Logique & Analyse **179–180** (2002), 269–281

MAXIMIZING PRINCIPLES AND MATHEMATICAL METHODOLOGY

ALAN BAKER

(A) *Introduction*

Philosophers interested in mathematical practice tend to emphasize the allegedly distinctive features of mathematics. Consider, for example, the mathematical naturalism put forward by Penelope Maddy. Maddy — following Quine — stipulates that to adopt naturalism toward a body of practice, X, is to repudiate the quest for external philosophical legitimization, or undermining, of X's methods. It entails adopting a respectful attitude to X, in the sense that X's methodology is deemed to be autonomous and self-supporting.

> "[T]he mathematical naturalist [asserts] that mathematics is not answerable to any extra-mathematical tribunal and not in need of any justification beyond proof and the axiomatic method."[1]

Thus mathematical naturalism involves a certain sort of respect for the methodology of mathematics. However, Maddy also makes explicit appeal to methodological considerations in order to *support* mathematical naturalism. Her starting point is the scientific naturalism of W.V.O. Quine. Maddy is sympathetic to Quine's stance as a way of approaching science, but she does not think that it can be extended to cover mathematics. The key problem, for Maddy, is that Quinean scientific naturalism insists on "identifying the proper methods of mathematics with the methods of science" and this distorts actual mathematical practice.[2] The reason why is that the methods of science and of mathematics are crucially distinct.

> "[I]t quickly becomes obvious that mathematics is central to our scientific study of the world and that the methods of mathematics differ markedly from those of natural science."[3]

[1] Maddy [1997, p. 184]. Later in the same passage, Maddy concludes that mathematics is "independent of both first philosophy and natural science."

[2] *op. cit.*, p. 184.

[3] *op. cit.*, p. 183.

How precisely do they differ? Maddy highlights an alleged contrast between the respective attitudes of the two disciplines towards ontology and ontological posits. She argues, in particular, that set theorists utilize

> "*maximizing principles of a sort quite unlike anything that turns up in natural science*: crudely, the scientist posits only those entities without which she cannot account for our observations, while the set theorist posits as many entities as she can, short of inconsistency."[4]

There are certainly places in Quine's writings where he recommends that mathematical decisions be made on grounds of parsimony. For example, he recommends that set theorists adopt the axiom $V = L$ since this provides an ample ontological underpinning for all conceivable applications of mathematics to science.[5] Maddy describes this as a stance which is "precisely opposite to that of the set theoretic community."[6]

My project in this paper is to investigate whether appeal to ontological maximization principles does in fact mark a clear methodological divide between mathematics and science, as Maddy claims. Mathematical naturalism rightly encourages philosophers to pay close attention to the details of mathematical practice. However, the philosophical *motivation* for this view turns out to be rooted in a caricature of mathematical practice which exaggerates the differences between mathematical and scientific methodology. I shall argue that once we play closer attention to the diversity of actual mathematical and scientific practice, it becomes clear that ontological maximization does not delineate a sharp methodological divide between the two. Part of the confusion here stems from the fact that philosophers have tended to focus their attention on pure, foundational mathematics on the one hand, and on experimental, non-foundational science on the other. My thesis is that — in the case of ontological maximizing principles — the crucial distinction is not between mathematics and science, but between the foundational and non-foundational areas of each practice. If I am right then this alleged methodological discontinuity between mathematics and science cannot legitimately be used to motivate Maddy's mathematical naturalism.

[4] *op. cit.*, p. 131, (my italics).

[5] "Considerations of simplicity, economy, and naturalness ... support Gödel's axiom of constructibility, $V = L$." (Quine [1990, p. 95]) Here V is the universe of sets and L is the constructible sets.

[6] Maddy [1997, p. 106].

(B) *MAXIMIZE in Set Theory*

Maddy's mathematical naturalism ought more properly to be labeled 'set theoretic naturalism' — indeed Maddy herself calls it this at various points in her writings. She explicitly restricts her attention to set theory (as is apparent from the third quote from section (A) above), and it is from this particular — albeit important — branch of mathematics that she draws her picture of mathematical methodology.[7] Within set theory, Maddy identifies a methodological principle which she calls MAXIMIZE. I want to start by trying to get clearer about what this principle amounts to.[8]

Maddy's initial characterization of MAXIMIZE, in her book *Mathematical Naturalism*, is as follows:

> "[T]he set theorist posits as many entities as she can, short of inconsistency."[9]

As it stands, this is clearly too crude. First off, the set-theorist (at least one pursuing pure set theory) only posits *sets*.[10] Set theory is not strengthened in the eyes of mathematicians by adding an axiom which asserts the existence of electrons, or of unicorns. Second, it is unclear just what 'as many as we can' means in radically transfinite contexts, contexts which are of course characteristic of the higher reaches of set theory.[11]

Later in her book, Maddy makes some more detailed remarks. She characterizes MAXIMIZE as a principle which states that the set-theoretic arena should be "as generous as possible," and that the set theoretic axioms should be "as powerful and fruitful as possible."[12] It is worth noting that these two goals do not always pull in the same direction. Mathematical axioms can expand ontology by making existence claims, but they can also contract ontology by imposing more stringent conditions for admission into the domain of a given theory. Restrictive axioms may turn out to be very fruitful, and

[7] In her [1997, p. 210], Maddy explicitly acknowledges this restriction to set theory, based on her own expertise, but sees no reason why a similar naturalistic account should not be extended to other branches of mathematics.

[8] cf. Maddy [2001, p. 26]; "[I]t is a delicate matter to tease out the exact content of MAXIMIZE."

[9] *op. cit.*, p. 131.

[10] And even impure set theory is indifferent to the (internal) properties of individual atoms.

[11] A third issue concerns what should be done when there are several, mutually inconsistent, equally maximal alternatives.

[12] *op. cit.*, p. 210.

powerful, despite (or maybe precisely *because*) they limit the set-theoretic arena. Another question concerns how to makes sense of the talk of possibility which is present in each of the above characterizations. Maddy's initial talk of inconsistency suggests that it is *logical* possibility which is operative here.

Maddy's final remark in her book concerning the nature of MAXIMIZE is the following claim:

> "One way in which [set theory] should MAXIMIZE is in the range of available isomorphism types."[13]

Maddy spends some time in the final chapter of the book exploring various subtleties in the application of MAXIMIZE, interpreted in the above sense, to the issue of whether to adopt $V = L$. Part of the difficulty is that there is no 'neutral' court of appeal from which to compare the range of isomorphism types of two alternative set theories.

Rather than press for a further sharpening of the precise content of MAXIMIZE, a project that has been pursued to some extent in a recent paper by Löwe,[14] I shall confine myself here to three more general remarks. First, MAXIMIZE is not a principle of 'anything goes' or 'postulate what you want.' In fact, as a principle, MAXIMIZE is quite restrictive. If there is a unique way to fully expand a theory's ontology, then MAXIMIZE gives the mathematician no choice but to pursue it. Second, MAXIMIZE is a tool for comparing theories, not a way of determining what to include in the ontology of a given theory. There are two separate ontological questions here, which need to be distinguished. On the one hand, what *theories* should we accept? On the other hand, what *entities* should we postulate, given our accepted theories? This latter question is generally more straightforward in mathematics than in science. What exists according to a given mathematical theory is what follows from its axioms. In the case of science the issue is more complicated; what exists according to a scientific theory depends on the interaction between that theory and empirical observations. Third, MAXIMIZE does not appear to be an overriding maxim of theory-choice (unlike, for example, CONSISTENCY). Rather it must be weighed against other methodological maxims, whose nature and strength may vary according to context.[15]

[13] *op. cit.*, p. 211.

[14] Löwe [2002].

[15] How MAXIMIZE might be weighed against other theoretical considerations is discussed by Maddy [1997, final chapter] and by Löwe [2002].

Putting together these three points, I propose the following rough characterization of MAXIMIZE:

> Other things being equal, if set theoretic axioms S_1 have a greater range of isomorphism types than set theoretic axioms S_2 then we ought to prefer S_1 over S_2.

In the context of set theory, then, the aim of this principle is to extend the range of 'large' cardinals upwards as far as possible. More generally, MAXIMIZE is an example of a type of principle of plenitude that I shall refer to as *supersizing plenitude*.

(C) *MAXIMIZE in Other Branches of Mathematics*

Let us grant, then, with Maddy that MAXIMIZE is indeed a methodological principle of set theory. Does MAXIMIZE (or something like it) play a methodological role in any branches of mathematics other than set theory? Once we move away from the foundational orientation of set theory and into other areas of pure mathematics, there does seem to be more of an openness to simultaneously pursuing various alternative, possibly mutually inconsistent, theories. Thus some group theorists study commutative groups, while others study non-commutative groups. This is closer to 'anything goes,' though not completely so since there is certainly room to criticize particular choices of theory as intractable, as trivial, or as mathematically uninteresting in some other way. However, there is no pressure to choose between these various alternative theories, no reason for example to fix either on commutative or on non-commutative groups as the 'proper' subject of group theory. Also, even when reasons are given for preferring one line of research over another, MAXIMIZE can cut both ways. Extending the ontology of a pure mathematical theory does not always make it more fruitful, valuable, or interesting.

When we move to areas of mathematics that are more closely tied to applications, a rather different scenario emerges. On the one hand, there is a desire not to use any mathematical theory that is unnecessarily strong for the task at hand. If a more powerful theory also brings with it greater complexity, then a weaker (but still adequate) theory will normally be preferred. For example, in order to measure lengths we need not just rational but real numbers, but there is no need — and hence no reason — to use complex numbers. On the other hand, there seems to be no 'cost' *per se* associated with extra mathematical ontology, even if it is not strictly necessary. Thus if expanding the domain of mathematical entities makes a theory simpler to apply, or unifies two otherwise distinct areas of science, then there is no reason not to prefer the expanded theory over its competitors in the context of applications.[16]

In short, it seems clear that the role of MAXIMIZE in many areas of mathematics is marginal at best. While there may be no cost associated with extra ontology in mathematics, as perhaps there is in empirical science, nor is there any significant value attached to extra mathematical ontology in and of itself. Assuming that this initial assessment is correct, relying solely on MAXIMIZE does not suffice to ground a mathematical naturalism that extends beyond set theory and related foundational areas.

(D) *Parsimony and Plenitude in Empirical Science*

Even if a broad-based mathematical naturalism is untenable, it may still be possible to hang on to a set theoretic naturalism based on appeal to MAXIMIZE. However, this will only work if maximizing principles do not operate in any parts of empirical science. The standard view, as expressed in the previous remarks by Quine and Maddy, is that it is minimizing principles — not maximizing principles — which have force in science. In other words, scientific methodology utilizes principles of parsimony but not principles of plenitude.

I agree with the first part of this claim but not the second. I have argued elsewhere, based on case studies from experimental physics, biology, and geology, that parsimony is a theoretical virtue in many areas of science, and that Occam's Razor and other related 'principles of parsimony' are utilized both explicitly and implicitly by working scientists.[17] However, once we move to the more foundational areas of theoretical physics, the role of parsimony is considerably less clear-cut. In particular, the rise of particle physics and quantum mechanics in the 20th Century has seen various principles of plenitude being appealed to as an essential part of the theoretical framework.

A particularly clear-cut illustration of an appeal to a principle of plenitude in modern physics is provided by the case of magnetic monopoles. Maxwell's theory of electromagnetism, developed in the 19th Century, postulates numerous analogies between electric charge and magnetic charge. One experimental difference, however, is that magnetic charges always come

[16] Colyvan [2002] gives an example where the move from real to complex numbers allows a unified approach to solving differential equations of the form $y' - y'' = 0$, and $y' + y'' = 0$. For a more detailed discussion of domain extension and its potential benefits within mathematics see Manders [1989].

[17] See Baker [1999, Chapter 4].

in oppositely-charged pairs, called "dipoles,"[18] whereas single electric charges, or "monopoles," can exist in isolation. Physicists began to wonder whether there was some theoretical reason why *magnetic* monopoles could not exist. In the early decades of this century it was thought that the newly developed theories of quantum mechanics ruled out the possibility of monopoles, and this is why none had ever been detected. However, in 1931 the physicist Paul Dirac showed that the existence of monopoles is consistent with quantum mechanics, although it is not required by it. Despite the inconclusive nature of this theoretical result, Dirac went on to assert the existence of monopoles, arguing that their existence is not ruled out by theory and that "under these circumstances one would be surprised if Nature had made no use of it."[19] This appeal to plenitude was widely — though not universally — accepted by other physicists.

> One of the elementary rules of nature is that, in the absence of laws prohibiting an event or phenomenon it is bound to occur with some degree of probability. To put it simply and crudely: anything that *can* happen *does* happen. Hence physicists must assume that the magnetic monopole exists unless they can find a law barring its existence.[20]

A more recent example of an appeal to a principle of plenitude involved the postulation of tachyons, which are 'superluminal' particles that move faster than light. Much as the existence of magnetic monopoles was originally thought to be ruled out by quantum mechanics, the existence of tachyons was thought to be ruled out by special relativity. Then, in 1969, Bilaniuk and Sudarshan showed that tachyons could be consistently described within the framework of special relativity. Having proved this result, the authors then argued as follows:

> If [tachyons'] existence would not lead to any contradictions [with laws of physics] one should be looking for them. There is an unwritten precept in modern physics, often facetiously referred to as Gell-Mann's totalitarian principle which states that in physics anything which is not prohibited is compulsory.[21]

[18] As in the North and South poles of a bar magnet.

[19] Dirac [1930, p. 71, note 5].

[20] Ford [1963, p. 122].

[21] Bilaniuk and Sudarshan [1969, p. 44].

These examples indicate that principles of plenitude do play some role in scientific methodology and hence that such principles are not the exclusive domain of mathematics.

(E) *Foundations*

My central thesis is that ontologically expansive principles, such as MAXIMIZE and plenitude, are more crucially linked to foundational role than to mathematics.[22] This thesis is tentative, and awaits in particular more data concerning mathematical practice. My present goal is more modest, and divides into two distinct parts. First, I shall defend the empirical claim *that* this link between plenitude and foundational role does occur. Second, I shall offer some philosophical reasons for *why* such a link might be expected.

In defense of the empirical claim, I appeal in the first instance to the historical record. Plenitude-style principles have tended to arise in areas that have been foundational in some sense, and to disappear from such areas if and when their foundational role ceases. Thus principles of plenitude appeared during the late medieval and early modern period in theology and in metaphysics, at a time when these areas were considered to constitute the basis of all rational inquiry. Perhaps the best-known version is associated with Leibniz, according to whom God created the best of all possible worlds with the greatest number of possible entities. As the natural sciences gradually took over this foundational role from theology and metaphysics in the 17th and 18th Centuries, principles of plenitude began to pop up in science, first in biology and later in chemistry. For example, plenitude was appealed to in 18th-century biology to argue for the existence of fantastical creatures such as mermaids. In fact there were two sorts of plenitude-based arguments that were offered. First, since the concept of mermaid was neither self-contradictory nor in conflict with the known laws of biology, some thinkers argued for their existence based on what might be termed *modal plenitude*.[23] Robinet, for example, writes:

[22] It should be noted that on p. 211 of her [1997], Maddy agrees that MAXIMIZE in set theory follows fairly directly from the intended foundational role of set theory for mathematics as a whole.

[23] For a comparison with Dirac's defense of magnetic monopoles, see Kragh [1981, p. 149]. Kragh's criticism is slightly unfair to Dirac since the 18th-century biologists were making a very restricted existence claim, that mermaids exist now somewhere on earth, whereas Dirac's claim is unrestricted.

> "I have formed so vast an idea of the work of the Creator that from the fact that a thing can exist I infer readily enough that it does exist."[24]

Other thinkers, most prominently among them John Locke, defended the existence of mermaids based on *gap-filling plenitude*.[25] Thus Locke writes:

> "In all the visible corporeal world we see no chasms or gaps. ... Amphibious animals link the terrestrial and aquatic together; ... not to mention what is confidently reported of mermaids or sea-men."[26]

Appeals to gap-filling plenitude surface later, in the 19th Century, in the context of chemistry to argue for the existence of specific hitherto undiscovered elements based on filling in gaps in the newly conceptualized periodic table. Finally, as we saw in the previous section, 20th-century physicists have used principles of plenitude — especially modal plenitude — in drawing ontological consequences from theories such as general relativity and quantum mechanics.

This progression in the areas of application of plenitude principles, from biology through chemistry and then to physics, can be made sense of from the perspective of foundations. As long as biology and chemistry were assumed to provide their own internal foundations, plenitude played a methodological role. But once biologists started looking to chemistry to reveal the building-blocks of biological systems (i.e. biochemistry), and once chemists similarly turned to physics to reveal the building blocks of which the chemical elements are made (i.e. atomic physics), principles of plenitude in these areas fell by the wayside. Thus there does seem to be a sense in which plenitude has historically traced a path through the sciences which tracks the shifting search for foundations. Meanwhile we also find plenitude-style principles in the foundational areas of mathematics, in particular set theory and category theory, but not in other non-foundational mathematical theories.

So much for the empirical claim, that plenitude-style principles are closely linked to foundational theories. What is the *explanation* for this link? It should be noted that we have encountered three (potentially distinct) types of plenitude principle in our discussion thus far — gap-filling plenitude, modal plenitude, and supersizing plenitude — and that their respective links

[24] J.B. Robinet, quoted in Lovejoy [1957, p. 272]. Note that the theological origins of plenitude are quite apparent here!

[25] A mathematical example of gap-filling plenitude would be the extension of the rational numbers to the real numbers.

[26] Locke [1959, pp. 67–8]. Analogous arguments for the existence of angels, based on filling in the gap between mankind and God, were also popular in the 18th Century. (cf. Lovejoy [1957, pp. 189–190]).

to foundations may vary. I will suggest several factors which contribute, in varying degrees, to this link between plenitude principles and foundational goals.

First, foundational theories are further removed from direct observation and application. There are typically more intermediate levels between theory and data, thus insulating the theory from the details and messiness of the data. This allows 'aesthetic' principles of theory choice to hold more sway, with emphasis placed on theoretical virtues such as elegance, simplicity, symmetry, and uniformity. Principles of plenitude seem harder to maintain in situations more closely linked to observation, since the question 'If all these things exist, why can't we *observe* them?' becomes more pressing and harder to deflect. One might wonder to what extent our three types of plenitude principle can be viewed as aesthetic principles on a par with principles of simplicity and elegance. The case seems easiest to make for gap-filling plenitude, since there is a clear link here with the aesthetic virtue of symmetry. Modal plenitude might be viewed in terms of uniformity; it is a principle which enjoins theorists to make existence claims across all possible entities. What about supersizing plenitude? The extent to which this is an aesthetic principle of theory choice is trickier to gauge. Perhaps it can be thought of as an 'anti-arbitrariness' principle, counseling against placing *ad hoc* limits on the size of the domain of what there is.

A second feature of their remoteness from observation is that foundational theories tend to be more general in scope. Principles of plenitude make more sense in such contexts since it is less implausible to postulate the existence of entities which we have not actually observed. Conversely, plenitude principles whose scope is restricted either spatially or temporally seem correspondingly less compelling. We may join with physicists in postulating the existence of tachyons, given their consistency with the known laws of physics. But this in itself gives us no reason to postulate the existence of tachyons in this room, at this moment in time. Generality links up naturally with the anti-arbitrariness aspect of supersizing plenitude mentioned above. It also makes sense in terms of modal plenitude; the more general the scope of the existence claim, the more time and space there is for each possible outcome to be realized.

A third feature of foundational theories is that they often function as a storehouse of frameworks and techniques which can be brought to bear (maybe indirectly) on specific problems as and when they are needed. In this context, a principle such as MAXIMIZE can be seen as ensuring that the foundational theory has available all potential resources. We cannot predict just what will turn out to be applicable, or where it will be applied, so we should avoid at all costs placing arbitrary limits on the range of our foundational theories.

But given the above remarks, why should we adopt MAXIMIZE — or some such principle — rather than taking the 'anything goes' approach? Maddy herself gives a plausible answer to this question, which connects up to a third feature of foundational theories, which is that they often aim to *unify* the various more specific theories they underpin.[27] Encouraging an 'anything goes' approach at the level of foundations is liable to produce a fragmented collection of theories which lacks unity, indeed maybe *cannot* be unified because different theories contradict one another.

(F) *Conclusions*

When discussing methodology, most philosophers of science have focused on experimental and non-foundational science, while most philosophers of mathematics have focused on pure and foundational mathematics.[28] This has fostered the illusion of a sharp methodological distinction between these two disciplines. What I hope to have shown is that this methodological contrast may largely be a product of the foundational / non-foundational divide rather than the mathematics / science divide.

What can we learn from shifting our focus to the foundational / non-foundational divide? It is ironic that the term 'pragmatic' can be applied both at the foundational and at the non-foundational ends of the spectrum, but with quite different meanings. At the foundational end, more attention tends to be paid to 'pragmatic' features of theories — in the Quinean sense of non-empirical features, such as simplicity, fruitfulness, and the like.[29] At the non-foundational end, and particularly once scientists are dealing with applications, considerations of a rather different sort come into play which are 'pragmatic' in the sense of being linked to immediate practicality. Here the focus is on what works, even if what works turns out to be ugly or non-rigorous.

Mark Steiner has described the methodology of the mathematician as "closer to that of the artist than the explorer."[30] I would argue that this metaphor more aptly describes the foundationalist than the mathematician. Thus the

[27] See Maddy [2001] for more details on this point.

[28] Note that philosophers of physics have paid attention to quantum mechanics, but they have mostly focused on interpretation of the formalism rather than on methodological issues.

[29] Recall that, for Quine, *all* features that systematically influence scientific theory choice are to some degree pragmatic, and are no less rational for being so.

[30] Steiner [1998, p. 154].

methodology of the theoretical physicist may be closer to that of the artist, and that of the applied mathematician closer to that of the explorer. The foundationalist emphasis on rigor and lack of patience for 'fudges' and 'rules of thumb' arises out of a concern for global, long-term ends over local, short-term ones. One of Maddy's chief criticisms of the Quinean indispensability argument is that a mathematics driven by considerations of applicability would be stifling. There is something right about this, but I would argue that it is foundational theories in general — and not only mathematical foundational theories — which would suffer in this way. A fundamental physics driven by short-term, foreseeable applications would likely be just as stifled.

The theme of this volume is mathematical practice. The thesis I have defended in this paper is that paying attention only to selected areas of practice can be just as philosophically misleading as paying no attention to practice at all. The tendency for philosophers to focus methodologically on pure, foundational mathematics on the one hand, and on experimental, non-foundational science on the other, has encouraged philosophical views — notably mathematical naturalism — which take the methodologies of mathematics and of science to be quite distinct. I have argued that — in the case of maximization — the real distinction is not between maximizing mathematics and minimizing science, but between the applied and foundational areas of each practice. Note that I am *not* claiming that there are no methodological differences between mathematics and science. It is just that the presence or absence of plenitude principles such as MAXIMIZE does not yield an adequate criterion for distinguishing the two areas of practice.

Department of Philosophy
Xavier University, Cincinnati OH
E-mail: Bakerar@xavier.edu

BIBLIOGRAPHY

Baker, A. (1999) 'Indispensability and the Existence of Mathematical Objects,' Princeton University Ph.D. Thesis [unpublished].

Bilaniuk, O-M & Sudarshan, E. (1969) 'Particles Beyond the Light Barrier,' *Physics Today*, 22, 43–52.

Colyvan, M. (2002) 'Mathematics and Aesthetic Considerations in Science,' *Mind*, 111, 69–78.

Dirac, P. (1930) 'The Proton,' *Nature*, 126, 606.

Ford, K. (1963) 'Magnetic Monopoles,' *Scientific American*, 209, 122–131.

Kragh, H. (1981) 'The Concept of the Monopole: a Historical and Analytic Case Study,' *Studies in History and Philosophy of Science*, 12(2), 141–172.

Locke, J. (1959) *Essay Concerning Human Understanding*, *Volume II*, New York: Dover.

Lovejoy, A. (1957) *The Great Chain of Being*, Cambridge, Mass.: Harvard University Press.

Löwe, B. (2002) 'A First Glance at Non-Restrictiveness,' *Philosophia Mathematica*, 3(10), 347–354.

Maddy, P. (1997) *Mathematical Naturalism*, Oxford: Clarendon Press.

Maddy, P. (2001) 'Some Naturalistic Reflections on Set Theoretic Method,' *Topoi*, 20, 17–27.

Manders, K. (1989) 'Domain Extension and the Philosophy of Mathematics,' *Journal of Philosophy*, 553–562.

Quine, W. (1990) *Pursuit of Truth*, Cambridge, Mass.: Harvard University Press.

Steiner, M. (1998) *The Applicability of Mathematics as a Philosophical Problem*, Cambridge, Mass.: Harvard University Press.