- **Peter Koellner, Harvard University**

On the Prospect of Bifurcation in Set Theory

The current project involves two connected themes from Gödel: The nature of reason in mathematics and the question of realism about the higher infinite. To motivate the details of the project it will be useful to situate it within a larger, more general project, which also concerns themes from Gödel. This larger project involves three stages of the hierarchy of axioms of infinity, each stage of which is discussed in the trilogy of papers: "On Reflection Principles", "On the Question of Absolute Undecidability", and "Incompatible Ω-Complete Theories" (with Hugh Woodin). Since the development is cumulative I will discuss the subjects of these papers in order. The final paper is of particular importance since it contains the first positive result in the current project.

The motivation for the larger project begins with the incompleteness phenomenon in set theory, which provides us with natural statements that cannot be settled on the basis of the standard axioms of set theory, ZFC. Two classic examples are PU (the statement that all projective sets admit of a projective uniformization) and CH (Cantor's continuum hypothesis).

This leads to the program of seeking and justifying new axioms settling the undecided statements. This program has both a mathematical component and a philosophical component. On the mathematical side one must find axioms which are sufficiently strong to do the work. On the philosophical side one must determine what would count as a justification and determine whether the axioms in question are justified. As will become apparent, the two components are intertwined and have contact points with a number of traditional philosophical questions, such as the nature of reason (Are there "absolutely undecidable" statements?) and the question of realism about the infinite (Is there an unambiguous notion of the higher infinite?)

1 Intrinsic Justifications: 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 involves features which do not involve the basic 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. It is of interest then 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 contenders for new axioms in set theory that admit of an intrinsic justification. These axioms aim to articulate the idea (arguably grounded in our conception of set) 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 yield large cardinal axioms beyond those provable in ZFC. In this regard they can be used to substantiate Gödel's claim that like the axioms of extent of ZFC there are certain small large cardinal axioms that merely unfold the content of the (iterative) concept of set and hence are intrinsically justified. 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 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 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.

In "On Reflection Principles" I undertake this investigation, focusing on the strongest and most general known reflection principles, namely, those 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 a number of theorems that collectively show that there are severe limitations on the power of reflection principles. The first theorem carves out a class of 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 dichotomy is sharp.

These results have a number of interesting consequences: First, the inconsistency result casts doubt on the security of intrinsic justifications. Second, one can use these results to provide a rational reconstruction of Gödel's early view that $V=L$ is "absolutely undecidable". Finally, the results show that intrinsic justifications do not yield reflection principles that effect a significant reduction in incompleteness; for example, they do not yield reflection principles that settle $V=L$, PU or CH.

2 Extrinsic Justifications: Large Cardinal Axioms and Axioms of Definable Determinacy

The above limitative results motivate the move to broaden the investigation and examine extrinsic justifications. Gödel came to endorse extrinsic justifications and, in fact, on the basis of such a broad notion of justification (which included a strong notion of intrinsic justification) he came to believe that one could justify axioms that settle all questions of set theory. As a particular approach he advanced his famous *program for large cardinals.* The aim of this program is to show (on the mathematical side) that certain large cardinal axioms settle undecided statements and (on the philosophical side) that these large cardinal axioms are justified.

In "On the Question of Absolute Undecidability" I provide a classification of the current forms of extrinsic justification in set theory. Following Martin, Steel, Woodin and others, I argue that a cluster of results in set theory make for a compelling extrinsic case for strong large cardinal axioms and axioms of definable determinacy.

The case for large cardinal axioms is in part based on the fruitfulness of their consequences. In a sense which can be made precise, Gödel's program for large cardinal axioms has been a complete success "below" CH. However, it has not been a success at the level of CH and there are some reasons (stemming from work of Levy and Solovay) for thinking that this is the final word on the matter. Thus, in choosing CH as a test case for his program, Gödel appeared to have put his finger precisely on the point where it fails.

The case for axioms of definable determinacy is in part based on the remarkable fact that they follow from large cardinal axioms and hence inherit the extrinsic justifications of the latter. However, there are independent extrinsic justifications for axioms of definable determinacy, some of which are extremely strong. One of the strongest is based on the "inevitability" of these axioms, that is, the phenomenon whereby axioms of definable determinacy appear to be implied by any "natural" theory of sufficiently strong interpretability power.

The hope is that although current large cardinal axioms cannot settle CH perhaps some of the *kinds* of extrinsic justifications supporting large cardinal axioms and (especially) axioms of definable determinacy can secure *other* axioms which settle CH and other undecided statements, or that there are other kinds of extrinsic justifications of an unanticipated nature which secure other, much stronger, principles.

3 The Prospect of Bifurcation: CH and Beyond

The question then remains of whether there are extrinsic justifications of axioms which settle CH or whether CH is "absolutely undecidable" or whether there is a "bifurcation" at the level of CH.

At this stage we need to stand back and ask (1) what it would take to have a convincing case for axioms settling CH and (2) what it would take to have a convincing case for the claim that CH is "absolutely undecidable" and (3) what it would take to have a convincing case for the claim that there is a "bifurcation" in set theory at the level of CH. The questions are interconnected. Woodin and I are currently investigating the first. In the present project I wish to focus on the second and third.

The goal is to examine mathematically precise scenarios in which it would be reasonable to say that a given statement is "absolutely undecidable" or signals a "bifurcation" and then to establish mathematical results on the tenability of these scenarios. This approach to the prospect of bifurcation, and more generally the question of realism in mathematics, finds its inspiration in theoretical developments in physics. In special relativity we have a clear case where developments in physics have led us to switch from being factualists about statements of the form "A and B are (absolutely) simultaneous" to being non-factualists about them. Are there analogous results in set theory?

Here is a sample result of what I have in mind. It concerns an optimistic scenario for extending the axioms of ZFC in conjunction with large cardinal axioms. One can ask for recursively enumerable axioms A such that relative to large cardinal axioms ZFC + A is Ω-complete for all of third-order arithmetic. Going further, for each specifiable segment V_λ of the universe of sets (for example, one might take λ to be the least huge cardinal), one can ask for recursively enumerable axioms A such that ZFC + A is Ω-complete for the theory of V_λ, relative to large cardinal axioms. If such theories exist, extend one another, and are unique in the sense that any other theory A' with the same level of Ω-completeness as A is actually Ω-equivalent to A, then this would make for a very strong case for new axioms that settle the theory of V in Ω-logic. One would have a unique Ω-complete picture of each such V_λ.

In "Incompatible Ω-Complete Theories" Woodin and I show uniqueness must fail. In particular, we show that if there is one such theory that Ω-implies CH then there is another that Ω-implies $\neg$CH. This is just a sample. One can replace CH by anything that can be forced with a definable, homogeneous partial order. Thus, if there is one such Ω-complete picture of such a level V_λ then there is of necessity a broad array of incompatible Ω-complete pictures of V_λ. In this sense one has a "bifurcation" into incompatible Ω-complete pictures. Furthermore, this is a case where the question of bifurcation (as regimented along the above lines) is something which can get mathematical traction since whether it is possible hinges on an outstanding conjecture in set theory, the Ω Conjecture.

There are open questions in this area and there are other candidates for rendering precise the view of bifurcation in set theory. In general the advocate of bifurcation holds that instead of a single, unambiguous universe of sets there is instead a "multiverse" of sets. This is accompanied by a "multiverse view of meaning and truth" according to which a statement is meaningful if and only if it has the same truth-value in all universes in the multiverse and true if and only if it is true in all universes of the multiverse. The earliest multiverse conception was based solely on the possibility of forcing over models of ZFC. However, this view is rendered untenable by the existence of strong justifications of axioms settling statements like PU since such a multiverse conception would deem these statements meaningless. Still, there are modern variants of the view, such as Woodin's "generic multiverse", which presuppose the existence of large cardinals. Is such a multiverse conception tenable? There are a number of difficulties with such a view. For example, it rests on set forcing and it is hard to see how one could incorporate class forcing. Perhaps one can prove a general theorem limiting the possibilities of such a conception in such a way as to render the prospect untenable.

The above is just to provide a flavor of the kind of results that I seek. In general the aim is to explore the space of possibilities in which one might maintain that certain statements are "absolutely undecidable" or signal a "bifurcation in set theory."

I plan to carry out the project at UC Berkeley and Harvard University.