Proceedings of the AMS Centennial Symposium
August 8–12, 1988

The Incompleteness Phenomena

HARVEY FRIEDMAN

The incompleteness phenomena have been a principal topic of research of the foundations of mathematics since the work of Kurt Gödel in the 1930's. Incompleteness refers to the following property of most, but not all, formal systems (i.e., set of axioms and rules of inference): that there remain sentences expressed within its language that are neither provable nor refutable within that formal system. Such a sentence is said to be *independent* of the given formal system. The incompleteness phenomena discussed here are distinguished by the variety of mathematical contexts and levels of abstraction represented by the independent sentences, as well as the scope or strength of the formal systems from which the sentences are independent.

To put the incompleteness phenomena in some historical perspective, note that two of the most celebrated revelations in the history of mathematics can be couched in its terms. The irrationality of $\sqrt{2}$ corresponds to the fact that $(\exists x)(x^2 = 2)$ is independent of the order field axioms, and the existence of non-Euclidean geometries corresponds to the independence of the parallel postulate from a suitable formal system for Euclidean geometry in which the parallel postulate is not present.

However, the incompleteness phenomena in the modern sense of the term, relates to formal systems surrounding those strictly mathematical concepts that are currently viewed as the basic notions from which all others are defined. Thus the focus has been on formal systems for natural numbers, and for sets, and also for restricted concepts of set.

The modern incompleteness phenomena obviously have the potential for forcing a reassessment of the foundations of mathematics. However, such a forced reassessment by the mathematics community has not occurred, despite the presently known incompleteness phenomena. We give a brief indication of why this is so.

The currently accepted foundation for mathematics is in terms of the formal system referred to as Zermelo Frankel set theory with the axiom of

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 03E35; Secondary 03C62, 03D35.

This research was partially supported by NSF.

0-8218-0167-8 \$1.00 + \$.25 per page

choice, abbreviated ZFC. This system seems to contain all of the easily identified and intuitive axioms about sets that stem from the usual explanation (story) about the universe of sets (as represented by the so-called cumulative hierarchy). The axiom of choice is a bit of a sore point in that, unlike the other axioms, it unavoidably asserts the existence of sets without naming them explicitly in terms of given sets. But the axiom of choice still has a pretty reasonable story.

It seems that there are no such additional axioms meeting these stringent criteria. Unfortunately, at present there is no theorem to this effect. However, as we discuss later, there certainly are additional proposed axioms, but the stories are comparatively far-fetched. Even for one of the most mild of all such proposed axioms—that of the existence of inaccessible cardinals—the story is already pretty strained.

All but one of the axioms of ZFC hold in the universe of (hereditarily) finite sets. The exception is of course the axiom of infinity. In fact, if one writes down all the obvious natural and intuitive principles about the universe of hereditarily finite sets that do not directly contradict the axiom of infinity, you wind up with a system (more or less) equivalent to ZFC without the axiom of infinity. Thus, in some profound sense that is not yet understood, the usual axioms for set theory (ZFC) are a straightforward adaptation of the usual axioms for finite set theory to an infinite context.

On a more general note, the entire issue of what constitutes an axiom about, say, natural numbers or sets, versus what constitutes merely a fact, is shrouded in mystery. It seems clear that induction is an axiom about the natural numbers, yet "$x^3 + y^3 = z^3$ fails universally" is a fact and not an axiom.

Further evidence that mathematicians are complacently satisfied with ZFC as the foundation for mathematics (to the extent that they think about foundations) is that every proof put forth in mathematics to date by mainstream mathematicians is straightforwardly formalizable in even small portions of ZFC, with rather rare, minor, and easily removable exceptions. On those rare occasions when mathematicians use something outside of ZFC—most commonly the continuum hypothesis—they state the outside assertion as a *hypothesis* to the theorem in question, thus staying within ZFC.

On the other hand, the continuum hypothesis *is* independent of ZFC [**Go1**, **Co**]. Furthermore, the continuum hypothesis is just about the most basic and fundamental question that can be raised in the context of set theory. In fact, the question was raised by Georg Cantor early in his initiation of set theory, and appears as the first problem on Hilbert's famous problem list.

So why has the independence of such a fundamental question not caused a crisis in the foundations of mathematics, and rendered ZFC obsolete?

We believe that the fundamental reason is the relative intellectual distance from the continuum hypothesis to finitary problems in mathematics.

To make the point in an extreme way, suppose that, instead of the continuum hypothesis, the twin prime conjecture (or some similar question about the infinitude of the prime pairs) was shown to be independent from ZFC. The mathematical community would be thrown into a foundational crisis. If, as is likely, such a result would be accompanied by a proof (of the conjecture about prime pairs) from some understandable extension of the ZFC axioms, then great interest would attach to the question of whether the additional axioms should be adopted. Having different answers or no answers to such questions about prime pairs according to which extensions of the ZFC axioms you postulate, would be regarded as wholly undesirable, and a uniform response would be widely sought.

It is not simply a matter of the independent sentence being about finite objects such as natural numbers. Thanks to Kurt Gödel, we already know that for any system such as ZFC there are sentences which are independent and which are, in a sense, even more finitary than twin prime conjectures. In fact, the consistency of ZFC itself is one such. (The latter result, known as the Gödel second incompleteness theorem, has terribly profound meaning for the foundations of mathematics in another direction.) Furthermore, it is also known that in systems such as ZFC there are always Diophantine equations over the integers such that the existence of solutions is independent of ZFC, using work of Gödel and Matijacevic [**Mat**], yet any such known example is truly gargantuan in size. *It is clearly also a matter of subject matter.*

To encampsulate: The continuum hypothesis is too infinitary. The consistency of ZFC is not a basic mathematical question (though it is a basic metamathematical or logical question). No remotely reasonable Diophantine equation is anywhere near being shown to display any kind of independence. And no twin prime conjecture is anywhere near being shown to be independent of ZFC.

The fundamental issue is this: Is there a basic mathematical problem about standard finite objects such as, say, natural numbers or rational numbers or polynomial rings over finite extensions of the rationals, etc., with a clear and intuitive meaning, conveying interesting mathematical information, that is readily graspable, and which is independent of ZFC?

We speculate that sometime during the twenty-first century, someone will answer the above question in the affirmative, and there will be nearly universal agreement in the mathematics community that this has been accomplished. Furthermore, there will be proofs of such mathematical problems accompanying such independence results using some of the extensions of ZFC that have already been explored in the set theory community. The current state of the art regarding this conjecture is discussed in the Appendix.

Before beginning the detailed discussion of various incompleteness phenomena, we give an informal description of the axioms of ZFC for the reader's convenience.

In ZFC, every object is a set, and the primitive relations between sets are that of equality and membership.

Informally, the axioms are as follows.

(1) *Extensionality.* Two sets are equal if and only if they have the same elements.
(2) *Pairing.* For any two sets x, y, $\{x, y\}$ exists.
(3) *Union.* For any set x, $\bigcup x$ exists, which is the set of all elements of elements of x.
(4) *Power set.* For any set x, the power set $\mathscr{P}(x)$ exists, which is the set of all subsets of x.
(5) *Separation* (comprehension). For any set x, $\{y \in x : A(y)\}$ exists, where A is any set theoretically describable predicate; A is allowed to mention specific sets called *parameters.*
(6) *Infinity.* There are many equivalent forms this axiom can take but the following is customary: There is a set ω which contains the empty set $\varnothing$ as an element, and for every $x \in \omega$, we have $x \cup \{x\} \in \omega$.
(7) *Axiom of choice.* Again there are many equivalent forms this can take, and the following is customary: For every set of pairwise disjoint nonempty sets, there is a set which meets each of these nonempty sets at exactly one place.
(8) *Replacement.* This asserts that for any set x and any function from x into sets that is set-theoretically described (with parameters as in 5 above), the range of this function exists as a set.
(9) *Foundation.* Every nonempty set possesses an ε-minimal element, i.e., an element which is disjoint from the given set.

Mathematicians seldom use axioms (8) or (9).

Some good works on set theory include **[Je2]** and **[Levy]**.

An important line of research that goes in a different direction than that emphasized here is the work on the projective hierarchy of sets (of real numbers). This hierarchy begins with the Borel sets and then grows upward in complexity through the operations of projection and complementation. The goal is to understand the properties and structure of the projective sets as thoroughly as we understand the Borel sets (in the sense of classical descriptive set theory). After a couple of levels or so in the hierarchy, we know that ZFC is not sufficient to do anything interesting along these lines. However, under the additional axiom of constructibility (or in the constructible universe of sets), all appropriate questions about the projective sets are answered. Alternatively, using axioms for large cardinals, again all appropriate questions are answered, but typically with different answers. See **[Mar]** and **[MS]**.

1. General incompleteness phenomena. These are the incompleteness properties that apply to a very wide class of formal systems. The major such results are due to Kurt Gödel **[Go3, Smo]**:

FIRST INCOMPLETENESS THEOREM. *In any formal system such as ZFC, there always are sentences which are neither provable nor refutable.*

SECOND INCOMPLETENESS THEOREM. *In any formal system such as ZFC, the consistency of the system itself is not provable in the system.*

For the first incompleteness theorem, we need to know that the system is effectively axiomatized, is consistent (i.e., free of contradiction), and contains a small amount of basic integer arithmetic.

For the second incompleteness theorem, we additionally need to know that provability in the system can be adequately formalized within the system. In fact, the modern formulations of the theorem assert that no adequate formalization of the consistency of the system will itself be provable in the system, and give particular families of adequate formalizations, which include usual intuitively based ones.

As spectacular as these results are, they do not provide examples of mathematically motivated problems which cannot be proved or refuted in ZFC, as discussed in the introduction here. This came later.

But these results did put an end to Hilbert's program, one of whose goals was to secure the consistency of mathematics within weak principles of integer arithmetic.

However, one of the most interesting of all the *completeness* phenomena is the result of **[Tarski]** that the axioms of real closed fields are complete. These axioms augment the ordered field axioms by the axioms which assert that every single variable polynomial of odd degree (with leading coefficient 1) has a zero, and every positive element has a square root. The system is effectively axiomatized, and so appears to violate the first incompleteness theorem. However, note that the real closed field axioms do not contain basic integer arithmetic (only real number arithmetic).

Another general incompleteness phenomenon is obtained by combining work of Gödel and Turing with the result of **[Mat]** on the nonrecursiveness of the solvability of Diophantine equations over the integers (Hilbert's 10th problem). Also see **[DMR]**. The following result is from the folklore:

THEOREM. *In any formal system satisfying the usual hypotheses for the first incompleteness theorem such as ZFC, there always is a Diophantine equation over the integers which is unsolvable, yet cannot be proved to be unsolvable in the system.*

As mentioned in the introduction, this theorem does not give any reasonable example of such a Diophantine equation. It is an open question whether for any system such as ZFC, there is such a Diophantine equation that can be written down on a page or so with all coefficients and exponents written out in base 10.

2. Independence results via forcing. The forcing method was introduced in **[Co]** and provides a general method for obtaining new models of ZFC from given ones by adjoining new objects. The resulting new models will in general satisfy different sentences than the given models, and, therefore, one can show that certain sentences are independent of ZFC in this way.

More specifically, in the modern treatment of forcing one begins with a countable model of ZFC, called the ground model. One chooses a partially ordered set from this ground model, called the "notion of forcing." Then one defines a family of sets called the generic sets with respect to the notion of forcing. If the notion of forcing satisfies some minimal conditions, then there are continuumly many such generic sets. Then one proves that if any generic set is adjoined to the ground model, the resulting model is a model of ZFC. Models of ZFC obtained in this way are called generic extensions. Furthermore, there are methods which are useful in determining whether sentences hold in generic extensions in terms of whether related sentences hold in the ground model about the notion of forcing. To apply the method in order to show that some particular sentence of interest is consistent with ZFC, one chooses a suitable ground model and delicately adjusts the notion of forcing in order that the generic extensions satisfy the given sentence.

The method has been extensively developed, streamlined, and unified by set theorists since **[Co]**, and has been quite successful in establishing independence results of a certain kind. However, the method has inherent limitations, particularly for establishing the independence of sentences of more than a certain level of concreteness as we presently indicate.

Firstly, every generic extension has the same ordinals as the ground model. In particular they have the same integers, and in fact the same basic arithmetical operations. From this it is clear that any sentence about the natural numbers must hold in the generic extension if and only if it holds in the ground model. The same assertion holds for any sentence about finite objects. Therefore we cannot hope to establish the independence of any sentence about finite objects through (at least a direct application of) the method of forcing.

Secondly, every generic extension also has the same functions on ordinals defined by transfinite recursion as the ground model. It is known that every sentence about natural numbers and (possibly infinite) sets of natural numbers that is not too complicated can be reduced in ZFC to a corresponding sentence about ordinals and functions on ordinals defined by transfinite recursion. Therefore, in the same way as we saw in the previous paragraph, we cannot hope to establish the independence of any sentence about finite objects and sets of finite objects at or below a certain level of complexity, through (at least a direct application of) the method of forcing. The precise result we are using is that every Π^1_3 sentence that holds in the generic extension must also have held in the ground model, and so one cannot prove the consistency of a Π^1_3 sentence directly through the method of forcing (see, e.g., **[Je2**, pp. 530–531]).

The method of forcing in its original form was also used to obtain models of ZF in which the axiom of choice fails. This aspect of the forcing method has not been given as neat a streamlining and unification as it has in the context of ZFC models. Nevertheless, most of the important independence results about ZF without the axiom of choice can be proved by constructing appropriate generic extensions satisfying ZFC, and then taking a suitable submodel to obtain the desired model satisfying ZF in which the axiom of choice fails.

The first application of the forcing method was the independence of the continuum hypothesis which we state as follows:

PROPOSITION 2.1. *Every uncountable set of real numbers is in one-one correspondence with all of the real numbers.*

The consistency of the above with ZFC is from **[Go1]** and the consistency of the negation of the above with ZFC is from **[Co]**.

Another application is to Souslin's hypothesis which we formulate as follows. The consistency is from **[ST]** and the consistency of the negation is from **[Je1]**:

PROPOSITION 2.2. *Every nonseparable linearly ordered set has an uncountable subset in which every element is isolated.*

Another application is to Whitehead's group conjecture, which has been proved in the case of countable groups (see **[Fu]**). The independence is from **[Sh]**:

PROPOSITION 2.3. *If* $\operatorname{Ext}(G, Z) = 0$ *then* G *is free* (*for Abelian* G).

Another application is to Kaplansky's conjecture which we state as follows. (For the consistency of the negation, see **[Dales]**, where the negation is actually proved from the continuum hypothesis. For the consistency, see **[DW]**.)

PROPOSITION 2.4. *Any homomorphism from the Banach algebra* $C[0, 1]$ *into any* (*separable*) *Banach algebra is continuous.*

Yet another application is to a generalization of Fubini's theorem, in which the hypothesis of two-dimensional measurability is relaxed. The consistency of its negation is clear since the generalization is refutable using the continuum hypothesis (folklore). For the consistency of this and other strenghthenings see **[Ship]**.

PROPOSITION 2.5. *If* $F : [0, 1]^2 \to [0, 1]$ *has almost all* F_x, F^y *measurable, then* $\int(\int F(x, y)\, dx)\, dy = \int(\int F(x, y)\, dy)\, dx$.

All of the above examples should be regarded as set-theoretic in that they involve unrestricted selections from uncountable domains. For example, in the statement of the continuum hypothesis above, we refer to arbitrary sets of real numbers. We can be more specific about the kinds of sets of real numbers to be considered by imposing a regularity condition. The most common regularity conditions on subsets of separable metric spaces are that of measurability, Borel measurability, and various strengthenings of Borel measurability. Measurability works well in contexts in which measure 0 sets are regarded as equivalent. If they are not regarded as such, then obviously one really does not have a regularity condition per se, since the measure 0 sets are as badly behaved as arbitrary sets.

From the logical point of view, Borel measurability makes sense as a kind of minimal regularity condition to impose. The Borel measurable subsets of complete separable metric spaces form a very wide class of objects for the vast majority of mathematical purposes and in an appropriate sense constitute (or at least include all) those subsets which are constructed via sequential processes.

Thus, in essence, the imposition of the regularity condition of Borel measurability removes mathematically undesirable and irrelevant pathology from the context. In §4, we explore the effect this "Borel point of view" has on the incompleteness phenomena.

We mention an independence result involving set-theoretically definable sets of real numbers. The consistency of the negation with ZFC follows from **[Go1, Go3]**, and the consistency with ZFC is from **[Sol]**:

PROPOSITION 2.6. *Every cross section of every definable set of real numbers is measurable and has the property of Baire. If the cross section is uncountable then it has a perfect subset.*

We can consistently add measurability, the Baire property, and uncountability implies perfect subsets for *all* sets to ZF (i.e., ZFC without the axiom of choice). But this is not so interesting without also having some choice. Fortunately, in **[Sol]** dependent choice is added to ZF for this result, which is enough choice to prove the basic facts about measurability, the property of Baire, and uncountability.

3. The constructible point of view. Needless to say, the incompleteness phenomena involving ZFC are not desirable features of the commonly accepted foundation for mathematics. It is natural to explore possible remedies for the situation short of overhauling ZFC.

We have already hinted at one possible remedy which will be explored in the next section. That is the remedy of imposing the regularity of Borel measurability on the objects considered. Of course, this does not make the original sentences any less independent of ZFC than they were before, but it does give a general process for removing the offending pathology that might be responsible for the difficulties while preserving the essential mathematical content of the original sentences.

In this section we explore a different remedy. We consider the effect of placing a general regularity condition of a logical nature on the set concept itself.

The usual modern description of the universe of sets is in terms of the cumulative hierarchy. This hierarchy associates a family of sets to every ordinal α. The sets are just the sets that appear somewhere in the hierarchy. Of course there is a circularity here since it is also customary to define the ordinals as certain kinds of sets, but this is usually ignored since the hierarchy is used as an informal description to motivate the axioms of ZFC.

This cumulative hierarchy is given as follows: $V_0 = \varnothing$, $V_{\alpha+1} = \wp(V_\alpha)$, and $V_\lambda = \bigcup_{\beta<\lambda} V_\beta$ for limit ordinals λ. Here $\wp$ stands for the power set operation—the family of all subsets of the set to which it is being applied. It is provable in ZFC that every set appears somewhere in this hierarchy. The class of all sets is denoted by V.

Now observe that there are really two quite different operations that drive this cumulative hierarchy. One is the ordinals and the process of transfinite recursion, and the other is the power set operation.

From the constructible point of view it is the power set operation that is suspect. All objects should be constructed on the basis of some general form of transfinite recursion, where "events take place on the basis of earlier events." In this sense, the power set operation must be derived from something more fundamental; every set that exists must be constructed from earlier constructed sets in some way.

The constructible point of view originated with Kurt Gödel in his proof of the consistency of the continuum hypothesis, where he introduces the so-called constructible hierarchy. Although he did briefly hold at least some variant of the constructible point of view, he quickly renounced it in favor of a strongly Platonist point of view now common among specialists in set theory (see **[Go2]**).

The constructible hierarchy is given as follows: $L_0 = \varnothing$, $L_{\alpha+1} =$ the set of all subsets of L_α that are explicitly definable over L_α (allowing parameters for elements of L_α), and $L_\lambda = \bigcup_{\beta<\lambda} L_\beta$. The class of all constructible sets (i.e., sets that appear somewhere in this hierarchy) is denoted by L.

L has many desirable properties. Within ZFC, we can prove that L obeys all of the axioms of ZFC. We can prove this even within ZF. This latter fact is what allowed Gödel to conclude that the axiom of choice was consistent with ZF. Put somewhat differently: L obeys the axiom of choice for a good reason, whereas V obeys it by conventional wisdom.

If the constructible hierarchy is modified in small ways, then we still provably get the same class of sets.

If we start with the constructible hierarchy as the point of departure in motivating the axioms of ZFC, we can use the same story that we use for motivating ZFC from the cumulative hierarchy (in fact, the story for the axiom of choice is much improved), except for the power set axiom. This is not surprising since the power set axiom is explicitly part of the mechanism of the cumulative hierarchy. However, reasonable extensions of the story for the replacement axiom in the constructible hierarchy can be given which will motivate the power set axiom in the constructible hierarchy. To more fully clear up the philosophical issues here we need to develop an appropriate general theory of transfinite iteration which applies to contexts much more general than set theory. It is likely that this can be done.

Regardless of these philosophical niceties, we can now effectively regard constructibility (i.e., membership in L) as a kind of regularity condition on sets.

But what happens to our mathematical proofs if we restrict all mathematical objects to constructible objects?

From what we have said above, it is clear that if we start with a proof in ZFC, then the result of this uniform restriction to L is still a proof in ZFC. We just have to attach proofs of the L-restricted forms of the axioms of ZFC that are used in the original proof; these attached proofs can be themselves given in ZFC. And since these attached proofs have already been given by Gödel in **[Go1]**, there is no need for anyone to do anything other than what they are doing now.

Now that it is clear that restricting to constructible sets is not of any real operational consequence for mathematicians (other than some set theorists who operate outside of ZFC), conceptually speaking how much of a restriction is constructibility?

It follows from what has been said above that it is consistent with ZFC that $V = L$, i.e., all sets are constructible. Thus there is no way to construct a nonconstructible set within ZFC. This effectively reduces the level of restriction for mathematicians other than some set theorists to nil.

On the other hand, what is the advantage of everybody simply declaring that they are using only constructible sets, functions, numbers, etc.?

The advantage is that if, e.g., Propositions 2.1–2.6 are reinterpreted as being about constructible sets (functions, and numbers, etc.), then the independence results associated with them disappear.

More specifically, the following is proved in **[Go1]**:

THEOREM 3.1. *Proposition* 2.1 *holds in the constructible universe.*

The following is due to **[Jensen]** (and see **[Je2**, pp. 226–229**]**):

THEOREM 3.2. *Proposition* 2.2 *fails in the constructible universe.*

The following is proved in **[Sh]**:

THEOREM 3.3. *Proposition* 2.3 *holds in the constructible universe.*

For the following see **[Dales]** since the continuum hypothesis holds in L:

THEOREM 3.4. *Proposition* 2.4 *fails in* L.

The following is a consequence of the continuum hypothesis holding in L:

THEOREM 3.5. *Proposition* 2.5 *fails in the constructible universe.*

The following is proved in **[Go1, Go3]**:

THEOREM 3.6. *Proposition* 2.6 *fails in the constructible universe.*

In fact, virtually all sentences that have been proved to be independent from ZFC by a direct application of the forcing method have now been decided when restricted to the constructible universe.

The axiom of constructibility asserts that $V = L$, i.e., all sets are constructible. Although it is not as obvious as it sounds, Gödel proved that the axiom of constructibility holds in the constructible universe.

From a purely operational point of view, there is no functional difference between assuming the axiom of constructibility and deciding to restrict oneself to constructible sets only.

However, it is *not* a tenant of the constructible point of view that the axiom of constructibility is somehow evidently true, or is even true at all. This would be like saying that a mathematician who imposes the regularity condition of differentiability of functions in his work somehow believes that all functions are differentiable. Constructibility is merely intended to be a regularity condition.

This author is quite sympathetic to the constructible point of view. We would like to go even further. We believe that the usual description of the set-theoretic universe is not sufficiently clear to "determine" an answer to even such a set-theoretically fundamental question such as the continuum hypothesis. The unrestricted power set of infinite (and especially uncountable) sets become a vague blur when examined too intensely.

We also believe that the constructible point of view is not going to prove to be sufficiently powerful to avoid all of the foundational difficulties that we suspect will arise. In particular, it is obviously helpless in dealing with the status of sentences about finite objects since they are already constructible. It is also of no use in dealing with the status of not too complex sentences about sets of natural numbers (Π^1_2 sentences) since they are provably equivalent to their restrictions to the constructible universe. In fact, we later discuss examples of sentences about Borel functions on groups and graphs which remains independent of ZFC even when restricted to the constructible universe.

4. The Borel measurable point of view. The Borel measurable point of view is based on a quite natural mathematical regularity condition.

We start with a complete separable metric space. The Borel measurable sets constitute the least σ-algebra containing the open sets. (Henceforth we omit the word "measurable.") The Borel functions are those functions for which the inverse image of every open set is Borel.

The Borel functions can be arranged in a tower of length ω_1 where we start with the class of continuous functions, and at every nonzero ordinal we take the class of all everywhere defined sequential limits of functions from the earlier classes. These concepts and this construction have obvious generalizations to the case of Borel functions between two spaces, and also Borel functions of finite or even countably infinitely many arguments (using product constructions for metric spaces). It is natural to also consider partially defined Borel functions, which are merely the restrictions of Borel functions to Borel subsets of the relevant space. Recall that there is a one-one onto Borel function with Borel inverse between any two uncountable complete separable metric spaces.

Throughout mathematics, one works with structures in the sense of a nonempty set endowed with distinguished elements, relations, and (partial) functions. A structure is said to be a Borel structure if its domain is a Borel set of real numbers, and its relations and functions are all Borel. (Sometimes it is convenient to allow equality to be represented by a Borel equivalence relation.) Most of the important structures in mathematics are naturally isomorphic to Borel structures. Separable Banach spaces form a natural family of such structures.

The Borel point of view takes the position that all mathematical structures to be considered are Borel structures (or naturally equivalent to such) and all sets and functions to be considered are Borel sets and Borel functions in and between Borel structures.

How severe is such a regularity condition?

Upon examination, it appears not to be very restrictive. Virtually all of the more important and intensively studied mathematical structures are Borel and the same is true of particular examples of sets and functions.

The typical case of where one goes beyond Borel in mainstream mathematics is where one is developing a general theory, say of groups or fields. A lot of useful facts simply can be proved without restricting the algebraic objects in some nonalgebraic way, such as being Borel representable. However, if one imposes the Borel regularity condition, then the theory is not generally any easier, and no mathematical content is lost in the theory. In particular, all of the examples one normally applies the theory to that are of central interest are generally relatively concrete in nature and meet the Borel regularity condition.

What is the advantage of imposing Borel regularity conditions?

In many cases, sentences independent of ZFC have straightforward reinterpretations using Borel regularity. Typically, the resulting sentence is no longer independent of ZFC. This is the case for Propositions 2.1–2.6 as follows:

The following is implicit in, e.g., [**Luzin**]:

THEOREM 4.1. *Every uncountable Borel set of real numbers is in Borel one-one correspondence with the set of all real numbers.*

The following is proved in [**HMS**]:

THEOREM 4.2. *Every nonseparable Borel linear ordering of the reals has an uncountable Borel subset in which every element is isolated.*

The following is proved in [**Sp**] (that G is not Borel free is in the folklore):

THEOREM 4.3. *G is free, but G is not Borel free, where G is the group of bounded infinite sequences of integers.*

In the above, Borel free means that there exists a Borel set of independent generators. The connection with Proposition 2.3 is that a free group G always has $\mathrm{Ext}(G, \mathbf{Z}) = 0$. However, we can still ask whether Proposition 2.3 is true for Borel groups, with the usual (non-Borel) notion of free group.

The following is proved in **[Ajtai]**, but goes back to Laurent Schwartz:

THEOREM 4.4. *Any Borel homomorphism from one separable Banach space to another is continuous.*

The following is classical:

THEOREM 4.5. *Fubini's theorem for Borel functions from the square into itself.*

The following is classical (see, e.g., **[Luzin]**):

THEOREM 4.6. *Borel sets are measurable, have the property of Baire, and if uncountable have perfect subsets.*

We are also sympathetic to the Borel point of view. It raises an interesting issue as to the proper role of generality in mathematics.

The rest of the discussion of the incompleteness phenomena will almost exclusively focus on the independence of sentences that are admissible from the Borel point of view.

Many of the independence results discussed are not independence results from the full ZFC axioms, but rather from significant fragments of ZFC. In fact, these independence results from fragments of ZFC that are discussed here are in fact theorems of ZFC.

It is natural to inquire as to the significance of such independence results since mathematicians generally accept all of ZFC. We give two replies.

Firstly, as noted above, virtually all of mathematics done outside of set theory is easily formalizable in surprisingly weak fragments of ZFC. This immediately raises the important and interesting question of whether and to what extent the axioms of ZFC (beyond such weak fragments) are useful or relevant to mathematics.

Secondly, we are still very far from a really convincing mathematically basic and interesting example of a theorem of ZFC about finite objects which uses more than, say, the part of ZFC that applies to countable sets; e.g., such a theorem of ZFC which cannot be proved in ZFC with the power set axiom deleted. It is only since 1977 that we have had a pretty convincing such example which cannot be proved in finite set theory (ZFC without the axiom of infinity), and since 1981 that we have gone beyond significant parts of countable set theory. Such fragments of ZFC form significant barriers to progress towards ZFC and beyond, and also have intrinsic interest.

5. Cantor's theorem and the discrete topology. We now discuss a theorem from **[Fr1]** about Borel functions which arises from an examination of the proof of Cantor's fundamental theorem that the reals are uncountable. This is an example of a basic theorem about Borel functions whose proof rather noticeably must take one quite far from the context in which it is expressed. In particular its proof cannot be given in what may be called separable mathematics. In separable mathematics, all of the objects one works with are countable, or at least can be described completely in countable terms. This allows for complete separable metric spaces, since they can be specified by the restriction of the metric space to any countable dense set. Elements in complete separable metric spaces are also admissible since they can be specified by any sequence from the countable dense subset that converges to it. Continuous functions between complete separable metric spaces can be specified by their restriction to any countable dense set. Borel functions can be specified by the countable process from which they are built (Borel codes).

From the axiomatic point of view, an appropriate system that reflects the above conception of separable mathematics is obtained by deleting the power set axiom from ZFC. The resulting system is written as $\mathrm{ZFC}\setminus\wp$. The (hereditarily) countable sets form a model of this system.

Cantor's theorem can be stated as follows. Let $x_1, x_2, \ldots \in I$, where I is the closed unit interval. Then there exists $y \in I$, $y \neq x_1, x_2, \ldots$.

Standard methods for constructing Borel functions establish rather easily that there is a Borel $F : I^\infty \to I$ such that for all $x \in I^\infty$, $F(x)$ is not a coordinate of x. For example, the following function obeys this property and is Borel: Take $F(x) = \bigcap_n J_n$, where each J_{n+1} is the first closed dyadic rational interval of length at most 2^{-n} contained in J_n which is disjoint from $\{x_1, \ldots, x_n\}$, and $J_1 = I$. (Any listing of the closed dyadic rational intervals will do for this construction.)

However, note that the value of F at a sequence may depend on the order in which that sequence is given, not just on the image of the sequence (even if multiplicities are counted). This leads to the following question: Is there such a Borel function which is permutation invariant, i.e., obeying $F(\sigma x) = F(x)$, for all permutations σ?

The answer is no. The following is proved in **[Fr1]**:

THEOREM 5.1. *Every permutation invariant Borel $F : I^\infty \to I$ sends some point to a coordinate of itself.*

The proof uses the topology $(\mathbf{I})^\infty$, where $\mathbf{I}$ is the closed unit interval endowed with the *discrete* topology, i.e., the product of countably infinitely many copies of $\mathbf{I}$. The Baire category theorem can be stated and proved in this context. One can also prove a $0, 1$-law for Baire category which states that every permutation invariant Borel subset of $\mathbf{I}$ is meager or comeager. Since every Borel subset of $\mathbf{I}$ (with the usual separable topology) is also a Borel subset of $\mathbf{I}$, we see that every permutation invariant Borel subset of $\mathbf{I}$

is meager or comeager in the sense of **I**. One can then prove by standard techniques that there is a $c \in I$ such that the given function F is constantly c on a comeager set in the sense of **I**. (So far we have not really used the nonseparability of **I**.) But by heavy use of the nonseparability of **I**, we see that comeagerly many x contain c as a coordinate in the sense of **I**. (The latter is false for Baire category or measure on the usual I). Hence for at least one x, $F(x)$ is a coordinate of x.

The following is also proved in **[Fr1]**:

THEOREM 5.2. *Theorem* 5.1 *cannot be proved in ZFC*$\backslash\wp$. *Hence, in the appropriate sense, the theorem cannot be proved within "separable mathematics."*

We sketch some of the ideas in this proof.

It suffices to prove that there is a model of ZFC$\backslash\wp$ from the axioms of ZFC$\backslash\wp$ together with Theorem 5.1. For then, if ZFC$\backslash\wp$ were to prove Theorem 5.1, then ZFC$\backslash\wp$ would prove the existence of a model of ZFC$\backslash\wp$, and hence by Gödel's second incompleteness theorem, ZFC$\backslash\wp$ would be inconsistent, which it is not.

Next, we introduce a system called second-order arithmetic, and written as Z_2. Despite its name, it is an ordinary first-order formal system like all of the ones we have been discussing. It has variables over natural numbers and over sets of natural numbers, contains the usual arithmetic of addition and multiplication, the axiom scheme of induction, and most importantly, the comprehension scheme which asserts that each $\{n|\varphi(n)\}$ exists, where φ may mention numbers and sets of numbers as parameters, and have quantifiers over all numbers and over all sets of numbers. It is known how to build a model of ZFC$\backslash\wp$ from a model of Z_2 directly, and in particular within ZFC$\backslash\wp$. Models in general do not have to have only standard integers (they may have nonstandard ones), but if the original model of Z_2 has only standard integers then the resulting model of ZFC$\backslash\wp$ also has only standard integers.

Combining the above two paragraphs, we now see that it suffices to construct a model of Z_2 with only standard integers using only ZFC$\backslash\wp$ and Theorem 5.1.

There are still real difficulties in obtaining such a model relating to the parameters that are allowed in the comprehension axiom scheme above. So the crucial next step, carried out in detail in **[Fr1]**, is the consideration of $p - Z_2$, which is the same as Z_2 except no parameters are allowed in the comprehension scheme. It is shown in **[Fr1]** how to go from a model of $p - Z_2$ with only standard integers to a submodel with only standard integers obeying Z_2. Again this construction can be done explicitly within ZFC$\backslash\wp$.

Combining the above three paragraphs, it is clear that it suffices to construct a model of $p - Z_2$ with only standard integers using only ZFC$\backslash\wp$ and Theorem 5.1.

Now to every sequence $x \in I^\infty$ we can associate a family of sets of natural numbers $M(x)$, which can be viewed as an attempted model of $p - Z_2$ with only standard integers. We can use any Borel correspondence of I with $\wp(\omega)$ for this purpose, taking $M(x)$ to be the image of x under this correspondence. Of course we can assume that $M(x)$ is viewed as being equipped with numbers and arithmetic. Thus the only possible reason that $M(x)$ does not satisfy the desired $p - Z_2$ is that the parameterless comprehension axiom scheme might fail.

We now let $F(x)$ be obtained by looking up the first instance of parameterless comprehension that fails in $M(x)$ and taking the image of the missing set under the above chosen Borel correspondence to be $F(x)$. (If parameterless comprehension holds, i.e., if $M(x)$ satisfies Z_2, then we are done anyway, but in this case let $F(x)$ be 0 by default.)

Careful consideration of the construction of F reveals that it is a permutation invariant Borel function.

Applying Theorem 5.1, there is an x such that $F(x)$ is a coordinate of x. Tracing through the construction of F, we see that the only way this can happen is the default case above, and hence $M(x)$ must satisfy $p - Z_2$ as desired.

The following two related results are proved in **[Fr2]**: Let K be the Cantor space consisting of the infinite sequences of 0's and 1's. The important shift map is given by $s(x) = (x_2, x_3, \ldots)$, where $x = (x_1, x_2, x_3, \ldots)$, i.e., shift deletes the first term. We say that $F : K \to K$ is shift invariant if it obeys $Fsx = Fx$. We also let $x^{(2)} = (x_1, x_4, x_9, \ldots)$.

THEOREM 5.3. *Every shift invariant Borel function $F : K \to K$ is somewhere its "square," i.e., for some x, $F(x) = x^{(2)}$.*

THEOREM 5.4. *Theorem 5.3 cannot be proved in ZFC$\backslash\wp$. Hence in an appropriate sense, it cannot be proved within "separable mathematics."*

Theorems 5.1 and 5.3 can be proved just beyond ZFC$\backslash\wp$. For example, if we add the existence of $\wp(\omega)$ to ZFC$\backslash\wp$, then the resulting system is powerful enough to prove these two theorems.

Theorems 5.1 and 5.3 are examples of what we call Borel diagonalization theorems. Such theorems assert that there are no Borel diagonalization functions with certain invariance properties.

Looking at such theorems conversely, they illustrate the following general principle which we do not know how to formulate in anything like full generality:

GENERAL THEME. Every "invariant" Borel function from one "space" into another sends some element to a "simpler" element.

6. Borel diagonalization on equivalence relations, linear orders, groups, and graphs. In this section we present some more powerful Borel diagonalization theorems than the basic Theorem 5.1 above. They very clearly illustrate the General Theme.

These diagonalization theorems fall into three basic categories:

CLASS A. These are the theorems of ZFC which, like Theorems 5.1 and 5.3, can be proved just beyond countable set theory (e.g., in $\mathrm{ZFC}\setminus\wp + \wp(\omega)$ exists), but not within $\mathrm{ZFC}\setminus\wp$ (or separable mathematics).

CLASS B. These are the theorems of ZFC which can be proved just beyond ZC but not within ZC itself. For instance, they can be proved within systems such as ZC $+V(\omega+\omega)$ exists, or $\mathrm{ZFC}\setminus\wp + V(\omega+\omega)$ exists. Here ZC is Zermelo set theory with the axiom of choice, which is obtained from ZFC by the removal of the replacement axiom scheme (and optionally, removal also of the formulation axiom).

CLASS C. These are the theorems of ZFC which can be proved using uncountably many iterations of the power set operation, but not using any (explicitly given) countable number of such iterations. In particular they cannot be proved within Zermelo set theory with the axiom of choice, ZC.

The phrase "iterations of the power set operation" needs some explanation. Recall the cumulative hierarchy as presented in §3. The stages in the hierarchy represent iterations of the power set operation. The ordinal number of the stage represents the number of iterations. Thus when we say that we have uncountably many iterations of the power set operation, we mean that we have, for each countable ordinal α, the stage V_α. The system ZC is easily seen to correspond to having $\omega+\omega$ iterations of the power set operation, i.e., having each $V_{\omega+n}$, where n is finite.

We first consider a direct generalization of Theorem 5.1. Let E be any equivalence relation on I. We use [ ] for the equivalence classes under E. For $S \subseteq I$, we write $[S]$ for $\{[x] : x \in S\}$.

We say that the Borel diagonalization theorem holds for E if there is no Borel function $F : I^\infty \to I$ such that (a) if $[\mathrm{rng}(\overline{x})] = [\mathrm{rng}(\overline{y})]$ then $[F(\overline{x})] = [F(\overline{y})]$, and (b) $[F(\overline{x})] \notin [\mathrm{rng}(\overline{x})]$, for all $\overline{x}$.

The following is proved in **[Fr1]**:

THEOREM 6.1. *The Borel diagonalization theorem holds for any Borel equivalence relation* E. *Furthermore, this theorem is in class* C.

A set $E \subseteq I^n$ is called analytic if it is of the form $\{x : \text{for some } y, (x, y) \in S\}$ for some Borel set $S \subseteq I^{n+1}$. Analytic sets go well beyond Borel sets from a conceptual point of view since they are obviously not constructed by any countable limit process. Analytic sets form the next natural step up in abstraction or complexity from Borel sets in what is called the projective hierarchy, which we discussed briefly at the end of the introduction.

The relevance here of analytic sets is that Theorem 6.1 was extended in **[St]** as follows:

THEOREM 6.2. *The Borel diagonalization theorem holds for any analytic equivalence relation* E. *Furthermore, this theorem is in class* C.

The coanalytic sets are just the complements of the analytic sets. On the other hand, the following can be proved:

THEOREM 6.3. *There is a coanalytic equivalence relation on I for which the Borel diagonalization theorem fails.*

There are many interesting equivalence relations (on Borel subsets of complete separable metric spaces) that are analytic. Thus Theorem 6.2 applies to them. Theorem 6.2 was used in [St] to obtain the following: Let $S(Q)$ be the Cantor space of subsets of Q, where Q is the rational numbers. Clearly every set $A \subseteq Q$ can be viewed as a linear ordering inherited from the linear ordering of Q. We say that two elements of $S(Q)$ are isomorphic if they are isomorphic as linear orderings. We say that $F : S(Q) \to S(Q)$ is isomorphically invariant if isomorphic arguments are sent to isomorphic values.

THEOREM 6.4. *Every isomorphically invariant Borel function on $S(Q)$ sends some argument to an isomorphic copy of an interval in that argument. Furthermore, this theorem is in class C.*

Theorem 6.4 can be modified in many different minor ways while remaining in class C. For instance, we can insist that the interval in the argument have endpoints in the argument, or that the interval be bounded from above and below in the argument.

We now let $\boldsymbol{G}$ be alternatively the space of all binary operations, semigroups, or groups on the natural numbers N. (This just means that the field of points is N.) These are Borel subspaces of the Baire space N^N. We also let $\boldsymbol{G}_f$ be the subspace of, respectively, finitely generated operations, semigroups, or groups. We say that a subset of an operation on N is finitely equationally defined if it is the set of all solutions of some finite set of equations in one variable with parameters allowed from N. In the case of groups, we allow the inverse operation to be used in these equations.

The following is proved in [St]:

THEOREM 6.5. *Every isomorphically invariant Borel function on $\boldsymbol{G}$ sends some group (semigroup, operation) to an isomorphic copy of a subgroup (subsemigroup, suboperation). Furthermore, this theorem is in class A for each one of the three choices for $\boldsymbol{G}$.*

Mappings $F : \boldsymbol{G}_f^\infty \to \boldsymbol{G}_f$ are also considered in [St]. The following is proved there:

THEOREM 6.6. *Every isomorphically invariant Borel function $F : \boldsymbol{G}_f^\infty \to \boldsymbol{G}_f$ sends some sequence of finitely generated groups (semigroups, operations) to a finitely generated group (semigroup, operation) which is embeddable in one of its coordinates. Furthermore, this theorem is in class B for each one of the three choices for $\boldsymbol{G}$.*

For our purposes, a graph consists of a subset of N called vertices, and a set of unordered pairs of vertices called edges. Infinite graphs are allowed, but no multiple edges. The space of graphs is naturally a Cantor space.

The detached subgraphs of a graph are taken to be the unions of connected components of the graph.

We have been able to prove the following:

THEOREM 6.7. *Every isomorphically invariant Borel function on graphs sends some graph to an isomorphic copy of a detached subgraph. Furthermore, this theorem is in class C.*

All of the examples given thus far in this section clearly illustrate the general theme stated at the end of the previous section. We conclude this section with an example that does not really fit into the general theme, but which is closely tied up with the so-called axiom of determinacy, which figures so prominently in the work on the projective hierarchy discussed at the end of the introduction.

By way of background, the following is well known to be false:

PSEUDOTHEOREM. *Every Borel set $E \subseteq I \times I$ contains or is disjoint from the graph of a Borel function on I.*

However, the following is proved in **[Fr1]** (we call a set $E \subseteq I \times I$ symmetric if $(x, y) \in E$ if and only if $(y, x) \in E$):

THEOREM 6.8. *Every symmetric Borel set $E \subseteq I \times I$ contains or is disjoint from the graph of a Borel (or even left continuous) function on I. Every symmetric Borel set $E \subseteq K \times K$ contains or is disjoint from the graph of a continuous function on K (K is the Cantor set). Furthermore, both theorems are in class C.*

7. Strong Borel diagonalization on groups and graphs. In this section we present some extensions of Theorem 6.6 which are not provable in ZFC. They are, however, theorems of one of the most intensively studied extensions of ZFC by set theorists, i.e., ZFC + "there exists a measurable cardinal." We abbreviate this system by ZFM.

This additional axiom is most simply stated as follows: There exists a countably additive measure on the class of all subsets of some set where the measure of every set is either 0 or 1, and the measure of points is zero.

So clearly these extensions of Theorem 6.6 are consistent with ZFC if ZFM is consistent. But is ZFM consistent?

Unfortunately, this question has a confusing answer. It seems to be consistent in the sense that the set theorist's use of ZFM has not led to any inconsistencies. On the other hand, the number of man hours devoted to testing ZFM is insignificant compared to that devoted to general mathematics, and set theorists have had a vested interest in ZFM being consistent for many years now.

Of course, it would be best if one could prove that ZFM is consistent if and only if ZFC is consistent, and carry out this relative consistency proof within ZFC.

Unfortunately the second incompleteness theorem creates an obstacle to this ever happening. The reason is that ZFM itself proves that ZFC is consistent. Hence if we could carry out this desired proof within ZFC (or even within ZFM) then we would have a proof within ZFM that ZFM is consistent. The second incompleteness theorem says this is impossible unless ZFM is inconsistent! Such is the legacy of Kurt Gödel.

Should we accept the consistency of ZFM on faith? Or should we regard this question as not meaningful? Or perhaps meaningful but perhaps forever beyond our grasp to decide?

These are deep questions about which there is no consensus among logicians. There is the background question which in this context is critical. Is it important whether or not ZFM is consistent?

The importance of an extension of ZFC such as ZFM is dependent on what you can do with it that you cannot do in ZFC. An ultimate illustration of the importance of the consistency of ZFM would be afforded by a dramatic result such as the following, which is by no means ruled out at this point (but of course could be ruled out at any time): I am *not* making this as a conjecture.

Possible but wildly speculative. There is a specific simple variant of the twin prime conjecture which is true if and only if ZFM is consistent. This equivalence is provable well within ZFC.

If some result anywhere near this was obtained, then clearly questions about the status of systems like ZFM would assume central importance in the history of mathematics.

It has been our view for many years that a first step towards obtaining this kind of stunning result is to first obtain such a result for a statement that at least fits into the Borel point of view. This already proved to be a difficult obstacle and there is still the expectation of much better results along these lines that fit into the Borel point of view.

Recall the definition of graph and detached subgraph that we used in the previous section. We say that a graph is embeddable in another graph if there is a one-one map from the vertices of the first into the vertices of the second such that every edge in the first is sent to an edge in the second. We say that a graph is completely embeddable if the same holds with the additional requirement that two vertices are connected by an edge in the first graph if and only if their images are connected by an edge in the second graph. Also, we say that a graph is locally finite if every vertex is joined to at most finitely many vertices.

The following propositions are discussed in **[St]**:

PROPOSITION 7.1. *Every isomorphically invariant Borel* $F : \boldsymbol{G}_f^{\infty} \to \boldsymbol{G}_f$ *sends all of the infinite subsequences of some sequence* G *to a group* (*semigroup, operation*) *which is embeddable in one of the coordinates of* G.

PROPOSITION 7.2. *Every isomorphically invariant Borel* $F : \boldsymbol{G}_f^{\infty} \to \boldsymbol{G}$ *sends all of the infinite subsequences of some sequence* G *to a group* (*semigroup, operation*) *which is embeddable in some direct limit of* G.

The following is proved in **[St]**:

THEOREM 7.3. *Propositions* 7.1 *and* 7.2 *are provable in ZFM but not in ZFC. This is true for any of the three choices* (*groups, semigroups, operations*) *for* G.

Alternatively, graphs can be used in the following way instead of groups, semigroups, and operations:

PROPOSITION 7.4. *Every isomorphically invariant Borel function on the locally finite graphs sends all of the detached subgraphs of some* G *to graphs embeddable* (*completely embeddable*) *into* G.

THEOREM 7.5. *Proposition* 7.4 *is provable in ZFM but not in ZFC. This is true for both kinds of embeddability.*

THEOREM 7.6. *Propositions* 7.1, 7.2, *and* 7.4 *imply the consistency of ZFC. Furthermore this fact can be proved well within ZFC.*

We now discuss the implications that the results cited in this section have for the constructible point of view.

Recall that the axiom of constructibility is known to decide the set-theoretic propositions that have been shown to be independent of ZFC by direct use of the forcing method such as Propositions 2.1–2.6.

However, here the propositions in question are not decided by the axiom of constructibility. In fact, the axiom of constructibility has a clear meaning in the context of weaker systems than ZFC, and so the same point can be made with regard to the results cited in §§5 and 6. More specifically:

THEOREM 7.7. *Theorems* 5.1, 5.3, 6.1, 6.2, 6.4–6.8 *and Propositions* 7.1, 7.2, 7.4 *remain unprovable in the same systems in which they were originally stated to be unprovable, even if the axiom of constructibility is added to those respective systems.*

Also recall that the constructible point of view does not assert that the axiom of constructibility is true, but only proposes that all mathematical statements be relativized to the constructible sets, i.e., that the mathematical universe be taken to be the constructible sets in the sense of a regularity condition. What happens when the assertions cited in Theorem 7.7 are so relativized?

THEOREM 7.8. *If any of Theorems* 5.1, 5.3, 6.1, 6.2, 6.4–6.8 *and Propositions* 7.1, 7.2, 7.4 *are relativized to the constructible sets, then their metamathematical status as cited remains unchanged, i.e., the resulting statements are provable in the same systems in which they were stated to be provable, and remain unprovable in the same systems in which they were stated to be unprovable.*

This important point can be taken further. There are various natural short initial segments and fragments of the constructible hierarchy of sets that have been studied. One purpose of examining such fragments is that, to varying extents, they constitute more explicit universes of sets which do not depend on the acceptance of any concept of abstract ordinal which is necessary in the case of the full constructible hierarchy of sets. Aside from the smallest of these fragments, the sets of integers present are closed under the hyperjump operation. Most of Theorem 7.8 depends only on the closure of the constructible sets under this operation:

THEOREM 7.9. *If any of Theorems* 5.1, 5.3, 6.1, 6.5, 6.6, *and* 6.7 *and Propositions* 7.1, 7.2, 7.4 *are relativized to any given universe of sets closed under hyperjump, then their metamathematical status as cited remains unchanged.*

The original propositions about Borel functions that exhibit these strong metamathematical properties appeared in **[Fr1]**. The versions discussed here are more natural.

8. The predicative point of view. The comprehension axiom scheme in ZFC allows one to construct a set by writing down $\{x \in a : A(x)\}$, where $A(x)$ is any set-theoretic property of sets x that is expressible in the language of ZFC. Of course, $A(x)$ may have side parameters. Here we discuss some philosophical aspects of this set existence principle in case the set a is N, the set of all natural numbers. Thus we are concerned with proofs of the existence of sets of natural numbers.

The issue is this. Suppose we assert the existence of $\{n \in N : A(n)\}$. Suppose also that the property A refers to all sets of natural numbers in its expression in the language of set theory. Have we really constructed a set of natural numbers? Why do we accept the existence of such a set of natural numbers?

If we take the position that this set of natural numbers is constructed by writing down $\{n \in N : A(n)\}$ in the sense that it did not exist before anybody wrote this down (unless it coincidentally happened to have the same members as some such set that was written down earlier), then there is the real question of the meaning of, say, $A(1)$. Do the references in A to all sets of natural numbers refer to the set allegedly under construction? How about sets that have not been so constructed, but will be so constructed in the future? If it is not clear what sets are being referred to in A then in what sense is this a construction? In what sense is A meaningful?

The most natural position to take on such matters, assuming one wishes to accept this set existence principle, is that all sets of natural numbers exist independently of how humans construct them, view them, or understand them. They are just there, independently of our mental processes, and we use our mental processes to observe them, study them, and use them. Through our mental processes we have observed that $\{n \in N : A(n)\}$ exists, and was there before any human thought about it or thought about A.

A problem with this so-called Platonistic approach is that it is unclear how far it can be reasonably taken. If a purely external objective reality of all sets of natural numbers exists for us to observe and study, then why not such a reality of all sets of sets of natural numbers? But then we seem stuck with accepting the point of view that the continuum hypothesis is a matter of objective reality that simply awaits additional observation and study. As discussed earlier, the continuum hypothesis is not only independent of ZFC, but at this point the discovery of any new fundamental principles about the cumulative hierarchy of sets that would settle it seems very remote. It seems hard to merely accept that what is needed is simply some hard work or clever idea, as has proved to be the case for so many hard open mathematical problems that eventually get solved. Most mathematicians are quite uncomfortable with the concept of objective external reality when pushed as far as to include sentences such as the continuum hypothesis. They are even more uncomfortable in the context of such sentences as "there are measurable cardinals."

As discussed earlier, many specialists in set theory wish to take this Platonistic approach to the extreme; that the entire cumulative hierarchy of sets has an objective external reality awaiting our observation and study, and that any well-formed assertion about this hierarchy is objectively true or false.

But for those who do not accept this extreme view, the question of where the objective external reality ends and human intervention begins is a real issue.

It seems to us that, ultimately, there is no such good dividing line, and that a certain kind of relativism is emerging: That there is no such thing as an objective external reality anywhere outside the most extreme basic context (such as the study of the integers from 1 to 100). Instead, there are degrees or levels of external objective reality, running the spectrum from $\{1, 2, \ldots, 100\}$ to the entire cumulative hierarchy of sets (or even maybe beyond). On this view, the really interesting thing to do is to analyze the relationships between these contexts and to obtain definite mathematical results which shed light on these degrees or levels. The incompleteness results discussed in this manuscript do just that. Other types of results, such as consistency proofs, which are not discussed here, also contribute to this general aim.

Let us return to our discussion of $\{n \in N : A(n)\}$. The predicative point of view accepts an objective external reality of the totality of natural numbers, but rejects any objective external reality of the totality of sets of natural numbers. On this view, sets of natural numbers do not exist independently of their construction. All constructions of sets of natural numbers must in some sense be grounded in the natural numbers themselves. On the predicative point of view, all such constructions are admissible.

The most typical case of a construction of a set of natural numbers from the predicative point of view is that of $\{n \in N : A(n)\}$, where A is arithmetical. In other words, when all quantifiers in A range over the natural numbers. If side parameters exist for sets of natural numbers, then the construction is relative to those side parameters. If the side parameters have been constructed, i.e., given a predicative meaning, then the expression is then a predicatively meaningful construction. This amounts to what is called the arithmetical comprehension axiom scheme.

Life would be very simple if one could merely identify predicativity with the arithmetical comprehension axiom scheme. However, consider the following situation. One may have an explicit assignment to each natural number n of an arithmetical formula $A_n(k)$, say, with no side parameters. Then we may wish to construct, say, $\{n \in N : (\exists k)(A_n(k))\}$. This cannot be done within arithmetical comprehension, but seems to be arguably within the scope of predicativity.

There has been considerable effort devoted to codifying the predicative point of view into appropriate formal systems, with some theorems suggesting, in some way, that such formal systems completely capture predicativity. We do not believe that the point of view naturally lends itself to such characterization, although there clearly are constructions such as the ones cited above which obviously fall within the predicative, as well as constructions which obviously do not fall within the predicative. It is possible that one may be able to amplify on the usual description of the predicative point of view, maintaining its fundamental philosophical flavor, so that the view would naturally lend itself to such characterization. But even this has not been accomplished in any convincing way.

We think that, under these circumstances, the really fruitful investigation is to see what consequences the predicative point of view has on actual mathematics.

Fortunately, in nearly all known interesting mathematical situations, a given proof of a theorem is either obviously predicative or obviously impredicative. Usually, a given theorem either can be given a proof which is obviously predicative, or a recursion-theoretic result is known which implies that it obviously has no predicative proof. The typical case of the former is that the arithmetical comprehension scheme is enough, and the typical case of the latter is that the theorem is shown to be false in the universe of hyperarithmetical sets of natural numbers.

Typical cases of theorems which are known to not be predicatively provable by the above method are (1) the order comparability of well-orderings of the natural numbers, (2) the presence of perfect sets within uncountable closed sets of real numbers, and (3) the least upper bound principle for (even arithmetically defined) sets of real numbers.

However, notice that all three examples assert the existence of some set of natural numbers (perhaps disguised as a real number, a function on the natural numbers, or a perfect set as the complement of the union of a sequence of rational open intervals). It is perhaps not too surprising that there would be such basic examples, since the predicative point of view severely restricts the set existence axioms allowed.

A crucial issue about the predicative point of view is whether there are such basic mathematical theorems that do not assert the existence of infinite sets of natural numbers, even under disguise, yet can only be proved impredicatively. It is to be expected that mathematicians advocating the predicative point of view are likely to believe that there are no such significant examples.

However, in the next section we present such examples, which have only been discovered in the 1980's.

Strong advocates of the predicative point of view include such great mathematicians as Hermann Weyl and Henri Poincaré. It would have been interesting to see how their advocacy of predicativity would have been affected by the discovery of these examples.

For more discussion on predicativity, see work of S. Feferman; e.g. **[Fe1]** and **[Fe2]**. In the next section we use the formal system ATR_0 as a working model for the upper limit of predicativity. This is generally accepted in light of its connection with hyperarithmetic sets and the proof theoretic ordinal Γ_0.

9. Finite trees and finite graphs. In this section we present the examples mentioned at the end of §8 of theorems not involving the existence of infinite sets of natural numbers, yet which cannot be predicatively proved. Some of the examples have the stronger property that they do not even mention infinite sets of natural numbers, even in disguise.

The first such example was the celebrated theorem of J. B. Kruskal in 1960 concerning the embeddability of finite trees in infinite sequences of finite trees. It was not until 1981 that anyone observed that it cannot be predicatively proved, despite the fact that the original proof was blatantly predicative and Kruskal had called attention to the peculiar nature of the proof.

The proof was later greatly simplified and streamlined by Nash-Williams. This new proof spawned a whole new interesting field of combinatorics called *wqo theory*. The Nash-Williams proof is sufficiently simple and the crucial impredicative step is sufficiently easy to identify, that we give a sketch of it here.

A tree consists of a nonempty set V called vertices, together with a partial ordering $\leq$ on V such that (a) there is a (unique) least element called the root, and (b) the set of predecessors of every vertex under $\leq$ is linearly ordered under $\leq$. We have the obvious sup and inf operations on sets of vertices provided that the tree is finite (i.e., has a finite number of vertices).

The crucial notion of embedding h from one finite tree T_1 into another T_2 is this: h is a one-one mapping from the vertices of the first into the vertices of the second, and h is inf preserving in the sense that $h(a \inf b) = h(a) \inf h(b)$. These conditions imply that h is order preserving in the strong sense that $a \leq_1 b$ if and only if $h(a) \leq_2 h(b)$. We write $T_1 \leq T_2$ if and only if there exists such an embedding from T_1 into T_2.

The following is proved in **[Kr]**:

THEOREM 9.1. *In any infinite sequence $T_1, T_2, \ldots$ of finite trees, there are $i < j$ such that $T_i \leq T_j$. In any infinite set of finite trees, one of the elements is embeddable into another.*

The following is proved in **[Si1]**:

THEOREM 9.2. *Theorem* 9.1 *cannot be proved in the formal system ATR_0, and hence cannot be given a predicative proof. This holds for either of the two forms given.*

Before we sketch the Nash-Williams proof of Theorem 9.1, we give some other variants which may be a little more natural from a graph theorist's viewpoint.

We can alternatively define a tree to be a connected graph with no cycles. Note there is no root in this treatment. The relevant concept of embedding is that of a one-one mapping h from vertices in the first tree into vertices in the second tree such that if ab and ac are edges in the first tree, $a \neq b \neq c$, then the unique simple path from $h(a)$ to $h(b)$ in the second tree does not cross the unique simple path from $h(a)$ to $h(c)$ in the second tree (except of course at $h(a)$). Or, alternatively, we may view graphs as topological spaces (1-dimensional complexes), and we merely require that the embeddings be homeomorphic mappings (continuous and one-one). The latter does not require that vertices go to vertices. If we did require that, then it would be identical to the graph-theoretic definition we have just given.

We have looked into these alternative definitions and found that the differences are inessential from the metamathematical point of view:

THEOREM 9.3. *Theorems* 9.1 *and* 9.2 *hold for infinite sequences of graph-theoretic finite trees, under any of the notions of embedding discussed above.*

We now sketch the proof of Theorem 9.1 given in **[Na]**. The method is called the minimal bad sequence argument.

A quasiordering is merely a nonempty set under a transitive and reflexive relation (i.e., if $a \leq b$ and $b \leq c$ then $a \leq c$, and also $a \leq a$). A well quasi ordering is a quasi ordering with the crucial property that for all infinite sequences $a_1, a_2, \ldots$, there are $i < j$ such that $a_i \leq a_j$. It is interesting and well known that this is equivalent to the requirement that within any infinite set A there are $a, b \in A$ such that $a \leq b$.

Note that Theorem 9.1 can be restated as asserting that the finite trees under embeddability constitute a well quasi ordering.

Let $(Q, \leq)$ be a quasi ordering. Then we form the new quasi ordering $FIN(Q)$ consisting of the finite subsets of Q under the following quasi order: $A \leq^* B$ if and only if there is a one-one mapping $h : A \to B$ such that for each $a \in A$, $a \leq h(a)$.

Let us assume for the moment the following theorem from **[Hi]** known as Higman's lemma:

THEOREM 9.4. *If Q is a well quasi ordering then so is $FIN(Q)$.*

We continue the sketch of the proof of Theorem 9.1.

By way of contradiction we let $T_1, T_2, \ldots$ be a counterexample to (the first form of) Theorem 9.1. Such a counterexample is called an infinite bad sequence. We want to prove that there is no infinite bad sequence.

We first need to construct what is called a minimal bad sequence. Let S_1 be any finite tree of minimal possible size (as measured by the number of vertices) such that S_1 starts some infinite bad sequence. Let S_2 be any finite tree of minimal possible size such that S_1, S_2 starts some infinite bad sequence. Continue in this way to obtain the minimal bad sequence S_1, $S_2, \ldots$.

There are two not very explicit aspects to this construction. Firstly, the axiom of choice is used since we did not specify which of the several possible finite trees is to be chosen at each stage. But this is truly a minor point. We can enumerate all the finite trees up to isomorphism in some reasonable order before the construction begins in order to avoid this problem (using canonical representations from the equivalence classes in a standard and routine way). Secondly, there is a blatant impredicativity in the construction since at each stage we refer to unrestricted infinite sets of natural numbers (finite trees), including the one being constructed. This is just the kind of construction that is criticized by predicativists. This second point is the crux of the matter. Theorem 9.2 explains why this aspect is unavoidable.

Now that we have our minimal bad sequence, we let Q be the set of all upwardly closed subtrees of the S's whose roots lie right above the roots of the S's. In other words, each S_i is the joining together of several disjoint subtrees by the root of S_i; such subtrees are called the immediate subtrees. Q consists of all such immediate subtrees. We make Q into a quasi ordering by our notion of embeddability.

It is not difficult to see that because of the minimal badness of the S's, Q must be a well quasi ordering. For, if Q had an infinite bad sequence, then that sequence could be used to obtain a new infinite bad sequence which agrees with S for a while, and then uses subtrees of the S's; this would violate the minimality of the S's at the spot where the subtrees of the S's start, and where the copying of the S's themselves end.

Now by Higman's lemma, the finite sequences from Q are also well quasi ordered. From this we obtain $i < j$ such that the set of all immediate subtrees of S_i is $\leq^*$ the set of all immediate subtrees of S_j. But this immediately implies that $S_i \leq S_j$, which is the contradiction we have been seeking.

Higman's lemma itself (Theorem 9.4) can also be proved by a (simpler) minimal bad sequence argument. However, in contrast to Theorem 9.1, it can be given an alternative proof within the arithmetical comprehension scheme.

Note that although Theorem 9.1 does not state the existence of infinite mathematical objects, it does mention them (universally). We now discuss a finite reformulation of Theorem 9.1 which does not involve infinite mathematical objects at all.

The idea is simple and natural. We first weaken the statement by placing bounds on the number of vertices, $|T|$, of the trees T. Thus we can consider the following:

(*) For all k and finite trees $T_1, T_2, \ldots$, with each $|T_i| \leq k + i$, there are $i < j$ such that $T_i \leq T_j$.

Note that the collection of infinite sequences of finite trees satisfying this growth condition (for a fixed k) is a compact space. Hence as is standard in such situations, this is true for infinite sequences of such trees if and only if it is true for sufficiently long finite sequences of such trees. Thus we are naturally led to the following:

THEOREM 9.5. *For $r >> k$ and finite trees $T_1, \ldots, T_r$ obeying $|T_i| \leq k + i$, there are $i < j$ such that $T_i \leq T_j$.*

The following is proved in **[Si1]** and **[Smith]**:

THEOREM 9.6. *Theorem 9.5 cannot be proved in ATR_0 and hence cannot be given a predicative proof. This is true even if graph-theoretic trees are used.*

An obvious question is: how large must r be as a function of k in Theorem 9.5? It is clear that this is a recursive function, since we can just look for a big enough r and check to see that we have it. However, the following is proved in **[Si1]**, which obviously strengthens Theorem 9.6:

THEOREM 9.7. *No provably recursive function of ATR_0 is sufficient to bound the required size of r as a function of k in Theorem 9.5.*

We mention an alternative finite form of Theorem 9.5 that may be viewed as being even more natural (see **[Smith]**):

THEOREM 9.8. *If $r >> k$ then every sequence $T_1, \ldots, T_r$ of finite trees obeying $|T_i| \leq i$ contains an increasing subsequence of length k. This is not provable in ATR_0 and hence does not have a predicative proof. Furthermore, no provably recursive function of ATR_0 is sufficient to bound the required size of r as a function of k. Again, graph-theoretic trees can be used.*

In **[Smith]** yet another finite form is considered which involves only the growth condition $|T_i| \leq i$ as in Theorem 9.8. Only this time the k represents the number of labels. Kruskal also considered finite trees with labels from a finite set. The embeddability condition is strengthened to demand that the embedding be label preserving. The following is implicit in **[Smith]**:

THEOREM 9.9. *If $r >> k$ and $T_1, \ldots, T_r$ are finite trees with k labels obeying $|T_i| \leq i$, then there are $i < j$ such that $T_i \leq T_j$ (label preserving). This theorem has the same properties cited in Theorem 9.8.*

The independence results stated in Theorems 9.2–9.3 and 9.6–9.9 are understated in that they hold for systems somewhat stronger than ATR_0. The optimal system to use is the somewhat stronger system $\Pi^1_2 - BI_0$. However, certainly ATR_0 is a more natural system, representing, in a sense, the border of the usual formalisms for predicativity, and being equivalent to basic mathematical facts such as the comparability of well-orderings, as in reverse mathematics (see **[Si2]**).

It is clear that all of the theorems about trees discussed from Theorem 9.5 and beyond in this section are of the form $\forall\exists$ over the natural numbers. Actually, technically speaking, they are presented in form $\forall\exists\forall$, since they assert that for all k there is a t such that for all $r > t$, something holds. However, it is obvious for the statements under question that if any r works then trivially any larger r works. Thus the statements are actually of the form: for all k there is an r.

If we specialize the outermost quantifier k of an $\forall\exists$ statement, then we get an $\exists$ statement. Such a statement is always provable in any reasonable system if and only if it is true. But the interesting question, under these circumstances, is: how large is the least possible r? And, how large is the least possible proof that there is an r?

From the point of view of the incompleteness phenomena, the second question is what really is interesting. If one can show that the least possible proof that there is an r is ridiculously large, then one has exhibited an incompleteness phenomena that is different from what we have discussed up to this point.

Let $2^{[n]}$ be a stack of n two's iteratively exponentiated; e.g., $2^{[4]} = 2^{16}$.

The following is proved in **[Smith]**:

THEOREM 9.10. *In Theorem 9.9, if k is set to 6 then the resulting $\exists$ statement cannot be proved within ATR_0 without using at least $2^{[1000]}$ symbols. Hence it cannot be proved predicatively without using a humanly unreasonable number of symbols: in this sense, it is unprovable predicatively.*

Similar results can be given for all of the $\forall\exists$ statements considered in this section.

Kruskal's theorem with finitely many labels can be strengthened so as to obtain independence results such as the above from yet stronger systems. We did discover such natural strengthenings by adding an additional condition on the embeddings. The added condition is called the gap condition.

The idea is as follows. Suppose h is an inf preserving embedding from S into T, and assume that the trees are labeled from the finite set $\{1, \ldots, n\}$ and are label preserving. If b is an immediate sucessor vertex to a in S, then $h(b)$ may not be an immediate sucessor vertex to $h(a)$ in T. Of course, $h(b)$ does lie above $h(a)$ in T. But there might be a gap of vertices strictly in between. The additional condition asserts that all of the labels of the vertices in this gap in T must be numerically at least as large as the label of b (or $h(b)$). We write $\leq_n$ for this quasi ordering.

The following is proved in **[Si1]**:

THEOREM 9.11. *Each $\leq_n$ is a well quasi ordering. This theorem can be proved in $\Pi_1^1 - CA$ but not in $\Pi_1^1 - CA_0$, or in what is called finitely iterated inductive definitions.*

The proof that each $\leq_n$ is a well quasi ordering involves an iteration of the minimal bad sequence construction n times. One way of looking at the proof is as follows: assume that the result is false, and then construct an appropriate minimal bad sequence. From this sequence, construct a new quasi ordering and another minimal bad sequence through that. Iterate this procedure n times until finally one obtains a bad sequence through some quasi ordering which ostensibly is a well quasi ordering, obtaining the desired contradiction. Such a proof would involve (roughly) n iterated inductive definitions. The union of n iterated inductive definitions corresponds to $\Pi_1^1 - CA_0$.

We conjectured that $\leq_\omega$ is a well quasi ordering, where the domain is the finite tree labeled from ω and the embedding is required to be inf preserving, nowhere label decreasing, and the gap condition holds (in the gap, the labels are all numerically at least that of $h(b)$). In fact, we made the more general conjecture that this was true for each $\leq_\alpha$ for any ordinal α. This conjecture has been recently proved in **[Kriz]**. It is interesting to observe that the proof is given in $\Pi_2^1 - CA$, even for $\alpha = \omega$. It is known that for each α this must take at least about α iterated inductive definitions to prove that $\leq_\alpha$ is well quasi ordered. So the lower and upper bounds are wildly far apart at this time.

The fact that the $\leq_n$ is a well quasi ordering was subsequently used several places in the very lengthy proof that the finite graphs are well quasi ordered under the relation of minor inclusion, written $\leq_m$ (see **[RS]**). We say that G is *minor included* in H if G can be obtained from H by successive applications of the following operations: (1) removing an edge, (2) removing a vertex (and all edges coming out of it), and (3) contracting an edge to a vertex.

THEOREM 9.12. *The relation $\leq_m$ on the finite graphs is a well quasi ordering, i.e., for all $G_1, G_2, \ldots$, there are $i < j$ such that $G_i \leq_m G_j$.*

The following is proved in **[FRS]** by showing that "$\leq_m$ is a well quasi ordering" implies "each $\leq_n$ is a well quasi ordering."

THEOREM 9.13. *Theorem 9.12 cannot be proved within* $\Pi^1_1 - CA_0$. *It can be proved in* $\Pi^1_1 - CA + BI$. *In particular, there is no predicative proof of Theorem 9.12.*

We can give a number of finite forms of this graph minor theorem which also cannot be proved in such systems just as we did for Kruskal's theorem. Let $|G|$ be the sum of the number of vertices and edges of G. We give two such forms as discussed in **[FRS]**:

THEOREM 9.14. *If* $r >> k$ *and* $G_1, \dots, G_r$ *are finite graphs with each* $|G_i| \leq k + i$, *then there are* $i < j$ *such that* $G_i \leq_m G_j$. *If* $r >> k$ *then every sequence* $G_1, \dots, G_r$ *of finite graphs obeying* $|G_i| \leq i$ *contains an increasing subsequence of length* k *under minor inclusion. Neither of these theorems are provable in* $\Pi^1_1 - CA_0$ *and hence do not have predicative proofs. Furthermore, no provably recursive function of* $\Pi^1_1 - CA_0$ *is sufficient to bound the required size of* r *as a function of* k.

10. Finite Ramsey theory. Ramsey theory has become an established branch of combinatorics of extensive scope (see **[GRS]**). In this section we will be discussing the original Ramsey theorems that form the basis of the subject. In **[PH]** a modified form of the original finite Ramsey theorem was given and shown to be unprovable within Peano arithmetic (or finite set theory). It is provable from the original infinitary Ramsey theorem, which in turn is provable by, for instance, augmenting Peano arithmetic with functions defined by arithmetical recursion. Putting it more simply, the modified finite Ramsey theorem is not provable in finite set theory or Peano arithmetic, but can be proved just beyond them.

The modified finite Ramsey theorems were the first examples of interesting mathematical theorems about finite objects which were shown to have substantial independence properties. They predate the earliest results of this kind from §9 by four years (1977 versus 1981). There was considerable expectation that the examples would blossom into further examples which would exhibit much stronger independence properties such as having no predicative proof. However, for this purpose the direct approach via Ramsey theory turns out to be an apparant dead end.

Here is the original infinitary Ramsey theorem from **[Ramsey]**:

THEOREM 10.1. *If all of the* k*-element subsets of a countably infinite set are colored from a finite set, then there is an infinite subset all of whose* k*-element subsets are assigned the same color.*

The proof of this theorem is by induction on k. The case $k = 1$ is obvious. The case $k + 1$ is reduced to case k as follows. Observe that if we fix any element x then we obtain an induced coloring of the k-element subsets of the set without x. Thus we fix x_1. We choose A_1 to be any infinite set excluding x_1 such that all k-element subsets of A_1 are assigned the same color induced by x_1. Then choose x_2 to be any element of A_1 and A_2 to be any infinite subset of A_1 excluding x_2 such that all n-element subsets of A_2 are assigned the same color induced by x_1. Continue in this way indefinitely. This results in an infinite sequence of x's. It is clear that the color assigned to any subset of the x's of size $k+1$ depends only on the identity of the earliest x in the subset. Since there are only finitely many colors, there is an infinite set E of x's such that all subsets of the x's whose first x is from E must be assigned the same color. In particular, clearly every subset of E of size $k+1$ must be assigned the same color.

The original finite form of Theorem 10.1 is as follows [**Ramsey**]:

THEOREM 10.2. *If $r >> k, n, m$ and all k-element subsets of an r-element set are colored from an n-element set, then there is an m-element subset all of whose k-element subsets are assigned the same color.*

The easiest proof of this theorem is to derive it from Theorem 10.1. Fix k, n, and m, and assume Theorem 10.2 is false. Then for each r there is a counterexample coloring of the r-element set $\{1, 2, \ldots, r\}$. One can now construct counterexample colorings C_r for each r, such that each coloring is extended by the next coloring. The union of the C's form a coloring of the n-element subsets of all of N. Applying Theorem 10.1, we obtain an infinite set all of whose n-element subsets are assigned the same color by the C's. But note that the first m elements of this infinite set satisfies the condition in Theorem 10.2 for the coloring C_t, where t is the last of these first m elements. Hence C_t was not a counterexample coloring after all. This is the desired contradiction.

Note that this proof gives no information about how large r must be relative to k, n, and m. It is clear that the proof is highly inexplicit.

However, in this case there is an explicit proof. In fact, Ramsey's original proof in [**Ramsey**] was explicit and gave iterated exponential bounds for how large r must be relative to k, n, and m.

In [**PH**] an additional clause is added to the conclusion of Theorem 10.2. We say that a set $A \subseteq N$ is *relatively large* if the number of elements in A is numerically at least as large as the minimum element of A.

The following is proved and studied in [**PH**]:

THEOREM 10.3. *If $r >> k, n, m$ and all k-element subsets of $\{1, \ldots, r\}$ are colored from an n-element set, then there is a $\geq m$-element subset of $\{1, \ldots, r\}$ all of whose k-element subsets are assigned the same color, and which is relatively large.*

Note that this is just as easy a corollary of Theorem 10.1 as is Theorem 10.2, since, trivially, every infinite subset of N contains arbitrarily large finite relatively large subsets.

It is shown in [**PH**] that Theorem 10.3 is not provable in Peano arithmetic (PA), but can be proved just beyond it in, e.g., arithmetic comprehension

(ACA). Furthermore, no provably recursive function of PA is sufficient to bound the required size of r as a function of k, n, m.

The concept of relatively large uses integers simultaneously in the role of "elements" and of "number of elements." This dual role is sufficiently unusual in mathematics as to prompt a search for alternatives to Theorem 10.3 that do not use concepts such as relatively large. A particularly attractive alternative is through what we call a *function value theorem*. We first give the infinitary form.

THEOREM 10.4. *Let F be a function from all $\leq k$-element subsets of N into N and $m \in N$. Then there is an infinite set $A \subseteq \{m, m+1, \dots\}$ such that F takes on at most $k+1$ values $\leq \min(A)$.*

And here is the straightforward finite form.

THEOREM 10.5. *Let $r >> k, n, m$ and F be a function from all $\leq k$-element subsets of $\{1, \dots, r\}$ into $\{1, \dots, r\}$. Then there is a $\geq n$-element $A \subseteq \{m, \dots, r\}$ such that F takes on at most $k+1$ values $\leq \min(A)$.*

Theorem 10.5 has the same metamathematical properties as Theorem 10.3 (in fact, the two can be shown to be equivalent within a weak fragment of PA).

We close the discussion by presenting some *function congruence theorems*.

THEOREM 10.6. *For any $F\colon N^k \to N^k$ there are $x_1 < x_2 < \cdots < x_{k+1}$ with $F(x_1, x_2, \dots, x_k) \equiv F(x_2, x_3, \dots, x_{k+1}) \bmod 2$. Furthermore, a bound can be placed on x_{k+1} which depends on k but not on F.*

THEOREM 10.7. *For any $F\colon N^k \to N^k$ there are $1 < x_1 < x_2 < \cdots < x_{k+1}$ with $F(x_1, x_2, \dots, x_k) \equiv F(x_2, x_3, \dots, x_{k+1}) \bmod x_1$. Furthermore, a bound can be placed on x_{k+1} which depends on k but not on F.*

Theorem 10.6 has a bound involving approximately k iterated exponentials, and no fewer.

Theorem 10.7 cannot be bounded with a provably recursive function of PA.

Appendix. Progress towards the construction or discovery of basic mathematical problems about finite objects, with a clear and intuitive meaning, conveying interesting mathematical information, that is readily graspable, and which is independent of ZFC, has been incremental. Here we indicate the current state of the art.

The most convincing independence results in this vein are currently stated in terms of countably infinite functions or sets. Nevertheless, one proves, well within ZFC, that the independent sentences are equivalent to sentences involving only the ring of integers. Unfortunately, we do not know how to directly put the sentences into such finite terms without causing unacceptable complications.

Let $\otimes : A \times A \to Q$, where Q is the rationals. We say that $R \subseteq Q \times Q$ is $\otimes$-Boolean if R can be defined by a quantifier free formula in $(Q, <, \otimes)$; i.e., if R can be defined in terms of conjunction, disjunction, negation, and inequalities between expressions built up from $\otimes$, variables, and constants from Q.

PROPOSITION 1. *There is a $\otimes : Q \times Q \to Q$ such that for all $\otimes$-Boolean R $Q \times Q$ and $a \in Q$, there is a $b \otimes b < a \otimes a$ such that for all $x < b \otimes b$ and $y < a \otimes a$, if $R(x, y)$ then $R(x, (b \otimes b) \otimes x)$.*

The following is proved in **[Fr3]**.

THEOREM 2. *Proposition 1 can be proved in ZFC with the use of Mahlo cardinals of every finite order, but cannot be proved in ZFC. In fact, Proposition 1 is provably equivalent to the consistency of ZFC + $\{$there is a Mahlo cardinal of order $\bar{n}\}_n$, (within RCA_0).*

Let $F: N^k \to N$ and A be any set. We write $F_<[A]$ for $\{x : F(y_1, \dots, y_k) = x$ for some $y_1, \dots, y_k < x$ chosen from $A\}$.

By way of background, note that the following is easily provable and compactly expresses the fundamental principle of definition by induction on the natural numbers.

THEOREM 3. *For all $F: N^k \to N$ there is an $A \subseteq N$ with $N = A \Delta F_<[A]$. A is necessarily infinite.*

However, the following proposition is provably false (within RCA_0).

PROPOSITION 4. *For all $n >> k$ and $F: N^k \to N$, there is an infinite set $1, n \in A \subseteq N$ with $N = A \Delta F_<[A]$.*

Now consider the following weakening of Proposition 4.

PROPOSITION 5. *For all $n >> k$ and $F: N^k \to N$, there are infinite sets $1, n \in A_1 \subseteq A_2 \subseteq \cdots \subseteq A_k \subseteq N$ with $A_i + A_i \subseteq A_{i+1} \Delta F_<[A_{i+1}]$, $i < k$.*

It can be shown that Proposition 5 can be proved in ZFC with Mahlo cardinals of every finite order, but not in ZFC. In fact, Proposition 5 can be shown to be equivalent to the 1-consistency of ZFC + $\{$there is a Mahlo cardinal of order $n\}_n$, (within *ACA*). Proposition 5 can be proved for each fixed k (with *ACA*). The rate of growth associated with $n >> k$ is bounded by a recursive function but not any provably recursive function of ZFC (even with Mahlo cardinals of any given finite order).

It can also be shown (within *ACA*) that Propositions 1 and 5 are true if and only if they are true in the arithmetic sets.

REFERENCES

[Ajtai] M. Ajtai, *On the boundedness of definable linear operators*, Period. Math. Hungar. **5** (1974), 343–352.

[Co] Paul J. Cohen, *The independence of the continuum hypothesis*, Proc. Nat. Acad. Sci. U.S.A. **50** (1963), 1143–1148; **51** (1964), 105–110.

[Dales] H. G. Dales, *A discontinuous homomorphism from* $C(X)$, Amer. J. Math. **101** (1986), 647–734.

[DMR] M. Davis, Y. Matijasevic, and J. Robinson, *Hilbert's tenth problem. Diophantine equations: Positive aspects of a negative solution*, Proc. Sympos. on the Hilbert Problems (DeKalb, IL, May 1974), vol. 28, Amer. Math. Soc., Providence, RI, 1976, pp. 323–378.

[DW] H. G. Dales and W. H. Woodin, *An introduction to independence for analysts*, London Math. Soc. Lecture Notes, no. 115, Cambridge Univ. Press, London and New York, 1986.

[Fe1] S. Feferman, *Systems of predicative analysis*, J. Symbolic Logic **29** (1964), 1–30.

[Fe2] ——, *A more perspicuous formal system for predicativity*, Konstruktionen versus Positionen I, Walter de Gruyter, Berlin, 1979, pp. 68–93.

[Fr1] H. Friedman, *On the necessary use of abstract set theory*, Adv. in Math. **41** (1981), 209–280.

[Fr2] ——, *Unary Borel functions and second order arithmetic*, Adv. in Math. **50** (1983), 155–159.

[Fr3] ——, *Necessary uses of abstract set theory in finite mathematics*, Adv. in Math. **60** (1986), 92–122.

[FRS] H. Friedman, N. Robertson, and P. Seymour, *The metamathematics of the graph minor theorem*, Logic and Combinatorics (S. G. Simpson, ed.) Contemp. Math., vol. 65, Amer. Math. Soc., Providence, RI, 1987, pp. 229–261.

[Fu] L. Fuchs, *Infinite Abelian groups*, vol. II, Academic Press, 1973, pp. 178–181.

[Go1] Kurt Gödel, *The consistency of the axiom of choice and of the generalized continuum hypothesis*, Ann. Math. Stud., no. 3, Princeton Univ. Press, Princeton, NJ, 1940.

[Go2] ——, *What is Cantor's continnuum problem?*, Philosophy of Mathematics (P. Benaceraf and H. Putnam, eds.), Cambridge Univ. Press, London and New York, 1983.

[Go3] ——, *Collected works*, vols. I, II, Oxford Univ. Press, New York, 1986, 1990.

[GRS] R. Graham, B. Rothschild, and J. Spencer, *Ramsey theory*, Wiley, New York, 1980.

[HML] *Handbook of mathematical logic*, Studies in Logic and the Foundations of Mathematics (Jon Barwise, ed.), vol. 90, North-Holland, Amsterdam 1977.

[HMS] L. Harrington, D. Marker, and S. Shelah, *Borel orderings*, Trans. Amer. Math. Soc. **310** (1988), 293–302.

[HMSS] *Harvey Friedman's research on the foundations of mathematics*, Studies in Logic and the Foundations of Mathematics (L. A. Harrington, M. D. Morley, A. Scedrov, and S. G. Simpson, eds.), vol. 117, Elsevier, Amsterdam, 1985.

[Hi] G. Higman, *Ordering by divisibility in abstract aglebras*, Proc. London Math. Soc. **2** (1952), 326–336.

[Je1] T. Jech, *Nonprovability of Souslin's hypothesis*, Comment. Math. Univ. Carolin. **8** (1967), 291–305.

[Je2] ——, *Set theory*, Academic Press, 1978.

[Je3] ——, *Multiple forcing*, Cambridge Tracts in Math., vol. 88, Cambridge Univ. Press, London and New York, 1986.

[Jensen] R. Jensen, *Souslin's hypothesis is incompatible with* $V = L$, Notices Amer. Math. Soc. **15** (1968), 935.

[Kriz] I. Kriz, *Well-quasi ordering finite trees with gap-condition. Solution of Harvey Friedman's conjecture*, Ann. of Math. (2) **130** (1989), 215–226.

[Kr] J. B. Kruskal, *Well-quasi-ordering, the tree theorem, and Vázsonyi's conjecture*, Trans. Amer. Math. Soc. **95** (1960), 210–225.

[Levy] A. Levy, *Basic set theory*, Springer-Verlag, 1979.

[Luzin] N. Luzin, *Leçons sur les ensembles analytiques*, Gautheir-Villars, Paris, 1930.

[Mar] D. A. Martin, *Descriptive set theory: projective sets*, in **[HML]**, pp. 783–815.

[Mat] Yu. Matijacevic, *Enumerable sets are Diophantine*, Soviet Math. Dokl. **11** 354–357.

[MS] D. A. Martin and J. R. Steel, *A proof of projective determinacy*, J. Amer. Math. Soc. **2** (1989), 71–125.

[Na] C. St. J. A. Nash-Williams, *On well-quasi-ordering finite trees*, Proc. Cambridge Philos. Soc. **59** (1963), 833–835.

[PH] J. Paris and L. Harrington, *A mathematical incompleteness in Peano arithmetic*, in **[HML]**, pp. 1133–1142.

[Ramsey] F. P. Ramsey, *On a problem of formal logic*, Proc. London Math. Soc. **30** (1930), 264–286.

[RS] N. Robertson and P. Seymour, ongoing series of papers in J. Combin. Theory Ser. B, 1983–.

[Sh] S. Shelah, *A compactness theorem for singular cardinals, free algebras, Whitehead problem and transversals*, Israel J. Math. **21** (1975), 319–349.

[Ship] J. Shipman, *Cardinal conditions for strong Fubini theorems*, Trans. Amer. Math. Soc. (to appear).

[Si1] S. G. Simpson, *Nonprovability of certain combinatorial properties of finite trees*, in **[HMSS]**, pp. 87–117.

[Si2] ——, *Reverse mathematics*, Proc. Sympos. Pure Math., vol. 42, Amer. Math. Soc., Providence, RI, 1985, pp. 461–471.

[Smith] R. L. Smith, *The consistency strengths of some finite forms of the Higman and Kruskal theorems*, in **[HMSS]**, pp. 119–135.

[Smo] C. Smorynski, *The incompleteness theorems*, in **[HML]**, pp. 821–865.

[Sol] R. Solovay, *A model of set theory in which every set of reals is Lebesgue measurable*, Ann. of Math. (2) (1970), 1–56.

[Sp] E. Specker, *Additive gruppen von Folgen ganzen Zahlen*, Portugal. Math. **9** (1950), 131–140.

[ST] R. Solovay and S. Tennenbaum, *Iterated Cohen extensions and Souslin's problem*, Ann. of Math. (2) **94** (1971), 201–245.

[St] L. J. Stanley, *Borel diagonalization and abstract set theory: recent results of Harvey Friedman*, in **[HMSS]**, pp. 11–86.

[Tarski] A. Tarski, *A decision method for elementary algebra and geometry*, 2nd rev. ed., Berkeley and Los Angeles, 1951.

Department of Mathematics, Ohio State University, Columbus, Ohio 43210