Research Statement

Peter Koellner

My research concerns the search for and justification of new axioms in mathematics. The need for new axioms arises from the independence results. Let me explain. In reasoning about a given domain of mathematics (or, in fact, any domain) the question of justification is successively pushed back further and further until ultimately one reaches principles that do not admit more fundamental justification. The statements at this terminal stage are elected as *axioms* and the subject is then organized in terms of derivability from the base of axioms. In the case of arithmetic, this led to the axiom system PA (Peano arithmetic) and, in the case of set theory, it led to the axiom system ZFC (Zermelo-Fraenkel set theory with the Axiom of Choice). Set theory is of particular interest since it is a sufficiently rich framework to incorporate all areas of mathematics (number theory, analysis, function theory, etc.) and the system ZFC is the generally accepted background framework of mathematics in the sense that a proof of a mathematical theorem is only regarded as legitimate if it can in principle be implemented in ZFC. Now, it turned out that there were some statements that were difficult to resolve. This is true, for example, of the statements PU (the statement that all projective sets admit of a projective uniformization) and CH (Cantor's continuum hypothesis). These statements were intensively investigated during the early era of set theory but little progress was made. The explanation for the lack of success was ultimately provided by results of Gödel and Cohen. For they showed that it is *in principle impossible to resolve these statements on the basis of* ZFC. And a similar situation prevails with regard to many other statements from diverse areas of mathematics. This is quite surprising. It means that if we are to have a hope of answering these questions we need new axioms.

This project has both a mathematical component and a philosophical component. On the mathematical side, one must find axioms that are sufficiently strong to do the work. On the philosophical side, one must determine what would count as a justification and whether the axioms in question are justified. These two components are intertwined and have contact points with a number of traditional philosophical questions, such as the question of the nature of reason (Are there "absolutely undecidable" statements?) and the question of pluralism (Are there incompatible yet equally legitimate mathematical frameworks?). Some have maintained that no reasons can be provided for axioms that settle statements like PU and CH. Others have gone further, maintaining that there is really a multitude of equally legitimate yet incompatible frameworks, some which settle CH positively, others which settle it negatively. These are the questions that animate my work.

It will be helpful to begin with a general comment on my approach and here it is useful to draw a comparison with the philosophy of physics. In the philosophy of physics, there are two kinds of work. First, there is work that proceeds in relative independence of contemporary developments in physics. Second, there is work that is closely tied to contemporary developments in physics. Both fields are quite active. In the philosophy of mathematics, one can draw a similar distinction. But in the case of the philosophy of mathematics, most contemporary work proceeds in relative independence of the current developments in mathematics. Although much excellent work has been done in this spirit, I think that there is much to be gained from pursuing a variety of philosophy of mathematics that is more intimately connected with current work in mathematics. Not only is there much to be gained from *assessing* the philosophical significance of existing results; there is even more to be gained through the *interaction* of the philosophy of mathematics with developments in mathematics. One of my goals is to prove theorems that are motivated by and shed light on philosophical questions.

It is sensible to approach the search for new axioms in a stepwise fashion, seeking first axioms that resolve certain low-level questions and then proceeding upward to questions of greater complexity. In what follows I will describe my past and current work on three stages of this program: (1) reflection principles, (2) axioms of definable determinacy and large cardinal axioms, and (3) the continuum hypothesis and beyond. In the first two cases (which cover second-order arithmetic (which is where PU resides)), the mathematical landscape with regard to the implications of new axioms has stabilized and we now have many of the resources needed to address the issue between the pluralist and the non-pluralist. In the third case (which concerns third- and higher-order arithmetic (which is where CH and everything else resides)), the subject awaits further mathematical developments. That is where my current research lies.[1]

1 Reflection Principles

Gödel drew a distinction between *intrinsic* and *extrinsic* justifications. An *intrinsic* justification of a statement concerning a given domain is one which is grounded in principles implicit in the conception of the domain. For example, mathematical induction is grounded in the conception of the natural numbers. In contrast an *extrinsic* justification of a statement is one that is not grounded solely in principles implicit in the conception of the domain. For example, a justification of a statement in terms of its fruitful consequences would count as an extrinsic justification. Intrinsic justifications are certainly more in line with traditional conceptions of mathematics. For this reason it is of interest to determine how far they can take us and, in particular, whether in the end we must resort to extrinsic justifications.

Reflection principles are the best current contender for new axioms in set theory that admit of an intrinsic justification. Indeed, some (most notably, Tait) have argued that intrinsic justifications are *exhausted* by reflection principles. Reflection principles aim to articulate the idea that the extent of the universe of sets cannot be described from below. The distinctive axioms of extent of ZFC (namely, Infinity and Replacement) are derivable from reflection principles and, furthermore, reflection principles imply the existence of large cardinals that are beyond the reach of ZFC. In this regard they can be used to substantiate Gödel's claim that certain small large cardinal axioms are as justified as the standard axioms of set theory. The hope, of course, is to go further and show that reflection principles imply large cardinal axioms which are sufficiently strong to effect a significant reduction in incompleteness. And, indeed, it is often maintained that reflection principles are capable of securing very strong large cardinal axioms. If this were indeed the case, then reflection principles (and hence intrinsic justifications, assuming that they did indeed secure reflection principles) would be capable of taking us quite far in effecting a significant reduction in incompleteness.

To asses just how far reflection principles can take us, we need to do

[1]For a related account of my research interests, see the table of contents and introduction to [7]. The body of that work contains a systematic account of my work in the area, though [2], [3], [5], and [6]—which are closely connected—are operating largely in the background.

two things: First, we need a precise explication of the notion of a "significant reduction in incompleteness." Second, we need to examine both the philosophical thesis that reflection principles are intrinsically justified and the mathematical situation concerning their strength.

I undertook this investigation in [4], focusing on the general reflection principles of Tait. On the philosophical side, I argue that intrinsic justifications are quite limited in terms of the strength of the reflection principles which they can secure. On the mathematical side, I prove three limitative results concerning general reflection principles. The first theorem carves out a class of general reflection principles and shows that they are consistent relative to a weak large cardinal axiom, namely, the axiom asserting that the Erdös cardinal $\kappa(\omega)$ exists. The second theorem shows that the remaining reflection principles are inconsistent. The third theorem shows that this is a sharp dichotomy in the sense that under a very fine stratification of the hierarchy of general reflection principles, as soon as one strengthens the consistent principles one arrives at inconsistent principles.

These results have a number of interesting philosophical consequences. First, the inconsistency result shows that serious problems can arise even when one is embarked on the project of unfolding the content of a conception. It should give us pause in placing too much confidence in the security of so-called intrinsic justifications. Second, the consistency result shows that intrinsic justifications, insofar as they are exhausted by the general reflection principles discussed above, cannot yield a significant reduction in incompleteness. Finally, these results can be used to provide a rational reconstruction of Gödel's early view that $V = L$, PU, and CH are "absolutely undecidable." For, if one has a conception of set theory which admits only intrinsic justifications and if one thinks that these are exhausted by reflection principles, then the above results make a case for the claim that these statements really are "absolutely undecidable". Fortunately, extrinsic justifications go a long way and I believe that one can make a strong extrinsic case for axioms that settle $V = L$ and PU.

2 Definable Determinacy and Large Cardinals

The above limitative results motivate the need to broaden the investigation and examine extrinsic justifications, which, recall, are justifications that involve more than what can be said to be implicit in the underlying conception. There are really two general approaches to new axioms—the *local* approach, which aims to understand the universe of sets level by level, and the *global* approach, which aims to understand the global structure of the universe of sets.

Faced with the independence of many of the questions of descriptive set theory—which includes the statements of second-order arithmetic and the statements concerning $L(\mathbb{R})$—set theorists took two approaches to finding new axioms—a local approach based on *axioms of definable determinacy* and a global approach based on *large cardinal axioms*. Both of these approaches are examined from a contemporary perspective in [1]. (See also Section 4 of [5] and Chapter 5 of [7].) I argue that a network of theorems (many of which are contained in [9]) enables one to make a strong case for axioms of definable determinacy and large cardinal axioms. These axioms are "effectively complete" with respect to statements of second-order arithmetic and $L(\mathbb{R})$. In fact, in a sense that can be made precise, they are "effectively complete" with respect to all statements of complexity strictly "below" that of CH. For example, they imply $V \neq L$ and PU.

There is not sufficient space here to describe the case in detail. Let me just make a few points, concentrating, for definiteness, on the axiom $\mathrm{AD}^{L(\mathbb{R})}$, which asserts that all sets of reals in $L(\mathbb{R})$ are determined. First, this axiom leads to a "complete" analysis of the structure $L(\mathbb{R})$ in much the way that $V = L$ leads to a "complete" analysis of L. One sign of this is the fact (discovered by Woodin) that under large cardinal assumptions (at roughly the level of $\mathrm{AD}^{L(\mathbb{R})}$) the theory of $L(\mathbb{R})$ cannot be altered by forcing. Second, it turns out that there is a very close connection between the local approach based on axioms of determinacy and the global approach based on large cardinal axioms. For large cardinal axioms *imply* axioms of definable determinacy. In fact, axioms of definable determinacy are *equivalent* to axioms asserting the existence of inner models of large cardinals. Finally, this inner model-theoretic connection leads to one of the strongest arguments for $\mathrm{AD}^{L(\mathbb{R})}$: This axiom is *inevitable* in the sense that it appears to be implied by *all* "natural" theories of sufficiently strength. This includes theories that are incompatible with one another. In short, $\mathrm{AD}^{L(\mathbb{R})}$ lies in the "overlapping consensus" of all sufficiently strong "natural" theories.

As noted above, in a sense that can be made precise, the above case extends to all statements of complexity strictly "below" that of CH. However, it has not been a success at the level of CH and there are reasons (stemming from work of Levy and Solovay) for thinking that this is the final word on the matter. With CH, it appears that one reaches a "phase transition" in the search for new axioms.

3 The Continuum Hypothesis and Beyond

This brings me to my current and future research, which concerns the prospect of resolving statements at the level of CH and beyond. The project has two closely related components. The first involves exploring the space of possibilities under which it would be reasonable to say that statements like CH are "absolutely undecidable" or that there is a plurality of equally legitimate systems, some of which resolve CH positively, others which resolve it negatively. The second involves exploring the space of possibilities under which it would be reasonable to say that CH has been solved. I have taken initial steps in both directions and will discuss them in turn.

3.1 Prospects for Bifurcation

There are many approaches to the question of pluralism that treat the question in a manner that is insufficiently sensitive to developments in mathematics. I discuss and criticize several such approaches in [2] and [5]. In [5] I suggest that there is room (and historical precedent) for a more meaningful engagement between philosophy and the exact sciences and, moreover, that it is through such an engagement that we can properly approach the question of pluralism in mathematics. I lay the groundwork for this new orientation by drawing a structural parallel between physics and mathematics. Einstein's work on special relativity is taken as an exemplar of the kind of meaningful engagement I have in mind. For it is with Einstein that we come to see, for reasons at the intersection of philosophy and science, that statements once thought to have absolute significance (such as "A is simultaneous with B") are ultimately relativized. Our question then is whether something similar could happen in arithmetic or set theory, that is, whether for reasons at the intersection of philosophy and mathematics—reasons sensitive to actual developments in mathematics—we could come to see that statements once thought to have absolute significance (such as CH) are ultimately relativized.

Let me discuss one scenario in which one could plausibly make a case for pluralism. It involves the failure of a very optimistic scenario for finding axioms that mitigate the effects of our current independence machinery.

It is well known that in contrast to the theory of $L(\mathbb{R})$, traditional large cardinal axioms cannot settle CH, the reason being that they are invariant under small forcing and yet one can alter the truth-value of CH by small forcing. Nevertheless, although one cannot have generic absoluteness at the level of Σ^2_1 (the level of CH), Woodin showed that one *can* have *conditional* generic absoluteness; more precisely, assuming that there is a proper class of measurable Woodin cardinals, any two generic extensions that satisfy CH agree on *all* Σ^2_1-statements. In this sense CH "freezes" the Σ^2_1-theory. Moreover, CH is the unique such Σ^2_1-statement, in that any other Σ^2_1-statement with this feature "freezes" the Σ^2_1-theory in the *same way*.

This can be conveniently reformulated in terms of a strong logic known as Ω-logic as follows: Assuming that there is a proper class of measurable Woodin cardinals, ZFC + CH is Ω-*complete* for Σ^2_1 in the sense that for any Σ^2_1-sentence φ, either

$$\mathrm{ZFC} + \mathrm{CH} \models_\Omega \varphi \quad \text{or} \quad \mathrm{ZFC} + \mathrm{CH} \models_\Omega \neg\varphi.$$

Moreover, CH is unique in that any other Σ^2_1-sentence A with this feature is Ω-equivalent to CH.

This result motivates the following optimistic scenario, which assumes large cardinal axioms: There may be a (recursively enumerable) sequence of axioms $\vec{A}$ such that for some large cardinal axiom L the theory $\mathrm{ZFC} + L + \vec{A}$ is Ω-complete for *all* of third-order arithmetic. Going further, it could be the case that for each specifiable fragment V_λ of the universe of sets there is a large cardinal axiom L and a (recursively enumerable) sequence of axioms $\vec{A}$ such that $\mathrm{ZFC} + L + \vec{A}$ is Ω-complete for the theory of V_λ. Moreover, it could be the case that any other theory with this feature, say $\mathrm{ZFC} + L + \vec{B}$, agrees with $\mathrm{ZFC} + L + \vec{A}$ on the computation of the theory of V_λ in Ω-logic. This would mean that there is a *unique* Ω-complete picture of the universe of sets up to V_λ. Furthermore, it could be the case that all of these Ω-complete pictures cohere. This would give us a unique Ω-complete picture of (the successive layers of) the entire universe of sets.

One could argue that such an Ω-complete picture of the entire universe of sets is the most that one could hope for. Should uniqueness hold, that would be the end of the story from the perspective of Ω-logic for there would be nothing about the (specifiable fragments of the) universe of sets that could not be settled on the basis of Ω-logic in this unique Ω-complete theory.

Unfortunately, uniqueness *must* fail. In joint work, Woodin and I proved that if there is one Ω-complete picture of such a level of the universe (and hence of arbitrarily large such levels), then there is necessarily an *incompatible* Ω-complete picture. For example, given an Ω-complete picture, say $\mathrm{ZFC} + L + \vec{A}$, one can manufacture another, say $\mathrm{ZFC} + L + \vec{B}$, such that the two differ on CH. It should be stressed that the choice of CH is just for illustration. Given one such theory, one has a great deal of control in generating others.

Suppose then that there is one such theory (or a sequence of such theories for higher and higher levels, all of which extend one another). The advocate of pluralism might argue as follows: Such an Ω-complete picture is the most that theoretical reason could hope to achieve. Given the fact that from one such picture (say $\mathrm{ZFC} + L + \vec{A}$) we can generate others (say $\mathrm{ZFC} + L + \vec{B}$) that are incompatible, and given that there is great flexibility (in the choice of $\vec{B}$) and that we can pass from one to another (by altering $\vec{B}$), this shows that the choice is merely one of expedience. The choice of $\vec{B}$ is analogous to the choice of a timelike vector in Minkowski spacetime. The choice of such a vector in Minkowski spacetime induces a foliation of the space relative to which one can ask whether "A is simultaneous with B" but there is no absolute significance to such questions independent of the choice of a timelike vector. Likewise, the choice of the sequence $\vec{B}$ (in conjunction with $\mathrm{ZFC} + L$) provides an Ω-complete picture relative to which one can ask whether "CH holds" (and many other such questions) but there is no absolute significance to such questions independent of the choice of $\vec{B}$. As in the case of special relativity we need to change perspective. There is no sense in searching for "the correct" Ω-complete picture, just as there is no sense in searching for "the correct" foliation. Instead of the naïve picture of the universe of sets with which we started we are ultimately driven to a new picture, one that deems questions we originally thought to be absolute to be ultimately relativized.

I do not want to endorse this position. I am merely presenting it on behalf of the pluralist as the best scenario that I can think of where one could arguably maintain that we had been driven by the primary data of mathematics to shift perspective and regard certain questions of mathematics (such as CH) as based on choices of mere expedience.

A key virtue of this scenario is that it is sensitive to future developments in mathematics. For to rule it out it suffices to prove the Ω Conjecture (an outstanding conjecture in set theory) and to establish it it suffices to find one such Ω-complete theory. In this way, by presenting mathematically precise scenarios that are sensitive to mathematical developments, the pluralist and non-pluralist can give the question of pluralism "mathematical traction" and, through time, test the robustness of mathematics.

What can we say at present concerning the above pluralist scenario? We can say this: Although it is an open mathematical possibility, there are reasons to think that such a scenario cannot obtain. For there is growing evidence for the Ω Conjecture and, as noted above, this conjecture rules out the existence of one (and hence many) such Ω-complete theories. Thus, one way to definitively rule out the above pluralist scenario is to prove the Ω Conjecture.[2]

3.2 Prospects for Resolution

We turn now to the positive approach. Here the goal is to examine mathematically precise scenarios in which it would be reasonable to say that CH and other statements about higher levels of the universe of sets have been settled. There are a number of approaches that have appeared in recent years and the task of exploring the space of possibilities is certainly more than one person can hope to achieve. I have had the good fortune of joining up with Hugh Woodin (the discoverer of much of the relevant mathematics) to embark on this project. Our findings are contained in the monograph [7]. Here I can give but the briefest sketch.

There are two *local* approaches to CH, one arguing in favour of CH, another arguing in favour of $\neg$CH. These are presented in Chapter 6 of [7]. Unfortunately, at present, there is no strong local case for one alternative over the other. Fortunately, there are two promising *global* approaches—one in terms of inner model theory and one in terms of structural theory.

[2]Should the Ω Conjecture be true, the proponent of pluralism would have to retreat and make another proposal. Two such proposals are discussed at the end of [8], but in each case one can show that the symmetry between the alternatives is broken and that there is one alternative that stands out. Once again, pluralism is undermined.

Inner model theory aims to produce "L-like" models that contain large cardinal axioms. For each large cardinal axiom Φ that has been reached by inner model theory, one has an axiom of the form $V = L^{\Phi}$. This axiom has the virtue that (just as in the simplest case of $V = L$) it provides an "effectively complete" solution regarding questions about L^{Φ} (which, by assumption, is V). Unfortunately, it turns out that the axiom $V = L^{\Phi}$ is incompatible with *stronger* large cardinal axioms Φ'. For this reason, axioms of this form have never been considered as plausible candidates for new axioms.

But recent developments in inner model theory (due to Woodin) show that everything changes at the level of a supercompact cardinal. These developments show that if there is an inner model N which "inherits" a supercompact cardinal from V (in the manner in which one would expect, given the trajectory of inner model theory), then there are two remarkable consequences: First, N is close to V (in, for example, the sense that for sufficiently large singular cardinals λ, N correctly computes λ^+). Second, N inherits all known large cardinals that exist in V. Thus, in contrast to the inner models that have been developed thus far, an inner model at the level of a supercompact would provide one with an axiom that could *not* be refuted by stronger large cardinal assumptions.

The issue, of course, is whether one can have an "L-like" model (one that yields an "effectively complete" axiom) at this level. There is reason to believe that one can. There is now a candidate model L^{Ω} that yields an axiom $V = L^{\Omega}$ with the following features: First, $V = L^{\Omega}$ is "effectively complete." Second, $V = L^{\Omega}$ is compatible with all large cardinal axioms. Thus, on this scenario, the ultimate theory is the (open-ended) theory ZFC$+$ $V = L^{\Omega} + \text{LCA}$, where LCA is a schema standing for "large cardinal axioms." The large cardinal axioms will catch instances of Gödelian independence and the axiom $V = L^{\Omega}$ will capture the remaining instances of independence. This theory would imply CH and settle the remaining undecided statements. Independence would cease to be an issue.

It turns out, however, that there are other candidate axioms that share these features, and so the spectre of pluralism reappears. For example, there are axioms $V = L^{\Omega}_{S}$ and $V = L^{\Omega}_{(*)}$. These axioms would also be "effectively complete" and compatible with all large cardinal axioms. Yet they would resolve various questions differently than the axiom $V = L^{\Omega}$. For example, the axiom, $V = L^{\Omega}_{(*)}$ would imply $\neg$CH. How, then, is one to adjudicate between them?

This brings me to the second global approach, one that promises to select the correct axiom from among $V = L^{\Omega}$, $V = L^{\Omega}_S$, $V = L^{\Omega}_{(*)}$, and their variants. This approach is based on the remarkable analogy between the structure theory of $L(\mathbb{R})$ under the assumption of $\mathrm{AD}^{L(\mathbb{R})}$ and the structure theory of $L(V_{\lambda+1})$ under the assumption that there is an elementary embedding from $L(V_{\lambda+1})$ into itself with critical point below λ. This *embedding assumption* is the strongest large cardinal axiom that appears in the literature.

The analogy between $L(\mathbb{R})$ and $L(V_{\lambda+1})$ is based on the observation that $L(\mathbb{R})$ is simply $L(V_{\omega+1})$. Thus, λ is the analogue of ω, λ^+ is the analogue of ω_1, and so on. As an example of the parallel between the structure theory of $L(\mathbb{R})$ under $\mathrm{AD}^{L(\mathbb{R})}$ and the structure theory of $L(V_{\lambda+1})$ under the embedding axiom, let us mention that in the first case, ω_1 is a measurable cardinal in $L(\mathbb{R})$ and, in the second case, the analogue of ω_1—namely, λ^+—is a measurable cardinal in $L(V_{\lambda+1})$. This is just one instance from among many examples of the parallel.

Now, we have a great deal of information about the structure theory of $L(\mathbb{R})$ under $\mathrm{AD}^{L(\mathbb{R})}$. Indeed, as we noted above, this axiom is "effectively complete" with regard to questions about $L(\mathbb{R})$. In contrast, the embedding axiom on its own is not sufficient to imply that $L(V_{\lambda+1})$ has a structure theory that fully parallels that of $L(\mathbb{R})$ under $\mathrm{AD}^{L(\mathbb{R})}$. However, the existence of an already rich parallel is evidence that the parallel extends, and we can supplement the embedding axiom by adding some key components. When one does, something remarkable happens: the supplementary axioms become *forcing fragile*. This means that they have the potential to erase independence and provide non-trivial information about $V_{\lambda+1}$. For example, these supplementary axioms might settle CH and much more.

The difficulty in investigating the possibilities for the structure theory of $L(V_{\lambda+1})$ is that we have not had the proper lenses through which to view it. The trouble is that the model $L(V_{\lambda+1})$ contains a large piece of the universe—namely, $L(V_{\lambda+1})$—and the theory of this structure is radically underdetermined. The results discussed above provide us with the proper lenses. For one can examine the structure theory of $L(V_{\lambda+1})$ in the context of ultimate inner models like L^{Ω}, L^{Ω}_S, $L^{\Omega}_{(*)}$, and their variants. The point is that these models can accommodate the embedding axiom and, within each, one will be able to compute the structure theory of $L(V_{\lambda+1})$.

This provides a means to select the correct axiom from among $V = L^{\Omega}$, $V = L^{\Omega}_S$, $V = L^{\Omega}_{(*)}$, and their variants. One simply looks at the $L(V_{\lambda+1})$ of each model (where the embedding axiom holds) and checks to see which has the true analogue of the structure theory of $L(\mathbb{R})$ under the assumption of $\mathrm{AD}^{L(\mathbb{R})}$. It is already known that certain pieces of the structure theory *cannot* hold in L^Ω. But it is open whether they can hold in L^Ω_S.

In our monograph we map out various scenarios for how the future might unfold. Let me mention one such (very optimistic) scenario: The true analogue of the structure theory of $L(\mathbb{R})$ under $\mathrm{AD}^{L(\mathbb{R})}$ holds of the $L(V_{\lambda+1})$ of L^Ω_S but not of any of its variants. Moreover, this structure theory is "effectively complete" for the theory of $V_{\lambda+1}$. Assuming that there is a proper class of λ where the embedding axiom holds, this gives an "effectively complete" theory of V. And, remarkably, part of that theory is that V *must* be L^Ω_S. This (admittedly very optimistic) scenario would constitute a very strong case for axioms that resolve all of the undecided statements.

I do not wish to place too much weight on this particular scenario. It is just one of many that we discuss. My main point is that we are now in a position to write down a list of definite questions with the following features: First, the questions in this list will have answers—independence is not an issue. Second, if the answers converge then one will have strong evidence for new axioms settling the undecided statements; while if the answers oscillate, one will have evidence that these statements are "absolutely undecidable" or that we must embrace pluralism. In this way the questions of "absolute undecidability" and pluralism are given mathematical traction. I conjecture that the answers will converge.

4 Other

1. In the above discussion I have considered a large segment of the hierarchy of mathematical systems. But there are other regions of this hierarchy that I wish to investigate. Let me mention the two extremes. At the very low level there is the question of finding the appropriate system corresponding to the philosophical position known as *strict finitism*. Typically, advocates of this view (such as Nelson) accept Q but reject the totality of exponentiation. However, Visser has shown that exponentiation is interpretable in $\mathrm{Q} + \mathrm{Con}(\mathrm{Q})$. Thus, this stance would appear to violate the principle of reflective closure. The question then arises: What statements beyond Q is the strict finitist committed to? Is there even a stable foundational position here?

Moving to the other extreme, I am interested in the hierarchy of large cardinal axioms that are inconsistent with AC, the first being the axiom (due to Reinhardt) that asserts the existence of a non-trivial elementary embedding of the universe into itself. Are these axioms consistent in ZF? If so is there any interesting structure theory? Supposing that there is, one might consider the foundational position that regards AC as a limitative axiom (much like $V = L$) and as something that holds only in a certain context. Is such a foundational position tenable?

2. Closely related to this last question is the question of absolute definability. Gödel proposed *ordinal definability* as a candidate for the notion of absolute definability. One reason for this particular proposal is that any candidate for an absolute notion of definability should be such that it cannot be transcended via diagonalization. Moreover, any notion of definability that does not render all of the ordinals definable can be so transcended (since, by considering the least ordinal that is not definable according to the notion, we exhibit a richer notion of definability). This suggests ordinal definability as the appropriate notion. But is it really "absolute"? There are other criteria of absoluteness. For example, one would expect that if the notion of ordinal definability is truly absolute then there should be no non-trivial embedding from HOD into HOD, where HOD is the class of sets that are hereditarily ordinal definable. It is open whether there can be such an embedding.

3. There has been a great deal of work showing that certain "natural" statements of first-order arithmetic require strong assumptions for their positive resolution. There is a wealth of opportunities in the context of second-order arithmetic. In fact, in this realm one can find statements that were arguably of "prior mathematical" interest and yet require very strong assumptions for their positive resolution. For example, the early analysts intensively studied the projective sets and they considered the uniformization problem. The projective sets are the subsets of $\mathbb{R}^n$ (where we treat the reals as ω^ω, as is customary in descriptive set theory) that can be obtained from closed sets by alternating the operations of projection and complementation. Consider the third level of this hierarchy, that is, the sets that can be obtained from the closed sets by applying projection and complementation three times. Call these sets the Π^1_3-sets. Let Π^1_3-uniformization be the statement that every Π^1_3-subset of the plane admits of a Π^1_3-uniformizing function. This statement was the kind of statement that the early analysts investigated. They also considered such statements as PM and PB asserting, respectively, that all projective sets are Lebesgue measurable and that all projective sets have the property of Baire. Now take these three statements

together. It is a corollary of work of Steel that these three statements imply the consistency of ZFC + "There is a Woodin cardinal." Thus, we have here a case where statements that were of prior mathematical interest require a theory of astonishing power for their positive resolution. I suspect that there are many more examples in this vein.

4. Another interesting line of investigation involves the exploration of analogies and disanalogies between the kinds of independence one finds in set theory and the kinds of independence one finds in arithmetic. For example, in the set theoretic case, there are "natural" examples of Orey sentences, for example CH. In the case of arithmetic, one can use Gödelian techniques to manufacture "artificial" Orey sentences but there are no known "natural" examples. The key difference, of course, is that in set theory one has the technique of forcing and there is no known analogue of this technique for arithmetic. Are there "natural" Orey sentences for arithmetic? The search for such a sentence could lead the a new independence technique.

This is just one of many questions in an area that looks for parallels between the kinds of independence in arithmetic and the kinds of independence in set theory. I have undertaken a preliminary study of this subject in [3].

5. Thus far I have concentrated on my research interests in the foundations of mathematics. But I also have strong interests in other fields, most notably, early analytic philosophy, the history of the relationship between philosophy and the exact sciences, and the philosophy of physics.

In early analytic philosophy, most of my research has concentrated on Russell, Carnap, and the logical positivists. I have had a long standing preoccupation (one might say obsession) with Russell's early work on logic and the foundations of mathematics. I have started to write a book on the subject but I expect that it will be some time before it is completed. In [2] I critically analyze Carnap's views on logic and the foundations of mathematics from a contemporary perspective. I have done some work on the engagement of the logical positivists with the special theory of relativity and will explore this topic further in a forthcoming course.

My work in the history of the exact science and the philosophy of physics has, for the most part, been confined to teaching courses and supervising theses on these subjects. I am interested in these subjects both as subjects in their own right and because I think that some of the problems I have struggled with in the foundations of mathematics have cousins within these disciplines (on which see [5]). I hope one day to make contributions that fall squarely in these fields, especially in the philosophy of physics.

References

[1] Peter Koellner. On the question of absolute undecidability. *Philosophia Mathematica*, 14(2):153–188, 2006. Revised and reprinted with a new postscript in *Kurt Gödel: Essays for his Centennial*, edited by Solomon Feferman, Charles Parsons, and Stephen G. Simpson. Lecture Notes in Logic, 33. Association of Symbolic Logic, 2009.

[2] Peter Koellner. Carnap on the foundations of logic and mathematics. Submitted, 2009.

[3] Peter Koellner. Independence in arithmetic and set theory. Preprint, 2009.

[4] Peter Koellner. On reflection principles. *Annals of Pure and Applied Logic*, 157(2), 2009.

[5] Peter Koellner. Truth in mathematics: The question of pluralism. In Otávio Bueno and Øystein Linnebo, editors, *New Waves in Philosophy of Mathematics*, New Waves in Philosophy. Palgrave Macmillan, 2009. Forthcoming.

[6] Peter Koellner. Strong logics of first and second order. *Bulletin of Symbolic Logic*, 2010. Forthcoming.

[7] Peter Koellner and W. Hugh Woodin. *Foundations of Set Theory: The Search for New Axioms*. 2009.

[8] Peter Koellner and W. Hugh Woodin. Incompatible Ω-complete theories. *The Journal of Symbolic Logic*, 74(4), December 2009.

[9] Peter Koellner and W. Hugh Woodin. Large cardinals from determinacy. In Matthew Foreman and Akihiro Kanamori, editors, *Handbook of Set Theory*. Springer-Verlag, 2009.