, non-recursive function would refute the possibility of such a system.

II.(b).(i).(α) shows that Gödel's incompleteness results prove not the impossibility of such a system, but only that we cannot have mathematical evidence of its adequacy.

II.(b).(i).(γ) discusses empirical, non-mathematical evidence, in particular stability of actual practice: systems such as *Principia Mathematica* are as adequate for number theory or analysis today as they were in the 19th century. K. notices that knowledge of Gödel's incompleteness proof would, at least naively, be expected to spoil the adequacy of such systems, but in practice it does not.[16]

[15] A technical result sketched in [1971c] and proved in Part I of [1972] shows that the same would hold in the weaker hypothesis that "mathematical reasoning is encompassed not by a single formal system, but by a recursive progression on a Π^1_1 path through Kleene's $\mathcal{O}$". See Note 4.(b).(iii) of [1987a] for an account of Gödel's role in prompting this result.

[16] He intriguingly remarks that also "knowledge of Freud's interpretation of the

3.3 Mind

According to the end of [1980], "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, ... the laws of thought are not mechanical (that is, cannot be programmed even on an idealized computer)".

As stressed there, "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 minds. Nor there is the slightest hint of any computer programs which simulate (even in outline) actual proof search; not even for solving problems which do have a mechanical decision procedure (for example, propositional algebra)".

In §4 of [1966] K. proposes a variant to a favourite twist of Gödel's, brought up in conversation: "either there are mathematical objects external to ourselves or they are our own constructions and mind is not mechanical".

The variant differs from Gödel's formulation in two respects: first, K. makes no assumption that "if mathematical objects are our own constructions we must be expected to be able to decide their properties";[17] second, he would "like to use an abstract proof of the non-mechanical nature of mind ... for the specific purpose of examining particular biological theories".

For the latter purpose, granted a negative result about the mechanical nature of mind, 4.(a).(iii) of [1966] points out that one needs to make specific assumptions about such theories (in addition to the general ones stated below, at the end of 4.1). In particular: "mathematical behavior is regarded as an integral part of the experience to be explained, and not as some corner far removed from the principal activities of the organism" (an assumption whose rejection implies an acceptance of the division between mental and 'ordinary' biologi-

dream symbolism would be expected to produce new symbols (to deceive the superego) but, apparently, it does not".

[17]K. has a gift for provocative comparisons, and displays it in this case by adding: "I do not see why one should expect so much more control over one's mental products than over one's bodily products – which are sometimes quite surprising" and, as added at the end of [1980], "can have painfully unexpected properties". According to the last source, "Gödel remained unsympathetic to [this] admittedly tasteless comparison".

As another example, in §8 of [1985] he attacks the "blithe talk about 'natural' notions" by reminding one of "the obvious parallel from botany where perfectly natural and often pretty mushrooms can be addictive or poisonous in other ways".

cal phenomena), and it is "to be explained in terms of the basic laws themselves"; moreover, "the basic laws are such that the laws for co-operative phenomena, i.e. interaction of organisms such as involved in mutual teaching of mathematics, are also recursive".

K. proposes the following as a debating point: "compare the place of mathematical behaviour among biological phenomena to the place of astronomical behaviour among mechanical phenomena; the former is far removed from ordinary life, exceptionally predictable, exceptional both in the sense that the predictions are precise, and also that they were the first to be noted; since astronomical phenomena played an important part in building up physical theories, should one not expect the analogue too?"

As to the role of Gödel's incompleteness theorem, K. states in 4.(b) of [1966] that he does not think that the result "establishes the non-mechanistic character of mathematical activity even under [the assumptions] above without [Gödel's own] assumption that we can decide all properties of our (mental) productions. For, what it establishes is the non-mechanistic character of the laws satisfied by, for instance, the natural numbers: and the theory of the behavior of arithmeticians mentioned above may well be wrong!"

Actually, there is the possibility that "the natural tendency of mathematicians to be finitist or predicativist is significant for the psycho-physical nature of reasoning". And if finitism or predicativism turned out to be the correct description of the behavior of finitist or predicativist mathematicians, one could actually mention problems which neither can decide.

4 Physical Theories

The question of whether physically realizable functions are recursive was first raised by K. in 2.714 of [1965], and discussed in:

- *Mechanism and materialism* (4.(d) of [1966]),
- *Analogue versus Turing computers* (§3 of [1970a]),
- *A notion of mechanistic theory* ([1974]),
- *Theories in natural science: rational and computable laws* (§6 of [1987]).

On the positive side, II.(a).(v) of [1972] states that the (important and neglected) empirical evidence provided the fact that a large class of patently non-mechanical functions turn out to be equivalent to recursive ones should be taken as a sign of the importance of the notion of recursive function.

On the negative side, K. notes in 4.(d) of [1966] and at the end of [1980] that the possibility of physically realizable but non-recursive functions shows on the one hand that "the hypothesis that reasoning is not mechanistic is by no means anti-materialist or anti-physicalist", and suggests on the other hand the possibility that "the notion of machine is not adequate 'in principle' to separate mind and matter".

To avoid misunderstanding, [1974] stresses the fact that it is not *phenomena*, but *theories* about them that are considered here. Accordingly, K. defines a theory as mechanistic if "every sequence of natural numbers or every real number which is well defined (observable) *according to theory* is recursive or, more generally, recursive in the data (which, according to the theory, determine the observations considered)".

As a corollary to this position, "the reader should not allow himself to be confused ... by doubts about the validity of a theory with regard to the phenomena for which it is intended", although obviously "such doubts imply doubts about the relevance (to those phenomena) of any results about the mechanistic character of the theory".

4.1 Positive results

The obvious starting point for the search of mechanistic theories is, as the name implies, classical *mechanics*. In 2.714 of [1965] K. noticed that "(excepting collisions as in the 3-body problem, which introduce discontinuities) the theory of partial differential equations shows that the behavior of discretely described (finite) systems of classical mechanics is recursive".[18]

This is reiterated in Note 2 of [1970], where the possibility of "(finitely specified) physical systems whose most probable behavior is nonrecursive" is reconsidered, and the fact that "the theory of partial differential equations gives a negative answer for a general class of systems in classical mechanics" restated (this time, with the comment that "the result is not trivial since we are dealing with the mechanics of the continua and Turing machines are discrete mechanisms").

In §4 (Footnote 1) of [1966] attention is shifted to *probabilistic processes*, and a proof is given of the fact that "if in a stochastic process (with a finite number of states) the transition probabilities are

[18]'Discrete' means that *all* relevant parameters take discrete values, not only that the systems are finite.

recursive, any sequence of states with non-zero probability is automatically recursive".[19] This is sharpened in 3.(c) of [1970a], where it is shown that the result "can be extended to stochastic processes with an *infinite* number of discrete states and a recursive table of transition probabilities".[20]

Finally, in §4 of [1966] K. touches on *biological processes*, and claims that "the stable macroscopic properties of organisms would be expected to be recursive" if, as currently assumed, biological theories will be general schemas for the explanation of biological processes, based on "combinatorial basic steps iterated a (large) number of times" (a characteristic of recursive processes).

This is supplemented in §5 of [1987], where it is noted that "such characteristic aims of the logical tradition as unity by reduction to a few primitives may be misplaced here. Thoughtful biologists are sensitive to those aims, and tell us that they are not compatible with the process of evolution. It selects from a mass of random mutations those specific elements that are adapted to the surroundings in which they happen to be. Quite simply, the process doesn't have a logical feel, and so the laws could not expected to have such a feel either. At most, somewhere on the molecular level the laws might satisfy the idea(l)s of the logical tradition, though often they do not".

4.2 Evaluation of positive results

The result quoted in 4.1 provide *empirical* evidence for the mechanistic character of existing physical theories. III.3 of [1974] draws a parallel with the fact that by the end of the twenties "the huge bulk of the mathematical problems that were regarded as solved had formal, that is, mechanically computable, solutions", and that "even today we do not have any theorem in ordinary number-theoretic practices which cannot be proved in *Principia Mathematica*". Nevertheless, "the non-mechanistic nature of the axiomatic theory of natural numbers was discovered, not by sifting existing applications which accumulated in the course of nature (here: in number-theoretic practice) but by looking for unusual or neglected applications (here: to metamathematical questions): applications specifically chosen for their relevance to ques-

[19] The idea of the proof is the following. If the transition table is recursive, the tree of all possible sequences of states is recursive. If a sequence of states has non-zero probability, it is an isolated branch in such a tree. And any isolated branch of a finitely branching recursive tree is recursive.

[20] One has to use an appropriate definition of 'sequence of states with non-zero probability' to prove that the latter is recursive, since an argument as in note 19 would only prove it is hyperarithmetic.

tions of mechanization or, equivalently, formalization". This is, according to K., "the lesson to be learned from the experience with axiomatic theories of mathematical objects; for use with our present problem concerning the mechanistic character of (other) scientific theories". The points in 4.3 below have been raised throughout the years with this explicit lesson in mind.

4.3 Where to look for negative results

4.(a).(iii) of [1966] notices that, unlike discrete classical systems, *cooperative phenomena* are not known to have recursive behavior.

Note 2 of [1970] hints at the possible non-recursiveness of a collision problem related to the *3-body problem*,[21] and suggests as a possible consequence "an analog computation of a non-recursive function by repeating collision experiments sufficiently often". A more explicit discussion of this example is in IV.2 of [1974].[22]

But, as stated in 2.714 of [1965], the natural place where to look for non- recursive behavior is "the *quantum theory*, for example, of large molecules". Here are two examples proposed by K.:

[21]The critical question is whether or not the masses collide (during the interval of time considered). If they don't, it is obvious that their paths can be computed as precisely as one wants. A physically meaningful formulation of the computability of this matter of collision must then refer not to points in phase space (in other words, precise positions and velocities), but neighbourhoods.

More precisely, one does not ask for a recursive decision procedure to determine whether, for arbitrary times t, a collision occurs exactly at time t (or $\leq t$). Rather, one asks for such a procedure to determine, for arbitrary t, an interval $(t-t_0, t+t_0)$ with small t_0 (possibly depending on t), such that one of the following happens: either there is no collision before $t - t_0$, or there is a collision at some time after $t + t_0$.

[22]K. states in [1976] that the formulation of this problem in [1974] is "distinctly better" than in [1970]. Judgments of this sort, both positive and negative, abound in K.'s (a bit schizophrenic) self-reviews, and the following choice may give a flavour.

In [1971b] he describes the arguments and formulations of [1974] as "unnecessary, ... inconclusive ... and practically useless". He mocks himself, by noticing that "the author, who does not usually avoid self-reference, forgets to quote [one of] his own observation[s]", and that "the author's objections ... seem to the reviewer much stronger that the author can have realized".

In [1971c] he complains that "the title [of [1970a]] is misleading", and that "the discussion trails off feebly instead of referring to the relevant literature". Moreover, "the author [amazingly] fails to stress what is perhaps the most obvious relevant contribution", and "is unsympathetic to his own subject". However, he "has a number of very interesting concrete suggestions".

In [1972a] he depicts the discussions in [1971] as "hesitant and discoursive", and the explanations as "incompetent but convincing".

Finally, in [1976] he plays on his double role by noticing that "the author gives no reference ... and the reviewer does not know any reference either".

- 4.(d) of [1966] notes that "it is not known whether there exists a physical system with a Hamiltonian H such that, for instance, $\sigma(n)$ is the set of possible spins in the n-th energy state, $\sigma(n)$ finite for each n, and $\sigma(n)$ is not a recursive function of n".

- 3.(a) of [1970a] suggests the possibility of "large molecules whose spectrum (or: to have a dimensionless quantity, the ratio of the first spectral line to the second) is not recursive".

However, K. seems to have reached a negative impression about these examples, and in 3.(a) of [1970a] conjectures that "Kato's theorem could be used to give arbitrarily close recursive approximations".[23]

A different example, using sequences of eigenvalues, has been suggested by Pour El and Richards [1989]: the example has the flavour of a Hamiltonian, but does not seem to satisfy the conditions of Schroedinger's equation (in particular, Kato's theorem may not apply).

4.4 Physical relevance

In [1974], K. considers the step from obtaining constants by empirical (usually approximate) measurements to calculating them theoretically, and suggests the possibility that this extension of the notion of physical theories is "liable to introduce non-mechanistic elements in a perhaps non altogether trivial way".

But K. notes that the exhibition of a problem without recursive solutions would not be enough to show the non-mechanistic character of (related) physical theories. Further work would be needed: in particular, "it would be necessary to describe (an ensemble of) experiments and their statistical analysis for which the most probable outcome of the experiments is determined by the solution to the problem. In other words, if the problem has no recursive solution the most probable outcome of the experiments should be non-recursive too".

The wave equation

§IV of [1974] discusses examples of non-recursive objects with a physical look, but without physical relevance: for example, recursively continuous curves which do not attain their maximum at any recursive point.

A step forward in this direction was made by Pour El and Richards [1981]: they proved that for certain choices of recursive data (initial conditions) the wave equation has a unique, but not recursive solution.

[23]Kato's theorem provides upper and lower bounds for arbitrary Schroedinger's equations.

K.'s review [1982] discusses this result: on the positive side, the *equation* itself is provided by current physical theory; on the negative side, the *data* are not (known to be) generated by recursively describable phenomena.[24]

This criticism is reiterated in [1992]: "naturally, some of the operators considered appear in theoretical physics. But not all their formal properties have a physically sensible interpretation!"

Thus the further work quoted above, needed to step from mathematical to physical relevance, is still lacking.

Hadamard's principle

IV.2 of [1974] discusses Hadamard's principle, restricting the class of meaningful (physical) theories to ones providing functions *continuous* in their data. In particular, K. suggests a refinement, requiring functions to be *recursive* in their data, and presents the collision problem (related to the 3-body problem) quoted in 4.3 as an example satisfying Hadamard's principle, but not known (even today) to satisfy this refinement.

In Note 4.(a).(ii) of [1987a] K. says that he had actually thought for some time that even the refinement had been a tacit assumption for people working on problems in mathematical physics; forcing them, as a consequence, to miss non-recursive solutions.

The work of Pour El and Richards [1981] discussed above provided a refutation to this impression, and showed in particular that this had not been the case with Kirchhoff's solution of the wave equation.

Nevertheless, "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".

Acknowledgments. I wish to thank Sol Feferman and Georg Kreisel for their comments on a first draft of this paper.

[24]In passing, K. suggests in [1982] a possible shift of emphasis from recursiveness to subrecursiveness: "the realistic potential of (suitable) analogue computers for cheaply and reliably solving problems that are costly for Turing machines, is at least as significant as that of doing a recursively unsolvable job" (at issue here).

Bibliography

Barendregt, H.

[1981] *The lambda calculus, its syntax and semantics*, North Holland, 1981.

Fitting, M.C.

[1981] *Fundamentals of generalized recursion theory*, North Holland, 1981.

Girard, J.Y.

[1986] *Proof theory*, Bibliopolis, 1986.

Gödel, K.

[1972] Some remarks on the undecidability results, in [1989], pp. 305–306.

[1986] *Collected Works, volume I*, Oxford University Press, 1986.

[1989] *Collected Works, volume II*, Oxford University Press, 1989.

Kalmar, L.

[1952] An argument against the plausibility of Church's Thesis, in *Constructivity in mathematics*, Heyting ed., North Holland, 1959, pp. 72–80.

Kleene, S.K.

[1952] *Introduction to metamathematics*, North Holland, 1952.

Kreisel, G.

[1962] On weak completeness of intuitionistic predicate logic, *J. Symb. Log.* 27 (1962) 139–158.

[1965] Mathematical logic, in *Lectures on modern mathematics*, Saaty ed., Wiley, 1965, vol. 3, pp. 95–195.

[1966] Mathematical logic: what has it done for the philosophy of mathematics?, in *Bertrand Russell. Philosopher of the century*, Schoemann ed., Allen and Unwin, 1966, pp. 201–272.

[1970] Church's thesis: a kind of reducibility axiom for constructive mathematics, in *Intuitionism and proof theory*, Kino et al. eds., North Holland, 1970, pp. 121–150.

[1970a] 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 and Yates eds., North Holland, 1971, pp. 139–198.

[1971a] A survey of proof theory II, in *Proceedings of the second scandinavian logic symposium*, Fenstad ed., North Holland, 1971, 109–170.

[1971b] Review of [1970], *Zentr. Math. Grenz.* 199 (1971) 300-301.

[1971c] Review of [1970a], *Zentr. Math. Grenz.* 206 (1971) 277–278.

[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.

[1972a] Review of [1971], *Zentr. Math. Grenz.* 219 (1972) 17–19.

[1973] Review of [1972], *Zentr. Math. Grenz.* 255 (1973) 28–29.

[1974] A notion of mechanistic theory, *Synth.* 29 (1974) 11–26.

[1976] Review of [1974], *Zentr. Math. Grenz.* 307 (1976) 18–19.

[1980] Kurt Gödel, *Bibl. Mem. Fell. Royal Soc.* 26 (1980) 149–224.

[1982] Review of Pour El and Richards [1979] and [1981], *J. Symb. Log.* 47 (1982) 900-902.

[1985] Review of Fitting [1981], *Bull. Am. Math. Soc.* 13 (1985) 182–197.

[1987] Church's Thesis and the ideal of informal rigour, *Notre Dame J. Form. Log.* 28 (1987) 499–519.

[1987a] Gödel's excursions into intuitionistic logic, in *Gödel remembered*, Weingartner and Schmettered eds., Bibliopolis, 1987, pp. 77–169.

[1990] Review of Gödel [1989], *Notre Dame J. Form. Log.* 31 (1990) 602–641.

[1990a] Logical aspects of computations, in *Logic and Computer Science*, Odifreddi ed., Academic Press, 1990, pp. 205–278.

[1992] Review of Pour El and Richards [1989], *Jahres. deutsch. Math. Verein.* 94 (1992) 53–55.

Kreisel, G., and Troelstra, A.S.

[1970] Formal systems for some branches of intuitionistic analysis, *Ann. Math. Log.* 1 (1970) 229–387.

Odifreddi, P.G.

[1989] *Classical Recursion Theory*, North Holland, 1989.

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.

[1989] *Computability in analysis and physics*, Springer Verlag, 1989.

Turing, A.M.

[1936] On computable numbers with an application to the Entscheidungsproblem, *Proc. London Math. Soc.* 42 (1936) 230–265.

Some Critical Remarks on Definitions and on Philosophical and Logical Ideals

by Paul Weingartner

It is with great pleasure that I contribute an essay to this volume in honour of Georg Kreisel. His writings in philosophy, logic and the foundations of mathematics are of great importance not only as contributions but essentially also as critical challenges and corrections. In his essays and personal letters one of his specialities is to uncover blind spots and hidden assumptions in great and heroic traditions and in contemporary trends and idea(l)s. This essay picks out two areas in respect to which Kreisel has offered challenges, given warnings and stimulated the author with his writings and with personal letters. The first is the topics Truth and Usefulness especially applied to definitions in sciences. The second is the area of heroic ideals (here) exemplified with the idea of a comprehensive language for the sciences and with the properties of the language used in logic and mathematics. I am afraid that this essay does not contain anything new to the person honoured by this volume (except perhaps some citations from the history of philosophy and logic). Rather it is to be understood as an invitation to study the much more detailed elaborations of the problems touched here in some of the essays of the person honoured by this volume.

1 Truth (Validity) and Usefulness

Example: Definitions. Should they be true (valid) or useful or both?

1.1 Russell and Quine

According to Russell and Quine definitions are certainly useful but not true or false. Russell's view on definitions is expressed very briefly on one and a half pages of volume I of his work with Whitehead.[1] Nevertheless the short chapter is stuffed with important statements about definitions. They can be summarized in the following points:

1. "A definition is a declaration that a certain newly-introduced symbol or combination of symbols is to mean the same as a certain other combination of symbols of which the meaning is already known."[2]

2. Definitions are concerned with symbols, not with what they symbolize.

3. Definitions do not belong to the system. The sign '=Df' is not equivalent to any functor of *Principia Mathematica*.

4. Definitions are mere typographical conveniences.

5. Definitions are theoretically superfluous.

6. Definitions are not propositions, they are not true or false.

7. Definitions are expressions of a volition, not of a proposition.

Conditions (1) and (2) show that Russell and Whitehead did not confuse – at least not in this place – designation with meaning. What symbols symbolize is their designation and according to (2) definitions are not concerned with the designation of the signs occurring in the definition. From this it follows also that definitions in the sense of Russell and Whitehead - cannot be *real* definitions (in the traditional sense); i.e. the above conditions, especially (1)-(6), are more closely related to that what was called a nominal definition in the traditional sense. Though such definitions are not concerned with the designata, they are nevertheless concerned with the meaning of the symbols. This is one place more - besides those concerned with propositional functions - to

[1]Whitehead and Russell [1925], p 11.
[2]Ibid. p. 11.

show that intensions play an important role in the building blocks of *Principia Mathematica.*

That definitions in this sense are not to be understood as "real definitions" is supported again in (6) which underlines that they are not propositions.

This line is continued in (7) which can be interpreted as saying that definitions are rules (norms) to tell the reader a volition of the author (i.e. that the *definiendum* has to be understood to mean the same as the *definiens*). One is reminded of Wittgenstein where he refers to Russell in respect to derivation-rules:

> "Now Russell wants to say: '*This* is how I am going to infer, and it is *right*'. So he means to tell us how he means to infer: this is done by a it rule of inference. How does it run? That this proposition implies that one? - Presumably that in the proofs in this book a proposition like this is to come after a proposition like this. - But it is presupposed to be a fundamental law of logic that it is *correct* to infer in this way! - Then the fundamental law would have to run: "It is correct to infer ... from ..."; and this fundamental law will self-evidently be correct, or justified. "But after all this rule deals with sentences in a book, and that isn't part of logic!" - Quite correct, the rule is really only a piece of information that in this book only *this* move from one proposition to another will be used (as it were a piece of information in the index); ..."[3]

Condition (5) contains the criterion of eliminability which is stated more explicitly when Whitehead-Russell say:

> "We might always use the *definiens* instead, and thus wholly dispense with the *definiendum*."[4]

Condition (4) seems to say that definitions are syntactical abbreviations especially because it is followed by pointing out that without them

> "our formulae would very soon become so lengthy as to be unmanageable."[5]

Condition (3) finally states the view that the sign "= Df" does not belong to the signs of the object-language of *Principia Mathematica* though both *definiendum* and *definiens* belong to it. Thus

[3] Wittgenstein [1967], I, 20.
[4] Whitehead and Russell [1925], p. 11.
[5] Ibid.

"= Df" belongs to the metalanguage but still "connects" *definiendum* and *definiens* in the object language. Quine's view on definition is touched upon in several of his writings, one place where it is elaborated in more detail, is his *Truth by Convention*:

> "A *simple* definition introduces some specific expression, e.g., 'kilometer', or 'e', called the *definiendum*, as arbitrary shorthand for some complex expression, e.g., 'a thousand meters' or '$\lim_{n\to\infty}(1+\frac{1}{n})^n$' called the *definiens*. A *contextual* definition sets up indefinitely many mutually analogous parts of *definienda* and *definientia* according to some general scheme; an example is the definition whereby expressions of the form '$\frac{sin—}{cos—}$' are abbreviated as 'tan—'. From a formal standpoint the signs thus introduced are wholly arbitrary; all that is required of a definition is that it be theoretically immaterial, i.e., that the shorthand which it introduces admit in every case of unambiguous elimination in favour of the antecedent longhand.
>
> Functionally a definition is not a premiss to theory, but a license for rewriting theory by putting *definiens* for *definiendum* or vice versa. By allowing such replacements a definition transmits truth: it allows true statements to be translated into new statements which are true by the same token".[6]
>
> "...for it was noted earlier that definitions are available only for transforming truths, not for founding them."[7]

In a much later passage Quine [1963] adds to the kind of definition he described as an arbitrary license of rewriting another type which he calls discoursive definition. Also these are described as having the purpose of language instruction or "to show how some chosen part of language can be made to serve the purposes of a wider part"; i.e. questions of validity of definition are also not at stake here:

> "Definition, in a properly narrow sense of the word, is convention in a properly narrow sense of the word. But the phrase 'true by definition' must be taken cautiously; in its strictest usage it refers to a transcription, by the definition, of a truth of elementary logic. Whether such a sentence is

[6] Quine [1949], p. 250. The essay appeared originally in 1936.
[7] Ibid., p. 258.

> true by convention depends on whether the logical truths themselves be reckoned as true by convention. Even an outright equation or biconditional connecting the *definiens* and the *definiendum* is a definitional transcription of a prior logical truth of the form '$x = x$' or '$p \equiv p$'.
>
> Definition commonly so-called is not thus narrowly conceived, and must for present purposes be divided, as postulation was divided, into legislative and discursive. Legislative definition introduces a notation hitherto unused, or used only at variance with the practice proposed, or used also at variance, so that a convention is wanted to settle the ambiguity. Discursive definition, on the other hand, sets forth a pre-existing relation of interchangeability or coextensiveness between notations in already familiar usage. A frequent purpose of this activity is to show how some chosen part of language can be made to serve the purposes of a wider part. Another frequent purpose is language instruction.
>
> It is only legislative definition, and not discursive definition nor discursive postulation that makes a conventional contribution to the truth of sentences. Legislative postulation, finally, affords truth by convention unalloyed."[8]

1.2 The status of definitions in the empirical sciences

This was a field of interests of mine already in my student years. Especially concerning problems arising in physics in connection with lectures I heard from Arthur March. Is Newton's second law of motion a definition (of force)? Already at that time I became more and more convinced that there are two different types of what one calls definitions.

One type is this to which Russell's and Quine's description fit fairly well. Examples in the empirical sciences (to complement Quine's from mathematics) are: Definitions of 1 calory, DNA, 1 Ohm, ... etc. In such cases the *definiens* is known and one looks for a *definiendum* which functions as an abbreviation. This type seems - at least at first - not to raise difficult problems. The choice of words (expressions for the abbreviation) is rather arbitrary and conventional. But is everything conventional here? There is especially one point worth mentioning. Should we say the *definiens* "is known" or "was selected" and one

[8] Quine [1963], p. 394 ff.

looks for a *definiendum*? And concerning the *definiendum* does one look just for an expression or for an expression which has already some meaning (at least in common language) to serve for some associations or analogies concerning the intended scientific use? Thus it is the selection of just this item as a unit or as an important property or sequence or the selection of that particular *definiendum*. And the reasons for such a selection are certainly beyond pure convention.

A different type of definition is that where the *definiendum* is already used and applied in the particular language (common language or scientific discourse) with a more or less specific meaning. But what one is looking for is a more precise and more definite *definiens* for that *definiendum*. Examples from natural sciences would be:

- Oxygen (Copper) is the chemical element with ordinal number 8 (29) i.e. with 8 (29) electrons and atomic weight 16 (63).

To this type belong also traditional definitions like

- man = rational animal,
- living being = being with nutrition, growth and propagation, or
- circle = the set of all points equidistant form another point.

To these examples the characterization of definitions given by Russell and Quine does not fit at all.

First and most importantly to say that (these) definitions are not true or false, (valid or invalid) but just typographical conveniences is absurd: the *definiens* of the chemical elements is due to empirical investigation and discovery concerning elements which have been known - in this case: oxygen and copper - (for a part of their chemical behaviour) for centuries. Moreover the *definiens* has been revised because of the discovery of isotopes: Cu consists of 69% ^{63}Cu and 31% ^{65}Cu and O consists of $99,76\%$ ^{16}O, $0,04\%$ ^{17}O and $0,2\%$ ^{18}O. The atomic weight is then the average. That the respective isotope belongs to the original element is *not a convention* but is justified by the empirical fact that the number of electrons is not changed and so the chemical reactions of the element and its isotopes can hardly be distinguished.

Secondly, the fact that such definitions are revised (revisable in the light of new knowledge) tells us also that they (their *definientia*) are not settled (fixed) (conventionally?) at the beginning - i.e. in the sense that an ideal formal language is introduced as a basic framework - such that these descriptions would become analytically true statements from then on. Also this idea (of Carnap) does not fit at all to definitions used in the sciences. And the remarks made concerning the first type of definition make it also questionable whether it fits there not to mention the question that analyticity (in Carnap's sense) may not be of help or does not add a new aspect to that type of definition (for analyticity see 2.23).

1.3 The status of definitions in logic and mathematics

Before 1981 my view on definitions was roughly this. In the natural sciences the important definitions are all of the second type, mentioned above: they are true or approximately true and they are usually revised when new facts are discovered which are connected with them. In respect to formal sciences I thought that both views of definitions the one of Russell and Quine and the other of the Polish Logic School (Lesniewski, Tarski further elaborated by Beth and Suppes)[9] are suitable and can be successfully applied. Although I found the method of the Polish Logic School and their followers much more suitable in respect to specific points and closer to the actual usage.[10]

1.3.1

It was only after I read Kreisel's paper *Zur Bewertung mathematischer Definitionen* (On the evaluation of mathematical definitions) that I became convinced that also in mathematics

> "some important questions of the form 'what is X?' have been answered with great success by – of course: valid – definitions".[11]

Concerning logic, Russell's and Whitehead's definitions in *Principia Mathematica* are not good examples for their own doctrine of definition. To begin with the simplest:

$$p \supset q = \ \sim p \vee q \text{ Df.}^{12}$$

[9]Lesniewski [1931], Tarski, [1935/36], Beth [1953], [1965], Suppes [1957], ch. 8. The view about definitions according to these authors is roughly that definitions are true identities or equivalences satisfying the criteria of eliminability and non-creativity. Cf. my [1989].

[10]Cf. my [1965] and [1989] specific controversial points in Russell's theory of definitions are for example points (3), (7) and (2) in the list given in 1.1 above. Cf. my [1976], p. 237 ff.

[11]Kreisel [1981], p. 186. This paper contains a lot of interesting and informative details and corrections of common place views on definitions (mainly) in mathematics.

[12]Cf. Whitehead and Russell [1925], Vol. I, p. 11.

To say that this is just typographical convenience, an abbreviation without being true or false is inadequate, to say the least, because the semantic meaning determined by truth-tables is known. More subtle is the definition of identity (according to the idea of Leibniz):

$$x = y \ . = \ : (\phi) : \phi!x \ . \supset \phi!y \qquad \text{Df}\ ^{13}$$

Not only that Leibniz's principle – identity means agreement in all properties – is expressed with it, a lot of other discoveries are included, one especially of Russell-Whitehead: In order to avoid antinomies in the system of *Principia Mathematica* the predicate variable 'F' can run over the properties of one type level only but not over the properties on different type levels. One understands immediately that this is not 'typographical convenience'. Moreover Leibniz's principle is not a harmless abbreviation. Taken too literally the notion "all properties" is much too strong such that only self-identity is left which is too sterile for most purposes:

> "Identity in the literal sense is hardly of scientific - and according to a sermon of bishop Butler also not of moral - interest: 'Everything is what it is and not something else'."[14]

1.3.2

But the mentioned article has brought to my attention even a more unexpected point. The usual doctrine about definitions is that they are first of all useful, purposeful (*zweckmäßig*). This is certainly true if only some obvious goals nearby (which are often superficial) are meant. But if this question is understood in a deeper sense, if it is the question whether the definition (or the new concept defined) will be rewarding in that area of science for a long time then the usefulness of a new definition is never clear at the very beginning and is realized only on the basis of specific "experience not only doctrine."[15]

An obvious example of mathematics would be the concept (definition of) group. Though general it is specific enough to be the basis of rewarding and interesting mathematical structures. Since generality by all means is a dangerous goal too:

> "Neither Aristotle nor (most of) his disciples heeded his warnings, for example, in the following: *General categories*

[13]Ibid. p. 168.
[14]Kreisel [1981], p. 187.
[15]Cf. Kreisel's [1990].

> of things and thoughts; *existence* and *validity* or *truth* respectively. They are central to his favourite sciences of metaphysics (being-as-being) and logic (valid inference).
>
> *Home truth.* If the warning against generality is remembered at all, the emphasis (on generality) in those favourite trades is a prime candidate for a structural weakness."[16]

2 Heroic Ideals

There have been attempts in the philosophical tradition to construct an universal language which is a maximal precise instrument for all the sciences. One example is Leibniz. Another is the program of the Vienna Circle.

2.1 Leibniz: From the characteristica universalis to the calculus ratiocinator

The procedure is in several steps. First a formal language is built up from primitives. Second that language is completely mathematized. Third the mathematical language is applied to proofs such that they can be mathematically calculated. By this Leibniz hoped to have a combinatorial-mechanical method to find every truth.

2.1.1 First step: Analyzability

To be precise according to Leibniz a scientific term has to be analyzable, i.e. it must be possible to trace it back by definitional replacement to ultimate primitive notions (which are usually inborn according to him):

> "Analysis is as follows: Let any given term be resolved into its formal parts, that is, let it be defined. Then let these parts be resolved into their own parts, or let definitions be given of the terms of the (first) definition, until (one reaches) simple parts or indefinable terms."[17]
>
> "When definition pushes analysis until it reaches primitive notions, without presupposing anything whose possibility requires an a priori proof, the definition is perfect or essential."[18]

[16] Kreisel [1991], p. 577.
[17] [1975/90] vol. 4, pp. 64-65 (= *De arte combinatoria*, 64). Cf. Kneale [1962], p. 333.
[18] [1975/90], vol. 4, p. 450.

2.1.2 Second step: Mathematization

This step has three substeps according to Leibniz:

1. Every primitive term can be represented by a characteristic basic number.

2. Every compound term can be represented by a characteristic number which is equal to the result of applying a certain mathematical function to basic numbers (i.e. numbers which represent primitive terms).

It follows also that every Subject-term and every Predicate-term of a categorical proposition is representable by a characteristic number.[19]

In his last and successful application to syllogistics Leibniz represented each term of a categorical proposition not by just one number but by an ordered pair of numbers of which one is positive the other negative and both have no common divisor.[20]

3. The choice of these numbers (or pairs of numbers) is not arbitrary. The characteristic number of the compound term has to be chosen in such a way that in addition to condition (2) the following condition is satisfied:

   If the predicate (P) is contained in the subset (S)[21] the number representing the predicate term (t_p) has to be contained mathematically (i.e. where 'containing' is interpreted by a certain math-

[19]Cf. [1903], pp. 4950. This is also the opinion of Kauppi, cf. her [1960], p. 146.

[20]Cf. [1903], pp. 7879.

[21]Observe that according to Leibniz a categorical proposition is true iff the predicate is contained in the subject: "*Propositio vera est cujus praedicatum continetur in subjecto seu ei inest*" ([1903], p. 68. Cf. [1975/90], vol. 7, pp. 199-200 and 208). I have shown elsewhere that the traditional expression "the predicate is contained in the subject" (predicatum inest subjecto) can be interpreted as a kind of intensional inclusion which can be defined as $P \sqsubseteq S \leftrightarrow (M)[P \subseteq M \rightarrow S \subseteq M]$, where $\subseteq$ is the usual class-inclusion. When this definition plus an additional one for intensional intersection are introduced into the usual theory of classes (without the need of any set-theoretical axiom) one gets a dual theory which is an intensional class theory (i.e. extensions and intensions are interdefinable). The theorems of this theory - as for instance the theorem $P \sqsubseteq S \leftrightarrow S \subseteq P$ which expresses the antagonistic relation between extensional and intensional inclusion - give suitable interpretations of doctrines of traditional logic especially of those of Aristotle, the Scholastics, Leibniz and De Morgan. Cf. my [1976] and [1981]. [In these two contributions however the respective definitions D7 are creative such that the theorems T10 and those theorems for the proof of which D7 and T10 are used are too strong. For a dual theory of intensions within the usual theory of ("virtual") classes which can be successfully used to interpret the traditional notions D7 is not needed. The essential point for the interpretation is the definition of intensional inclusion (D3).]

ematical function F) in the number representing the subject term (t_s).

For instance for the categorical proposition which is universal and affirmative Leibniz requires

> "*Regulae usus characterum in propositionibus categoricis sunt sequentes: Si propositio Universalis Affirmative est vera, necesse est ut numerus subjecti dividi possit exacte seu sine residuo per numerum praedicati.*"[22]

2.1.3 Third step: Application to proofs (i.e syllogistic moods)

Leibniz has given quite a number of different proposals for a mapping from the forms of the different categorical propositions to characteristic numbers.[23] The difficulties are not only related to the categorical propositions alone but also to their role in syllogistic moods so as to be able to distinguish valid arguments from invalid ones by a method of mathematical calculation. But Leibniz's last proposal was successful. With this proposal Leibniz hoped to find a mathematical decision method for syllogistics, such that the moods can be mathematically calculated as to their validity. By this method he has found that all laws of conversion, of the square of opposition, and all the valid moods of assertoric syllogism can be proved to be valid. Moreover the method exhibits as valid Lukasiewicz' four axioms of assertion and also his axiom of rejection.[24]

There is however some difficulty which seems to have bothered also Leibniz. There are invalid syllogistic forms for which one can find numbers in such a way as to exhibit them as valid. But, on a closer look, this fact does not make Leibniz's mathematization of syllogistic incorrect. As Marshall points out correctly, the argument for rejection is as follows:

> "If invalidating instantiations can be produced, the mood is invalid; it is valid if they cannot."[25]

[22] [1903], p. 42. For the exact rules concerning the four categorical propositions of syllogistics cf. my [1983], chapter 3.1.3.1.

[23] Cf. [1903], pp. 40-84.

[24] Cf. Lukasiewicz [1957], pp. 126-129. Kauppi and Rescher seem not to be aware of the success of Leibniz's last proposal They discuss only earlier proposals where Leibniz represents terms by simple numbers (not pairs). In this respect they were correct with their doubts whether it works with all the syllogistic moods (though the simple method works for some, for example BARBARA, as they observe).

[25] Marshall [1977], p. 241. For the historically difficult question whether Leibniz had doubts about the correctness of his last proposal see Marshall, ibid., p. 241 ff.

2.1.4 Extension of the Method

Leibniz probably thought that this or a similar method of mathematization can be extended to the whole of mathematics and then also to metaphysics. In respect to physics (we may even say to empirical sciences) and to jurisprudence and ethics Leibniz was much more careful: These truths are "infinitely analytic", i.e. an infinite number of steps would be necessary to trace them back to first evident axioms. That is man cannot demonstrate them (only God knows the connection) although also for them the principle of sufficient reason (as a completeness-principle) holds: every truth has its proof from the respective axioms .

> "The laws of motion which actually occur in nature and which are verified by experiments are not in truth absolutely demonstrable, as a geometric proposition would be."[26]

But for logic, mathematics and metaphysics Leibniz not only thought that completeness of the respective axiomatic systems (built up *more geometrico*) holds, but also, when his method of mathematization is extended, a calculus ratiocinator, a combinatorical mechanical method is available to find every truth. I want to emphasize strongly that one should understand this heroic and far reaching hope in connection with his success in syllogistics and the three steps described above. This could of course not be known to historians of philosophy before the edition of Leibniz [1903] by Couturat. But even since Leibniz's proposals to mathematize syllogistics seem to be known very little. Although he was aware of some important shortcomings of syllogistics (for instance that it does not contain relations[27]) it still seems that his success there was an important motive for the claims and hopes he had in respect to a comprehensive system which includes all logical and mathematical truths.

The following passage seems even to suggest that this method is available for every science although other passages (like the one above) seem to show that he makes a clear distinction between logic, mathematics and metaphysics on one side and the empirical sciences and jurisprudence and ethics on the other:

> "In Philosophy, I have found a means of accomplishing in all sciences what Des Cartes and others have done in Arithmetic and Geometry by Algebra and Analysis, by the Ars

[26] [1975/90], vol. 3, p. 400. For more details see my [1983], ch. 2.5.
[27] Cf. Leibniz [1975/90], vol. 5, p. 461.

> Combinatoria. ... By this all composite notions in the whole world are reduced to a few simple ones as their Alphabet; and by the combination of such an alphabet a way is made of finding, in time, by an ordered method all things with their theorems and whatever is possible to investigate concerning them."[28]

> "Methodus inveniendi perfecta, si praevidere possimus, imo demonstrare antequam rem aggrediamur, nos ea via ad exitum perventuros; perfecta magis illa quae nullis utitur theorematis apud alios demonstratis, vel problematis ab aliis solutis."[29]

2.2 Towards an all embracing Logical Syntax

2.2.1

The idea of a universal language for all sciences was also very strong in the Vienna Circle. But contrary to Leibniz for whom it was an aim to include metaphysics and to apply the characteristica universalis and the method of mathematization in that area in such a way as to decide upon difficult metaphysical alternatives by "*calculemus*" the Vienna Circle and in particular Carnap wanted to prove that metaphysics consists of "*Scheinsätze*":

> "The supposititious sentences of metaphysics, of the philosophy of values, of ethics (in so far as it is treated as a normative discipline and not as a psycho-sociological investigation of facts) are pseudo-sentences."[30]

There is also another difference: Leibniz tried to show that scientific (and philosophical) concepts (predicates) can be mapped onto certain pairs of numbers, the relations between predicates (making up a sentence) to mathematical functions and so the inferential relations between sentences. Carnap's program was to build up a general logic of science which investigates scientific concepts and sentences by logical analysis but is itself nothing else but logical syntax (where philosophy will take the role of the logic of science):

> "*Philosophy is to be replaced by the logic of science* - that is to say, by the logical analysis of the concepts and sentences

[28] [1975/90], vol. 1, p. 57.
[29] [1903], p. 161.
[30] Carnap [1937], p. 72.

> of the sciences, for *the logic of science is nothing other than the logical syntax of the language of science.*"[31]

A third difference is this: Whereas Leibniz was searching for a true characteristica universalis (since compound terms are traced back to primitives in an absolute sense since they are inborn) and for an optimal method of mathematization Carnap tells us that the form of the universal language for the sciences is arbitrary:

> "In it [i.e. in the book *Logical Syntax* ...], the view will be maintained that we have in every respect complete liberty with regard to the forms of language. ... For language, in its mathematical form, can be constructed according to the preferences of any one of the points of view represented; so that no question of justification arises at all, but only the question of the syntactical consequences to which one or other of the choices leads, including the question of non-contradiction."[32]

A fourth difference is this: Leibniz believed that systems of truths of reason (*verités de raison*) can be built up more geometrico (as axiomatic systems) for the areas of logic, mathematics and metaphysics. And in respect to these three areas he seemed to believe that these systems are complete, i.e. that every truth has an a priori proof (from the true axioms) which is finitely analytic, that is which can be carried through in a finite number of steps. Carnap didn't believe in the completeness of an axiomatic system of mathematical truths in general (he knew Gödel's results) (except First Order Logic) let alone metaphysics which consisted for him of pseudo-sentences. But he seemed to believe in another heroic ideal: From Carnap's Logical Syntax one gets the impression that for him language or better to say the logical syntax built up in the way he does can be made complete, in the sense that every logical and (pure) mathematical predicate and relation and every meaningful philosophical predicate can be expressed in it in a formal way, that is without taking into consideration the meaning of the predicates, relations or respective sentences:

> "All questions in the field of logic can be formally expressed and are then resolved into syntactical questions ... *Logic is syntax.*"[33]

[31]Ibid., Foreword, XIII (Carnap's italics).
[32]Ibid., Foreword, XV.
[33]Carnap [1937], p. 259 (Carnap's italics).

> "The view we intend to advance here is rather that all problems of the current logic of science, as soon as they are exactly formulated, are seen to be syntactical problems."[34]

> "For anyone who shares with us the antimetaphysical standpoint it will thereby be shown that all philosophical problems which have any meaning belong to syntax."[35]

But logical syntax in its ultimate comprehensive form (that is a syntax which contains besides logico-mathematical, i.e. analytical sentences also synthetical sentences which are needed when mathematics is applied) can even include "the mathematical calculus". In this comprehensive syntax a unification of formalism and logicism is possible according to Carnap:

> "For on the one hand, the procedure is a purely formal one, and on the other, the meaning of the mathematical symbols is established and thereby the application of mathematics in actual science is made possible, namely, by the inclusion of the mathematical calculus in the total language ... The *requirement of logicism* is then formulated in this way: *the task of the logical foundation of mathematics is not fulfilled by a metamathematics (that is, by a syntax of mathematics) alone, but only by a syntax of the total language, which contains both logico-mathematical and synthetic sentences.*"[36]

Already in his *Principles of Mathematics* Russell claimed quite a strong reduction of mathematics to logic where analysis of Logic is required instead of analysis of Syntax as in Carnap:

> "The fact that all Mathematics is Symbolic Logic is one of the greatest discoveries of our age; and when this fact has been established, the reminder of the principles of mathematics consists in the analysis of Symbolic Logic itself."[37]

2.2.2

The first remark of Kreisel to all this might be expressed by a joke though it is not understood as being only a joke:

[34]Ibid., p. 282.

[35]Ibid., p. 280.

[36]Ibid., p. 326 ff. (Carnap's italics).

[37]Russell [1903], p. 5. Later in his [1919] he had a more modest view understanding that the axiom of infinity does not follow from logical axioms (cf. ibid., p. 202) and that assumptions about the existence of objects in the universe of discourse are not logical assumptions (cf. ibid., p. 203).

> "... For example it does not hold: The more comprehensive the language the better. Since according to the motto of supermarkets *less can be more*."[38]

As a second remark – among the many things one could comment on these heroic ideals one might mention his emphasis that important knowledge about language comes from outside logic:

> "*Manifesto.* In contrast to algebra with such words as 'field', traditionally logic has (cl)aimed to contribute to effective knowledge of the phenomena ordinarily labelled 'language'; pedantically, facts about its logical aspects are offered as such contributions. Realistically speaking, - and without gushing about the mystery or miracle of linguistic capacities - we simply have massive knowledge of those phenomena; ranging from the familiar kind to those found in the case of brain damage; massive albeit not primarily theoretical(ly interpreted). This knowledge becomes even wider if thoughts over and above languages are included. Obviously, some elementary properties of those logical aspects have to be remembered. But the assumption that further logical elaboration would contribute - rather than distract from more rewarding aspects - overlooks not only this risk concerning effective knowledge of the phenomena in question but also the *logical risk*: the tools used in establishing such elaborations are liable to be much more rewarding elsewhere, from which the focus on those familiar phenomena distracts."[39]

2.2.3

Let me just point out two difficulties with the (Carnap's) "idea(l)" of analyticity in connection with Logicism:

Consider the simple example (of Bernays)[40] of a formula containing only terms of First Order Predicate Logic:

$$(\forall x)\neg Rxx \wedge (\forall xyz)((Rxy \wedge Ryz) \rightarrow Rxz) \wedge (\forall x)(\exists y)Rxy.$$

This formula is not satisfiable in a finite domain but only in an infinite domain. To which domain should we refer upon judging whether it

[38] Cf. Kreisel [1989], p. 210 (my translation).
[39] Kreisel [1992], p. 36 ff.
[40] Cf. Hilbert and Bernays [1934], Vol. I, p. 122.

is analytic or contradictory? It seems that the wellknown distinction (between analytic and synthetic) is also an example of a heroic ideal and works only in some very restricted area or under some questionable presuppositions.

We may assume as usually that the assumption that there is at least one entity in the universe of discourse is still an assumption of logic (Standard First Order Logic) or agree with the later Russell (cf. note 37) that any assumptions about the existence of objects in the universe of discourse are not logical assumptions (on which the approaches of Free Logic and Empty Logic agree). In any case it is not the business of logic to assume an infinite number of entities. It seems therefore we have to relativize the analyticity-claims to the kind of domain.

A further obstacle is that in order to find out whether the negation of the above formula is analytic in the finite domain we have to be clear about the meaning of the negation and thus quantify over R (departing thereby from First Order). This leads to the next example. Simple number-theoretic statements when transformed into the language of logic (by reducing mathematical concepts to logical ones) do not only presuppose additional existence assumptions but seem also not to have an adequate form in respect to negation. Thus a possible transformation of $7 + 5 = 12$ (which is analytic according to Frege and Carnap) into the language of logic is:

$$[(E!_7x)GX \wedge (E!_5x)Hx \wedge \neg(\exists u)(Gu \wedge Hu)] \to (E!_{12}x)(Gx \vee Hx),$$

where $(E!_1x)Gx$ is defined as

$$(\exists x_1 \forall y)[Gx_1 \wedge (Gy \to y = x_1)].$$

Here it is presupposed again that there are enough entities in the universe of discourse. If there are not enough one can derive contradictions (cf. the remarks to the first example). If we try to negate the above sentence we run into difficulties in respect to the interpretation of G and H. If they are treated as free variables then there are existential quantifiers in the negation. And this would mean that $7 + 5 = 12$ be analytic but $7 + 4 \neq 12$ not analytic.[41]

2.3 A contemporary example of heroic ideals: Logical Closure

Whereas nowadays no mature philosopher or logician would believe anymore in these two traditional ideals, to unmask contemporary heroic

[41] Cf. the discussion in Wang [1974], p. 234.

ideals – a more difficult matter – is one of Kreisel's specialities (though not always understood by his readers).

If knowledge is based on first principles or on axioms assumed by convention, in both cases the ideal can be characterized by the following picture of Kreisel:

> "Growth of a tree from its seed corresponds to the idea(l) of developing or at least presenting a body of knowledge from first principles."[42]

2.3.1

For such a development a property in respect to both logical operations and inference rules is very important: logical closure.

> "In logic, logical closures are a, if not the, principal stock-in-trade. Without exaggeration, given a question about a bunch of relations its ideal form is assumed to concern the logical closure of the given bunch."[43]

A first warning is a comparison with the application of this concept in mathematics:

> "Within mathematics the category occurs rarely. (Since it occurs sometimes this fact is not merely a matter of principle.) Mathematicians know, but rarely say very well, that what is true in general is liable to be futile in particular cases of interest; for example, not even the propositional logical closure of - equations defining - conic sections (in the plane) is an object of geometric interest. *Some* logical closures have become members of the household of mathematics, above all, of p-adic fields. These not only admit quantifier elimination, but the quantifier-free (Boolean) part has striking algebraic properties; in contrast to the propositional closure above ."[44]

2.3.2

A rather drastic example of the ideal of logical closure is the usual Epistemic Logic. Because of maintaining logical closure in the usual systems of Epistemic Logic they have the (non-human) properties of

[42] Kreisel [1991], p. 573.
[43] Ibid., p. 586.
[44] Ibid.

deductive infallibility (i.e. one knows all the consequences of what one knows) and logical omniscience (i.e. one knows all the logically true sentences). Therefore such epistemic logics are rather (partial) models for the concept of provable instead of knowledge (or knowing).

The reason for those properties in (the usual) epistemic systems is that they are reconstructed out of modal systems (usually $S4$, $S5$ or some system between). The two crucial properties of such modal systems are these:

1. They have the "necessitation rule" in the sense that if p is logically true (a theorem in the underlying logic) then necessarily-p is also true.

2. They have the following closure condition in respect to the operator "necessary" (which is quite adequate in modal systems): If p is necessarily true and if q necessarily follows from p then also q is necessarily true.

The main reconstruction of the epistemic system out of the modal system consists in reinterpreting the necessity-operator as a knowledge-operator. Thus whereas it is quite acceptable that all necessary consequences of logically necessary sentences are again logically necessary sentences it follows by this interpretation from (2) that also all consequences of sentences which we know are also sentences which we know (i.e. deductive infallibility). And by (1) it follows that we know all logically true (valid) sentences (i.e. logical omniscience).[45] By containing deductive infallibility and logical omniscience the usual systems of Epistemic Logic are simply incompatible with human rationality. Why do logicians make difficult proofs of complicated theorems if they know all the logically true (valid) sentences?[46] The example shows that here *Hausverstand* was sacrificed for the ideal of Logical Closure.

2.3.3

It should be mentioned however that in the usual systems of Epistemic Logic the ideal of Logical Closure is intertwined with another widespread ideal: The one to have a (model-theoretic) semantics by all

[45] Observe however that it is a great difference when the modal operator 'necessary' is interpreted as "provable". For instance as in the interesting interpretation "provable in arithmetic" by Solovay [1976]. This successful interpretation shows once more - because of the obvious differences between 'know' and 'provable (in arithmetic)' - that an interpretation of '$\Box$' with 'know' is entirely inadequate.

[46] For further critical details and for a discussion of desiderata for an Epistemic Logic which meets conditions of human rationality cf. my [1982].

means. Since there is (was) no (for epistemic systems) and since there was one available for modal systems one reinterpreted modal systems epistemically and had at once a semantics: Using some analogy like that of $Kp \to p$ to $\Box p \to p$ and forgetting the many differences and disanalogies (cf. above). Analogous examples are widespread in the area called "Philosophical Logic".

References

Beth, E.

[1953] On Padoas Method in the Theory of Definition, *Indigationes Mathematicae*, 15 (1953) 330-339.

[1965] *The Foundations of Mathematic*, Amsterdam, 1965.

Carnap, R.

[1937] *The Logical Syntax of Language*, London, 1937.

Hilbert, D., and Bernays, P.

[1934] *Grundlagen der Mathematik*, Springer, Berlin 1934 (new edition 1968).

Kauppi, R.

[1960] Über die Leibnizsche Logik. Mit besonderer Berücksichtigung des Problems der Intension und der Extension, *Acta Philosophica Fennica* XII, Helsinki, 1960.

Kneale, W., and Kneale, M.

[1962] *The Development of Logic*, Oxford, 1962.

Kreisel, G.

[1981] Zur Bewertung mathematischer Definitionen, in Morscher, E., Neumaier, O., and Zecha, G. (eds.), *Essays in Scientific Philosophy (dedicated to Paul Weingartner)*, Comes, Bad Reichenhall, 1981, pp. 571-589.

[1989] Zu Wittgensteins Sensibilität, in Gombocz, W., Rutte, H., and Sauer, W. (eds.), *Traditionen und Perspektiven der Analytischen Philosophie Festschrift für Rudolf Haller*, Hölder-Pichler-Tempsky, Wien, 1989, pp. 203-223.

[1990] Logical Aspects of Computation: Contributions and Distractions, in Odifreddi, P. (ed.), *Logic and Computer Science*, Academic Press, 1990, pp. 205-278.

[1991] Suitable Descriptions for Suitable Categories, in Schurz, G., and Dorn, G. (eds.): *Advances in Scientific Philosophy. Essays in Honour of Paul Weingartner on the Occasion of the 60th Anniversary of his Birthday*, Rodopi, Amsterdam-New York, 1991.

[1992] On the Idea(l) of Logical Closure, in *Annals of Pure and Applied Logic*, 56 (1992) 19-41.

Leibniz, G.W.

[1975/90] *Die philosophischen Schriften von G. W. Leibniz*, C.I. Gerhardt (ed.), 7 Vols. Berlin, 1975-90, Georg Olms Verlag, Hildesheim.

[1903] *Opuscules fragmentes inédits de Leibniz*, L. Couturat (ed.), Paris, 1903, Georg Olms Verlag, Hildesheim.

Lesniewski, S.

[1931] Über Definitionen in der sogenannten Theorie der Deduktion, *Comptes Rendus des Séances de la Societé des Sciences et des Lettres de Varsovie*, 24 (1931) 169-177.

Lukasiewicz, J.

[1957] *Aristotle's Syllogistic from the Standpoint of Modern Formal Logic*, Oxford UP, 1957, London.

Marshall, D. Jr.

[1977] Lukasiewicz, Leibniz and the Arithmetization of the Syllogism, *Notre Dame Journal of Formal Logic*, 18 (1977) 235-242.

Quine, W.V.O.

[1949] Truth by Convention, in Feigl, H., and Sellars, W. (eds.), *Readings in Philosophical Analysis*, New York, 1949.

[1963] Carnap and Logical Truth, in Schilpp, P.A. (ed.): *The Philosophy of Rudolf Carnap. The Library of Living Philosophers*, La Salle, Illinois, 1963.

Russell, B.

[1903] *The Principles of Mathematics*, London, 1903.

[1919] *Introduction to Mathematical Philosophy*, London, 1919.

Solovay, R.

[1976] Provability Interpretations of Modal, *Israel Journal of Mathematics*, 25 (1976) 287-304.

Suppes, P.

[1957] *Introduction to Logic*, Princeton, 1957.

Tarski, A.

[1935/36] Einige methodologische Untersuchungen über die Definierbarkeit der Begriffe, *Erkenntnis*, 5 (1935/36) 80-100.

Wang, H.

[1974] *From Mathematics to Philosophy*, Routledge, London, 1974.

Weingartner, P.

[1965] Can One Say of Definitions that They Are True or False?, *Ratio*, 7 (1965) 61-93.

[1976] *Wissenschaftstheorie II, 1. Grundlagenprobleme der Logik und Mathematik*, Frommann-Holzboog, Stuttgart-Bad Cannstatt, 1976.

[1981] A New Theory of Intension, in Agassi, J., and Cohen, R. S. (eds.), *Scientific Philosophy Today. Essays in Honor of Mario Bunge*, Reidel, Dordrecht, 1981, pp. 439-464.

[1982] Conditions of Rationality for the Concepts Belief, Knowledge and Assumption, *Dialectica*, 36 (1982) 243-263.

[1983] The Ideal of Mathematization of All Sciences and 'More Geometrico' in Descartes and Leibniz, in Shea W.K.(ed.), *Nature Mathematized. Proceedings of the 3rd International Congress of History and Philosophy of Science (Montreal 1980).* Reidel, Dordrecht, 1983, pp. 151-195.

[1989] Definitions in Russell, in the Vienna Circle and in the Lvov-Warsaw School, in Szaniawski, K. (ed.), *The Vienna Circle and the Lvov-Warsaw School*, Kluwer, Dordrecht, 1989, pp. 225-247.

Whitehead, A. N., and Russell, B.

[1925] *Principia Mathematica*, Vol. I, second edition, Cambridge, 1925.

Wittgenstein, L.

[1967] *Remarks of the Foundations of Mathematics*, Oxford, 1967.

TECHNICAL TRIBUTE

On the Decidability of the Real Exponential Field

by Angus Macintyre
and A.J. Wilkie

1 Introduction

We are interested in the two structures $\langle\overline{\mathbf{R}}, \exp\rangle$ and $\langle\overline{\mathbf{R}}, e\rangle$, where $\overline{\mathbf{R}} = \langle\mathbf{R}, +, \cdot, -, 0, 1, <\rangle$ denotes the ordered field of real numbers (in the language of ordered rings), $\exp : \mathbf{R} \to \mathbf{R}$ is the usual exponential function and $e : \mathbf{R} \to \mathbf{R}, x \mapsto \exp((1+x^2)^{-1})$ is the *restricted* exponential function (so called because only information concerning $\exp(x)$ on *bounded* sets is available in $\langle\overline{\mathbf{R}}, e\rangle$). We denote the theories of these structures by $T_{\exp}$ and T_e respectively and we shall be concerned with questions of decidability as proposed by Tarski in [1951].

It is shown in Wilkie [199?] that T_e is model complete and in [199?a] it is deduced from this that $T_{\exp}$ is also model complete. In this paper we shall exhibit a recursive subtheory of T_e which is model complete. This, of course, immediately implies that given any formula of $L(T_e)$ one can *effectively* find an equivalent (in T_e) existential formula. One would like to show that the same thing holds for $T_{\exp}$ but unfortunately the arguments of [199?a] rely not only on the model completeness of

T_e but also on its *completeness* [1] so in the absence of a recursive axiomatization of T_e, the effective model completeness of $T_{\exp}$ remains an open problem. However, we shall prove the following

Theorem 1.1 *If Schanuel's Conjecture (for* $\mathbf{R}$*) is true then* $T_{\exp}$ *is a decidable theory (and hence certainly effectively model complete).*

Schanuel's Conjecture for R (SC) *If* $\alpha_1, \ldots, \alpha_n$ *are real numbers linearly independent over* $\mathbf{Q}$*, then the field*

$$\mathbf{Q}(\alpha_1, \ldots, \alpha_n, \exp(\alpha_1), \ldots, \exp(\alpha_n))$$

has transcendence degree at least n *(over* $\mathbf{Q}$*).*

(SC) is a famous conjecture of transcendental number theory which is generally thought to be out of reach at present. Actually we can formulate an apparently weaker conjecture (but still, it seems, hopelessly unattainable) which is *equivalent* to the decidability of $T_{\exp}$ and we shall discuss this in our concluding remarks.

Concerning the theorem we should remark that Ressayre has shown that $T_{\exp}$ is axiomatized over T_e (or, more correctly, over the obvious interpretation of T_e)by the sentences expressing the facts that *exp* is a strictly increasing isomorphism from $\langle \mathbf{R}, + \rangle$ to $\langle \mathbf{R}^{>0}, . \rangle$ which eventually dominates every polynomial. It follows that the decidability of $T_{\exp}$ is implied by that of T_e. This last result will also follow from our analysis (although we will not obtain such an elegant axiomatization as Ressayre's) and will be presented in the next section, as will the effective model completeness of T_e. The proof of the most difficult lemma needed here will be postponed to Section 3. In Section 4 we shall write down a recursive subtheory of $T_{\exp}$ and show that SC implies that this subtheory axiomatizes $T_{\exp}$. In fact SC will only appear at the end of Section 4, all previous results being unconditional. Finally, the weakening of SC mentioned above will be discussed in Section 5.

2 Effective Model Completeness

Our aim here is to write down axiom schemes (in $L(T_{exp})$, $L(T_e)$) such that the arguments of Wilkie [199?] (with which total familiarity is assumed in this section) can be carried out for models of these schemes rather than just models of $T_{\exp}$ or T_e. Of course we also require that

[1] One needs the fact that every model of T_e contains an elementary submodel which is Archimedean ordered.

these schemes express true statements about our intended structures (i.e. $\langle \mathbf{R}, exp\rangle, \langle \mathbf{R}, e\rangle$). It turns out that most of the properties we need can actually be expressed by schemes valid in arbitrary expansions of $\overline{\mathbf{R}}$, so let us fix such an expansion, $\overline{\overline{\mathbf{R}}}$ say, and let $\overline{\overline{L}}$ be its language. Let L be the language of ordered rings. We denote by $\mathcal{K}_1$ the class of $\overline{\overline{L}}$-structures K satisfying (A1)–(A4) below.

(A1) The reduct of K to L is an ordered field.

(A2) Every definable (i.e. $\overline{\overline{L}}$-definable with parameters from K, unless otherwise stated) subset of K with an upper bound has supremum.

(A1) and (A2) clearly imply that the reduct of K to L is a real closed ordered field. They also imply many elementary results from analysis and differential calculus for *definable* sets and *definable* functions (formulas for derivatives of sums and products, chain rule, maxima and minima criteria, intermediate value theorem, Rolle's theorem, Taylor's theorem ...) where the concepts are interpreted according to the usual ε-δ definitions. Perhaps they also imply the implicit function theorem and standard properties of solutions of linear, first-order differential equations, but it is safer to include these:

(A3) Let $r, n \geq 1, P = \langle p_1, \ldots, p_{n+r}\rangle \in K^{n+r}$ and suppose that $f_1, \ldots, f_r$ are definable C^1 functions from U to K where U is a definable open neighbourhood of P. Suppose further that $f_i(P) = 0$ for $i = 1, \ldots, r$ and

$$\det\left(\frac{\partial(f_1, \ldots, f_r)}{\partial(x_{n+1}, \ldots, x_{n+r})}\right)(P) \neq 0.$$

Then there exist open box neighbourhoods V_1, V_2 of $\langle p_1, \ldots p_n\rangle$, $\langle p_{n+1}, \ldots, p_{n+r}\rangle$ respectively such that for all $Q_1 \in V_1$ there exists a unique $Q_2 \in V_2$ such that $f_i(Q_2, Q_2) = 0$ for $i = 1, \ldots, r$. Further,

$$\det\left(\frac{\partial(f_1, \ldots, f_r)}{\partial(x_{n+1}, \ldots, x_{n+r})}\right)(Q_1, Q_2) \neq 0$$

and the (definable) function $V_1 \to V_2, Q_1 \mapsto Q_2$ is C^1 with derivative given by the usual rule.

(Notice that the last sentence here implies the corresponding version of (A3) for C^m (any $m \neq 1$) and C^∞ functions.)

(A4) Let $F : U \to K^n$ be a definable C^1 function with domain an open interval $U \subseteq K$. Let A be an $n \times n$ matrix of definable C^1 functions from U to K and suppose that $\frac{dF}{dx} = FA$ on U. Then either F is identically zero or else has no zero.

The reader may now easily verify that all the results of Section 4 of Wilkie [199?] hold good for any $K \in \mathcal{K}_1$. (A4 is exactly what is needed for lemma 4.5. The other results follow from this using (A3) and finitary algebraic manipulations.) These in turn imply that any $K \in \mathcal{K}_1$ also satisfies lemma 5.1 of Wilkie [199?]. (The comments immediately following (A2) above are used heavily in the final paragraph of the proof.)

Clearly $\mathcal{K}_1$ is also an elementary class of $\overline{\overline{L}}$-structures with a recursive axiomatization.

In order to proceed further through Wilkie [199?] we now assume that $\overline{\overline{L}}$ is just L together with an extra unary function symbol.

We of course have the interpretations $\langle \overline{\mathbf{R}}, \exp \rangle$ and $\langle \overline{\mathbf{R}}, e \rangle$ in mind so we let $\mathcal{K}_2$ be the class of those $K \in \mathcal{K}_1$ satisfying (A5):

(A5) (i) (unrestricted case) for all $x \in K$, E is differentiable at x and $E'(x) = E(x)$ and $E(0) = 1$, or

(ii) (restricted case) for all $x \in K$, E is differentiable at x and $E'(x) = \frac{-2x}{(1+x^2)^2} E(x)$ and $E(0) = E(1)^2 \neq 0$,

where E denotes the interpretation of the new unary function symbol in K. (Thus $\mathcal{K}_2$ stands for either the class of those $K \in \mathcal{K}_1$ satisfying (A5)(i) or for the class of those $\mathcal{K} \in K_1$ satisfying (A5)(ii).)

Now fix a structure $K \in \mathcal{K}_2$ and let k be any subfield of (the reduct to L of) K. For $n, r \in \mathbf{N}, r \leq n$, let $M_{n,r}(k)$ denote the following subring of the ring of all C^∞ (in the sense of K) definable functions from K^n to K:

$$M_{n,r}(k) = k[x_1, \ldots, x_n E(x_1), \ldots, E(x_r)]$$

(unrestricted case);

$$M_{n,r}(k) = k[x_1, \ldots, x_n, (1+x_1^2)^{-1}, \ldots, (1+x_r^2)^{-1}, E(x_1), \ldots, E(x_r)]$$

(restricted case).[2]

Clearly $M_{n,r}(k)$ is closed under taking partial derivatives and so the following version of 2.7 of Wilkie [199?] follows immediately from our

[2]Here we are employing the usual abuse of notation – x_i denotes the i'th projection function from K^n to K.

current version of 5.1 of [199?].

Theorem 2.1 *With K, k, n, r as above, suppose that $f \in M_{n,r}(k)$ and let $V = \{\overline{\alpha} \in K^n : f(\overline{\alpha}) = 0\}$. Suppose further that S is a non-empty definable subset of V which is both closed in K^n and open in V. Then there exist $f_1, \ldots, f_n \in M_{n,r}(k)$ such that S contains a non-singular zero of the function $K^n \to K^n, \overline{\alpha} \mapsto \langle f_1(\overline{\alpha}), \ldots, f_n(\overline{\alpha})\rangle$..*

We now need to axiomatize the results of Khovanski and van den Dries used in Section 3 of Wilkie [199?]. Concerning Proposition 3.1 of [199?] note that Khovanski has shown in [1991] that the natural number N in the conclusion can be found effectively from the following data: m, n, the degrees of the polynomials defining $g_1, \ldots, g_m$ and the degrees of the polynomials defining the derivatives for the Pfaffian chain $h_1, \ldots, h_\ell$. It follows that there is a recursive function $\Theta(n, r, m)$ such that any model of $T_{\exp}$ or T_e satisfies (A6) below. We therefore let $\mathcal{K}_3$ denote the class of those $K \in \mathcal{K}_2$ satisfying (A6).

(A6) For all $n, r, m \in \mathbf{N}$, $r \leq n$, and $f_1, \ldots, f_n \in M_{n,r}(K)$ such that each f_i has (total) degree at most m (considered as a polynomial in the $n+r$ (unrestricted case) or $n+2r$ (restricted case) generators of $M_{n,r}(K)$), there are at most $\theta(n, r, m)$ non-singular zeros in K^n of the system $f_1(\overline{x}) = \cdots = f_n(\overline{x}) = 0$.

We also need a version of (A6) for one dimensional (rather than just zero dimensional) zero-sets and there does not seem to be an easily quotable form in the literature. Further, the proof of 5.3 of Wilkie [199?] does not immediately adapt to our present needs. We therefore establish

Theorem 2.2 *Let $K \in \mathcal{K}_3$, k be a subfield of K and $n, r \in \mathbf{N}$, $r \leq n, n \geq 2$.*

Let $V = \{\overline{\alpha} \in K^n : f_i(\overline{\alpha}) = 0$ for $i = 1, \ldots, n-1\}$, where $f_1, \ldots, f_{n-1} \in M_{n,r}(k)$, and suppose that V is everywhere non-singular. Then the Boolean algebra $\mathcal{B}$ consisting of those definable subsets of V which are both open in V and closed in K^n is finite.

Proof. For $f \in M_{n,r}(k)$ and $\overline{\alpha} \in K^n$ let

$$d_{\overline{\alpha}} f = \langle \frac{\partial f}{\partial x_1}(\overline{\alpha}), \ldots, \frac{\partial f}{\partial x_n}(\overline{\alpha})\rangle.$$

Our hypotheses assert that for each $\overline{\alpha} \in V, \{d_{\overline{\alpha}} f_1, \ldots, d_{\overline{\alpha}} f_{n-1}\}$ is a linearly independent (over K) set of vectors in K^n. It follows (using

(A2), (A3) and elementary linear algebra) that for any $\overline{\alpha} \in V$ and $f \in M_{n,r}(K)$, if $\overline{\alpha}$ is a local minimum of the function $f \mid V$, then $d_{\overline{\alpha}}f$ lies in the subspace of K^n generated (over K) by $d_{\overline{\alpha}}f_1, \ldots, d_{\overline{\alpha}}f_{n-1}$ or, equivalently, that $f^*(\overline{\alpha}) = 0$, where $f^*(\overline{x})$ denotes the determinant of the Jacobian matrix $\frac{\partial(f_1,\ldots,f_{n-1}f)}{\partial(x_1,\ldots,x_n)}$. (Note that $f^* \in M_{n,r}(k)$.)

Now, for $\ell = 1, \ldots, n$ let

$$g_\ell(\overline{x}) = (x_\ell - 1)^2 + \sum_{i=1, i\neq\ell}^{n} x_i^2$$

and set

$$g_0(\overline{x}) = \sum_{i=1}^{n} x_i^2.$$

Then $g_\ell \in M_{n,r}(k)$ for $\ell = 0, \ldots, n$.

Clearly (using (A2)), for each $\ell = 0, \ldots, n$, every non-empty $S \in \mathcal{B}$ contains a local minimum of the function $g_\ell \mid V$ and hence a zero of g_ℓ^*. Further, since $d_{\overline{\alpha}}g_0, \ldots, d_{\overline{\alpha}}g_n$ generate the entire space K^n for any $\overline{\alpha} \in K^n$ (by direct calculation), it follows that for no $\overline{\alpha} \in V$ do we have $g_\ell^*(\overline{\alpha}) = 0$ simultaneously for all $\ell = 0, \ldots, n$. Now notice also, that for any $S \in \mathcal{B}$ and $f \in M_{n,r}(k)$, the set

$$\{\overline{\alpha} \in S : f \mid U \cap V \equiv 0 \text{ for some definable, open } U \subseteq K^n \text{ with } \overline{\alpha} \in U\}$$

is open in V and definable and hence, if it coincided with

$$\{\overline{\alpha} \in S : f(\overline{\alpha}) = 0\}$$

(which is closed in K^n) would actually lie in $\mathcal{B}$.

Putting all this together we have shown that for each non-empty $S \in \mathcal{B}$, there is some $\ell = \ell(S) \in \{0, \ldots, n\}$ and some $\overline{\alpha} = \overline{\alpha}(S) \in S$ such that $g_\ell^*(\overline{\alpha}) = 0$, but such that for all open definable $U \subseteq K^n$ with $\overline{\alpha} \in U$, $g_\ell^* \mid U \cap V \not\equiv 0$.

We now claim that there is a finite subset $\mathcal{F} \subseteq M_{n,r}(k)$ such that for each $S \in \mathcal{B}$ there exists $\overline{\alpha} \in S$ and $g \in \mathcal{F}$ such that $g(\overline{\alpha}) = 0$ and $g^*(\overline{\alpha}) \neq 0$. This clearly suffices to establish the theorem by (A6).

If the claim were false, then by the above and assuming K is $|\,k\,|^+$-saturated (as we may) there exist $\overline{\alpha} \in V$ and $\ell \in \{0, \ldots, n\}$ such that $g_\ell^*(\overline{\alpha}) = 0$, $g_\ell^* \mid U \cap V \not\equiv 0$ for all open, definable $U \subseteq K^n$ with $\overline{\alpha} \in U$ and, for all $g \in M_{n,r}(k)$, if $g(\overline{\alpha}) = 0$ then $g^*(\overline{\alpha}) = 0$. But this contradicts 4.9 of Wilkie [199?] (which, as we have observed, holds for all $K \in \mathcal{K}_1 \subseteq \mathcal{K}_3$). $\square$

Corollary 2.3 *Let $K \in \mathcal{K}_3$, k be a subfield of K and $n, r \in \mathbf{N}$, $r \leq n$, $n \geq 2$. Suppose that $f_1, \ldots, f_{n-1} \in M_{n,r}(k)$. Let*

$$V = \{\overline{\alpha} \in K^n : f_i(\overline{\alpha}) = 0 \text{ for } i = 1, \ldots, n-1\}$$

and suppose that for all $\overline{\alpha} \in V$, $\frac{\partial(f_1,\ldots,f_{n-1})}{\partial(x_2,\ldots,x_n)}(\overline{\alpha})$ is a non singular matrix. Then the following hold:

(i) *There exists $N \in \mathbf{N}$ such that for all $\alpha_1 \in K$ there are at most N points $\langle \alpha_2, \ldots, \alpha_n \rangle \in K^{n-1}$ such that $\langle \alpha_1, \ldots, \alpha_n \rangle \in V$.*

(ii) *Let $g \in M_{n,r}(k)$. For all $m \in \mathbf{N}, m \geq 1$, the set S_m of those $\alpha_1 \in K$ such that*

$$\mid \{g(\beta_1, \cdots, \beta_n) : \langle \beta_1, \ldots, \beta_n \rangle \in V \text{ and } \beta_1 = \alpha_1\} \mid \ \geq m$$

is a finite union of open intervals (with end-points in $K \cup \{\pm\infty\}$) and points.

Proof. (i) follows from (A6).

For (ii) consider the following system of $m(n + m - 2)$ equations:

$$f_1(x_1, x_2^s, \ldots, x_n^s) = 0,$$

for $1 \leq i \leq n-1, 1 \leq s \leq m$, and

$$(g(x_1, x_2^s, \ldots, x_n^s) - g(x_1, x_2^t, \ldots, x_n^t))y_{s,t} - 1 = 0,$$

for $1 \leq s, t \leq m, s \neq t$, in the

$$1 + (n-1)m + m^2 - m = m(n + m - 2) + 1$$

variables indicated. The functions here all lie in $M_{n',r'}$ for sufficiently large n', r' and, by an easy calculation, all solutions $\overline{\gamma}$ satisfy

$$\frac{\partial(\overline{F})}{\partial(\overline{z})}(\overline{\gamma}) \neq 0$$

where $\overline{F}$ is some row-listing of the functions and $\overline{z}$ is a row-listing of all variables except x_1. Further, the projection of the solution set, V_m say, onto the x_1-axis is clearly the set S_m. By Theorem 2.2 V_m contains only finitely many connected components. The result follows. □

Corollary 2.3 allows us to get started on the proof of 6.2 of Wilkie [199?] for structures in $\mathcal{K}_3$. The rest of that proof goes through easily by (A2) and (A3), as does the proof of lemma 6.3 of Wilkie [199?]. The results of Section 7 of [199?] now also follow using 6.3 and the local calculus and analysis provided by (A2). Thus we have the following

Theorem 2.4 *Suppose $k, K \in \mathcal{K}_3, k \subseteq K$, and that $n, r \in \mathbf{N}, r \leq n$. Suppose further that for all $n' \geq n$ and $f_1, \ldots, f_{n'} \in M_{n',r}(k)$ every non-singular zero of $f_1, \ldots, f_{n'}$ has coordinates bounded between elements of k. Then all such zeros actually lie in $k^{n'}$.*

It now follows from Theorems 2.1 and 2.4 (and Robinson's Test together with routine logical manipulations the details of which can be found in Section 2 of Wilkie [199?]) that to complete the proof of effective model completeness we need only find a recursively axiomatized class $\mathcal{K}' \subseteq \mathcal{K}_3$ such that the hypotheses of Theorem 2.4 hold for all $k, K \in \mathcal{K}'$ with $k \subseteq K$ and all $n, r \in \mathbf{N}$. Concerning the restricted case, we shall show in the next section that there is a recursive function $\tau : \mathcal{N}^3 \to \{s : s \subseteq \mathbf{Q}, s \text{ finite}\}$ such that any model of T_e satisfies (A7) below (which is the version of van den Dries's result mentioned above that we need here). We therefore let $\mathcal{K}_4$ denote the class of those $K \in \mathcal{K}_3$, in the *restricted* case, satisfying (A7).

(A7) For all $n, r, m \in \mathbf{N}, r \leq n, n \geq 2$ and $f_1, \ldots, f_{n-1}, f_n \in M_{n,r}(K)$ having (total) degree at most m in the $n + 2r$ generators of $M_{n,r}(K)$, and for all $a \in K$ and continuous definable functions $\phi_2, \ldots, \phi_n : \{\alpha \in K : \alpha > a\} \to K$ satisfying

$$f_i(\alpha, \phi_2(\alpha), \ldots, \phi_n(\alpha)) = 0$$

(for $i = 1, \ldots, n-1, \alpha \in K, \alpha > a$) and

$$\det \frac{\partial(f_1, \ldots, f_{n-1})}{\partial(x_2, \ldots, x_n)}(\alpha, \phi_2(\alpha), \ldots, \phi_n(\alpha)) \neq 0$$

(for $\alpha \in K, \alpha > a$), there exists $q \in \tau(n, r, m)$ such that, setting $h(x) = f_n(x, \phi_2(x), \ldots, \phi_n(x))$ we have either h is identically zero, or else $\alpha^q h(\alpha)$ tends to a finite non-zero limit in K as α tends to infinity (in the sense of K).

Clearly $\mathcal{K}_4$ has a recursive axiomatization and the reader may now easily verify that the proof contained in Section 8 of Wilkie [199?] goes through for structures in $\mathcal{K}_4$. We have therefore established the following

Theorem 2.5 *There exists a recursive subtheory of* T_e*, (which we shall denote by* T_{res}*), which is model complete.*

Turning now to the theory T_{exp} and the paper Wilkie [199?a], it is completely routine to check that the only properties of T_{exp} used there are the following:

(i) $\langle \overline{\mathbf{R}}, \exp\rangle \models T_{\exp}$;

(ii) if $\langle K, E\rangle \models T_{\exp}$ then $\langle K, x \mapsto E((1+x^2)^{-1})\rangle \models T_e$;
(K an L-structure)

(iii) Theorems 2.1 and 2.4 hold for all models k, K of $T_{\exp}$;

(iv) The Ressayre axioms (cf. Section 1) hold in all models of $T_{\exp}$.

Now for any ordered field K and $E : K \to K$ define $E^* : K \to K$ by $E^*(x) = E((1+x^2)^{-1})$. Also, for any sentence σ of $\overline{\overline{L}}$ (= the language of ordered rings together with an extra unary function symbol) let σ^* be the natural translation of σ (σ^* also a sentence of $\overline{\overline{L}}$) having the property that for any ordered field K and function $E : K \to K$,

$$\langle K, E\rangle \models \sigma^* \text{ if and only if } \langle K, E^*\rangle \models \sigma.$$

(Clearly the map $\sigma \mapsto \sigma^*$ and its range are recursive.) For T an $\overline{\overline{L}}$-theory let $T^* = \{\sigma^* : \sigma \in T\}$.

Let us denote by $\mathcal{E}_{res}, \mathcal{E}_{unres}$ the existential theories of $\langle \overline{\mathbf{R}}, e\rangle$ and $\langle \overline{\mathbf{R}}, exp\rangle$ respectively, and by T_3^{unres} a recursive axiomatization of the class $\mathcal{K}_3$ in the unrestricted case. Finally let

$$T_0 = T_3^{unres} \cup (T_{res})^* \cup (\mathcal{E}_{res})^*.$$

Clearly (i) above holds with T_0 in place of T_{exp} and so does (iii) (by Theorems 2.1 and 2.4). One may also easily check that the Ressayre axioms are a consequence of T_3^{unres}. Lastly, (ii) holds with T_0 in place of T_{exp} because $T_{res} \cup \mathcal{E}_{res}$ axiomatizes T_e (by Theorem 2.5). It therefore follows that T_0 is a model complete subtheory of T_{exp} and hence that T_{exp} is axiomatized by $T_0 \cup \mathcal{E}_{unres}$. But if σ is an existential sentence of $\overline{\overline{L}}$ then σ^* is also existential, so $\mathcal{E}_{unres} \supseteq (\mathcal{E}_{res})^*$ and therefore T_{exp} is axiomatized by $T_3^{unres} \cup (T_{res})^* \cup \mathcal{E}_{unres}$. Hence we obtain:

Theorem 2.6 *There is a recursive subtheory of* $T_{\exp}$*, which we denote by* T_{unres} *(=* $T_3^{unres} \cup (T_{res})^*$*), with the property that* $T_{unres} \cup \mathcal{E}_{unres}$

axiomatizes T_{exp}.

We remind the reader that the above argument does *not* show that T_{unres} is model complete.

3 The Proof of (A7) for Models of T_e.

Given $n, r, m \in \mathbf{N}, n \geq 2, r \leq n$ we must effectively find a finite set $\tau = \tau(n, r, m)$ of rationals such that (A7) holds for all models K of T_e. Clearly we need only consider the case $K = \langle \mathbf{R}, e \rangle$ since for each fixed n, r, m the result can be transferred to arbitrary $K \models T_e$ (as τ does not depend on parameters). To establish (A7) in this case we *claim* that it is sufficient to prove the following

Theorem 3.1 *Given* $n, r, m \in \mathbf{N}, r \leq n, n \geq 2$ *one can effectively find a finite set* $\tau' = \tau'(n, r, m)$ *of rationals such that whenever:*

(i) $f_1, \ldots, f_{n-1} \in \mathbf{R}[x_1, \ldots, x_n, exp(x_1), \ldots, exp(x_r)]$ *and each* f_i *has total degree at most* m *(as a polynomial in the exhibited generators),*

(ii) $\varepsilon \in \mathbf{R}, \varepsilon > 0$,

(iii) $\phi_2, \ldots, \phi_n : (0, \varepsilon) \to \mathbf{R}$ *are continuous,* $\langle \overline{\mathbf{R}}, e \rangle$*-definable functions such that*

$$f_i(\alpha, \phi_2(\alpha), \ldots, \phi_n(\alpha)) = 0,$$

($i = 1, \ldots, n-1$*) and*

$$\det \frac{\partial(f_1, \ldots, f_{n-1})}{\partial(x_2, \ldots, x_n)}(\alpha, \phi_2(\alpha), \ldots, \phi_n(\alpha)) \neq 0$$

for all $\alpha \in (0, \varepsilon)$*, and*

(iv) $\phi_j(\alpha) \to 0$ *as* $\alpha \to 0^+$ *for* $j = 2, \ldots, n$*,*

then there exists $q \in \tau'(n, r, m)$ *such that either* ϕ_n *is identically zero or else* $\alpha^q \phi_n(\alpha)$ *tends to a finite non-zero limit as* $\alpha \to 0^+$.

To establish the claim let τ' be as in Theorem 3.1 and define τ by

$$\tau(n, r, m) = \{\pm q : q \in \tau'(n + r + 1, n + r, 1 + m(n + 1))\} \cup \{0\}.$$

We show that this τ witnesses (A7).

For suppose $n, r, m, f_1, \ldots, f_{n-1}, f_n, a, \phi_2, \ldots, \phi_n$ are as in the hypotheses of (A7) (for $K = \langle \overline{\mathbf{R}}, e \rangle$). Now van den Dries has observed in [1986], using the Łojasiewicz-Gabrielov theory of semi-and sub-analytic sets that all functions from $\mathbf{R}$ to $\mathbf{R}$ that are definable in $\langle \overline{\mathbf{R}}, e \rangle$ are given in a neighbourhood of $+\infty$ by absolutely convergent Puiseux series. It follows that $\phi_2(\alpha), \ldots, \phi_n(\alpha), f_n(\alpha, \phi_2(\alpha), \ldots, \phi_n(\alpha))$ all tend to a (finite or infinite) limit as $\alpha \to \infty$. Clearly we may suppose that $f_n(\alpha, \phi_2(\alpha), \ldots, \phi_n(\alpha)) \to 0$ *or* $\pm\infty$ as $\alpha \to \infty$ for otherwise we take $q = 0 \in \tau$.

Let

$$s = \{1\} \cup \{j \in \mathbf{N} : 2 \leq j \leq n \text{ and } \phi_j(\alpha) \to \pm\infty \textit{ as } \alpha \to \infty\} \cup B$$

where

$$B = \begin{cases} \{n+r+1\} & \text{if } f_n(\alpha, \phi_2(\alpha), \ldots, \phi_n(\alpha)) \to \pm\infty. \\ \emptyset & \text{otherwise.} \end{cases}$$

For $j \in \{2, \ldots, n\} \setminus s$ set $r_j = \lim_{\alpha\to\infty} \phi_j(\alpha)$ and $u_j = (1 + r_j^2)^{-1}$, and let $u_j = 0$ for $j \in s$.

Now choose polynomials

$$\rho_1, \ldots, \rho_{n-1}, \rho_n \in \mathbf{R}[x_1, \ldots, x_n, y_1, \ldots, y_r, z_1, \ldots, z_r]$$

of degrees $\leq m$ such that

$$\begin{aligned} &f_i(x_1, \ldots, x_n) = \\ &\qquad \rho_i(x_1, \ldots, x_n, (1 + x_1^2)^{-1}, \ldots, (1 + x_r^2)^{-1}, \\ &\qquad\quad \exp((1 + x_1^2)^{-1}), \ldots, \exp((1 + x_r^2)^{-1})) \end{aligned}$$

for $i = 1, \ldots, n$.

Now for $j = 1, \ldots, n$ let

$$x_j^* = \begin{cases} x_j^{-1} & \text{if } \quad j \in s, \\ x_j + r_j & \text{if } \quad j \in s, \end{cases}$$

and let

$$x_{n+r+1}^* = \begin{cases} x_{n+r+1}^{-1} & \text{if } \quad B \neq \emptyset \\ x_{n+r+1} & \text{if } \quad B = \emptyset. \end{cases}$$

Now define for $i = 1, \ldots, n-1, n$:

$$\begin{aligned} &g_i(x_1, \ldots, x_{n+r}) = \\ &\qquad (\prod_{j \in s \setminus B} x_j^m) \rho_i(x_1^*, \ldots, x_m^*, x_{n+1} + u_1, \ldots, x_{n+r} + u_r, \\ &\qquad\qquad\qquad \exp(x_{n+1} + u_1), \ldots, \exp(x_{n+r} + u_r)). \end{aligned}$$

Now consider the system:

$$\mathcal{S}\begin{cases} g_i(x_1,\dots,x_{n+r})=0 & i=1,\dots,n-1\\ (1+x_i^2)x_{n+1}-x_i^2=0 & i\in s\cap\{1,\dots,r\}\\ (1+(x_i+r_i)^2)(x_{n+i}+u_i)-1=0 & i\in\{1,\dots,r\}\setminus s\\ x_{n+r+1}(x^*_{n+r+1}-g_n(x_1,...,x_{n+r}))=0. \end{cases}$$

This is a system of $n+r$ equations in the $n+r+1$ variables $x_1,\dots,x_{n+r+1}$. Also, each function appearing here can clearly be represented in the form $\rho(x_1,\dots,x_{n+r+1},\exp(x_1),\dots,\exp(x_{n+r}))$ where ρ is a polynomial of degree $\le m(n+1)+1$ (note that we may suppose $m\ge 1$ so that $m(n+1)\ge 3$).

Now by van den Dries's observation it follows that we may choose $b\in\mathbf{R}, b\ge a$ such that for all $\alpha\ge b$,

$$\phi_2(\alpha),\dots,\phi_n(\alpha),f_n(\alpha,\phi_2(\alpha),\dots,\phi_n(\alpha))$$

are non-zero whenever the corresponding function is not identically zero.

Now for $\beta\in(0,b^{-1})$ define

$$\psi_i(\beta)=\begin{cases}\phi_i(\beta^{-1})^{-1} \text{ for } i\in\{2,\dots,n\}\cap s,\\ \phi_i(\beta^{-1})-r_i \text{ for } i\in\{2,\dots,n\}\setminus s,\end{cases}$$

$$\psi_{n+i}(\beta)=\begin{cases}(1+\phi_i(\beta^{-1})^2)^{-1} \text{ for } i\in\{1,...,r\}\cap s,\\ (1+\phi_i(\beta^{-1})^2)^{-1}-u_i \text{ for } i\in\{1,\dots,r\}\setminus s,\end{cases}$$

and finally

$$\psi_{n+r+1}(\beta)=\begin{cases} f_n(\beta^{-1},\phi_2(\beta^{-1}),\dots,\phi_n(\beta^{-1})) \text{ if } B=\emptyset,\\ f_n(\beta^{-1},\phi_2(\beta^{-1}),\dots,\phi_n(\beta^{-1}))^{-1} \text{ if } B\neq\emptyset.\end{cases}$$

By our assumption on b we have only inverted non-zero quantities here so $\langle\beta,\psi_1(\beta),\dots,\psi_{n+r+1}(\beta)\rangle$ is well defined, $\langle\overline{\mathbf{R}},e\rangle$-definable and continuous for $\beta\in(0,b^{-1})$ and, by direct substitution, is a solution to the system $\mathcal{S}$. It is also easy to calculate the determinant of the Jacobian of $\mathcal{S}$ with respect to $x_2,\dots x_{n+r+1}$ at this solution. It is a product of a non-zero quantity with

$$\det\frac{\partial(f_1,\dots f_{n-1})}{\partial(x_2,\dots,x_n)}(\beta^{-1},\phi_2(\beta^{-1}),\dots,\phi_n(\beta^{-1})),$$

and is therefore non-zero.

Finally, we have clearly guaranteed that $\psi_i(\beta)\to 0$ as $\beta\to 0^+$ (for $i=2,\dots,n+r+1$) so it follows from Theorem 3.1 that there exists

$q' \in \tau'(n+r+1, n+r, 1+m(n+1))$ such that $\beta^{q'}\psi_{n+r+1}(\beta) \to$ a finite non-zero limit as $\beta \to 0^+$ (if ψ_{n+r+1} is not identically zero). The conclusion of (A7) now follows (by definition of ψ_{n+r+1}) with $q = \pm q'$.

The proof of 3.1 uses the following lemma which is essentially due to Khovanskii [1991].

Lemma 3.2 *There exists a (primitive) recursive function $\mu : \mathbf{N}^2 \to \mathbf{N}$ such that if $m, n \in \mathbf{N}\backslash\{0\}$ and $h_1(z_1, ., z_n), \ldots, h_n(z_1, ., z_n)$ are complex functions expressible as polynomials (with complex coefficients) of degrees $\leq m$ in $z_1, \ldots, z_n$, $\exp(z_1, \ldots, \exp(z_n)$, then the system $h_1(\overline{z}) = \cdots = h_n(\overline{z}) = 0$ has at most $\mu(n, m)$ non-singular solutions in the unit polydisc in $\mathbf{C}^n$.*

Proof. Given such $n, m, h_1, \ldots, h_n$, write

$$z_i = x_i + \sqrt{-1}y_i$$

and

$$h_i(\overline{z}) = f_i(\overline{x}, \overline{y}) + \sqrt{-1}g_i(\overline{x}, \overline{y})$$

for $i = 1, \ldots, n$, where the f_i and g_i are real functions of the $2n$ real variables $x_1, \ldots, x_n, y_1, \ldots, y_n$. Clearly the f_i and g_i can be expressed as (real) polynomials in x_j, y_j, $\exp(x_j)$, $\sin(y_j)$, $\cos(y_j)$ (for $j = 1, \ldots, n$) of degrees $\leq m$. By a result of Khovanskii [1991] (§1.4) there are at most $\mu'(2n, m)$ non-singular solutions to the system $f_i(\overline{x}, \overline{y}) = 0, g_i(\overline{x}, \overline{y}) = 0$ $(i = 1, \ldots, n)$ in the unit polydisc in $\mathbf{R}^{2n}$, where μ' is an explicitly given primitive recursive function. But it easily follows from the Cauchy-Riemann equations that any non-singular solution in $\mathbf{C}^n$ of our original system gives rise to a non-singular solution in $\mathbf{R}^{2n}$ of the new system. □

Now suppose $n, r, m, f_1, \ldots, f_{n-1}, \epsilon$ are as in the hypotheses of 3.1. We *claim* that whenever $\phi_2, \ldots, \phi_n$ are as in (iii) and (iv) of 3.1, there exist $c \in \mathbf{Z}$ and $d \in \mathbf{N}\backslash\{0\}$ with $d \leq \mu(n-1, m)$ (μ as in 3.2) such that $\alpha^{c/d} \cdot \phi_n(\alpha)$ tends to a finite, non-zero limit as $\alpha \to 0^+$ (provided that ϕ_n is not identically zero).

Proof of claim. Given such ϕ_j's it follows from van den Dries's observation and (iv) of 3.1 that there exist $\varepsilon_0 \in \mathbf{R}, 0 < \varepsilon_0 < \varepsilon$, real analytic functions $\theta_2, \ldots, \theta_n : (-\varepsilon_0, \varepsilon_0) \to \mathbf{R}$ and $\alpha \in \mathbf{N} \setminus \{0\}$ such that

$$\phi_j(\alpha) = \theta_j(\alpha^{1/\alpha})$$

for $\alpha \in (0, \varepsilon_0)$ and $j = 2, \ldots, n$ (where the positive d'th root is taken). Let us suppose we have chosen the smallest possible d here. It clearly suffices to show that $d \leq \mu(n-1, m)$.

Now since we may assume $\varepsilon_0 < 1$, we have

$$\phi_j(\alpha^d) = \theta_j(\alpha)$$

for $\alpha \in (0, \varepsilon_0), j = 2, \ldots, n$ and hence

$$f_i(\alpha^d, \theta_2(\alpha), \ldots, \theta_n(\alpha)) = 0$$

for $i = 1, ..., n-1$. Further,

$$J(\alpha^d, \theta_2(\alpha), \ldots, \theta_n(\alpha)) \neq 0$$

for $\alpha \in (0, \varepsilon_0)$, where

$$J(x_1, \ldots, x_n) = \det \frac{\partial(f_1, \ldots, f_{n-1})}{\partial(x_2, \ldots, x_n)}.$$

Now (by reducing ε_0 if necessary), the θ_j's can be extended to analytic functions, which we shall also denote by θ_j, on

$$D = \{z \in \mathbf{C} : \mid z \mid \ < \varepsilon_0\}$$

(by using the same Taylor series) and, of course, the f_i's have natural analytic extensions to $\mathbf{C}^n$. By analytic continuation it follows that

$$f_i(z^d, \theta_2(z), \ldots, \theta_n(z)) = 0$$

for $i = 1, \ldots, n-1$ and $z \in D$. Further, the (analytic) function

$$z \mapsto J(z^d, \theta_2(z), \ldots, \theta_n(z)) \ (z \in D)$$

is not identically zero on D and hence has only finitely many zeros on each compact subset of D. We may therefore suppose (by reducing ε_0 if necessary) that this function is non-zero throughout $D \setminus \{0\}$. Notice also that for any $w_0 \in D$ we clearly have

$$J(w_0^d, \theta_2(w_0), \ldots, \theta_n(w_0)) = \\ \left(\det \frac{\partial(f_1, \ldots, f_{n-1})}{\partial(z_2, \ldots, z_n)}\right) (w_0^d, \theta_2(w_0), \ldots, \theta_n(w_0)).$$

Now for $w \in D$ and $\nu = 0, \ldots, d-1$ define

$$\eta_\nu(w) = w \exp\left(\frac{2\pi\sqrt{-1}\nu}{d}\right) (\in D),$$

and for $i = 1, \dots, n-1$ define

$$h_i^w : \mathbf{C}^{n-1} \to \mathbf{C}, \langle z_1, \dots z_{n-1} \rangle \mapsto f_i(w^d, z_1, \dots, z_{n-1}).$$

By the above remarks it follows that for any $w \in D \backslash \{0\}$ and $\nu = 0, \dots, d-1$, the element

$$\lambda_\nu(w) =_{def} \langle \theta_2(\eta_\nu(w)), \dots, \theta_n(\eta_\nu(w)) \rangle$$

of $\mathbf{C}^{n-1}$ is a non-singular zero of the system

$$h_1^w(z_1, \dots, z_{n-1}) = \cdots = h_{n-1}^w(z_1, \dots, z_{n-1}) = 0.$$

Since the h_i^w's can clearly be expressed as polynomials in $z_1, \dots, z_{n-1}$, $\exp(z_1), \dots, \exp(z_{n-1})$ of degrees $\leq m$ (for any w and d), we shall be done, by 3.2, if we can find $w \in D \backslash \{0\}$ such that $\lambda_0(w), \dots, \lambda_{d-1}(w)$ are distinct. (Note also that since, by (iv) of 3.1, $\theta_j(0) = 0$ for $j = 2, \dots, n$, we can suppose that ε_0 has been chosen small enough so that $|\theta_j(z)| < 1$ for $z \in D$ and $j = 2, \dots, n$.)

Now suppose there is no such w as required above. Then since each function

$$w \mapsto \theta_j(\eta_\nu(w))$$

is analytic on D it follows that for some ν_0, ν_1 with $0 \leq \nu_0 < \nu_1 \leq d-1$,

$$\lambda_{\nu_0}(w) = \lambda_{\nu_1}(w) \text{ for } \textit{all } w \in D.$$

Hence, for all $p \in \mathbf{N}$ and $j = 2, \dots, n$

$$\frac{d^p}{dw^p}(\theta_j(\eta_{\nu_0}(w)) - \theta_j(\eta_{\nu_1}(w))) \mid_{w=0} = 0,$$

i.e.

$$\theta_j^{(p)}(0)(\exp\left(\frac{2\pi\sqrt{-1}p(\nu_0 - \nu_1)}{d}\right) - 1) = 0.$$

Let $d' = g.c.d.(\nu_0 - \nu_1, d)$ and $d'' = d/d' (\in \mathbf{N})$. Then $d'' > 1, d'' \mid d$ and for all $p \in \mathbf{N}$ and $j = 2, \dots, n, \theta_j^{(p)}(0) \neq 0$ implies $d'' \mid p$. Thus each $\theta_j(z)$ has Taylor series (around 0) of the form $\sum_{\ell=0}^{\infty} a_{j,\ell} \cdot z^{\ell d''}$, convergent for $\mid z \mid < \varepsilon_1$ say, where $0 < \varepsilon_1 \leq \varepsilon_0$. But then $\sum_{\ell=0}^{\infty} a_{j,\ell} z^\ell$ is convergent for $\mid z \mid < \varepsilon_1^{d''}$ so defines an analytic function, $\tilde{\theta}_j(z)$ say, here. Clearly

$$\tilde{\theta}_j(z^{d''}) = \theta_j(z)$$

for $\mid z \mid < \varepsilon_1$ and so

$$\phi_j(\alpha) = \tilde{\theta}_j(\alpha^{d''/d})$$

for all $\alpha \in (0, \varepsilon_1^d)$ and $j = 2, \ldots, n$, which contradicts the minimality of d and establishes the claim. $\Box$

Theorem 3.1 can now be established easily from the claim. For suppose $n, r, m, f_1, \ldots, f_{n-1}, \varepsilon$ are as in the hypotheses of 3.1 and that $\phi_2, \ldots, \phi_n$ are as in (iii) and (iv) of 3.1. By the initial observation in the proof of the claim we may suppose (by reducing ε if necessary) that (each $\phi_i(t)$ is analytic on $(0, \varepsilon)$ and) $\frac{d\phi_n}{dt}(t) \neq 0$ for all $t \in (0, \varepsilon)$. This implies, by the chain rule and elementary linear algebra, that

$$\det \frac{\partial(f_1, \ldots, f_{n-1})}{\partial(x_1, \ldots, x_{n-1})}(t, \phi_2(t), \ldots, \phi_n(t)) \neq 0$$

for all $t \in (0, \varepsilon)$. It also implies (by (iv) of 3.1) that ϕ_n is a bijection from $(0, \varepsilon)$ to $(0, \varepsilon')$ or $(-\varepsilon', 0)$ for some $\varepsilon' \in \mathbf{R}, \varepsilon' > 0$. We may suppose the former by replacing x_n with $-x_n$ in each f_i and multiplying by $\exp(mx_n)$ – which increases m to at most $2m$. Thus the map

$$(0, \varepsilon') \to \mathbf{R}^n, \qquad t \mapsto \langle \phi_n^{-1}(t), \phi_2(\phi_n^{-1}(t)), \ldots, \phi_{n-1}(\phi_n^{-1}(t)), t \rangle$$

is continuous and definable and also parameterizes a solution of the system

$$f_1(\overline{x}) = \cdots = f_{n-1}(\overline{x}) = 0$$

which is non-singular with respect to $x_1, \ldots, x_{n-1}$. It follows from the claim (by interchanging the variables x_1, x_n) that for some $c' \in \mathbf{Z}$ and $d' \in \mathbf{N} \backslash \{0\}$ with $d' \leq \mu(n-1, 2m)$, we have $\alpha^{c'/d'} \cdot \phi_n^{-1}(\alpha) \to$ a finite, non-zero limit as $\alpha \to 0^+$. But we also have, by the claim for the original ϕ_j's, that for some $c \in \mathbf{Z}$ and $d \in \mathbf{N} \backslash \{0\}$ with $d \leq \mu(n-1, m)$, that $\alpha^{c/d} \cdot \phi_n(\alpha) \to$ a finite, non-zero limit as $\alpha \to 0^+$. Necessarily $\frac{c}{d} = \frac{d'}{c'}$ (and $c' \neq 0$), so we may take

$$\tau' = \tau'(n, r, m) = \{\frac{u}{v} \in \mathbf{Q} : u, v \in \mathbf{Z}, v \neq 0, \mid u \mid, \mid v \mid \leq \mu(n-1, 2m)\}$$

to complete the proof of 3.1, and also of (A7) for models of T_e.

4 On the Decidability of $T_{\exp}$

Recall from Theorem 2.6 that $T_{\exp}$ is axiomatized by $T_{unres} \cup \mathcal{E}_{unres}$ where T_{unres} is recursive and $\mathcal{E}_{unres}$ denotes the existential theory of $\langle \overline{\mathbf{R}}, \exp \rangle$. The decidability of $T_{\exp}$ therefore follows from that of $\mathcal{E}_{unres}$ and, as stated in our introduction, we shall establish the latter under the assumption SC. We do not, however, assume SC for the moment.

For $n \in \mathbf{N}\backslash\{0\}$ let M_n denote the ring

$$\mathbf{Z}[x_1, \ldots, x_n, \exp(x_1), \ldots, \exp(x_n)]$$

(regarded as a ring of functions from $\mathbf{R}^n$ to $\mathbf{R}$). Let $\mathcal{E}'_{unres}$ denote the set of all sentences of the form $\exists x_1, \ldots, x_n f(x_1, \ldots x_n) = 0$, for $n \in \mathbf{N}\backslash\{0\}$ and $f \in M_n$, which are true in $\langle \overline{\mathbf{R}}, \exp\rangle$. (From now on we do not distinguish notationally between terms of the language $\overline{\overline{L}}$ and the functions they represent in $\langle \overline{\mathbf{R}}, \exp\rangle$.) Then it is easy to show (see Wilkie [199?], Section 2) that

$$T_{RCF} \cup \mathcal{E}'_{unres} \vdash \mathcal{E}_{unres}$$

(where T_{RCF} denotes the theory of real closed ordered fields formulated in the sublanguage L of $\overline{\overline{L}}$). Thus the whole problem reduces to that of effectively determining when an element f of M_n (for $n \in \mathbf{N}\backslash\{0\}$) has a zero in $\mathbf{R}^n$. We first show that this can be done in the case of non-singular zeros (various other cases have also been considered in the literature, see for example Vorobjov [1991]), for which we need the following effective version of Newton Approximation.

Theorem 4.1 *There exists an effective procedure which, given $n, N \in \mathbf{N}\backslash\{0\}$ and $f = \langle f_1, \ldots, f_n\rangle \in M_n^n$, produces $\theta = \theta(n, N, f_1, \ldots, f_n) \in \mathbf{N}\backslash\{0\}$ such that whenever $\overline{\alpha} \in \mathbf{R}^n, \| \overline{\alpha} \| < N, \| f(\overline{\alpha}) \| < \theta^{-1}$ and $| \det \frac{\partial(f_1,\ldots,f_n)}{\partial(x_1,\ldots,x_n)}(\overline{\alpha})| > N^{-1}$, then there exists $\overline{\gamma} \in \mathbf{R}^n$ such that $f(\overline{\gamma}) = 0$ (and $\| \overline{\alpha} - \overline{\gamma} \| < N^{-1}$).*

(Here, $\| \, . \, \|$ denotes the sup norm: $\| \langle \beta_1, \ldots, \beta_n\rangle \| = \max_{1 \le i \le n} | \beta_i |$.)

Given such a recursive function θ we consider the $\overline{\overline{L}}$- theory T_{NA} consisting of the sentences expressing the conclusion of 4.1 (one sentence for each choice of $n, N, f_1, \ldots, f_n$). We also let T_H consist of all those sentences of the form $f(q_1, \ldots, q_n) > 0$ (for $n \in \mathbf{N}\backslash\{0\}, f \in M_n$ and $q_1, \ldots, q_n \in \mathbf{Q}$) that are true in $\langle \overline{\mathbf{R}}, \exp\rangle$.

Theorem 4.2 *T_H is recursive.*

Granted 4.1 and 4.2 we now prove

Theorem 4.3 *$T_{RCF} \cup T_H \cup T_{NA}$ is a recursive subtheory of $T_{\exp}$. It has the property that whenever $n \in \mathbf{N}\backslash\{0\}, f_1, \ldots, f_n \in M_n$ and the system*

$$f_1(\overline{x}) = \cdots = f_n(\overline{x}) = 0$$

has a non-singular solution in $\mathbf{R}^n$*, then*

$$T_{RCF} \cup T_H \cup T_{NA} \vdash \exists \overline{x} \bigwedge_{i=1}^{n} f_i(\overline{x}) = 0.$$

Proof. The first assertion is clear.

Suppose $n \in \mathbf{N}\backslash\{0\}, \overline{\beta} \in \mathbf{R}^n, f_1(\overline{\beta}) = \cdots = f_n(\overline{\beta}) = 0$ and $\det \frac{\partial(f_1,\ldots,f_n)}{\partial(x_1,\ldots,x_n)}(\overline{\beta}) \neq 0$. Choose $N \in \mathbf{N}\backslash\{0\}$ such that $\| \overline{\beta} \| < N$ and $|\det \frac{\partial(f_1,\ldots,f_n)}{\partial(x_1,\ldots,x_n)}(\overline{\beta})| > N^{-1}$. Let $\theta = \theta(n, N, f_1, \ldots, f_n)$ and choose $\overline{q} \in \mathbf{Q}^n$ so close to $\overline{\beta}$ that $\| \overline{q} \| < N, \| \langle f_1(\overline{q}), \ldots, f_n(\overline{q}) \rangle \| < \theta^{-1}$, and $\left|\det \frac{\partial(f_1,\ldots,f_n)}{\partial(x_1,\ldots,x_n)}(\overline{q})\right| > N^{-1}$. The three inequalities here are all consequences of $T_{RCF} \cup T_H$ and hence the sentence $\exists \overline{x} \bigwedge_{i=1}^{n} f_i(\overline{x}) = 0$ is a consequence of $T_{RCF} \cup T_H \cup T_{NA}$. □

Proof of Theorem 4.1 We assume familiarity with the basics of the theory of differentiation in Banach Spaces, which can be found in, for example, Lang [1983].

For E_1, E_2 Banach spaces we denote by $L(E_1, E_2)$ the Banach space of all continuous, linear maps from E_1 to E_2. We shall need the following version of *Newton Approximation for Banach spaces (NABS)* (see Lang [1983]):

> Suppose $\langle E, \| \cdot \| \rangle$ is a (real) Banach space, $\delta \in \mathbf{R}$, $0 < \delta \leq \frac{1}{2}$, $x_0 \in E$, and let $B = \{x \in E : \| x_0 - x \| < \delta\}$. Suppose $f : B \to E$ is a C^2 function such that $\| f(x_0) \| < \delta^4$. Suppose further that for each $x \in B, f'(x)$ is invertible, $\| f'(x)^{-1} \|_1 < \delta^{-1}$ and $\| f''(x) \|_2 < \delta^{-1}$, where $\| \cdot \|_1$ and $\| \cdot \|_2$ denote the norms in $L(E, E)$ and $L(E, L(E, E))$ respectively. Then there exists $y \in B$ such that $f(y) = 0$.

Now let $n, N \in \backslash\{0\}$ and $f = \langle f_1, \ldots, f_n \rangle \in M_n^n$ be given. We shall apply NABS to the Banach space $\mathbf{R}^n$ (with the sup norm) and the (C^∞) function $f : \mathbf{R}^n \to \mathbf{R}^n$.

Let

$$g(\overline{x}) = \det \frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)},$$

so that $g \in M_n$ and g can be effectively found from $f_1, \ldots, f_n$. By formal inversion of $n \times n$ matrices we can effectively find $g_{i,j} \in M_n$ (for $1 \leq i, j \leq n$) such that

$$g \cdot I_n = \frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)} \cdot (g_{i,j})_{1 \leq i,j \leq n}$$

(identically in $\overline{x}$), where I_n denotes the identity $n \times n$ matrix. The assertion that $f'(\overline{\beta})$ is invertible (for $\overline{\beta} \in \mathbf{R}^n$) is equivalent to the assertion that $g(\overline{\beta}) \neq 0$, and when this holds we have, by the usual estimates

$$\| f'(\overline{\beta})^{-1} \|_1 \leq |\, g(\overline{\beta}) \,|^{-1} \cdot n \cdot \max\{ |\, g_{i,j}(\overline{\beta}) \,| \; : 1 \leq i, j \leq n\}.$$

Now, by a crude estimation, we can effectively find $\theta_0 \in \mathbf{N}\backslash\{0\}$ such that

$$n \cdot |\, g_{i,j}(\overline{\beta}) \,| \leq \theta_0$$

for $1 \leq i, j \leq n$ and all $\overline{\beta} \in \mathbf{R}^n$ with $\| \overline{\beta} \| \leq N + 1$. Similarly, we can effectively find $\theta_1 \in \mathbf{N}\backslash\{0\}$ such that

$$n \cdot \left| \frac{\partial g}{\partial x_i}(\overline{\beta}) \right| \leq \theta_1$$

for $1 \leq i \leq n$ and all $\overline{\beta} \in \mathbf{R}^n$ with $\| \overline{\beta} \| < N + 1$. (Note that $\frac{\partial g}{\partial x_i} \in M_n$ and that $\frac{\partial g}{\partial x_i}$ can be effectively found from $f_1, \ldots, f_n$ for $1 \leq i \leq n$.) Now it follows from the mean value theorem that

$$|\, g(\overline{\alpha}) - g(\overline{\beta}) \,| \leq \| \overline{\alpha} - \overline{\beta} \| \cdot \theta_1$$

and hence that

$$|\, g(\overline{\beta}) \,| \geq |\, g(\overline{\alpha}) \,| - \| \overline{\alpha} - \overline{\beta} \| \cdot \theta_1$$

for all $\overline{\alpha}, \overline{\beta} \in \mathbf{R}^n$ with $\| \overline{\alpha} \|, \| \overline{\beta} \| < N + 1$. Hence we obtain that if $\overline{\alpha}, \overline{\beta} \in \mathbf{R}^n$, $\| \overline{\alpha} \|, \| \overline{\beta} \| < N + 1$ and $|\, g(\overline{\alpha}) \,| - \| \overline{\alpha} - \overline{\beta} \| \cdot \theta_1 > 0$, then $f'(\overline{\beta})$ is invertible and

$$\| f'(\overline{\beta})^{-1} \|_1 \leq \theta_0 \cdot (|\, g(\overline{\alpha}) \,| - \| \overline{\alpha} - \overline{\beta} \| \cdot \theta_1)^{-1}.$$

Finally, effectively find $\theta_2 \in \mathbf{N}\backslash\{0\}$ such that

$$n^2 \left| \frac{\partial^2 f_\ell}{\partial x_i \partial x_j}(\overline{\beta}) \right| \leq \theta_2$$

for all $\overline{\beta} \in \mathbf{R}^n$ with $\| \overline{\beta} \| < N + 1$. This guarantees that

$$\| f''(\overline{\beta}) \|_2 \leq \theta_2$$

for all such $\overline{\beta}$. Now set

$$\theta = (2N \max\{\theta_0, \theta_1, \theta_2\})^4.$$

Then θ is effective in the initial data $n, N, f_1, \ldots, f_n$. We claim that the conclusion of 4.1 holds.

For suppose $\overline{\alpha} \in \mathbf{R}^n, \| \overline{\alpha} \| < N, \| f(\overline{\alpha}) \| < \theta^{-1}$ and $\mid g(\overline{\alpha}) \mid \; > N^{-1}$. Define

$$\delta = (2N \max\{\theta_0, \theta_1, \theta_2\})^{-1}.$$

Then $0 < \delta \leq \frac{1}{2}$ and $\| f(\overline{\alpha}) \| < \delta^4$. Suppose $\overline{\beta} \in \mathbf{R}^n$ satisfies $\| \overline{\alpha} - \overline{\beta} \| < \delta$. Then $\| \overline{\beta} \| < N + 1$, so

$$\| f''(\overline{\beta}) \|_2 \leq \theta_2 < \delta^{-1}.$$

Further,

$$\mid g(\overline{\alpha}) \mid - \| \overline{\alpha} - \overline{\beta} \| \cdot \theta_1 > N^{-1} - \delta\theta_1 \geq (2N)^{-1}$$

so, by the above, $f'(\overline{\beta})$ is invertible and

$$\| f'(\overline{\beta})^{-1} \|_1 < 2N\theta_0 \leq \delta^{-1}.$$

It now follows from NABS that f has a zero $\overline{\gamma} \in \mathbf{R}^n$ satisfying

$$\| \overline{\alpha} - \overline{\gamma} \| < \delta < N^{-1}. \quad \square$$

Turning now to Theorem 4.2 it is clear enough that the problem reduces to effectively determining the sign of real numbers of the form

$$(^*) \qquad \sigma = \sum_{i=0}^{m} a_i e^{i/q}$$

from the data $m \in \mathbf{N}, q \in \mathbf{N} \backslash \{0\}$ and $a_0, \ldots, a_m \in \mathbf{Z}$. (We now use e to denote $\exp(1)$.) But it is more or less obvious that this can be done, because either e is algebraic, in which case the statements $\sigma > 0, \sigma = 0, \sigma < 0$ can be expressed as L-sentences and exactly one is a consequence of T_{RCF}, or else it is not, in which case $\sigma = 0$ if and only if $a_0 = \cdots = a_m = 0$ and if $\sigma \neq 0$ then the sign of σ can be determined by successively approximating $e^{1/q}$ using a recursive sequence of rationals converging effectively to $e^{1/q}$. Of course, we know by Hermite's work that e is transcendental and it seems appropriate to extract an algorithm from a proof of this fact. (We also get a *primitive* recursive procedure this way.) The following is based on Baker's account in [1975].

Proof of Theorem 4.2 Let $m, q, a_0, \ldots, a_m$ be given as in $(^*)$ above. Clearly we may suppose that $a_0 \neq 0$ and $m \geq 1$.

Let p be the least prime number greater than

$$q + 36m^{2(m+1)} + \sum_{i=0}^{m} |a_i|$$

and define

$$\begin{aligned} f(t) &= t^{p-1}(t-\frac{1}{q})^p \cdots (t-\frac{m}{q})^p, \\ r &= \deg(f) = (m+1)p-1, \\ \tilde{f}(t) &= \sum_{j=0}^{r} f^j(t), \text{and} \\ H &= \sum_{i=0}^{m} a_i \tilde{f}(\frac{i}{q}). \end{aligned}$$

Then $H, \tilde{f}(0)$ are rational numbers which can be found effectively from the initial data. We *claim* that they are both non-zero and that σ has the same sign as $H/\tilde{f}(0)$. By the remarks before $^{(*)}$ above, this completes the proof of Theorem 4.2. □

To prove the claim consider the function

$$I(x) = e^x \int_0^x f(t)e^{-t}dt.$$

Successive integration by parts yields

$$I(x) = \tilde{f}(0)e^x - \tilde{f}(x),$$

and hence

$$(\dagger) \qquad \sum_{i=0}^{m} a_i I(\frac{i}{q}) = \tilde{f}(0)\sigma - H.$$

Now it is easy to see that for $0 \le j \le r$ and $0 \le i \le m, q^r f^{(j)}(\frac{i}{q})$ is an integer divisible by $(p-1)!$ and that it is also divisible by p if and only if $j = p-1$ and $i = 0$ (note that $p > q$). It follows that p does not divide the integer $q^r \tilde{f}(0)$, and hence $\tilde{f}(0) \neq 0$, and also that p does not divide the integer $q^r H$ (note that $p > a_0$) but that $(p-1)!$ does, and hence $\mid q^r H \mid \ \ge (p-1)!$ (so $H \neq 0$).

Now by crudely estimating we see that for $0 \le x \le \frac{m}{q}$,

$$\mid I(x) \mid \ \le 3^{m/q} \cdot \left(\frac{m}{q}\right)^{r+1}.$$

Therefore, using also the definition of p, we obtain

$$\begin{aligned} \left| q^r \sum_{i=0}^{m} a_i I(\frac{i}{q}) \right| &\leq 3^{\frac{m}{q}} \cdot m^{r+1} \cdot \sum_{i=0}^{m} \mid a_i \mid \\ &\leq (3 \cdot m^{m+1})^p \cdot p \\ &\leq (6 \cdot m^{m+1})^p \\ &< p^{\frac{p}{2}} \leq (p-1)! \\ &\leq \mid q^r H \mid . \end{aligned}$$

It follows that

$$\left| \sum_{i=0}^{m} a_i I(\frac{i}{q}) \right| < \mid H \mid$$

which, together with (†), proves the claim. □

Now consider an arbitrary $g \in M_n$ having a zero in $\mathbf{R}^n$. Since the structure $\langle \overline{\mathbf{R}}, \exp \rangle$ lies in the class $\mathcal{K}_2$ (in the unrestricted case - see Section 2) it follows from Theorem 2.1 (with $k = \mathbf{Q}, n = r$ and $S = V = \{\overline{\alpha} \in \mathbf{R}^m : g(\overline{\alpha}) = 0\}$) that there exist

$$f_1, \ldots, f_n \in M_{n,n}(\mathbf{Q}) (= \mathbf{Q}[x_1, \ldots, x_n, e^{x_1}, \ldots, e^{x_n}])$$

and $\overline{\alpha} \in \mathbf{R}^n$ such that $f(\overline{\alpha}) = 0$ *and* $\overline{\alpha}$ is a non-singular solution of the system

$$f_1(\overline{x}) = \cdots = f_n(\overline{x}) = 0.$$

Clearly we may suppose that $f_1, \ldots, f_n \in M_n$ so it follows from Theorem 4.3 that this system has a solution in K^n for any model K of $T_{RCF} \cup T_H \cup T_{NA}$. In fact, one may even suppose that there is a unique such solution. But why should it also be a zero of g? This, of course, is where we need Schanuel's conjecture.

Proof of Theorem 1.1 We show that SC implies that the recursive theory $T_{unres} \cup T_H \cup T_{NA}$ axiomatizes $T_{\exp}$.

Let $n \in \mathbf{N} \backslash \{0\}$ and $g \in M_n$ and suppose that g has a zero in $\mathbf{R}^n$. It is sufficient to show, by the comments at the beginning of this section, that g has a zero in K^n for any model K of $T_{unres} \cup T_H \cup T_{NA}$. We may suppose that all zeros of g in $\mathbf{R}^n$ have coordinates linearly independent over $\mathbf{Q}$. For if $\overline{\alpha} = \langle \alpha_1, \ldots, \alpha_n \rangle \in \mathbf{R}^n$, $g(\overline{\alpha}) = 0$ and, say

$$c_n \alpha_n = c_1 \alpha_1 + \cdots + c_{n-1} \alpha_{n-1}$$

with $c_1, \ldots, c_n \in \mathbf{Z}$ and $c_n \neq 0$, then define

$$h(x_1, \ldots, x_{n-1}) = \\ \Pi_{i=1}^{n-1} e^{\mid c_i \mid x_i} \cdot g(c_n x_1, \ldots, c_n x_{n-1}, c_1 x_1 + \cdots + c_{n-1} x_{n-1}).$$

Clearly $h \in M_n$, $\langle \alpha_1 c_n^{-1}, \ldots, \alpha_{n-1} c_n^{-1} \rangle$ is a zero of h and, further, any zero of h in K^n, for $K \models T_{unres} \cup T_H \cup T_{NA}$, gives rise to a zero of g in that model. We may continue this reduction until the supposition holds.

We now introduce the following notation. For $h(x_1, \ldots, x_n) \in M_n$ let $\tilde{h}(x_1, \ldots, x_{2n})$ be the (unique) element of $\mathbf{Z}[x_1, \ldots, x_{2n}]$ satisfying

$$h(x_1, \ldots, x_n) \equiv \tilde{h}(x_1, \ldots, z_n, e^{x_1}, \ldots, e^{x_n}).$$

For $\overline{\beta} = \langle \beta_1, \ldots, \beta_n \rangle \in \mathbf{R}^n$ let

$$\tilde{\beta} = \langle \beta_1, \ldots, \beta_n, e^{\beta_1}, \ldots, e^{\beta_n} \rangle (\in \mathbf{R}^{2n})$$

and let $\tilde{x}$ be the $2n$-tuple of variables $\langle x_1, \ldots, x_{2n} \rangle$. (The n-tuple $\langle x_1, \ldots, x_n \rangle$ is still denoted $\overline{x}$.).

It is easy to show that for any $h_1, \ldots, h_n \in M_n$ and $\overline{\beta} \in \mathbf{R}^n, \overline{\beta}$ is a non-singular solution of the system

$$h_1(\overline{x}) = \cdots = h_n(\overline{x}) = 0$$

if and only if $\tilde{\beta}$ is a non-singular solution of the system

$$\tilde{h}_1(\tilde{x}) = \cdots = \tilde{h}_n(\tilde{x}) = 0.$$

Now, returning to our g above, choose $f_1, \ldots, f_n \in M_n$ and $\overline{\alpha} \in \mathbf{R}^n$ as described in the remarks preceding this proof. Then $\tilde{\alpha}$ is a non-singular solution of the system

$$\tilde{f}_1(\tilde{x}) = \cdots = \tilde{f}_n(\tilde{x}) = 0,$$

from which it follows (using elementary differential algebra - see Lang [1965], chapter 10, §7) that the field $\mathbf{Q}(\tilde{\alpha})$ has transcendence degree (over $\mathbf{Q}$) at most n. By SC this degree is exactly n (since $g(\overline{\alpha}) = 0$, so $\alpha_1, \ldots, \alpha_n$ are linearly independent over $\mathbf{Q}$).

Define

$$P = \{h(\tilde{x}) \in \mathbf{Z}[\tilde{x}] : h(\tilde{\alpha}) = 0\}.$$

Then P is a prime ideal of $\mathbf{Z}[\tilde{x}]$ and the field of fractions of the domain $\mathbf{Z}[\tilde{x}]/p$ is isomorphic to $\mathbf{Q}(\tilde{\alpha})$. It follows that there exists $p_0 \in \mathbf{Z}[\tilde{x}] \backslash P$ such that the ideal $p_0 P$ is generated by n elements, $p_1, \ldots, p_n$ say. (This is a consequence of the fact, easily proved by induction on $m \in \mathbf{N}\backslash\{0\}$, that if Q is a prime ideal of $\mathbf{Z}[x_1, \ldots, x_m]$ such that $Q \cap \mathbf{Z} = \emptyset$ and such that the field of fractions of $\mathbf{Z}[x_1, \ldots, x_m]/Q$ has transcendence degree r, then for some $h \in \mathbf{Z}[x_1, \ldots, x_m]\backslash Q, hQ$ is generated by $m - r$ elements.)

Now choose $a_{i,j}(\tilde{x}) \in \mathbf{Z}[\tilde{x}]$ $(1 \leq i, j \leq n)$ such that

$$p_0(\tilde{x})\tilde{f}_i(\tilde{x}) \equiv \sum_{j=1}^{n} a_{i,j}(\tilde{x})p_j(\tilde{x}) \; (1 \leq i \leq n).$$

By differentiating these identities and using the fact that $p_0(\tilde{\alpha}) \neq 0$, it easily follows that α is a non-singular solution of the system

$$p_1(\tilde{x}) = \cdots = p_n(\tilde{x}) = 0$$

and hence that $\overline{\alpha}$ is a non-singular solution of the system

$$g_1(\overline{x}) = \cdots = g_n(\overline{x}) = 0,$$

where

$$g_i(x_1, \ldots, x_n) =_{\text{def}} p_i(x_1, \ldots x_n, e^{x_1}, \ldots, e^{x_n}),$$

i.e. $\tilde{g}_i = p_i$, for $0 \leq i \leq n$.

Now consider the system

$$g_1(\overline{x}) = \cdots = g_n(\overline{x}) = x_{n+1}g_0(\overline{x}) - 1 = 0.$$

The functions here lie in M_{n+1} and by a routine check $\langle \overline{\alpha}, g_0(\overline{\alpha})^{-1} \rangle$ is a non-singular solution. Now if $K \models T_{unres} \cup T_H \cup T_{NA}$ it follows from Theorem 4.3 (as $T_{RCF} \subseteq T_{unres}$) that this system has a solution, $\langle \overline{\gamma}, \gamma_{n+1} \rangle$ say, in K^{n+1}. So

$$K \models g_i(\overline{\gamma}) = 0$$

for $1 \leq i \leq n$ and

$$K \models g_0(\overline{\gamma}) \neq 0.$$

However, recall that our original $g \in M_n$ satisfies $g(\overline{\alpha}) = 0$. So $\tilde{g} \in P$ and hence there exist $a_1(\tilde{x}), \ldots, a_n(\tilde{x}) \in \mathbf{Z}[\tilde{x}]$ such that

$$p_0(\tilde{x}) \cdot \tilde{g}(\tilde{x}) \equiv \sum_{j=1}^{n} a_j(\tilde{x}) \cdot p_j(\tilde{x}).$$

But this (polynomial) identity must also hold in K, so setting $x_i = \gamma_i$ and $x_{n+i} = e^{\gamma_i}$ for $1 \leq i \leq n$ we see that $g(\overline{\gamma}) = 0$. Thus g has a zero in K^n for any model K of $T_{unres} \cup T_H \cup T_{NA}$ as required. $\square$

5 Concluding Remarks

We consider the following conjecture which can be regarded (unhelpfully, it seems) as a zero-dimensional version of Theorem 3.1.

Weak Schanuel's Conjecture (WSC) *There exists an effective procedure which, given $n \in \mathbf{N}\backslash\{0\}$ and*

$$f_1, \ldots, f_n, g \in M_n (= \mathbf{Z}[x_1, \ldots x_n, e^{x_1}, \ldots, e^{x_n}]),$$

produces

$$\eta = \eta(n, f_1, \ldots, f_n, g) \in \mathbf{N}\backslash\{0\}$$

with the property that whenever $\overline{\alpha} \in \mathbf{R}^n$ is a non-singular solution of the system

$$f_1(\overline{x}) = \cdots = f_n(\overline{x}) = 0,$$

then either $g(\overline{\alpha}) = 0$ or $\mid g(\overline{\alpha}) \mid\ > \eta^{-1}$.

It is easy to see that the decidability of T_{exp} implies WSC. For suppose $n, f_1, \ldots, f_n, g$ are given as in the conjecture. Then by Khovanskii's theorem mentioned earlier there are only finitely many non-singular solutions to the system

$$f_1(\overline{x}) = \cdots = f_n(\overline{x}) = 0$$

and hence for some $\ell \in \mathbf{N}\backslash\{0\}$ the $\overline{\overline{L}}$-sentence $\Phi_\ell(n, f_1, \ldots, f_n)$:

$$\forall \overline{x} \left[(\bigwedge_{i=1}^{n} f_i(\overline{x}) = 0 \ \wedge \ \det \frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)}(\overline{x}) \neq 0) \right.$$
$$\left. \rightarrow (g(\overline{x}) = 0 \vee \mid g(\overline{x}) \mid\ > \ell^{-1}) \right]$$

is in T_{exp}. Now let $\eta =$ the least $\ell \in \mathbf{N}\backslash\{0\}$ s.t. $\Phi_\ell(n, f_1, \ldots, f_n) \in T_{\mathrm{exp}}$.

Notice that this also shows that SC implies WSC – a result that does not seem too obvious without going through the intermediary of the decidability of T_{exp}.

We now show that WSC implies the decidability of T_{exp}. Naturally we consider the recursive theory

$$T^* =_{\mathrm{def}} T_{unres} \cup T_H \cup T_{NA} \cup T_{wsc}$$

where

$$T_{wsc} = \{\Phi_\eta(n, f_1, \ldots, f_n, g) :$$
$$n \in \mathbf{N}\backslash\{0\}, \ f_1, \ldots, f_n, g \in M_n, \ \eta = \eta(n, f_1, \ldots, f_n, g)\}$$

(for some recursive η witnessing WSC), and show that it axiomatizes $T_{\exp}$.

By the remarks at the beginning of Section 4 and those immediately preceding the proof of Theorem 1.1 we must consider $n \in \mathbf{N}\backslash\{0\}$, $f_1, \ldots, f_n$, $g \in M_n$ and $\overline{\alpha} \in \mathbf{R}^n$ such that $g(\overline{\alpha}) = 0$ and $\overline{\alpha}$ is a nonsingular solution of the system

$$f_1(\overline{x}) = \cdots = f_n(\overline{x}) = 0.$$

We must show that g has a zero in K^n for an arbitrary model K of T^*.

Consider the system

$$f_1(\overline{x}) = \cdots = f_n(\overline{x}) = x_{n+1}h(\overline{x}) - 1 = (1 - \eta^2 g(\overline{x}))x_{n+2}^2 - 1 = 0,$$

where

$$h(\overline{x}) = \det \frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)}(\overline{x}) \text{ and } \eta = \eta(n, f_1, \ldots, f_n g)$$

(as used in the definition of T_{wsc}). The functions here lie in M_{n+2} and it is easy to verify that $\langle \overline{\alpha}, h(\overline{\alpha})^{-1}, 1\rangle$ is a nonsingular solution. Hence, by Theorem 4.3, this system has a solution, $\langle \overline{\gamma}, \gamma_{n+1}, \gamma_{n+2}\rangle$ say, in K^{n+1}. Clearly the last two equations force $h(\overline{\gamma}) \neq 0$ and $\mid g(\overline{\gamma}) \mid\ < \eta^{-1}$ so, since $K \models \Phi_\eta(n, f_1, \ldots, f_n, g)$, we obtain $g(\overline{\gamma}) = 0$ as required.

There are several equivalent formulations of WSC, all intractable at present, and we end by mentioning two such which, at first sight, do not appear connected with transcendence theory. They concern the last root problem for exponential polynomials which was a major preoccupation in the early days of the model theory of $\langle \overline{\mathbf{R}}, \exp\rangle$. We leave the proofs that they are equivalent to the decidability of $T_{\exp}$ to the reader.

Last root conjecture. *Given $n \in \mathbf{N}\backslash\{0\}$ and $f_1, , \ldots, f_n \in M_n$ one can effectively find $\eta' = \eta'(n, f_1, \ldots, f_n) \in \mathbf{N}$ such that any nonsingular solution $\overline{\alpha} \in \mathbf{R}^n$ to the system $f_1(\overline{x}) = \cdots = f_n(\overline{x}) = 0$ satisfies $\parallel \overline{\alpha} \parallel < \eta'$.*

First root conjecture. *Given $n \in \mathbf{N}\backslash\{0\}$ and $f \in M_n$ one can effectively find $\eta'' = \eta''(n, f)$ such that if f has a zero in $\overline{R}^n$ then it has one, $\overline{\alpha}$ say, satisfying $\parallel \overline{\alpha} \parallel < \eta''$.*

References

Baker, A.

[1975] *Transcendental Number Theory*, CUP, 1975.

Khovanskii, A.G.

[1991] *Fewnomials*, Translations of Mathematical Monographs, vol. 88, AMS, 1991.

Lang, S.

[1983] *Real Analysis*, 2nd ed., Addison-Wesley, London, 1983.

[1965] *Algebra*, Addison-Wesley, Reading, Mass., 1965.

Tarski, A.

[1951] *A decision method for elementary algebra and geometry*, 2nd ed., Berkeley and Los Angeles, 1951.

van den Dries, L.

[1986] A generalization of the Tarski-Seidenberg theorem, and some non-definability results, *Bull. Amer. Math. Soc.*, 15 (1986) 189-193.

Vorobjov, N.N. Jr.

[1991] Deciding consistency of systems of polynomial in exponent inequalities in subexponential time, in *Effective Methods in Algebraic Geometry*, ed. T. Mora and C. Traverso, Progress in Mathematics, vol. 94, Birkhäuser, Boston, 1991.

Wilkie, A.J.

[199?] Model completeness results for expansions of the real field I: restricted Pfaffian functions, to appear in *Journal of the Amer. Math. Society*.

[199?a] Model completeness results for expansions of the real field II: the exponential function, to appear in *Journal of the Amer. Math. Society*.

Normal Forms for Sequent Derivations

by Gregory Mints

1 Introduction

Two constructions suggested in this paper deal with derivations in the intutitionistic predicate logic and its fragment, linear logic, operating with sequents $A_1, \ldots, A_n \Rightarrow B$ or $\Gamma \Rightarrow B$, and possessing means for operations with assumptions (antecedent formulas) $A_1, \ldots, A_n$ and goal B. Formalization of this kind was put forward by Gentzen [1934] who established cut-elimination theorem, the first normal form result for one of these formulations, L-type systems LK and LJ. Corresponding results by D. Prawitz [1965] for the natural deduction (N-type) formulation was a starting point for many developments in proof theory and adjacent fields. I have read Prawitz [1965] rather soon after it appeared, and recognized connection of some of its constructions with the work of Leningrad group (Davydov et al. [1965]) on intertranslations between L-type and N-type formulations, but my own involvement into this line of investigation was influenced by G. Kreisel. Many of the ideas and results accumulated by the year 1971 were collected in Fenstad [1971], especially in the papers Prawitz [1971], Girard [1971], Martin-Löf [1971], and Kreisel [1971]. The papers of Girard, Martin-Löf and Prawitz acknowledge influence of Kreisel. Main instrument of this influence was personal correspondence supplemented by the published reviews of papers and books, and by his own papers. Majority of subsequent proof-theoretic publications on normalization and normal forms of proofs also refers to his work. My own work [1974], [1975], [1977], [1979], [1979a] dealing with enrichments of derivations and their transformation in the process of normalization contained implementation of some suggestions from Kreisel [1970], [1971] as well as refutations of some conjectures stated there and in other writings by G. Kreisel.

The topic of the present paper can be contrasted with Kreisel's remark in the Section 1 of Kreisel [1971] (*Discussion* in the subsection (a)) concerning normalization of proofs in the calculus of sequents (L-formulation): in Kreisel [1971] it is suggested to consider various cut elimination procedures, and we begin with a cut-free proof.

Our starting point is the following difference between normal natural deductions and cut-free L-derivations (in the intuitionistic predicate logic). Normal form of a natural deduction is unique, but L-derivation can have different cut-free forms depending of the order of cut-elimination. On the other hand, one and the same normal natural deduction corresponds to these cut-free forms, and they can be obtained from each other by permutation of some antecedent rules. It was not clear how to pick up the distinguished permutation. We establish here that it is sufficient to require "consistency" in analyzing formulas by antecedent rules (cf. Definition 4). This requirement defines unique normal form for L-derivations in the intuitionistic logic. However, it turns out to be too strong for the linear logic (Girard [1987]): the derivation in the Example 2 after Theorem 5 below does not possess normal form. Considerations of this kind led to the introduction of proof nets in Girard [1987] but proof nets have other drawbacks. This difficulty is overcome here in the manner of Kelly and MacLane [1971]: derivations are treated modulo suitable equivalence relation, and this allows to obtain unique normal form.

One can get an idea of this normal form using Prawitz translation (Prawitz [1965], Pottinger [1977]) N of GJ-derivations d into natural deductions $N(d)$ preserving normal form: if d is cut-free then $N(d)$ is normal. Let $G(d)$ be Prawitz translation ([1965], pp. 91-93) of natural deduction d into GJ-derivation preserving normality. Then the normal form of a cut-free derivation d is close to $G(N(d))$.

Although we do not use explicitly results from Zucker [1974] and Pottinger [1977] on connection between cut elimination and normalization, they underlie all subsequent developments.

Acknowledgments. The author is grateful to the members of logic seminar at Stanford (especially to G. Schwarz) and to H. Barendregt for useful discussion.

2 Intuitionistic Predicate Logic

We consider derivations in a formulation GJ of the intuitionistic predicate calculus with structural rules which is close to Gentzen's system LJ [1934] and is even closer to the Kleene's system G [1952a]. Our language contains constant $\perp$, but no negation, and sequents have the form $\Gamma \Rightarrow D$ where Γ is a list of formulas and D is a formula. Main differences between GJ and LJ are as follows:

(a) Permutation rule is included into the formulations of all other rules.

(b) Parametric formulas in all premises of two-premise rules are the same.

(c) We do not have negation or inference rules for the constant $\perp$ (false), and the only axiom for $\perp$ has atomic succedent.

(d) There is no rule for weakening in the succedent.

Our main tool will be Curry-Howard (CH) isomorphism assigning λ-terms (or *deductive terms*) $T(d)$ to derivations d (see below). We use almost the same language for deductive terms as Troelstra and van Dalen [1988]. It is understood that the typed variables used in deductive terms are distinct from individual variables which can be bound by quantifiers in predicate formulas. Let us list the axioms and inference rules of the system GJ together with assignment of deductive terms to them (cf. Pottinger [1977]).

- Axioms.

$$\begin{array}{lc} A \Rightarrow A & x^A \\ \perp \Rightarrow A & \perp_A(x^\perp), \end{array}$$

  where A is atomic and $A \neq \perp$.

- Inference rules

$\Rightarrow \&$	$\dfrac{\Gamma \Rightarrow A \qquad \Gamma \Rightarrow B}{\Gamma \Rightarrow A\&B}$	$\dfrac{t_0 \qquad t_1}{p(a,b)}$
$\& \Rightarrow$	$\dfrac{A, \Gamma \Rightarrow D}{A\&B, \Gamma \Rightarrow D}$	$\dfrac{t[x^A]}{t[p_1 x^{A\&B}]}$
$\& \Rightarrow$	$\dfrac{B, \Gamma \Rightarrow D}{A\&B, \Gamma \Rightarrow D}$	$\dfrac{t[x^B]}{t[p_2 x^{A\&B}]}$
$\to\Rightarrow$	$\dfrac{\Gamma \Rightarrow A \qquad B, \Gamma \Rightarrow D}{(A \to B), \Gamma \Rightarrow D}$	$\dfrac{a \qquad t[x^B]}{t[x^{A\to B}(a)]}$
$\Rightarrow\to$	$\dfrac{A, \Gamma \Rightarrow B}{\Gamma \Rightarrow (A \to B)}$	$\dfrac{t[x^A]}{\lambda x^A t[x^A]}$
$\Rightarrow \vee$	$\dfrac{\Gamma \Rightarrow A}{\Gamma \Rightarrow (A \vee B)}$	$\dfrac{t}{k_0 t}$
$\Rightarrow \vee$	$\dfrac{\Gamma \Rightarrow B}{\Gamma \Rightarrow (A \vee B)}$	$\dfrac{t}{k_1 t}$
$\vee \Rightarrow$	$\dfrac{A, \Gamma \Rightarrow D \qquad B, \Gamma \Rightarrow D}{A \vee B, \Gamma \Rightarrow D}$	$\dfrac{t_0[x^A] \qquad t_1[x^B]}{D_{u,v}(x^{A\vee B}, t_0[u], t_1[v])}$
$\Rightarrow \exists$	$\dfrac{\Gamma \Rightarrow A[t_0]}{\Gamma \Rightarrow \exists x A}$	$\dfrac{t_1}{p(t_0, t_1)}$
$\exists \Rightarrow$	$\dfrac{A[y], \Gamma \Rightarrow D}{\exists x A[x], \Gamma \Rightarrow D}$	$\dfrac{t_1[y, z^{A[y]}]}{E_{u,v}(z^{\exists x A}, t_1[u,v])}$
$\Rightarrow \forall$	$\dfrac{\Gamma \Rightarrow A[x]}{\Gamma \Rightarrow \forall y A[y]}$	$\dfrac{t[x]}{\lambda y.t[y]}$
$\forall \Rightarrow$	$\dfrac{A[t'], \Gamma \Rightarrow D}{\forall y A[y], \Gamma \Rightarrow D}$	$\dfrac{t[x^{A[t']}]}{t[z^{\forall y A[y]}(t')]}$
Weakening	$\dfrac{\Gamma \Rightarrow D}{A, \Gamma \Rightarrow D}$	$\dfrac{t}{t}$
Contraction	$\dfrac{A, A, \Gamma \Rightarrow D}{A, \Gamma \Rightarrow D}$	$\dfrac{t[x^A, y^A]}{t[x^A, x^A]}$

All inference rules are by the definition invariant under permutation of formulas in the sequents. For example,

$$\frac{G, F, A, E \Rightarrow D}{E, A\&B, F, G \Rightarrow D}$$

is the application of the rule $\& \Rightarrow$.

The term $T(d)$ is defined by induction on the derivation d according to the rules listed above.

The formula explicitly shown in the conclusion of a logical rule (under the line) and containing connective introduced by the rule is the *main* formula. Subformulas of the main formula explicitly shown in the premises of the rule (above the line) are *side* formulas. Main formulas of Weakening and Contraction and side formulas of Contraction are defined similarly. Main formulas and side formulas are *active* formulas. Remaining formulas are *passive or parametric.*

Often we replace typed variable x^A by the formula A itself to simplify notation. For example, terms corresponding to the axioms above can be written A and $\perp_A(\perp)$ respectively. The following is equivalent to the standard definition via natural deduction corresponding to a Gentzen-type derivation (cf. Notes 5,7 below).

Definition 1 *Two derivations d, d' of the same sequent are* equivalent *(written $d \equiv d'$) iff $T(d) = T(d')$ (under the same choice of bound variables and of free variables for assumptions).*

When we say that a weakening 'immediately precedes' a sequent or 'introduces' a formula occurence into this sequent, we mean that at most other weakenings can occur between this weakening and the sequent in question.

Definition 2 *A derivation d in GJ is W-normal (for weakening) if weakenings occur only as follows (with only other weakenings intervening):*

(a) immediately preceding the endsequent,

(b) to introduce the antecedent side formula of $\Rightarrow\rightarrow$, or of $\vee\Rightarrow$, or of $\exists\Rightarrow$,

(c) to introduce a certain one of the antecedent parametric formulas into one premise of a two-premise inference, this formula not being introduced into the antecedent of the other premise by a weakening.

Note 1 Main difference with the pruning property of the Lemma 4 in Kleene [1952a] is the exception for rules $\vee\Rightarrow, \exists\Rightarrow$ in the clause (b), since pruning of these rules can change deductive term assigned to a derivation.

Definition 3 *A derivation d in GJ is C-normal (for contraction) if contractions occur in it only just preceding (with only other contractions, followed by weakenings, intervening):*

(a) the endsequent,

(b) $\Rightarrow\rightarrow$, $\vee\Rightarrow$, or $\exists\Rightarrow$ for which the formula C of contraction is the antecedent side formula.

Note 2 Main difference with the similar property of the Lemma 12 in Kleene [1952a] is the exclusion of the rule $\vee \Rightarrow$, since moving contraction down this rule can change deductive term assigned to a derivation.

The next definition is central for the Section 2.

Definition 4 *A derivation in GJ is M-normal (for main formula) if any antecedent side formula of an antecedent rule for* &, $\forall$, $\rightarrow$ *is itself a main formula of an inference rule or axiom.*

Note 3 Nothing is required of the side formulas of the antecedent rules for $\vee, \exists$.

Definition 5 *A derivation in GJ is normal if it is cut-free, W-,C- and M-normal. A normal form of a derivation d is any normal derivation equivalent to d.*

Note 4 Complete description of the inferences in WC-normal derivations is presented in Section 5. W-normality plus C-normality has the same effect as allowing (but not requiring) preservation of main formulas in the premises of all antecedent rules, as well as allowing contraction of passive formulas in two-premise rules (i.e. mix of multiplicative and additive forms (Girard [1987]) of these rules).

Lemma 1 *Every cut-free derivation d can be transformed into an equivalent W-normal derivation by pruning, i.e. permuting weakening downward and deleting some formulas and whole branches of the derivation. This preserves C- and M-normality.*

Proof. Weakening permutes down any rule (cf. Kleene [1952a], Lemma 4 and Note 1) except the cases mentioned in the definition of W-normality. Preservation of equivalence of derivations follows from the fact that deductive term assigned to a sequent does not depend of the free variable v^A assigned to an antecedent formula A introduced by a weakening. $\Box$

Lemma 2 *Every cut-free derivation d can be transformed into an equivalent C-normal derivation by permuting contraction downward preserving C- and M-normality.*

Proof. Exactly as in Kleene [1952a], Lemma 12, contractions are permuted downward. Preservation of the equivalence follows from the fact that identification of typed variables commutes with non-exceptional rules. Consider main case when the main formula of the contraction is the side formula of a rule different from $\vee \Rightarrow, \exists \Rightarrow, \Rightarrow\to$. Take for example the rule $\& \Rightarrow$. Relevant part of the original derivation is as follows:

$$\dfrac{\dfrac{A, A, \Gamma \Rightarrow D}{A, \Gamma \Rightarrow D}}{A\&B, \Gamma \Rightarrow D} \qquad \dfrac{\dfrac{t[x^A, y^A]}{t[x^A, x^A]}}{t[p_1 x^{A\&B}, p_1 x^{A\&B}]}$$

Permutation of Contraction inference results in the figure:

$$\dfrac{\dfrac{\dfrac{A, A, \Gamma \Rightarrow D}{A\&B, A, \Gamma \Rightarrow D}}{A\&B, A\&B, \Gamma \Rightarrow D}}{A\&B, \Gamma \Rightarrow D} \qquad \dfrac{\dfrac{\dfrac{t[x^A, y^A]}{t[p_1 x^{A\&B}, y^A]}}{t[p_1 x^{A\&B}, p_1 y^{A\&B}]}}{t[p_1 x^{A\&B}, p_1 x^{A\&B}]}$$

and final deductive terms are the same. □

Theorem 1 *Every cut-free derivation d can be put into a normal form d' by permuting some inferences violating normality restriction. These permutations preserve equivalence of derivations; in particular, $d \equiv d'$.*

Example 1 A normal form of the derivation

$$\dfrac{\dfrac{\dfrac{B \Rightarrow B}{A\&B \Rightarrow B} \quad \dfrac{A \to A \quad \dfrac{\dfrac{C \Rightarrow C \quad D \Rightarrow D}{C \to D, C \Rightarrow D}}{C \to D, C\&(B \to (A \to (C \to D))) \Rightarrow D}}{A, (A \to (C \to D)), C\&(B \to (A \to (C \to D))) \Rightarrow D}}{\dfrac{A, A\&B, (B \to (A \to (C \to D))), C\&(B \to (A \to (C \to D))) \Rightarrow D}{A, A\&B, C\&(B \to (A \to (C \to D))) \Rightarrow D}}}{A\&B, C\&(B \to (A \to (C \to D))) \Rightarrow D}$$

(where some applications of the structural rules are not shown explicitly) is

$$\dfrac{\dfrac{\dfrac{B \Rightarrow B}{A\&B \Rightarrow B} \quad \dfrac{\dfrac{A \Rightarrow A}{A\&B \Rightarrow A} \quad \dfrac{\dfrac{C \Rightarrow C}{C\&(B \to (A \to (C \to D))) \Rightarrow C} \quad D \Rightarrow D}{C \to D, C\&(B \to (A \to (C \to D))) \Rightarrow D}}{A\&B, (A \to (C \to D)), C\&(B \to (A \to (C \to D))) \Rightarrow D}}{A\&B, (B \to (A \to (C \to D))), C\&(B \to (A \to (C \to D))) \Rightarrow D}}{A\&B, C\&(B \to (A \to (C \to D))) \Rightarrow D}$$

Let us prove correctness part of the Theorem 1.

Lemma 3 *Permuting antecedent $\&, \to, \forall$-inferences up any other inference preserves the equivalence of the derivations.*

Proof. General schema (for permutation with two-premise rules) is given below.

- **&-case.** Original derivation with corresponding deductive terms:

$$\frac{\dfrac{B, \Gamma' \Rightarrow D' \quad B, \Gamma'' \Rightarrow D''}{B, \Gamma \Rightarrow D}}{A\&B, \Gamma \Rightarrow D} \qquad \frac{\dfrac{t'[x^B] \quad t''[x^B]}{t[x^B]}}{t[p_2 x^{(A\&B)}]}$$

The derivation after the permutation:

$$\frac{\dfrac{B, \Gamma' \Rightarrow D'}{A\&B, \Gamma' \Rightarrow D'} \quad \dfrac{B, \Gamma'' \Rightarrow D''}{A\&B, \Gamma'' \Rightarrow D''}}{A\&B, \Gamma \Rightarrow D}$$

$$\frac{\dfrac{t'[x^B]}{t'[p_2 x^{(A\&B)}]} \quad \dfrac{t''[x^B]}{t''[p_2 x^{(A\&B)}]}}{t[p_2 x^{(A\&B)}]}$$

The final deductive term is the same.

- **∀-case.** Treated in the same way.

- **→-case.** Original derivation:

$$\frac{\Gamma \Rightarrow A \qquad \dfrac{B, \Gamma' \Rightarrow D' \quad B, \Gamma'' \Rightarrow D''}{B, \Gamma \Rightarrow D}}{(A \to B), \Gamma, \Rightarrow D} \qquad \frac{a \qquad \dfrac{t'[x^B] \quad t''[x^B]}{t[x^B]}}{t[x^{(A\to B)}(a)]}$$

where a is the term assigned to the derivation of $\Gamma \Rightarrow A$. Result of permuting $\to\Rightarrow$ upward:

$$\frac{\dfrac{\Gamma \Rightarrow A \quad B, \Gamma' \Rightarrow D'}{A \to B, \Gamma' \Rightarrow D'} \quad \dfrac{\Gamma \Rightarrow A \quad B, \Gamma'' \Rightarrow D''}{A \to B, \Gamma'' \Rightarrow D''}}{A \to B, \Gamma \Rightarrow D}$$

$$\frac{\dfrac{a \;\; t'[x^B]}{t'[x^{A\to B}(a)]} \quad \dfrac{a \;\; t''[x^B]}{t''[x^{A\to B}(a)]}}{t[x^{A\to B}(a)]}$$

Again the deductive terms of two derivations coincide. □

Proof of Theorem 1. To transform arbitrary derivation into an equivalent normal form first use Lemmas 1, 2 to make it WC-normal.

Then permute upward all $\&, \rightarrow, \forall$-antecedent inferences violating M-restriction, beginning with the uppermost ones. To ensure termination of this process, make permutations in blocks, putting given inference into the place where it no more violates M-restriction. More precisely, use induction on the size $n(d)$ of non-normal part of given WC-normal derivation d, that is on the number n of inferences which are below any premise of an inference violating M-restriction. Pick up one of such inferences. Assume (to save notation) that it is $\&$-inference with main formula $A\&B$ and side formula A. Since the derivation d is W-normal, the side formula A is traceable up the derivation to main formulas of some inferences or axioms:

$$\begin{array}{ccc}
\dfrac{B', \Delta' \Rightarrow E' \quad B'', \Delta'' \Rightarrow E''}{B, \Delta \Rightarrow E} & & B \Rightarrow B \\
| & & / \\
\dfrac{B, \Gamma \Rightarrow D}{A\&B, \Gamma \Rightarrow D}\; R & &
\end{array}$$

where the part ending in $B, \Gamma \Rightarrow D$ is normal. Permuting inference R upward to the places shown explicitly results in the following derivation:

$$\begin{array}{ccc}
\dfrac{B', \Delta' \Rightarrow E' \quad B'', \Delta'' \Rightarrow E''}{R \quad \dfrac{B, \Delta \Rightarrow E}{A\&B, \Gamma \Rightarrow E}} & & \dfrac{B \Rightarrow B}{A\&B \Rightarrow B}\; R \\
| & & / \\
& A\&B, \Gamma \Rightarrow D &
\end{array}$$

where the part ending in $A\&B, \Gamma \Rightarrow D$ is normal, and the induction parameter is decreased at least by 1. $\Box$

Theorem 2 *Normal form is unique: d is equivalent to e iff their normal forms coincide.*

Strictly speaking, d and e still can differ in the order of weakenings and contractions preceding one and the same rule or endsequent. Precise statement should refer to derivations d, e according to the rules of the Section 6.

Let us prove first several lemmas.

Deductive term $T(d)$ will be used to establish some properties of the derivation d. We use standard permutative and β-reductions for CH-terms.

- β-reductions:

$$\begin{array}{rll}
\lambda x.t[x](t') & \mathit{conv} & t[t'] \\
p_i p(t_0, t_1) & \mathit{conv} & t_i \qquad\qquad i = 0, 1 \\
D_{x,y}(k_i t, t_0[x^A], t_1[y^B]) & \mathit{conv} & t_i[t] \qquad\quad\; i = 0, 1 \\
E_{x,y}(p(t_0, t_1), t_2[y, z]) & \mathit{conv} & t_2[t_0, t_1] \\
\bot_\bot(t) & \mathit{conv} & t.
\end{array}$$

- Permutative reductions.

- D-reductions:

$$\begin{array}{rll}
D_{x,y}(t, t_1, t_2)(t') & \mathit{conv} & D_{x,y}(t, t_1(t'), t_2(t')) \\
p_i D_{x,y}(t, t_1, t_2) & \mathit{conv} & D_{x,y}(t, p_i t_1, p_i t_2) \\
D_{u,v}(D_{x,y}(t, t_1, t_2), t_3, t_4)) & \mathit{conv} & D_{x,y}(t, D_{u,v}(t_1, t_3, t_4), \\
 & & \qquad D_{u,v}(t_2, t_3, t_4)) \\
E_{u,v}(D_{x,y}(t, t_1, t_2), t_3) & \mathit{conv} & D_{x,y}(t, E_{u,v}(t_1, t_3), \\
 & & \qquad E_{u,v}(t_2, t_3)) \\
\bot_A(D_{x,y}(t, t_1, t_2)) & \mathit{conv} & D_{x,y}(t, \bot_A(t_1), \bot_A(t_2)).
\end{array}$$

- E-reductions:

$$\begin{array}{rll}
E_{u,v}(t, t_1)(t') & \mathit{conv} & E_{u,v}(t, t_1(t')) \\
p_i E_{u,v}(t, t_1) & \mathit{conv} & E_{u,v}(t, p_i t_1) \\
D_{x,y}(E_{u,v}(t, t_1), t_2, t_3) & \mathit{conv} & E_{u,v}(t, D_{x,y}(t_1, t_2, t_3)) \\
E_{x,y}(E_{u,v}(t, t_1), t_2) & \mathit{conv} & E_{u,v}(t, E_{x,y}(t_1, t_2)) \\
\bot_A(E_{u,v}(t, t_1)) & \mathit{conv} & E_{u,v}(t, \bot_A(t_1)).
\end{array}$$

The left hand side of the conversion is called its *redex*. A term is *normal* if it does not contain redexes. It is known (Troelstra and van Dalen [1988]) that the reduction relation defined by these conversions is strongly normalizing: every sequence of conversions terminates in the unique normal form.

Let us list deductive terms substituted in the antecedent rules for $\&, \forall, \rightarrow$.

Definition 6 *Terms of the form $p_i x, x(t)$ where x is a variable are called* head-terms.

Lemma 4 *If s is a normal head-term and q is a deductive term, then for any redex in $q[s]$ there is a redex in q.*

Proof. Consider all possible redexes $r_w[s]$ in $q_w[s]$ or shorter $r[s]$ in $q[s]$. Note that $r[s] \neq s$ since s is normal.

1. $r[s] = (\lambda x.t)(t_1)$. Since s is a head-term, $s \neq \lambda x.t$. Hence s is a subterm of t, t_1, hence

   $$r[w] = (\lambda x.t'[w])(t'_1[w])$$

   i.e. r is a redex in q.

2. $r[s] = p_i p(t_0, t_1),\ i = 1, 2$. Again s is a subterm of t_0, t_1, hence

   $$r = p_i p(t'_0, t'_1)$$

   In all remaining β-conversions s also cannot be introduction component $k_i t, p(t_0, t_1)$ since it 'ends in elimination rule', and the structure of a redex is preserved when s is replaced by a variable w.

   Redexes in permutative conversions are stable under any (type preserving) changes in the components $t, t', t_1, t_2, t_3, t_4$. It remains to notice that s cannot be a 'composite elimination component' of the form $D_{x,y}(t, t_1, t_2)$ or $E_{x,y}(t, t_1)$. $\Box$

The following proposition is given in Pottinger [1977] without detailed proof.

Theorem 3 *If d is a derivation in GJ, then $T(d)$ is normal*

Proof. Note that our formulation of GJ is cut-free. Use induction on d and consider all cases. For each rule and each redex in the deductive term for its conclusion, it is possible to find a redex in deductive terms of one of the premisses. This is easy for the succedent rules and $\vee, \exists$-antecedent rules. Remaining antecedent rules are covered by the previous Lemma. $\Box$

Note 5 Sketch of an alternative proof. Following Prawitz [1965] and Pottinger [1977], consider a translation N of GJ-derivations d into natural deductions $N(d)$ preserving normal form: if d is cut-free then $N(d)$ is normal. Assign deductive terms $T(d)$ to natural deductions d in the natural way (according to Curry-Howard isomorphism). Then $T(d) = T(N(d))$, and the right hand side is normal.

In the following we freely use for deductive terms standard notation used for natural deductions. For example, we say that a term ends in the $\vee$-elimination rule if it is of the form

$$D_{u,v}(t, r[u], s[v]), \qquad (*)$$

i.e. corresponds to a natural deduction ending in $\vee$-elimination:

$$\frac{\Gamma \Rightarrow A \vee B; \quad \Sigma, A \Rightarrow C; \quad \Pi, B \Rightarrow C}{\Gamma, \Sigma, \Pi \Rightarrow C}$$

A term ending in elimination rule will be called *e-term*.

Head subterm of e-term is the subterm corresponding to the main premise, for example t in e-term (*) above.

Head variable of e-term is the first free variable encountered in going to head subterm of head subterm etc. Normal e-term always has a head variable which corresponds to the axiom $A \Rightarrow A$ on the top of the main branch of corresponding natural deduction.

Definition 7 *Main branch of a normal GJ-derivation d is the branch ending in the final sequent of d and closed under antecedent side formulas of &, $\forall$, $\rightarrow$-inferences. More precisely, the main branch contains a premise of any such inference if it contains its conclusion. In the case of $\rightarrow$-antecedent the right premise is chosen.*

Lemma 5 *Let d be a normal derivation. Then main branch of d always ends in the final sequent, and it consists of this sequent only if d ends in the rule different from &, $\forall$, $\rightarrow$-antecedent. Otherwise it begins with the axiom or conclusion of an antecedent $\exists, \vee$ rule. The variable v^A corresponding to the last (lowermost) main formula A in the main branch of d is exactly the main variable of the term $T(d)$, and the steps of construction of $T(d)$ from v^A correspond exactly to &, $\forall$, $\rightarrow$-inferences in the main branch.*

Note 6 Exact correspondence mentioned in the Lemma means: $T(d)$ is obtained from v^A by the operations p_i, and application $t(u)$ corresponding to the antecedent rules for &, $\forall$ and $\rightarrow$ in the order opposite to their order in the main branch. The term u in $\forall$-term is the term substituted in the corresponding inference. For $\rightarrow$-antecedent the term u is deductive term assigned to the left premise.

Example 2 If d is the normal derivation from the example 1, then

$$T(d) = (p_2 M)(b)(a)(p_1 M)$$

where M is the typed variable corresponding to the main formula $C\&(B \rightarrow (A \rightarrow (C \rightarrow D)))$ of the lowermost rule in d, and a, b are obvious derivations of $A\&B \Rightarrow A$, $A\&B \Rightarrow B$.

Proof of Lemma 5. Routine bottom-up induction on d.

Lemma 6 *Let d be a normal derivation. Then T(d) is e-term iff d ends in an antecedent rule.*

Proof. If d is an axiom or ends in the succedent rule, then $T(d)$ is not e-term. If d ends in an antecedent rule, then $T(d)$ ends in elimination rule by the Lemma 5.

Proof of Theorem 2. Let $d, d' : \Gamma \Rightarrow G$ be normal, and $T(d) = T(d')$. We prove that d and d' coincide by induction on the sum of lengths of d, d'. If one of them, say d, is an axiom or ends in a succedent rule, then $T(d)$ has corresponding form, so d' ends in the same rule with the same partition of the antecedent formulas, and so the induction hypothesis is applicable. If d ends in the succedent rule, we can reason in the same way by the Lemma 5. This concludes the proof of the Theorem 2. $\square$

Note 7 The remark on connection of normal forms of derivations and Prawitz transformations (made at the end of the introduction) is justified by the equations

$$T(G(N(d))) = T(N(d)) = T(d)$$

3 Normal Deductions in Multiplicative Linear Logic

To fix ideas consider positive multiplicative fragment (intuitionistic linear logic) without negation and constant I. The difficulty here begins already with a natural deduction formulation, since the tensor product $\otimes$ does not have projections, and so standard &-elimination rules are not valid for $\otimes$, and are replaced by the rule $\otimes E$ shown below. It is similar to implication-elimination (cf. Mints [1977], Babaev [1980], Lincoln and Mitchell [1992]). One solution is proposed in Troelstra [1993]: to treat $\otimes E$ like $\vee$-elimination in the intuitionistic case.

We suggest another treatment base on special property of the multiplicative linear logic: each derivation is a substitution instance of a 'balanced' derivation, where each propositional variable occurs exactly in one axiom. This is intended to contribute to discovering a viable alternative to proof nets.

Normalization theorem was established in Mints [1977] and used in a standard way, for example to prove coherence results. Still $\otimes E$ has some of the drawbacks of the antecedent rules: it can be freely moved up and down the derivation without affecting its (categorical) content.

Absence of a natural deduction formulation possessing unique normal form was one of the stimuli for the introduction of proof nets in Girard [1987]. It turns out that moving $\otimes$ upward produces required normal form.

System NLL (Natural Deduction for Linear Logic)

- Derivable objects: *sequents* $A_1, \ldots, A_n \Rightarrow B$
- Axioms: $A \Rightarrow A$
- Structural rule: Permutation.

  (So antecedent is treated up to permutations).
- Logical rules:

$$\rightarrow I \quad \frac{A, \Gamma \Rightarrow B}{\Gamma \Rightarrow (A \rightarrow B)} \qquad \rightarrow E \quad \frac{\Gamma \Rightarrow (A \rightarrow B) \qquad \Sigma \rightarrow A}{\Gamma, \Sigma \Rightarrow B}$$

$$\otimes\, I \quad \frac{\Gamma \rightarrow A \qquad \Sigma \rightarrow B}{\Gamma, \Sigma \Rightarrow (A \otimes B)} \qquad \otimes\, E \quad \frac{\Gamma \Rightarrow A \otimes B \qquad A, B, \Sigma \Rightarrow C}{\Gamma, \Sigma \Rightarrow C}$$

Definition 8 *Rules marked with I (with E) are introduction (elimination) rules. Main formula of an elimination rule is one containing corresponding connective, and main premise is one containing the main formula (i.e. the leftmost one). Minor premise is the other one. Segment in a derivation is a branch consisting of minor premises of $\otimes$-eliminations. Cut, or main segment is as usual a segment which begins with an introduction rule and ends in main formula of an elimination rule.*

We write $d : \Gamma \Rightarrow G$ to indicate that d is a deduction of the sequent $\Gamma \Rightarrow G$.

Definition 9 *Natural deduction in NLL is cut-free iff it does not contain cuts, i.e. main segments.*

Definition 10 *To every deduction d we assign corresponding deductive term $T(d)$ in a standard way. Rule $\rightarrow I$ corresponds to λ-abstraction λxt, rule $\otimes I$ corresponds to pairing $p(t,u)$, rule $\rightarrow E$ corresponds to application $t(u)$, and rule $\otimes E$ corresponds to projections (p_0 for left, p_1 for right) plus substitution:*

$$\otimes E \quad \frac{\Gamma \Rightarrow A \otimes B \qquad A, B, \Sigma \Rightarrow C}{\Gamma, \Sigma \Rightarrow C} \qquad \frac{t \quad u[x^A, y^B]}{u[p_0 t, p_1 t]}$$

Note that any term of our system (i.e. of the form $T(d)$ for some d) is a lambda term with paring in the standard sense. So one can use standard syntactic properties of such lambda terms. Notation $t : \Gamma \Rightarrow G$ means that t is a term of type G, and its free variables exactly correspond to Γ, i.e. $\Gamma = A_1, \ldots, A_n$, and the free variables of t are $x^{A_1}, \ldots, x^{A_n}$.

Proper conversions P1. These are standard conversions for implication $\to$ and more or less standard reduction for $\otimes$:

$$\dfrac{\dfrac{\Gamma \Rightarrow A \quad \Gamma' \Rightarrow B}{\Gamma, \Sigma \Rightarrow A \otimes B} \qquad \begin{array}{c} A \Rightarrow A \quad B \Rightarrow B \\ | \qquad\quad / \\ A, B, \Gamma' \Rightarrow C \end{array} \; \Gamma, \Gamma', \Sigma \Rightarrow C}{}$$

conv $\begin{array}{c} \Gamma \Rightarrow A \quad \Gamma' \Rightarrow B \\ | \qquad\quad / \\ \Gamma, \Gamma', \Sigma \Rightarrow C \end{array}$ *whereasalwaysallpredecessorsoftheantecedent*A,B*arereplacedb*

and the topmost axioms by the subderivations of the main premise.

Permutative conversions P2. Permuting elimination rules upward in a segment.

$$\dfrac{\dfrac{\Gamma \Rightarrow A \otimes B \quad A, B, \Sigma \Rightarrow C}{\Gamma, \Sigma \Rightarrow C} \qquad C', \Delta \Rightarrow D}{\Gamma, \Sigma, \Delta \Rightarrow D'}$$

conv

$$\dfrac{\Gamma \Rightarrow A \otimes B \qquad \dfrac{A, B, \Sigma \Rightarrow C \quad C', \Delta \Rightarrow D}{A, B, \Sigma, \Delta \Rightarrow D'}}{\Gamma, \Sigma, \Delta \Rightarrow D'}$$

It is possible that reduction relation (determined by) P1, P2 is strongly normalizing. At least the following is evident.

Theorem 4 *Every deduction can be reduced to a cut-free form by a series of P1, P2.*

Proof. Induction on the size (number of inferences) of the deduction. Since contraction rule is absent, each proper reduction P1 reduces the size of the derivation, and it is sufficient to apply permutative reductions P2 only in conjunction with the proper reductions. □

From now on all derivations are assumed to be cut-free. Nevertheless this normal form is not satisfactory: different normal forms can have one and the same associated deductive term.

Example 1 Consider derivation d:

$$\dfrac{a\otimes b \Rightarrow a\otimes b \quad \dfrac{\dfrac{\dfrac{a\Rightarrow a \quad c\Rightarrow c}{a,c\Rightarrow a\otimes c} \quad \dfrac{b\Rightarrow b \quad d\Rightarrow d}{b,d\Rightarrow b\otimes d}}{a,b,c,d\Rightarrow (a\otimes c)\otimes(b\otimes d)}}{a,b,c\otimes d\Rightarrow (a\otimes c)\otimes(b\otimes d)}}{a\otimes b, c\otimes d\Rightarrow (a\otimes c)\otimes(b\otimes d)}$$

$$T(d) = p(p(p_0x^{a\otimes b}, p_0x^{c\otimes d}), p(p_1x^{a\otimes b}, p_1x^{c\otimes d}))$$

One can try to define normal form of deductions in such a way that normal derivations d, e are equal if $T(d) = T(e)$, but there is no evident reason to prefer the derivation d in this Example to the result of permuting two lowermost $\otimes$-inferences. The solution is to treat them "at the same level" and make them maximal. Consider two more reductions permuting $\otimes E$ up to the right:

Conversions P3.

$$\dfrac{\Gamma\Rightarrow A\otimes B \quad \dfrac{A,B,\Delta'\Rightarrow C' \quad \Delta''\Rightarrow C''}{A,B,\Delta\Rightarrow C}}{\Gamma,\Delta\Rightarrow C}$$

is converted into

$$\dfrac{\dfrac{\Gamma\Rightarrow A\otimes B \quad A,B,\Delta'\Rightarrow C'}{\Gamma,\Delta'\Rightarrow C'} \quad \Delta''\Rightarrow C''}{\Gamma,\Delta\Rightarrow C}$$

Similar conversions are postulated for the case when A, B are in the upper right sequent:

$$\dfrac{\Gamma\Rightarrow A\otimes B \quad \dfrac{\Delta'\Rightarrow C' \quad A,B,\Delta''\Rightarrow C''}{A,B,\Delta\Rightarrow C}}{\Gamma,\Delta\Rightarrow C} \tag{2}$$

is converted into

$$\dfrac{\Delta'\Rightarrow C' \quad \dfrac{\Gamma\Rightarrow A\otimes B \quad A,B,\Delta''\Rightarrow C''}{\Gamma,\Delta''\Rightarrow C''}}{\Gamma,\Delta\Rightarrow C} \tag{3}$$

and there are similar conversions for one-premise rule ($\to I$ in our case).

Now it should be easy to prove (see below) that every two derivations with the same assigned term are *interconvertible*. However (3) can convert back to (2), and hence loops are possible. In particular, normal form does not always exist: the derivation from the example 1 gives rise to infinite sequence of P3-permutations. We introduce a new equivalence relation for derivations d, d':

> $d \sim d'$ iff d, d' can be obtained from each other by a sequence of conversions P3 and the conversions inverse to P3.

Lemma 7 *If $d \sim d'$ then $T(d) = T(d')$*

Proof. Verify this for conversions P3. □

Note 8 In fact P2-conversions also preserve λ-terms assigned to derivations.

Our main tool will be the following observation: $\otimes$-inference can be moved down the derivation provided necessary antecedent members are present.

Lemma 8 *Let $d : \Gamma, \Delta \Rightarrow G$ be a derivation such that $T(d)$ contains subterms $p_0 t, p_1 t$ for some term $t : \Gamma \Rightarrow A \otimes B$ (with Γ traceable to the Γ in the last sequent). Then there is $d' \sim d$ and $e : \Gamma \Rightarrow A \otimes B$ such that $T(e) = t$ and*

$$\frac{e:\ \Gamma \Rightarrow A \otimes B \quad A, B, \Delta \Rightarrow G}{d:\ \Gamma, \Delta \Rightarrow G}$$

In other words, t is assigned to some derivation ending in the $\otimes$- inference corresponding to $p_0 t, p_1 t$, and this inference can be made the lowermost in d.

Proof. By induction on the derivation d. Induction base is obvious since variable (corresponding to a derivation consisting of the axiom) does not contain projections. Induction step is proved by cases depending of the last inference L in d:

1. L is $\Rightarrow I$

$$\frac{f:\ C, \Gamma, \Delta \Rightarrow D}{d:\ \Gamma, \Delta \Rightarrow C \to D}$$

By the induction hypothesis there are e, f'

$$\frac{e:\ \Gamma \Rightarrow A \otimes B \qquad A, B, C, \Delta \Rightarrow D}{f \sim f' : C, \Gamma, \Delta \Rightarrow D}$$

with $T(e) = t$, hence

$$\frac{e:\ \Gamma \Rightarrow A \otimes B \quad \dfrac{A, B, C, \Delta \Rightarrow D}{A, B, \Delta \Rightarrow C \to D}}{d \sim d' : \Gamma, \Delta \Rightarrow C \to D}$$

2. $L = \otimes I$

$$\frac{f:\ \Gamma', \Delta' \Rightarrow C \qquad \Gamma'', \Delta'' \Rightarrow D \quad g}{d:\ \ \Gamma, \Delta \Rightarrow C \otimes D}$$

where $\Gamma = \Gamma', \Gamma''$ and $\Delta = \Delta', \Delta''$. Then $T(d) = p(T(f), T(g))$, or slightly abusing notation,

$$T(d[\Gamma, \Delta]) = p(T(f[\Gamma', \Delta']), T(g[\Gamma'', \Delta''])) \qquad (4)$$

If Γ', Γ'' are both nonempty, then the only subterm of $T(d)$ containing all of Γ is $T(d)$ itself, hence $t = T(d)$, but then $T(d)$ does not contain p_0t. So one of Γ', Γ'' is empty, and induction hypothesis is applicable as in the Case 1.

3. $L = \Rightarrow I$. Similarly to the case 2.

4. $L = \otimes E$

$$\frac{f:\ \Gamma', \Delta' \Rightarrow C \otimes D \qquad g:\ C, D, \Gamma'', \Delta'' \Rightarrow G}{d:\ \ \Gamma, \Delta \Rightarrow G}$$

Again, if one of Γ', Γ'' is empty, then the induction hypothesis is applicable, so assume they are both nonempty:

$$T(d[\Gamma, \Delta]) = \\ T(g[\Gamma'', \Delta''])[C := p_0T(f[\Gamma', \Delta']), D := p_1T(f[\Gamma', \Delta'])] \qquad (5)$$

If Δ' is nonempty, then every subterm of $T(d)$ containing Γ should also contain $T(f)$, and so Δ', hence Δ' is empty and $\Delta'' = \Delta$.

Since

$$f:\ \Gamma' \Rightarrow C \otimes D \qquad \text{and} \qquad e : \Gamma', \Gamma'' \Rightarrow A \otimes B$$

both contain Γ', and all occurrences of Γ' in $T(d)$ come from $T(f)$, the terms $T(f)$ and $T(e)$ should overlap in $T(d)$. Hence one of them is contained in the other, and since Γ'' is nonempty, $T(e)$ contains $T(f)$. Applying induction hypothesis to this pair, we get

$$\frac{f:\ \Gamma' \Rightarrow C \otimes D \qquad h:\ C, D, \Gamma'' \Rightarrow A \otimes B}{e \sim e' : \Gamma', \Gamma'' \Rightarrow A \otimes B} \qquad (6)$$

and

$$T(e) = T(h)[C := p_0 T(f), D := p_1 T(f)].$$

Comparing free variables in the left hand side and the right hand side of (5) we see that $T(g)[C, D, \Gamma'', \Delta]$ contains $p_0 T(h), p_1 T(h)$, since the r.h.s. contains $T(e)$. [Recall that our terms are also lambda terms in the standard sense].

Applying induction hypothesis to $T(g)$ and $T(h)$, and permuting with (6) we get:

$$\frac{\dfrac{f:\ \Gamma' \Rightarrow C \otimes D \qquad h:\ C, D, \Gamma'' \Rightarrow A \otimes B}{\Gamma', \Gamma'' \Rightarrow A \otimes B} \qquad A, B, \Delta \Rightarrow G}{e \sim e' : \Gamma, \Delta \Rightarrow G}$$

that concludes the proof. □

Corollary 1 *$d \sim d'$ iff $T(d) = T(d')$.*

Proof. Induction on d. The case when d ends in $\otimes E$ is treated using Lemma 8. Cf. the proof of the Theorem 6 below. □

Definition 11 *Let $d : \Gamma, \Delta \Rightarrow G$, $e : \Gamma \Rightarrow A \otimes B$ be deductions and $p_0 T(e)$, $p_1 T(e)$ occur in $T(d)$. We say that $T(e)$ is* maximal *in $T(d)$ (and e is* maximal *in d) if no strictly bigger subterm satisfies the same conditions. The list Γ of formulas is* maximal *for d if there is a maximal subterm of the type $\Gamma \Rightarrow C \otimes D$ for some C, D.*

Theorem 5 *For any deduction $d : \Pi \Rightarrow G$ there are deductions $d_1, \ldots,$ d_n maximal for d such that*

$$\frac{d_1 : \Gamma_1 \Rightarrow A_1 \otimes B_1 \ \cdots \ e : A_1, B_1, \ldots, A_n, B_n, \Delta \Rightarrow G}{d \sim d' : \Gamma_1, \ldots, \Gamma_n, \Delta \Rightarrow G}$$

(order of $\otimes E$-inferences being arbitrary), and Δ does not contain maximal lists for d. Derivations $d_1, \ldots, d_n$ are unique up to the order of d_i.

Proof. List all maximal subterms of $T(d)$:

$$T(d_1):\ \Gamma_1 \Rightarrow A_1 \otimes B_1; \ldots T(d_n):\ \Gamma_n \Rightarrow A_n \otimes B_n \qquad (7)$$

By maximality, they do not overlap in $T(d)$, hence the listing (7) is unique and the lists $\Gamma_1, \ldots, \Gamma_n$ are disjoint. Again by maximality, the remaining part of the antecedent Π does not contain maximal lists. Now we use Lemma 8 to obtain required representation of d. □

Definition 12 *Sequence $d_1, \ldots, d_n, e$ with the properties listed in Theorem 5 is called final segment of d'. Deduction is* normal *if it is cut-free and every $\otimes E$-inference is a part of the final segment of corresponding subdeduction.*

Theorem 6 *If d, d' are normal and*

$$T(d) = T(d') \qquad (8)$$

then $d = d'$ up to the order of $\otimes E$-inferences in maximal segments.

Proof. Induction on d. If d is axiom, then d' is axiom and $d = d'$. If d ends in $\Rightarrow I$ then induction hypothesis is applicable. If d ends in $\otimes I$, then d' also ends in $\otimes I$, and antecedent is divided between premisses in the same way: otherwise (8) is violated. Use induction hypothesis. The same argument works for $\Rightarrow E$. If d ends in $\otimes E$, the maximal segment is the same for d, d' except possibly the rightmost derivations $e, e' : A_1, B_1, ..., A_n, B_n, \Delta \Rightarrow G$. We have $T(e) = T(e')$ since (variables for) $A_i, B_i, \Delta, \Gamma_j$ are distinct. Apply induction hypothesis. □

4 Order of Cut-Elimination and Permutation of Rules in a Cut-Free Derivation

We recall here some observations referred to previously (Zucker [1974]).

It is possible to move rules & $\Rightarrow$, $\forall \Rightarrow$ to any place up the derivation inserting and then eliminating "almost trivial" cuts.

Original derivation:

$$\dfrac{\begin{array}{c} A, \Gamma' \Rightarrow G' \\ | \\ A, \Gamma \Rightarrow G \end{array}}{A\&B, \Gamma \Rightarrow G}$$

Insert cut with $A\&B \Rightarrow A$.

$$\dfrac{\dfrac{A \Rightarrow A}{A\&B \Rightarrow A} \qquad \dfrac{\begin{array}{c} A, \Gamma' \Rightarrow G' \\ | \\ A, \Gamma \Rightarrow G \end{array}}{A\&B, \Gamma \Rightarrow G}}{A\&B, \Gamma \Rightarrow G}$$

Make standard cut elimination steps moving cut up the right hand side premise (substituting $A\&B$ for A):

$$\begin{array}{c} \dfrac{\dfrac{A \Rightarrow A}{A\&B \Rightarrow A} \qquad A, \Gamma' \Rightarrow G'}{A\&B, \Gamma' \Rightarrow G'} \\ | \\ A\&B, \Gamma \Rightarrow G \end{array}$$

Finally, eliminate the cut by moving it up the left hand side premise (axiom reduction):

$$\begin{array}{c} \dfrac{A, \Gamma' \Rightarrow G'}{A\&B, \Gamma' \Rightarrow G'} \\ | \\ A\&B, \Gamma \Rightarrow G \end{array}$$

5 Equivalent Form for LJ Free of the Structural Rules

Let us explain Note 4 and the title of this section. It may seem that several of the formulations in the literature, especially Pliuskevicius [1965], Vorob'ev [1970], and Dickoff [1992], already have all necessary properties. The problem is that the transformation of an arbitrary LJ-derivation into a derivation in these formulation does not in general preserve equivalence of derivations. Lemmas 1 and 2 above allow to single out some modifications of the standard logical LJ-rules such that every LJ-derivation can be transformed into an equivalent derivation by these rules (without any structural rules). Recall that the sequents are treated up to the permutation of the formulas to the left of $\Rightarrow$.

- Axioms. The same as previously.
- Inference rules.

  $\Rightarrow \&$

$$\dfrac{\Gamma, \Pi \Rightarrow A \quad \Sigma, \Pi \Rightarrow B}{\Gamma, \Sigma, \Pi \Rightarrow A\&B}$$

$\& \Rightarrow$: the same as before plus

$$\frac{A, A\&B, \Gamma \Rightarrow D}{A\&B, \Gamma \Rightarrow D} \qquad \frac{B, A\&B, \Gamma \Rightarrow D}{A\&B, \Gamma \Rightarrow D}$$

The standard and new versions of the rule $\& \Rightarrow$ can be presented together as follows:

$$\frac{A, (A\&B)^\circ, \Gamma \Rightarrow D}{A\&B, \Gamma \Rightarrow D} \qquad \frac{B, (A\&B)^\circ, \Gamma \Rightarrow D}{A\&B, \Gamma \Rightarrow D}$$

where the superscript $\circ$ shows the possibility that the formula is absent.

The rule $\forall \Rightarrow$ is similar:

$$\forall \Rightarrow \quad \frac{A[t'], (\forall y A[y])^\circ, \Gamma \Rightarrow D}{\forall y A[y], \Gamma \Rightarrow D}$$

The rule $\to\Rightarrow$:

$$\to\Rightarrow \quad \frac{(A \to B)^\circ, \Gamma, \Pi \Rightarrow A \quad (A \to B)^\circ, B, \Sigma, \Pi \Rightarrow D}{(A \to B), \Gamma, \Sigma, \Pi \Rightarrow D}$$

where all four combinations of absence or presence of the main formula $A \to B$ in the premises are possible.

Next three rules are self-explaining:

$$\Rightarrow\to \quad \frac{A^\circ, \Gamma \Rightarrow B}{\Gamma \Rightarrow (A \to B)}$$

$$\Rightarrow \vee \quad \frac{\Gamma \Rightarrow A}{\Gamma \Rightarrow (A \vee B)} \qquad \Rightarrow \vee \quad \frac{\Gamma \Rightarrow B}{\Gamma \Rightarrow (A \vee B)}$$

Since the side formulas of the rules $\vee \Rightarrow$, $\exists \Rightarrow$ can be introduced by a weakening, these rules can be presented as follows:

$$\vee \Rightarrow \quad \frac{A^1, (A \vee B)^0, \Gamma, \Pi \Rightarrow D \quad B^1, (A \vee B)^0 \Sigma, \Pi \Rightarrow D}{A \vee B, \Gamma \Rightarrow D}$$

$$\exists \Rightarrow \quad \frac{(A[y])^1, (\exists x A[x])^0, \Gamma \Rightarrow D}{\exists x A[x], \Gamma \Rightarrow D}$$

All 4 combinations of the absence and presence of $A[y], \exists x A[x]$, and all 16 combinations of the absence and presence of $A, B, A \vee B$ in both premises are possible. Next two rules are not changed.

$$\Rightarrow \exists \quad \frac{\Gamma \Rightarrow A[t_0]}{\Gamma \Rightarrow \exists x A} \qquad \Rightarrow \forall \quad \frac{\Gamma \Rightarrow A[x]}{\Gamma \Rightarrow \forall y A[y]}$$

References

Babaev, A.

[1980] Equality of Canonical Maps in Closed Categories (in Russian), *Izvestija Azerb. Akad. Nauk. Ser. Matem.*, 1980, n. 6.

Davydov, G., Maslov, S., Mints, G., Orevkov, V., Shanin, N., and Slisenko, A.

[1965] An Algorithm for a Machine Search of a Natural Logical Deduction in a Propositional Calculus, in *The Automation of Reasoning*, Ed. J. Sieckmann, G. Wrightson, Springer Verlag, 1983 (Russian original: 1965).

Dickoff, R.

[1992] Contraction-free sequent calculi for intuitionistic logic, *J. Symbolic Logic*, 57 (1992) 795–807.

Gentzen, G.

[1934] Untersuchungen über das logische Schliessen, *Mathematische Zeitschrift*, 39 (1934) 176–210, 405–431.

[1936] Die Widerspruchsfreiheit der reinen Zahlentheorie, *Mathematische Annalen*, 112 (1936) 493–565.

Fenstad, J. (Ed.)

[1971] *Proceedings of the Second Scandinavian Logic Symposium*, North Holland, 1971.

Girard, J.Y.

[1971] Un extension de l'interpretation de Gödel a l'analyse, in Fenstad [1971], pp. 63–92.

[1987] Linear Logic, *Theoretical Computer Science*, 50 (1987) 1–102.

Kleene, S.C.

[1952] *Introduction to Metamathematics*, North-Holland, 1952.

[1952a] Two Papers on the Predicate Calculus, *Memoirs of the American Mathematical Society*, 10, 1952.

Kelly, G.M., and MacLane, S.

[1971] Coherence in Closed Categories, *Journal of Pure and Applied Algebra*, 1 (1971) 97–140.

Kreisel, G.

[1970] Church's Thesis: a Kind of Reducibility Axiom for Constructive Mathematics, in *Intuitionism and Proof Theory*, North-Holland, 1970, pp. 121–150.

[1971] A Survey of Proof Theory II, in Fenstad [1971], pp. 109–170.

Lincoln, P., and Mitchell, J.

[1992] Operational aspects of linear lambda calculus, *Proc. IEEE Symp. on Logic in Computer Science*, 1992, pp. 235–247.

Martin-Löf, P.

[1971] Hauptsatz for the Theory of Species, in Fenstad [1971], pp. 179-216.

Mints, G.

[1974] On E-theorems, in [1992], pp. 105–116 (Russian original: 1974).

[1975] Finite Investigation of Transfinite Derivations, in [1992], pp. 17-72 (Russian original: 1975).

[1977] Closed Categories and the Theory of Proofs, in [1992], pp. 183–212 (Russian original: 1977).

[1979] Stability of E-theorems and Program verification, in [1992], pp. 117–122 (Russian original: 1979).

[1979a] Normalization of Natural Deductions and the Effectivity of Classical Existence, in [1992], pp. 123–146 (Russian Original: 1979).

[1992] *Selected Papers in Proof Theory*, North Holland & Bibliopolis, 1992.

Pliuskevicius, R.

[1965] On a variant of the constructive predicate calculus without structural deduction rules, *Soviet Math. Doklady*, 6 (1965) 416–419.

Pottinger, H.

[1977] Normalization as a Homomorphic Image of Cut Elimination, *Annals of Pure and Applied Logic*, 12 (1977) 323–357.

Prawitz, D.

[1965] *Natural deduction*, Almquist and Wiksell, 1965.

[1971] Ideas and Results in Proof Theory, in Fenstad [1971], pp. 235–308.

Troelstra, A.

[1993] *Natural Deductions for Intuitionistic Linear Logic*, ILLC Prepublication Series, ML-93-09, Amsterdam, 1993.

Troelstra, A., and van Dalen, D.

[1988] *Constructivism in Mathematics, An Introduction*, 2 volumes, 1988, North-Holland.

Vorob'ev, N.

[1970] A new algorithm for derivability in a constructive propositional calculus, *American Math. Society Translations*, ser. 2, 94 (1970) 37–71.

Zucker, J.

[1974] The Correspondence between Cut-elimination and Normalization, *Annals of Pure and Applied Logic*, 7 (1974) 1–156.

The Authors

Henk **Barendregt**
Informatica
Toernooiveld 1
6525 ED Nijmegen
The Netherlands
henk@cs.kun.nl

Jon **Barwise**
Department of Philosophy
Indiana University
Bloomington, Indiana 47405
USA
barwise@phil.indiana.edu

Carlo **Cellucci**
Dipartimento di Studi Filosofici
Università la Sapienza
Via Nomentana 118
00161 Roma
Italy

Francis **Crick**
The Salk Institute
P.O. Box 85800
San Diego, CA 92138-9216
USA

John **Crossley**
Department of Mathematics
Monash University
Clayton, Victoria 3168
Australia
jnc@bruce.cs.monash.edu.au

Charles **Delzell**
Department of Mathematics
Louisiana State University
Baton Rouge, LA 70803
USA
delzell@lsuvax.sncc.lsu.edu

Freeman **Dyson**
Institute for Advanced Study
Princeton, NJ 08540
USA

Anita Burdman **Feferman**
883 Lathrop Drive
Stanford, CA 94305
USA

Solomon **Feferman**
Department of Mathematics
University of Stanford
Stanford, CA 94305
USA
sf@csli.stanford.edu

William **Howard**
Department of Mathematics
University of Illinois at Chicago
Box 4348
Chicago, IL 60680
USA

Verena **Huber-Dyson**
R.R.1
Pender Island
British Colombia, V0N 2M0
Canada

Carl **Jockusch**
Department of Mathematics
University of Illinois
310 Altgeld Hall
Urbana, IL 61801
USA
`jockusch@symcom.math.uiuc.edu`

Hoarst **Luckhardt**
Mathematical Institute
University J.W. Goethe
Robert-Mayer Strasse 6210
D-6000 Frankfurt Main 11
Germany

Angus **MacIntyre**
Mathematical Institute
24–29 St. Giles
Oxford OX1 3LB
Great Britain
`ajm@vax.oxford.ac.uk`

David **McCarty**
Department of Philosophy
Indiana University
Bloomington, Indiana 47405
USA
`lmccarty@silver.ucs.indiana.edu`

Gregory **Mints**
Department of Philosophy
Stanford University
Stanford, CA 94305
USA
`mints@csli.stanford.edu`

Michael **Morley**
Department of Mathematics
Cornell University
Ithaca, NY 14853
USA
`morley@math.cornell.edu`

Anil **Nerode**
Department of Mathematics
Cornell University
Ithaca, NY 14853
USA
`anil@math.cornell.edu`

Piergiorgio **Odifreddi**
Dipartimento di Informatica
Corso Svizzera 185
10149 Torino
Italy
`piergior@di.unito.it`

Rohit **Parikh**
Department of Computer Science
CUNY Graduate Center
33 West 42nd Street
New York, NY 10036
USA
`ripbc@cunyvm.cuny.edu`

Richard **Platek**
Department of Mathematics
Cornell University
Ithaca, NY 14853
USA
`richard@oracorp.com`

Gerald **Sacks**
Department of Mathematics
Harvard University
Cambridge, MA 02138
USA
`sacks@math.harvard.edu`

Helmut **Schwichtenberg**
Mathematisches Institut
der Universität München
Theresienstrasse 39
8000 München 2
Germany
`schwicht@`
`mathematik.uni-muenchen.dbp.de`

Gaisi **Takeuti**
Department of Mathematics
University of Illinois
310 Altgeld Hall
Urbana, IL 61801
USA
`takeuti@symcom.math.uiuc.edu`

Paul **Weingartner**
Institut Für Wissenschaftstheorie
Mönchsberg 2
A-5020 Salzburg
Austria

A.J. **Wilkie**
Mathematical Institute
4–29 St. Giles
Oxford OX1 3LB
Great Britain