Some Naturalistic Reflections on Set Theoretic Method

Penelope J. Maddy

Set theory has been subject to sharp methodological disputes from its very beginnings, but over the years most of these – like those concerning impredicative definitions or the axiom of choice – have been resolved to the general satisfaction of the set theoretic community. So many themes have run through these debates, so many theses have been raised, challenged, and defended, that it's difficult to sort out which of these considerations actually brought about the stable outcome. Some observers think these matters were decided on philosophical grounds; some pessimists see only sociological forces at work. These questions are important for the practice of contemporary set theory, because methodological questions remain to this day: what is the status of independent questions like the Continuum Hypothesis? Should they be abandoned or pursued? If the latter, what means are appropriate? In particular, how are new axiom candidates to be evaluated?

On these issues, I disagree with both the philosophers and the pessimists. In the now-settled methodological disputes, I think consensus was eventually reached for good mathematical reasons, that is, for reasons integrally connected to the mathematical goals the developing theory hoped to attain: for example, a classical theory of real numbers, in the case of impredicative definitions; a staggeringly wide range of particular benefits in the case of the axiom of choice (a well-behaved theory of infinite cardinals, to take just one example). In other words, I think one can argue for the rationality of these methodological decisions of the past and illuminate the underlying justificatory structure of debates of the present by strict attention to the detailed mathematical considerations in play.[1] This is what I call 'naturalized methodology', and I have tried to apply it in some particular cases.[2]

One factor often cited in philosophical and methodological discussions of mathematics is the widespread and impressive applicability of mathematics in the natural sciences. Quineans hold that mathematics is confirmed by these applications, that only that part of mathematics with application is justified. Defenders of pure mathematics naturally disagree, finding justification for unapplied mathematics in various places, perhaps in some largely unspecified aesthetic virtues or in some philosophy of mathematics.

My ultimate goal in this paper is to illuminate, from a naturalistic point of view, the significance of the application of mathematics in the natural sciences for the practice of contemporary set theory. But it will take me a while to get there. In section I, I give a sketch of the naturalistic approach by outlining an assessment of the role and rationality of two methodological maxims; in section II, I suggest that this way of seeing things undermines some misguided analogies familiar from discussions of set theoretic method. This leads back, in section III, to the issue at hand: the significance of applied mathematics for mathematics in general, and for the practice of set theory in particular.

I. Two contrasting maxims

Let me illustrate the naturalistic approach to set theoretic methodology with a simple example. When it comes to conflicting new axiom candidates – say $V = L$ vs. various Large Cardinals – set theorists want to choose; they want a single theory to enshrine at the beginning of their textbooks. This is a very basic methodological decision: we want one fundamental theory, not an array of options.[3] The methodologist naturally wonders if there is any justification for this preference.

One attempt at justification stands out, namely, the claim that there is an objective world of sets which it is the set theorist's job to investigate, a world in which

Topoi **20**: 17–27, 2001.

V = L and Large Cardinal axioms cannot both be true. Though focused on a different example, this is the philosophical realism made famous by Gödel:

> The set theoretic concepts and theorems describe some well-determined reality, in which Cantor's conjecture must be either true or false. (Gödel, 1964, p. 260)

The pessimists, on the other hand, see no rational justification, only sociological explanations.

The naturalist, as advertised, sides with neither the philosopher nor the pessimist; instead she looks for garden-variety mathematical considerations that might bear on the rationality of the methodological preference for a single fundamental theory of sets. Above, I suggested that such a garden-variety study might begin with an investigation of the goals, or a goal, of the mathematical practice and go on to assess the rationality of a method in terms of its efficacy toward achieving those goals. I think this strategy plays out in a particularly straightforward way for this example.

Now I have no doubt that plausible, even conclusive arguments could be given for a range of different goals and subgoals of any mathematical practice as rich and varied as contemporary set theory, but I do think we can agree to at least one fairly uncontroversial motivation: in particular, it seems that set theory hopes to provide a foundation for classical mathematics. This means at least that set theory aims to provide a dependable and perspicuous mathematical theory that is rich enough to include (surrogates for) all the objects of classical mathematics and strong enough to imply all the classical theorems about them. In this way, set theory aims to provide a court of final appeal for existence claims of classical mathematics – the vague question 'is there a mathematical object of such-and-such description?' is replaced by the precise question 'is there a set like this?' Likewise, the vague question 'can so-and-so be proved mathematically?' is replaced by the precise question 'is there a proof from the axioms of set theory?' Set theory aims to provide a single arena in which the objects of classical mathematics are all included, where they can be compared and contrasted and manipulated and studied side-by-side.

This gives some indication of the positive sense I attach to the claim that set theory aims to provide a foundation for classical mathematics. But given the history of foundationalism, especially its philosophical history, I should also emphasize what I do *not* mean. In a metaphysical or ontological sense, I surely do not claim that set theory reveals the true nature of mathematical objects, that it reveals numbers, ordered pairs, functions, algebraic structures, topological and geometric spaces, etc., etc., to have been sets all along. On this point, my set theoretic foundationalist need take no position; she simply claims that satisfactory set theoretic surrogates can be found for the objects of classical mathematics and that the classical theorems about those objects can be proved from the axioms of set theory. No deeper metaphysical or ontological claim is required.

I should also emphasize that no strong epistemic claim is involved. It might have been nice if the original dream of Frege and others had been realized, if classical mathematics could have been derived by transparent steps from self-evident, absolutely certain truths, but this is obviously not a style of foundation that set theory can provide. Indeed, we now know that we can't even be fully confident that our theory of sets is consistent. But it's just as important to recognize that diminished expectations are far from vanishing expectations: the ability of set theory to encompass the entire range of mathematicalia devised by classical mathematicians to our day and to prove the entire range of theorems proved in those efforts is a far from trivial achievement, even if the epistemic security blanket envisioned by the early foundationalists is not forthcoming.

In fact, I think there is another discomfort lurking behind the reluctance of some mathematicians to acknowledge the foundational role of set theory, a discomfort not unrelated to the overblown epistemic ambitions of the early foundationalists. Mathematicians in fields other than set theory often feel that set theoretic thinking doesn't capture the special sensibility that is essential to their subject, that it doesn't capture the way an algebraist, a topologist, or a geometer thinks. This seems quite true. Set theory is an individual branch of mathematics in its own right, with distinctive approaches and insights and methods of its own, approaches, insights and methods often quite different from those of algebra, topology or geometry.

Now if set theory were claimed to provide a foundation in the sense that the reduction of a branch of classical mathematics to set theory reduced all the methods of that branch to set theoretic methods, if the reduction were claimed to remove the need to pursue the particular branches of mathematics in their own right, if the upshot of the reduction were claimed to be that all mathematics *is just set theory* in the sense that all of mathematics would be developed simply by developing set theory – if any of these ridiculously strong claims were made, then the methodological independence of the particular branches of mathematics would undermine set theoretic foundations. But it would be both silly and unnecessary to make any of these claims. Set theory *does* provide an arena rich enough to include surrogates for all classical mathematical objects and axioms strong enough to prove all classical theorems, but it seems only reasonable to admit that the (surrogates for) objects of the particular branches of mathematics might not be salient from a purely set theoretic point of view and that the proof techniques of those branches might not be discovered in the ordinary pursuit of set theory itself. None of these eminently reasonable observations casts the slightest shadow on set theory as a foundation in the sense that I'm proposing.

It seems to me that there can be no doubting the mathematical benefits of a foundation for mathematics even in this circumscribed sense: it's important to have a canonical answer to questions of existence and proof, and to have an arena in which (surrogates for) the various objects of classical mathematics can all be considered and compared. But, just as obviously, there are those who hold that even this sort of a foundation is a wrong-headed goal or that it can be better served by, say, category theory. Fortunately, our purposes don't require us to enter into these controversies. For the purposes of our investigation of set theoretic methods, all that matters is that set theoretic practice does in fact have as one of its goals that of providing a foundation of this sort for classical mathematics. Whatever we may think about the goal itself or the unique suitability of set theory to it, the simple fact that set theory has this goal makes it rational for set theorists to adopt methods that are effective ones for reaching it.

All that said, it's now easy to see that this foundational goal does in fact underwrite a general methodological maxim I call UNIFY, that is, the admonition to settle on one official theory of sets. If you want set theory to provide a final court of appeal for existence and proof, if you want to provide a single arena in which (surrogates for) all classical mathematical objects can be manipulated and compared, then you must provide a single, fundamental theory. This is not to say that alternative set theories could not or should not be studied, but their models would be viewed as residing in the one true universe of sets, V. Given this foundational goal, our naturalistic analysis reveals that it is fully rational for set theorists to follow the methodological maxim UNIFY, quite apart from any controversial philosophical considerations.

It might clarify the structure of this naturalistic justification for UNIFY if we contrast this case with that of another, perhaps more familiar maxim, namely GENERALIZE. In set theoretic axiomatics, GENERALIZE is often cited in discussions of large cardinal axioms. So, for example, many large cardinals are presented as generalizations beginning from properties of $\aleph_0$: inaccessible and measurable cardinals generalize properties of $\aleph_0$ to uncountable levels; measurability itself has been generalized to supercompactness. But the notion of generalization is a common one that often appears in other branches of mathematics. For example, the historian Morris Kline describes the notion of a topological space as 'a generalization of a metric space' (Kline, 1972, p. 1160). What I want to suggest here is that the broad application of GENERALIZE beyond set theory isn't the only way in which it differs from UNIFY, but first we should take a moment to compare its application in set theory with its application elsewhere in mathematics, for example, in topology.

At one level, they seem quite different. In the topological case, we take an established notion – metric space – and we weaken the conditions: the metric space has a metric which determines a family of neighborhoods; a topological space is required to have the family of neighborhoods but not the metric that was once used to generate them. So the new notion is broader; every metric space is a topological space, but some topological spaces don't admit a metric that generates the requisite neighborhoods. In contrast, in the set theoretic case, we strengthen the conditions: $\aleph_0$ is a regular, strong limit cardinal; an inaccessible is an *uncountable*, regular, strong limit cardinal. The new notion is narrower; all inaccessibles are regular, strong limits, but the regular, strong limit $\aleph_0$ is not inaccessible.

This disanalogy strikes me as superficial, generated by the logical differences between generalizing to a new definition and generalizing to a new existential axiom. In the first case, when the conditions of the old definition are weakened, the original metric and function spaces that inspired the generalization then take their place in the broader class of topological spaces. In the second case, when the axioms of the old system are strengthened, the original regular, strong limit cardinal that inspired the generalization, namely $\aleph_0$, takes its place in the broader class of inaccessible cardinals;

likewise, the merely measurable cardinals that inspired the generalization take their place among the broader class of λ-supercompacts.[4] In this sense, it seems reasonable to view both moves as generalizations.

But the more important analogy comes at a higher remove, where we see both topological and set theoretic practices as engaged in concept formations in pursuit of their own particular mathematical goals. Topology has roots in such earlier studies as Felix Klein's classification of geometries in terms of invariants under transformations and Fréchet's work bringing Cantorian ideas to bear on the study of function spaces. As Kline puts it

> The basic task of point set topology is to discover properties that are invariant under continuous transformations and homeomorphisms. (Kline, 1972, p. 1161)

It is in pursuit of this goal that the notion of a topological space is isolated. Similarly set theory develops notions like inaccessible, measurable and supercompact cardinals as ways of developing a notion of the set theoretic universe that more effectively meets its particular mathematical goals, UNIFY among them.[5] What we've noted so far is that both, in the course of this development, employ a principle that deserves the name GENERALIZE.

Notice, also, that in both cases there are many, many ways to generalize: there are many ways to weaken the requirements on metric spaces and many ways to strengthen the requirements on regular limit or measurable cardinals. After his historical and mathematical presentation of the development of strongly compact and supercompact cardinals, Akihiro Kanamori concludes:

> In a few years . . . supercompactness came to be accepted as the proper generalization of measurability in the emerging hierarchy of large cardinals. (Kanamori, 1994, p. 308)

This type of observation is familiar from other branches of mathematics as well; it isn't enough to generalize, one must find the *right* generalization. And, as naturalists, we would expect rightness to be judged in terms of the goals of the particular practice.

This hardly serves to distinguish GENERALIZE from UNIFY in the practice of set theory, as there are various ways to UNIFY as well, and these are also to be judged in terms of the goals of the practice. But there is a further methodological point worth noting: generalization is not a good in itself; when a mathematical notion has reached its 'proper' level, further generalizing is actually counterproductive. For our purposes, this means: the fact that X is a generalization of an axiom that helps set theory better meet its goals doesn't by itself give us reason to suppose that X might help set theory better meet its goals. The methodological moral is that the maxim GENERALIZE doesn't play the same role as UNIFY: UNIFY is so directly tied to the particular goals of set theory that to motivate a decision on the basis of UNIFY is to lend support to the rationality of that decision at the same time (even if it ultimately turns out to be the wrong way to UNIFY, in light of other goals). GENERALIZE is not directly connected to the particular goals of set theory and thus provides no such justificatory force. So, for example, GENERALIZE, applied to large cardinal axioms, may well yield promising new large cardinal axioms, but the actual support for those new axioms must come from other sources. This is not to say that GENERALIZE isn't a good maxim, but to note that it serves a heuristic rather than a justificatory role.[6]

I hope this sketchy analysis of these two methodological maxims helps give the flavor of naturalistic methodology. Let me now turn to some commonplaces in ongoing discussions of set theoretic method to see what illumination our naturalistic approach can provide.

II. Two false analogies

There is a familiar line – found both in print and (perhaps more frequently) in the folklore – that aims to undermine the set theorist's effort to settle on answers to the CH and other independent questions by pressing analogies with algebra or geometry. 'To try to decide the CH', supporters of this line might say, 'is like trying to decide if groups are commutative'. In set theory and in group theory, these thinkers insist, we are studying all models of a given set of assumptions; some of those models will be commutative, some not; some will satisfy CH, some not; to try to figure out which models are 'right' or 'intended' is a nonsensical undertaking. Another version likens settling the CH with deciding the parallel postulate: clearly we have Euclidean geometry with the parallel postulate and non-Euclidean geometries without; just so there are models of set theory with CH and models without. Short of physical interpretation, this is all there is to the story.

Now I've sketched a naturalistic argument that the set theorists' efforts to settle CH and the rest, and to add new axioms, are rational in light of the goals of set theoretic practice. This suggests that the proposed analogies are deceptive, but we can get a much better idea of why this is so by taking a naturalistic look at the parts of algebra and geometry in question. Let me begin with group theory.

The origins of the group notion are often traced to Galois' work (around 1830) on the solvability of equations by algebraic means. Though his methods are presented to modern audiences in terms of substitution groups and their subgroups, Galois himself never explicitly isolated the group concept. About twenty years later, Cayley, inspired by Galois, did define the abstract group, and in addition to Galois' substitution groups, he mentions matrices with multiplication and quaternions with addition. Perhaps it's surprising to us nowadays, accustomed as we are to the importance and centrality of the concept of group, but at the time, Cayley's idea passed unnoticed. Kline explains this way:

> Cayley's introduction of the abstract group concept attracted no attention at this time, partly because matrices and quaternions were new and not well known and the many other mathematical systems that could be subsumed under the notion of groups were either yet to be developed or were not recognized to be so subsumable. (Kline, 1972, pp. 769–770)

About ten years later, Dedekind derived the abstract notion from his work on permutation groups, again without much influence.

It wasn't until the 1870s that the group concept found its audience. During that decade, Dedekind identified finite groups in his work on algebraic number theory, Kroneker isolated the group concept while working on Kummer's ideal numbers, Cayley wrote several more papers emphasizing that the abstract structure goes beyond substitution groups, Frobenius and Stickelberger extended the concept to congruences and Gauss's composition of forms, Netto introduced the notions of isomorphism and homomorphism between groups, Klein used infinite transformation groups to classify geometries, and Lie considered infinite transformation groups in connection with differential equations. Kline summarizes:

> By 1880 four main types of groups were known. There are the discontinuous groups of finite order, exemplified by substitution groups; the infinite discontinuous (or discrete) groups, such as occur in the theory of automorphic functions; the finite continuous groups of Lie exemplified by the transformation groups of Klein and the more general analytic transformations of Lie; and the infinite continuous groups of Lie defined by differential equations. (Kline, 1972, p. 1140)

At this point, von Dyck, influenced by Cayley and his teacher Klein, brought together the threads of 'theory of equations, number theory, and infinite transformation groups' under the abstract notion of a group, and the theory flourished from there (Kline, 1972, p. 1141).

The moral of this story, I suggest, is that the notion of group only came into mathematical prominence when it began to serve a particular mathematical goal. Speaking of abstract algebra in general, Kline puts it this way:

> Its concepts were formulated to unify various seemingly diverse and dissimilar mathematical domains as, for example, group theory did. (Kline, 1972, p. 1157)

If this is a fair analysis of the goal served by the group concept – and it is the burden of Kline's historical analysis to show that it is – then we can easily see why Galois didn't bother to draw it out, why Cayley's initial definition fell on deaf ears, and why group theory leapt to the fore in the 1880s. The purpose of the notion of a group is to call attention to similarities between a broad range of otherwise dissimilar structures. In doing so, it not only provides an elaborate and detailed general theory that can be applied again and again; it also more accurately isolates the features responsible for the particular phenomena ('that x has feature y isn't due to its idiosyncrasies z or v or w, but only to its group structure'). Until enough such structures had been examined and explored, the concept wasn't doing any work, wasn't serving any mathematical purpose. Only when it began to serve that purpose was it embraced by the mathematical community.

Now let's return to the purported analogy between group theory and set theory: trying to settle CH is like trying to decide if a group must be commutative; both theories have many different models, some of which satisfy CH or the commutative law and some of which don't; there's no reason to try to select out one of these models as 'right' or 'intended'. If we now reconsider this argument while bearing in mind the contrasting goals of set theory and group theory, it no longer sounds so persuasive. Given that group theory is *designed* to bring together a wide range of disparate mathematical structures, it would make no sense to try to rule out some of those structures as groups. (We might go on to consider rings and fields, of course, but that doesn't remove the importance of the underlying group structure.) On the other hand, given that set theory is (at least partly) *designed* to provide a foundation for classical mathematics, to provide a single arena for mathematical existence and mathematical proof, it does make sense to try to make our theory of sets as decisive as possible, to try to choose between alternative axioms, to try to rule out models that do this foundational job less well than others, in short, to UNIFY. There's nothing wrong with group theory or set theory, but they are aimed at different mathematical goals.

But if we reject this common analogy between set theory and group theory, the one posed as an objection to UNIFY, I think we can still see the two as analogous at a somewhat higher remove: though they are aimed at different goals, the two practices are both aimed at identifiable mathematical goals, and the large scale structure of their efforts to meet those goals *are* analogous. What I mean is this. Part of the historical story Kline tells is how the mathematical and conceptual developments gradually converged on the 'right' formulation of the group concept. For example, he reports that

> Lie recognized during the course of his work that one should postulate as part of the definition of a group the existence of an inverse to each element. (Kline, 1972, p. 1140)

There are, of course, many perfectly consistent concepts in the general vicinity of the group concept, but the vast majority of these would not serve the relevant mathematical purposes nearly as well as the concept of group. So a large part of the real mathematical work in this case was isolating precisely the underlying structure that was responsible for so many important features of the particular examples at hand, and codifying that structure in the concept of a group.

What I want to suggest is that at this level of generality, set theoretic axiomatics is involved in a similar process: zeroing in on the best notion of set, that is, the notion best suited to the mathematical goals that notion is intended to serve. I've argued that one of those goals is providing a foundation, and that that goal recommends in favor of one fundamental theory of sets that is as determinate as possible, that is, in favor of the methodological maxim, UNIFY. So the lower level disanalogy between set theory and group theory – the one aims for a single model, the other aims to draw dissimilar structures together – is accompanied by an important analogy at a higher level – both practices choose methods of concept formation that are best suited to achieving their particular goals.

This leads to one last remark about this set theory/group theory analogy. Philosophers are often tempted to say that definitions need no justification, that for example we could have defined the word 'group' in any number of (consistent) ways, that our particular choice is a matter of pure convention. The analogous line of thought in the case of set theory is that we could extend ZFC in any number of (relatively consistent) ways, that our particular choice is a matter of pure convention. What I'm suggesting is that this position is incorrect, that there are constraints, legitimate mathematical constraints on how best to define 'group' and on how best to extend ZFC, and that these constraints are determined by the goals of the particular practices in which those decisions are taking place.

Finally, let's return to the proposed geometric analogy: trying to settle CH is like trying to decide the parallel postulate; some models of set theory satisfy CH, some don't, just as some spaces are Euclidean and some aren't; in either case it's silly to try to select out one of these models or spaces as 'correct'. Once again, I think this purported analogy fails to acknowledge the differing goals of the two practices. A glance at the origins and development of non-Euclidean geometry should make this clear.[7]

Though Euclidean geometry was regarded as the true theory of physical space from the time of Euclid until around 1800, dissatisfaction with the parallel postulate is perhaps just as old. Kline describes the situation this way:

> The axioms adopted by Euclid were supposed to be self-evident truths about physical space. . . . However, the parallel axiom . . . was believed to be somewhat too complicated. No one really doubted its truth and yet it lacked the compelling quality of the other axioms. Apparently even Euclid himself did not like his own version of the parallel axiom because he did not call upon it until he had proved all the theorems he could without it. (Kline, 1972, p. 863)

Under the circumstances, it's only natural that generation after generation of geometers would attempt to prove the parallel postulate from the other axioms. One common method was indirect: consider the alternatives to the parallel postulate – that there is no line through a given point parallel to a given line or that there are many lines through a given point parallel to a given line – and try to rule them out by deriving contradictions from them. Sacchieri employed a version of this strategy and eventually, in 1733, published a book called *Euclid Vindicated from All Faults*. What this book in fact contains is a proof that one of the options does lead to contradiction, while the other leads to conclusions so unacceptable as to rule it out as well.

Thirty years later, Klügel suggested that the parallel postulate is known to be true only by experience, and that the conclusions repugnant to Sacchieri were only contrary to experience; soon Lambert envisioned a range of logically possible geometries that might have little to do with physical reality. By 1818, Schweikart proposed an alternative to Euclid called 'astral geometry', which he thought might be the correct geometry for the stars. Taurinus also pursued astral geometry, and though he considered Euclidean geometry to be true, he insisted that astral geometry was logically consistent. All these mathematicians believed that the parallel postulate could not be proved, that alternative geometries were logically consistent, and Euclidean geometry was the true theory of physical space (with the possible exception of the stars).

The turning point came with Gauss. In the 1790s, Gauss was engaged in the attempt to prove the parallel postulate, but in 1799, he wrote to a fellow mathematician

> . . . the path I have chosen does not lead at all to the goal which we seek [a proof of the parallel postulate]. . . . It seems rather to compel me to doubt the truth of geometry itself. (Quoted in Kline, 1972, p. 872)

This sentiment grew stronger as he developed his non-Euclidean geometry; in 1817, he wrote to another colleague:

> I am becoming more and more convinced that the [physical] necessity of our [Euclidean] geometry cannot be proved. . . . Perhaps in another life we will be able to obtain insight into the nature of space. . . . Until then we must place geometry not in the same class with arithmetic, which is purely *a priori*, but with mechanics. (Op. cit.)

A sometimes-disputed anecdote has it that Gauss went so far as to measure the sum of the angles of a triangle formed by three nearby mountains, to see if it differed from the Euclidean 180°, only to conclude that the disparity fell within the margins of experimental error. It seems beyond dispute that Gauss considered alternative geometries to be candidates for application to the physical world.

This line of thought reached its full flowering when Gauss set the foundations of geometry as the topic for the qualifying lecture of his student, Riemann.[8] In his famous work of 1854 (finally published in 1868), Riemann presents a general formalism for a range of geometries: parameters for determining the local metric, and hence the local curvature, of the space are left open; Euclidean geometry then appears as a special case among an abundance of non-Euclidean possibilities. Which of these best represents physical space is left as a matter for empirical study. Perhaps most striking to the contemporary eye is Riemann's suggestion that the local variations in the metric of physical space might depend on the actual spatial distribution of mass! This idea appears again in Clifford, in 1870, but otherwise rests unexplored till Einstein.

I'll pick up the thread of this remarkable episode in the next section, but for now our interest is in the purported analogy between CH and the parallel postulate. Clearly non-Euclidean geometry became, in the hands of Gauss and Riemann, a device for investigating the possibilities for the mathematical structure of physical space. Given that the goal, quite explicitly in Riemann, was to provide a range of models, leaving the choice between them to natural science, it would be counterproductive for the mathematician to limit that range by coming to a prior decision on the parallel postulate. Once again, as in the case of group theory, the goals of the geometry practice differ so dramatically from those of set theory that it is unreasonable to expect the same methodology to be rational in both cases. While UNIFY makes perfect sense for the set theoretic community, in its efforts to provide a single foundational theory, it would be madness for the geometric community, in its efforts to provide a broad range of models for the scientist.

We have seen that the force of these analogies is an illusion, that we are persuaded by them only so long as we fail to examine the realities of the mathematical practices in question. Our investigation of this question has led us to the doorstep of our original topic: the application of mathematics in natural science. Let us turn, at last, to that problem.

III. Why more is better

Until recently, the relations between mathematics and the natural sciences were understood to be quite simple: the two were hardly distinguished at all. In Kline's words:

> . . . the Greeks, Descartes, Newton, Euler, and many others believed mathematics to be the accurate description of real phenomena . . . they regarded their work as the uncovering of the mathematical design of the universe. (Kline, 1972, p. 1028)

Over the course of the 19th century, this picture changed dramatically:

> ... gradually and unwittingly mathematicians began to introduce concepts that had little or no direct physical meaning. (Ibid., p. 1029)

Citing the rise of negative numbers, complex numbers, *n*-dimensional spaces, and non-commutative algebras, he remarks that 'mathematics was progressing beyond concepts suggested by experience', but that 'mathematicians had yet to grasp that their subject ... was no longer, if it ever had been, a reading of nature' (ibid., p. 1030). By mid-century, the tide had turned:

> ... after about 1850, the view that mathematics can introduce and deal with rather arbitrary concepts and theories that do not have immediate physical interpretation but may nevertheless be useful, as in the case of quaternions, or satisfy a desire for generality,[9] as in the case of n-dimensional geometry, gained acceptance. (Ibid., p. 1031)

This movement continued with such studies as abstract algebra, pathological functions, and transfinite numbers.

The heady new view of mathematics that accompanied this change is perhaps best expressed by Cantor:

> Mathematics is entirely free in its development. ... The essence of mathematics lies in its freedom. (Quoted in Kline, 1972, p. 1031)

This sentiment appears in the thinking of many of the most innovative mathematicians of the late 19th century; today, it is standard orthodoxy. Mathematics progresses by its own lights, independent of ties to the physical world. Legitimate mathematical concepts and theories need have no direct physical interpretation.

This line of development suggests that apart from isolated pockets of applied mathematics, the bulk of modern mathematics has severed all ties with natural science to strike out on its own. This picture of the brave new outlook might seem to fully answer our question about the relations between mathematics and natural science: mathematics grew out of science, but it is now an entirely separate enterprise; its products may occasionally be useful to the natural scientist, but this is not the concern of mathematics itself, which pursues its own ends.

In fact, I think this analysis is far too quick, because it overlooks this simple fact: one among the many goals of modern mathematics as a whole remains the goal of providing tools useful to natural science, theories applicable to the physical world. Before the great shift in the 19th century, nearly all mathematics was done in pursuit of this single goal,[10] and all mathematical concepts and theories were required to enjoy direct physical interpretation. The mathematical landscape is much broader and more varied these days, but this traditional goal has by no means entirely disappeared.

Of course, individual bits of mathematics are pursued for many, many different reasons, some reasonable, some frivolous, and some loosely or closely tied to specific applications or potential applications. What I'm describing here takes place at a much more precarious level of global analysis; I'm claiming that one of the completely general goals of modern mathematics as a whole is to provide concepts and theories useful in natural science. Again, this is one goal among many, but I won't attempt to identify any others here.

What's intriguing for our purposes is that we seem to have uncovered a mismatch. On the one hand, we find a methodological maxim – call it 'FREEDOM' – that mathematics should be pursued as the mathematician sees fit, for mathematical reasons, unconstrained by physical interpretation or application. On the other hand, we find a goal of the practice – call it 'APPLICATION' – which is to produce mathematical models and theories of use to the natural scientist. These would seem a poor fit. Wouldn't APPLICATION be more effectively achieved if that FREEDOM were reined in, if it were tempered with some CONSTRAINT? Perhaps mathematicians should be admonished to concentrate on providing means for particular physical problems, or to direct their attention towards areas of likely application.

Here the recent history of science and mathematics yields a real surprise: it seems the answer to this question is no. To illustrate, we need look no further than the two examples introduced in the previous section: non-Euclidean geometry and group theory. Let me explain.

Following the 1868 publication of Riemann's work of 1854, his approach to non-Euclidean geometry was pursued by several mathematicians into the 1880s. At this point, interest in non-Euclidean geometries began to decline, partly because the subject seemed to have reached its full development, but partly also for a more fundamental reason:

> Another reason for the loss of interest in the non-Euclidean geometries was their seeming lack of relevance to the physical world. It is curious that the first workers in the field, Gauss, Lobatchevsky, and Bolyai, did think that non-Euclidean geometry might prove applicable when further work in astronomy had been

done. But none of the mathematicians who worked in the later period believed that these basic non-Euclidean geometries would be physically significant. . . . In fact, most mathematicians regarded non-Euclidean geometry as a logical curiosity. (Kline, 1972, p. 921)

Kline also reports that interest in Riemannian geometry, in particular, was overshadowed by the dominant pursuit of projective geometry (ibid., pp. 922–923).

So matters stood until 1912, when Einstein moved to Zurich. Up to that point, he had described the statics of gravitation, but the dynamics eluded him. In June and July of that year, he wrote:[11]

> The further development of the theory of gravitation meets with great obstacles. . . . The generalization [of the static case] appears to be very difficult . . . it cannot yet be grasped what form the general space-time equations could have. I would ask all colleagues to apply themselves to this important problem!

The supreme difficulty was that Euclidean geometry would not do.

> If all [accelerated] systems are equivalent, then Euclidean geometry cannot hold in all of them. To throw out geometry and keep [physical] laws is equivalent to describing thoughts without words. We must search for words before we can express thoughts. What must we search for at this point?

Early in 1912, Einstein realized that

> . . . Gauss's theory of surfaces holds the key for unlocking this mystery . . . I realized that the foundations of geometry have physical significance.

But despite these insights, the problem remained so troublesome that Einstein, on arrival in Zurich, told his mathematician friend Marcel Grossman, 'You must help me or else I'll go crazy'.

As history records, Grossman did help. In the words of Abraham Pais, Einstein's co-worker and biographer:

> . . . he told Grossman of his problems and asked him to please go to the library and see if there existed an appropriate geometry to handle such questions. The next day Grossman returned (Einstein told me) and said that there indeed was such a geometry, Riemannian geometry. (Pais, 1982, p. 213)

Einstein himself reports arriving in Zurich

> . . . without being aware at that time of the work of Riemann, Ricci, and Levi-Civita. This [work] was first brought to my attention by my friend Grossman when I posed to him the problem . . . (Pais, 1982, p. 212)

Grossman and Einstein immediately set to work on their collaboration, during which Einstein wrote:

> At present, I occupy myself exclusively with the problem of gravitation and now believe that I shall master all difficulties with the help of a friendly mathematician. (Ibid., p. 216)

This was premature, as the theory of general relativity didn't reach its final form until a few years later, after the collaboration with Grossman, but the introduction of Riemannian geometry was, nevertheless, a fundamental breakthrough.

The moral I'd like to draw from this story is: if Riemann had been constrained by the contemporary wisdom that Euclidean geometry was the physical truth, or that only three dimensional geometry was worth pursuing, or that geometry is *a priori* (and thus independent of the distribution of mass), he would most likely not have developed his non-Euclidean geometry. This suggests that if your old goal is APPLICATION, it is best to allow mathematicians their FREEDOM. Feynman tells the general story like this:

> Mathematicians like to make their reasoning as general as possible. If I say to them, 'I want to talk about ordinary three dimensional space', they say, 'If you have a space of n dimensions, then here are the theorems'. 'But I only want the case 3', 'Well, substitute $n = 3$.'! . . . The physicist is always interested in the special case; he is never interested in the general case. . . . So a certain amount of reducing is necessary, because the mathematicians have prepared these things for a wide range of problems. This is very useful, and later on it always turns out that the poor physicist has to come back and say, 'Excuse me, when you wanted to tell me about four dimensions . . .' (Feynman, 1965, p. 56)

And if Einstein needed four dimensions, modern string theorists talk of at least ten! FREEDOM looks better and better.

Still, Riemann did have application in mind when he developed his geometry, even if he dismissed many of the common assumptions of his contemporaries.[12] Perhaps the pursuit of APPLICATION would be best achieved if mathematicians were required to focus their attention on applications, even if they are not constrained by the prevailing views of what such a focus demands. Alas, history suggests that even this weaker CONSTRAINT would be a bad idea for a subject in pursuit of APPLICATION. Consider this time the theory of groups.

We've seen that group theory arose as a means of unifying a wide variety of mathematical phenomena from different areas, hardly in response to any direct physical application. The physicist Dyson tells this story:

> In 1910 the mathematician Oswald Veblen and the physicist James Jeans were discussing the reform of the mathematical curriculum at Princeton University. 'We may as well cut out group theory,' said Jeans. 'That is a subject which will never be of any use in physics.' It is not recorded whether Veblen disputed Jeans's point, or whether he argued for the retention of group theory on purely mathematical grounds. All we know is that group theory continued to be taught. (Dyson, 1964, p. 249)

Some years later, Wigner took up the problem of the interaction of more than two identical particles. Pais recounts that

> He rapidly mastered the case $n = 3$ (without spin). His methods were rather laborious; for example, he had to solve a (reducible) equation of degree six. It would be pretty awful to go on this way to higher n. So, Wigner told me, he went to consult his friend the mathematician Johnny von Neumann. Johnny thought a few moments then told him that he should read certain papers by Frobenius and by Schur which he promised to bring the next day. As a result Wigner's paper on the case of general n (no spin) was ready soon and was submitted for publication in November 1926. It contains an acknowledgement to von Neumann, and also the following phrase: 'There exists a well-developed mathematical theory which one can use here: the theory of transformation groups which are isomorphic with the symmetric group (the group of permutations)'. (Pais, 1986, pp. 265–266)

Pais concludes:

> Thus did group theory enter quantum mechanics. (Ibid., p. 266)

Dyson echoes the momentus tone of this remark in the epilogue to his account of Jeans's gaffe:

> By an irony of fate group theory later grew into one of the central themes of physics, and it now dominates the thinking of all of us who are struggling to understand the fundamental particles of nature. (Dyson, 1964, p. 249)

In this case, it seems even the most attenuated form of CONSTRAINT would have undermined the successful pursuit of APPLICATION.

I have argued here only by anecdote, but I hope to have generated some sympathy for the conclusion that the most effective way we know of for pursuing the mathematical goal of APPLICATION is to follow the methodological maxim of FREEDOM.[13] If this is correct, we can expect APPLICATION to provide one strand in a complex and varied defense of the overall rationality of FREEDOM. And this would provide the beginnings of our naturalistic analysis of the significance of the application of mathematics for the philosophy, or at least, for the methodology of mathematics. Let me conclude, as promised, by indicating how this conclusion about mathematics in general can be seen to connect back to the practice of set theory in particular.

The connection is quite simple. Our naturalistic analysis of set theory focused on one goal (among many) of that practice, namely to provide a foundation (in a certain limited sense) for classical mathematics. In the typography of this section, let's call this FOUNDATION. We also hold that the methodology of classical mathematics is governed (in part) by the maxim FREEDOM. To prevent conflict here, we will want to pursue the goal FOUNDATION so as not to interfere with the method of FREEDOM, so as not to interpose any constraint. We still want to provide a court of final appeal for existence and proof, a single arena for (surrogates) of all classical mathematical objects, but we want to do so in such a way as to preserve the mathematician's FREEDOM to pursue any structures or theories she chooses. This means that the set theoretic arena should be as generous as possible, should provide the broadest possible range of surrogates, so that the mathematician will remain FREE at every turn. This produces another set theoretic maxim, instructing the set theorist to seek just this sort of generosity, a maxim I call MAXIMIZE.

Now it is a delicate matter to tease out the exact content of MAXIMIZE and to balance its counsel with that of UNIFY, but I discuss these matters elsewhere.[14] My purpose here is simply to point out how this naturalistic analysis sees the application of mathematics as bearing on the practice of set theory: the application of mathematics is reflected in the mathematical goal APPLICATION; APPLICATION is one support for the mathematical maxim FREEDOM; FREEDOM plus the set theoretic goal FOUNDATION yields the set theoretic maxim MAXIMIZE. From this point of view, there is a real connection between the application of mathematics and the practice of set theory, but it is indirect and far more attenuated than the Quinean would have it. I leave the reader to ponder which view seems closer to the truth.[15]

Notes

[1] This description leaves out a host of difficult questions: e.g., how are the mathematical considerations to be separated from the philosophical or sociological? How are the goals of a particular mathematical practice to be determined? These and related issues are discussed in some detail in my (1997), §III.4, so I won't go into them here.

[2] See my (1997), §§III.5 and III.6.

[3] Of course, this is not to say that various alternative set theories should not be pursued – they should – but that we prefer to view these as theories of non-standard models studied within the single, fundamental theory of sets.

[4] A measurable cardinal κ is κ-supercompact.

[5] UNIFY, as we've seen, is at work here in the goal of a single theory of sets, so that a choice must be made between Large Cardinals and alternatives like V = L. A naturalistic argument that Large Cardinals are the better choice would probably begin from the maxim MAXIMIZE, discussed in §III, below.

[6] In (1997), I discuss the heuristic role of extra-mathematical philosophical positions; what I'm now suggesting is that some intra-mathematical principles also function as heuristics rather than as justifications.

[7] The following three paragraphs follow Kline (1972), chapter 36.

[8] For this paragraph, see Kline (1972), chapter 37, §3.

[9] As becomes clear in note 13, Kline is *not* suggesting here that – contrary to the claims of the previous section – generalizations are justified by their status as generalizations alone. He is well aware of the sound mathematical reasons for pursuing n-dimensional geometry.

[10] The more esoteric reaches of number theory would seem to be an exception, but even if they were not pursued in the interests of applications, the concept 'number' on which it focused has an especially direct physical interpretation.

[11] This paragraph and the next follow Pais (1982), chapter 12. The quotations in this paragraph all come from pp. 211–212.

[12] For completeness, it should be noted that the mathematicians who developed his ideas after Riemann were not concerned with applications.

[13] I should mention two other works of Kline (1968 and 1980, chapter XIII) in which he argues strenuously (and I think correctly) against those who would describe these cases as showing that 'arbitrarily chosen themes . . . proved subsequently to be just what the scientist needed' (1968, p. 232). (Non-Euclidean geometry and group theory are hardly 'arbitrarily chosen themes'!) Here Kline also notes, in agreement with the line of thought in the text, that it would be dangerous to try to legislate in advance what counts as non-arbitrary: 'The great mathematicians of the past often transcend the immediate problems of science' (1980, p. 279). Despite the tone of some of his remarks, I think what Kline opposes is not FREEDOM or its connection with APPLICATION, but what I call 'glib formalism' and related 'glib' views (1997, e.g., p. 202), all of which hold that any consistent system is as good as any other. (I am grateful to Don Fallis for this way of putting the point.) It takes sound mathematical judgment to tell the worthwhile projects from the rest; in the words of Felix Klein, quoted approvingly by Kline, 'whoever has the privilege of freedom should also bear responsibility' (1972, p. 1037).

[14] See my (1997), III.5 and III.6 for some efforts on this project.

[15] Thanks to Don Fallis for helpful comments on an earlier draft.

References

Dyson, Freeman: 1964, 'Mathematics in the Physical Sciences', in M. Kline (Ed.), *Mathematics in the Modern World*, San Francisco: W. H. Freeman and Co., pp. 248–257.

Feynman, Richard: 1965, *The Character of Physical Law*, Cambridge, MA: MIT Press.

Gödel, Kurt: 1964, 'What is Cantor's Continuum Problem?', reprinted in his *Collected Works*, vol. II, S. Feferman et al., eds., New York: Oxford University Press, 1990, pp. 254–270.

Kanamori, Akihiro: 1994, *The Higher Infinite*, Berlin: Springer Verlag.

Kline, Morris: 1968, 'The Import of Mathematics', in M. Kline (Ed.), *Mathematics in the Modern World*, San Francisco: W. H. Freeman and Co.

Kline, Morris: 1972, *Mathematical Thought from Ancient to Modern Times*, New York: Oxford University Press.

Kline, Morris: 1980, *Mathematics: the Loss of Certainty*, New York: Oxford University Press.

Maddy, Penelope: 1997, *Naturalism in Mathematics*, Oxford: Oxford University Press.

Pais, Abraham: 1982, *'Subtle is the Lord': The Science and Life of Albert Einstein*, Oxford: Oxford University Press.

Pais, Abraham: 1986, *Inward Bound: of Matter and Forces in the Physical World*, Oxford: Oxford University Press.

Department of Logic & Philosophy of Science
University of California
Irvine, CA 92697-5100
U.S.A.