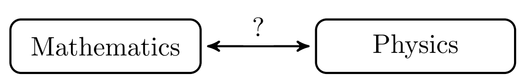
Mathematical Metaphysics

Clark Glymour Luke Serafin

April 30, 2015

0 Introduction

The fundamental question of metaphysics is what exists, not in any particular structure, but in general. To answer this question requires determination of the nature of existence, or more concretely, what it means for something to exist. Thus a worthwhile metaphysics should provide an explicit criterion for existence. A wide variety of such criteria have been proposed, and these can be divided into broad categories based on how they handle abstracta, in particular mathematical objects. Platonistic metaphysical accounts incorporate physical objects and mathematical objects as disjoint categories, but require an account of how these two categories of objects interact, which is a vexing philosophical question [6]. One way to handle this issue is to eliminate one of these two categories, and this is precisely what is done in nominalistic accounts, which admit the existence of physical objects but not of mathematical objects.



Platonism (tradition of Plato)

Nominalism (tradition of Aristotle)

Mathematical metaphysics (possibly tradition of Pythagoras)

A third option immediately presents itself, which is a metaphysical account that admits the existence of mathematical objects but not of physical objects. Because it is intuitively obvious that physical objects exist, this appears absurd, and so it should not be surprising that few philosophers have considered it. The absurdity is only prima facie, however, because just as mathematical objects can be viewed as mental constructions, etc. on a nominalistic account, so on a purely mathematical account physical objects can be seen as identical with certain mathematical objects (which from a Platonistic viewpoint could be thought of as perfect mathematical models of these objects.) It is possibly reasonable to interpret the Pythagorean metaphysical position as one in which all objects, including physical ones, are purely mathematical [2], and more recently such thinkers as Tegmark [5] have noted virtues of this position and even advocated it. However, such advocacy is relegated to a tiny minority, and it is the author's intention to show that mathematical metaphysics is indeed tenable from a philosophical perspective, and in fact provides novel solutions to a variety of perennial philosophical problems.

As the reader surely recognizes, the position taken up in this paper is very radical, and a few words on why its pursuit is worthwhile may be in order. First, many problems remain with the platonistic and nominalistic metaphysical accounts despite centuries of intense study. Mathematical metaphysics provides novel solutions to a number of these problems, which suggests it should be given at least some consideration alongside platonism and nominalism. Even if a philosopher's goal is to defend either platonist or nominalist metaphysics (or other positions regarding mathematical objects), knowledge of mathematical metaphysics and its means of solution of platonist and nominalist problems can certainly be valuable for inspiring means by which to approach these problems from less radical (or other radical) perspectives. In addition, mathematical metaphysics can be seen as an extreme form of some more common traditions (for example ontic structural realism [1]), and consequently promises to shed light on these positions, perhaps through use as a toy metaphysics. However, the author does not wish mathematical metaphysics to be relegated to the status of a toy, and will treat the position as a serious alternative to platonism and nominalism.

The structure of this paper is as follows. Section 1 states the position of mathematical metaphysics in detail. Section 2 examines a litany of objections to this position, and answers them. Section 3 examines how a proponent of mathematical metaphysics might solve various perennial philosophical problems. Section 4 describes possible consequences of accepting the position of mathematical metaphysics for a human individual or society.

1 What Mathematical Metaphysics Is

Quite simply, the position of mathematical metaphysics is that an object exists if and only if it is an element of some mathematical structure. To be is to be a mathematical object. The terms employed in this characterization appear unproblematic to a working mathematician, but to the philosopher of mathematics it is clear that much work remains before this position has been described precisely. First, let us examine the notion of a mathematical structure. The most common precise explication of this notion is set-theoretic: A relational structure by definition consists of a tuple of relations, some of which are unary and describe domains of various sorts of individuals (most commonly just one sort), and the others of which are relations of sorts constructed from the sorts of individuals. This explication of the notion of structure is adequate for most mathematical purposes, but it is easy to argue that it is both too specific and too broad. The precise definition is too specific because it rules out some intuitively appealing structural descriptions, in particular the characterization of a simple undirected graph as a pair (V, E) where V is a set (of vertices) and E is a collection of unordered pairs of elements of V (the edges of the graph). The problem here is that E need not be a relation, in the standard set-theoretic sense that a relation is a set of (Kuratowski) ordered pairs. The common definition of structure is also too broad because it admits multiple distinct characterizations of what should intuitively be the same structure, for example the two tuples $(G, *, e, {}^{-1})$ and $(G, *)$ both representing (what is intuitively) the same group.

The former problem can be ignored given a sufficiently strong set theory, because any collection of sets $\mathcal{A}$ can be coded by a relational structure which is interdefinable with $\mathcal{A}$. The latter can perhaps be solved by regarding two structures as equivalent if both can be extended to the same structure via addition of definable relations (this can be checked by adding all definable relations to both structures and verifying the results agree), and regarding rearrangements of a tuple and relabellings of individuals as representing the same structure (we do not employ unordered collections of relations so as to not trivialize duality relations like $(A, \vee, \wedge)$ as the dual of $(A, \wedge, \vee)$ in lattice theory). Since what is definable depends strongly on the language used for definitions and its semantics, this obviates the importance of specifying both the language of the structure and its semantics. For example, in general far more is definable with the infinitary logic $L_{\omega_1\omega}$ than with ordinary first-order logic, and definitions employing the standard Kripke semantics for intuitionistic predicate logic have quite different content from definitions utilizing the canonical semantics for first-order logic. Thus to fully specify a structure we must specify a tuple of relations, a language built from these relations and logical symbols, and a semantics for that language. Again there is a problem because different languages can be equivalent (e.g. propositional logic with connectives $\{\neg, \wedge\}$ vs. propositional logic with connectives $\{\neg, \vee\}$), and different means of defining semantics can be equivalent (e.g. differences in how models are defined, for example classical propositional models may be taken to be simply sets (of true atomic propositions), or two-valued truth functions defined for atomic propositions). This problem is solved as usual by regarding structures with definitionally equivalent languages or semantics to be the same. Thus we see that a reasonable definition of a structure is a tuple of relations together with a language built from those relations and logical symbols plus a semantics for that language, with all of this taken modulo definitional equivalence (of both relations and the logical language), relabelling of individuals and reordering of relations in the tuple, and equivalence of semantics. Having proposed this definition for the term 'structure,' we shall refer to a simple tuple of relations as a relational structure, and to a relational structure representing a structure just in case it is an element of the equivalence class which is the structure. For readers concerned that the equivalence classes which we are here calling structures can be too big to be sets, observe that Scott's trick can be employed.

A problem remains, because in order to make this definition of structure precise, we must be working in some background theory (a metatheory in logical parlance). How are we to define this metatheory precisely, and its language, semantics, models, etc.? The most reasonable way to do this appears to be embedding the metatheory in a still larger theory (call it a metametatheory, or 2-metatheory), but clearly this only pushes the problem back a step. To truly solve the problem we need to keep pushing: Have an n-metatheory for each natural number n, an ω-metatheory serving as metatheory for all of n-metatheories, an $(\omega + 1)$-metatheory serving as metatheory for the ω-metatheory, and indeed an α-metatheory for each ordinal α. In fact, any model of set theory has bounded ordinals in some larger model (regard the universe of such a model as a set and keep building the cumulative hierarchy), so we can continue this process even further, etc. Thus we see that this hierarchy of metatheories has no end in a very strong sense. However, everything receives a precise meaning from each higher level, so such a scheme nicely bootstraps the notion of meaning. Traditionally this position would be rejected because it leads to infinite regress in the definition of meaning and infinite regressions are considered unacceptable. However, I see no reason an infinite regression should be unacceptable unless it leads to a contradiction (there is no infinite descending sequence of natural numbers, for example), and it is clear that no contradiction arises here provided all the metatheories are consistent.

One should worry, however, about which metatheories are being used. Type theory (higher-order logic) and set theory provide quite different general metatheories, with set theory generally being stronger in mathematical practice (in terms of consistency strength, though certainly very strong type theories are also studied). The metamathematics of intuitionistic logic is quite different depending on whether it is developed with a classical or an intuitionistic metatheory [3]. Fortunately, dealing with this problem in the context of mathematical metaphysics is simple: All metatheories are valid, and all possible ways of building up a tower of metatheories should be considered, with particular ones being specified as necessary. Any structure at any level of this branching tree of metatheories counts as a 'mathematical structure' for purposes of the mathematical metaphysics existence criterion. Note that any collection of subtowers of this tower of metatheories can be embedded into a sufficiently strong version of any of the standard foundations for mathematics (set theory, type theory, or category theory, in either classical or intuitionistic form), and so the various towers of theories form a directed system.

One might be reasonably concerned about where this hierarchy of metatheories starts. Our definition of structure associates an object language with each structure, so an individual structure might be regarded as the zeroth level. However, what the first level should be is unclear. An intuitive answer is that it should be a foundation for mathematics such as simple type theory with the natural numbers, or Zermelo-Fraenkel set theory, but I see no need to be that restrictive (which has the added benefit of avoiding the need to define the term 'foundation for mathematics'). The level above a given structure $\mathfrak{A}$ can be any other structure $\mathfrak{B}$ which is rich enough to contain the relations of some relational structure representing $\mathfrak{A}$ and to define the language and semantics of $\mathfrak{A}$. Note that this entails that new languages with respect to which structures can be defined can appear as we ascend the hierarchy of metatheories; it is not the case that all languages must appear at the first level.

Now that we have explored what the notion of 'mathematical structure' means, we should ask what it means to be an element of a mathematical structure. A straightforward interpretation is that an element of a mathematical structure is just an individual from that structure. A more generous interpretation counts all the relations of a structure as 'elements' as well. Fortunately, given our understanding of the term 'mathematical structure,' it does not matter which interpretation we adopt: All the relations of a structure $\mathfrak{A}$ are elements of some structure (say the transitive closure of the tuple $\mathfrak{A}$ in a set-theoretic metatheory), and hence satisfy the existence criterion of mathematical metaphysics. Thus we see that this existence criterion is very precise and at the same time very lenient (in that it admits the existence of a broad range of objects, perhaps the broadest imaginable). Note that because the empty structure is a structure, literally any collection of axioms in any logic is witnessed by some structure (namely the empty structure). This is because the empty structure satisfies classical contradictions (in the standard model-theoretic sense of satisfaction), and hence can accommodate as true any assertion from any logic (one can simply take a classical metatheory for an arbitrary logic and conclude anything about that logic in the empty structure, whose theory can serve as a metatheory for itself). Also, a structure is nonempty iff its theory is consistent; so any object which can be characterized in a consistent theory exists. Since it is natural to interpret mathematical structures as possible worlds, we see that mathematical metaphysics entails the existence of a Lewis-like assemblage of possible worlds.

Having given a precise characterization of mathematical metaphysics, it is worthwhile to pause briefly to compare it further to Platonism and nominalism. We have already seen how mathematical metaphysics eliminates the mysterious Platonic bridge between the mathematical and physical realms (physical objects simply are certain mathematical objects). In addition, the plethora of possible worlds in mathematical metaphysics gives plenty of material to support any plausible definition of what it means for a (mathematical) object to be physical, whereas Platonism and nominalism traditionally support only one possible physical world, raising the vexing question of why certain physical objects exist, while others which clearly could exist do not. Mathematical metaphysics, in rendering this question irrelevant by accomodating a huge array of possible worlds, also solves the problem of how to deal with the intuitive notion of possiblility in standard Platonist and nominalist accounts where there is only one physical ('concrete' or 'real') world. Finally, the elimination of abstracta demanded by nominalism has proved to be technically extremely difficult in the context of modern scientific theories, and mathematical metaphysics clearly eliminates this problem by embracing abstracta as first-class entities existing in the same sense as physical objects (which are, after all, particular mathematical objects). To briefly illustrate the difficulty with eliminating abstracta in nominalism, consider the question of whether a composite physical object such as an atom exists on a nominalistic account. Intuitively it should, and forbidding such composites complicates the formal statements of physical theories immensely. However, if an atom does exist, it should be a structured collection of elementary particles. Furthermore collections of atoms can form objects which plausibly exist (such as molecules), and collections of these can form still larger objects which we should like to say exist (such as threads, which may constitute a pocket, which may be a component of a pair of trousers, etc.). Thus it appears that we can start with a physical collection and iterate the powerset operation to obtain the existence of all hereditarily finite sets, which is anathema to the nominalist stance. This is also a problem for Platonism, since the same argument can be used to import some mathematical objects (specifically hereditarily finite sets) into the physical realm. One possible solution from the Platonist and nominalist perspective on this issue is to allow the existence of aggregate objects and acknowledge that some objects stand in a 'part of' relation to others (the back is part of the chair). Of course this relation does not exist on a nominalist account, but nevertheless nominalists do allow objects to bear relations to other objects. However, it remains unclear which aggregate objects should be admitted existence in this case: If we consider 267 randomly chosen electrons from the table I am writing at, do they form an aggregate which is part of the table? The cleanest solution which I perceive is to simply reject the traditional metaphysical systems and pursue mathematical metaphysics.

2 What Mathematical Metaphysics Is Not

Now that we have precisely characterized the position of mathematical metaphysics, we consider a range of objections and how these can be met by a proponent metaphysician. The objections likely account for the fact that this metaphysical position has hardly ever been seriously considered, but the answers open novel conceptual avenues for approaching a broad range of philosophical issues.

2.1 Objections concerning physical objects

Perhaps the most obvious objection to the existence criterion of mathematical metaphysics is that it clearly omits an entire category of extant objects: the physical objects. Intuitively, physical objects are not mathematical objects, and so this objection appears to be decisive. However, philosophers are in the business of challenging intuition, and in fact this intuition is easily challenged. Physicists have been increasingly successful in precisely characterizing the behaviour of physical systems through mathematical models, and it is not too difficult to imagine that there might be perfect mathematical models for physical systems. Given a perfect model for a physical system, what is the difference between the model and the system? This is a difficult question for which to provide a precise answer, but mathematical metaphysics provides an enticingly simple answer: There is no such difference; a physical system is identical with its perfect mathematical model. In order to sustain the objection that physical objects are not mathematical objects, the objector must provide some means of distinguishing physical objects from mathematical objects. We now consider several possible means, and find each inadequate.

(i) Physical entities always fail to perfectly instantiate mathematical objects.

Though it is common to call a physical object a triangle, such objects are always imperfect representations of true mathematical triangles: the sides of a physical triangle always have some thickness, its angles never add to precisely 180°, etc. The problem here is what is meant by 'physical triangle.' Given a standard mathematical definition of a (non-degenerate) triangle as the union of three geodesics between three points (one between each pair of points), none of which lies on a geodesic between the other two, we find that so-called 'physical triangles' are not triangles at all. Instead, they are complex configurations of atoms (which are themselves complex configurations of elementary particles). It is plausible that a precise definition of a 'physical approximate triangle' could be given with sufficient patience (and knowledge of a perfect mathematical model of the physical world), in which case the triangles we see in the physical world would perfectly instantiate the concept of a physical approximate triangle. See section 3.6 below on soft concepts for related discussion.

One might still ask whether a mathematical triangle in the standard sense could exist in the physical world, and given the quite general mathematical definition of a triangle in the preceding paragraph, this certainly seems plausible. In particular, if spacetime is a Riemannian manifold as postulated by the general theory of relativity, then the spacetime points lying on any of three geodesics none of whose endpoints lie on a geodesic between the other two constitute a triangle which is plausibly called physical. It is certainly possible to challenge the assertion that spacetime points are physical entities, or that sets of physical entities need be physical, but we do see that under a sufficiently broad understanding for what it means for an entity to be physical, physical triangles likely do exist.

This objection can be given an epistemic twist by noting that apparently humans can have perfect knowledge of mathematical objects, but cannot have perfect knowledge of physical objects (there is always measurement error, uncertainty about the veracity of our sense-data, etc.). The solution is to note that were we to know a perfect mathematical model of the physical universe, we could have certain knowledge about physical objects by proving theorems about the perfect mathematical model which on mathematical metaphysics is identical to the physical world. This would certainly give us perfect knowledge about physical objects, but one might then ask whether we have any knowledge of physical objects at all. If by knowledge of a proposition p we mean possession of a proof of p, then we have very little knowledge indeed about the physical world. However, mathematical metaphysics is also perfectly capable of supporting less strict definitions of knowledge, provided everything about it is made precise (which should be possible assuming mathematical metaphysics is true). For example, defining knowledge as justified true belief is valid (even though it has unintuitive consequences) provided we can give precise definitions of what it means for a proposition to be justified to an agent and what it means for an agent to believe a proposition. Section 3.6 is again relevant to this discussion.

(ii) Mathematical objects are atemporal, whereas physical objects exist in time.

Here we must consider carefully what is meant by 'temporal.' Given a mathematical model of time, there are certainly mathematical objects which exist 'in time' (for example, if we model time as $\mathbb{R}$ and space as $\mathbb{R}^3$, then the origin of $\mathbb{R}^3$ exists at every time). The fact that an unchanging structure can support dynamics is familiar to physicists (spacetime is a fixed mathematical structure, and the static worldlines of particles determine the time-evolution of the particle's positions relative to a given frame of reference), and is the basis for tenseless theories of time in philosophy (McTaggart's B-series). Thus we have that physical objects simply exist in a structure equipped with time, and hence certainly can be mathematical.

(iii) Mathematical objects are causally inert, while physical objects are capable of causal relations.

Working with precise mathematical definitions and models of causality is now a major area of inquiry in philosophy and statistics (among other fields), and so we can handle this objection in a manner analogous to the response to the temporality objection. In particular, given a general definition of causality, one can find the causal structure of any mathematical structure for which causality makes sense, and thus obtain mathematical objects which have causal relations to each other and hence are not causally inert. Given a perfect mathematical model of the physical universe and assuming it is the sort of mathematical object for which causality makes sense (otherwise we must have gotten either the definition of causality or the model for the physical universe wrong), one can deduce the causal structure of the physical universe and conclude what physical events cause which other physical events, with everything still being a mathematical object.

(iv) Mathematical objects are deterministic, whereas physical objects can behave nondeterministically.

Since mathematicians frequently discuss processes which are nondeterministic, we should do well to examine what they mean. A stochastic process, studied in probability theory, is a sequence (say $(X_n : n \in \mathbb{N})$) of random variables, which are themselves measurable functions from a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ into a measurable space $(S, \mathcal{G})$. The intuitive interpretation is that for $\omega \in \Omega$, $X_n(\omega)$ is the value of the random variable X_n at the possible world ω, and the probability that X_n is an element of A, where $A \in \mathcal{G}$, is $\mathbb{P}\{\omega \in \Omega : X_n(\omega) \in A\}$. Thus we see that a stochastic process is nondeterministic in the sense that it can proceed differently in different possible worlds (the sequence $(X_n(\omega) : n \in \mathbb{N})$ depends on which world ω one observes it in). Since mathematical stochastic processes have been very successful in modelling of statistical and quantum mechanical systems (to name two physical use-cases), it seems very plausible that the nondeterminism of the physical world can be perfectly mathematically modelled using probability spaces, complete with their possible worlds.

A significant problem when modelling the nondeterministic behaviour of physical systems using a probability space of possible worlds is determining the significance of the probability measure. Intuitively, the probability measure of an event E (set of worlds) should give the probability that E occurs, but it is a difficult philosophical problem to determine what this means. Let us look at an example: If a photon encounters a perfect unbiased beamsplitter, the probability that it passes through is 1/2 and the probability that it is reflected is 1/2. Consequently, under our mathematical interpretation of probability, the measure of the set of possible worlds where the photon passes through the beamsplitter is 1/2 and the same is true of the set of worlds where the photon is reflected. Less precisely and more colloquially, the photon passes through the beamsplitter in half of the possible worlds, and is reflected in half of them. Thus if we were to direct a photon into this beamsplitter, we should not be at all surprised if it were transmitted, and also would be completely unsurprised if it were reflected. If the beamsplitter only reflected photons with probability 1/3, we should be mildly surprised if it were reflected, while if it reflected photos with probability 1/1000 we should be quite surprised if the photon were reflected. Furthermore, if we repeatedly directed photons into a beamsplitter which reflects them with probability p, we should expect to find the proportion of photons reflected approaching p (the rate of convergence we should expect can be derived from the strong law of large numbers).

Although the preceding discussion of the photons seems intuitively innocuous, we must ask what it means to be surprised if an event happens, and why we are more surprised when events with lower probability do occur. To illustrate why there is a problem here, consider choosing a real number uniformly at random from the interval $(0, 1)$. Clearly, with probability one (in fact, certainly) we shall obtain some real number (call it x), but for any real number y fixed in advance (including x), the probability that we will choose y is zero. Thus with probability one (in fact, certainly) an event with probability zero (namely choosing some specific real number) must occur, and so we may plausibly argue that we should not be surprised when an event of probability zero occurs. In a sense this is true: If we examine the result of a random process and find that it only occurs with probability zero, this is not in itself reason for surprise. What would be surprising is if we correctly predicted that an outcome which has probability zero will occur. More generally, if we predict in advance that an event with small probability will occur, this is reason for surprise.

We have phrased the interpretation of probability in terms of the mental condition of surprise because it will be argued below that minds have associated probability measures which determine the level of surprise which would be experienced by the mind were it to find itself in a world or future with certain properties. This argument is given in the discussion of 'worlds imagined' in the following subsection; the essential purpose is to explain why some possible futures are far more plausible than others (for example, it would be absurd to think that you may wake up tomorrow with an elephant's nose, but according to mathematical metaphysics that is possible since it results in no contradiction). To conclude, we have seen that once again an apparently non-mathematical property of physical objects (nondeterminism) can be seen as mathematical by using the mathematical tools which have been developed by physicists, mathematicians, and others to cope with it.

(v) There can be multiple physical objects with precisely the same structure, but mathematical objects with precisely the same structure are identical.

First, I should point out that on the mathematical metaphysics view of structure as independent of embedding into a larger structure (discussed in section 1), it is true that objects with exactly the same structure are identical. However, a given structure may be embedded into a larger structure in many ways: For example, the naturals can be embedded into the reals by any of the maps $n \mapsto kn$ for $k \in \mathbb{N}$, and a structure $\mathfrak{A}$ can be represented by many set-theoretic tuples of relations. This meets the objection (likely noted by mathematicians) that there exist isomorphisms between distinct structures; this is true in a set-theoretic sense, but from the perspective of mathematical metaphysics the different representations of $\mathfrak{A}$ have additional, set-theoretic structure which distinguishes them (one representation may use reals which are Dedekind cuts, while another employs Cauchy sequences), and we still have that there is only one object with exactly the structure of $\mathfrak{A}$, nothing more.

The mathematical discussion of the preceding paragraph is directly relevant to the objection we are examining: For if multiple physical objects have exactly the same internal structure, they must be distinguished by distinct embeddings into the physical universe. Indeed, a collection of photons which apparently all have the same structure are distinguished by positions in spacetime (if two were to inhabit the same point in spacetime, their wavefunctions would combine to produce a different object), and the spacetime location of a photon adds additional structure to the photon, so that the photons may very well be identical in every other respect and hence literally identical if the relations of the photons to the external world are forgotten (mathematically this would correspond to applying a forgetful functor, which we see is non-injective).

2.2 Objections concerning fictitious worlds

In this section we examine the consequences of the Lewis-like assemblage of possible worlds whose existence is entailed by mathematical metaphysics (as seen in section 1), and why this does not lead to absurdity.

(i) Worlds imagined

According to the mathematical metaphysician, there is a possible world where all the events of Shakespeare's The Twelfth Night occur exactly as described in the play (which is not very precisely, but we can sidestep that issue by specifying one of the many mathematical models which intuitively describe a world where all the events of The Twelfth Night occur). The same remarks apply to any work of fiction that is not inconsistent, whether or not it has actually been written, and this strikes one intuitively as absurd. Surely there is a sense in which our world is real, and all these fictitious worlds are, well, fictitious. Indeed there is, but it is a matter of perspective. On mathematical metaphysics, any intelligence S, being a mathematical object, will find itself embedded in some possible world w which we might call its world of experience, and will think of worlds disagreeing with that experience as fictitious. However, if a world w' disagreeing with w in ways which can be experienced by S contains an intelligence S', then from the perspective of S', w' is real and w is fictitious. There is no logical problem here; merely a need to adjust our intuition.

However, one may still wonder how we know which world is the real world for us. We imagine a single, distinguished world which is 'real,' but it is extremely difficult to give an explicit definition of that world. In modal logic models, the distinguished real world is often just assumed as part of the structure of the model. Any reasonable definition of the real world will make it a domain of our experience, but there are parts of the real world which are not in the domain of our direct experience (e.g. the center of the sun), and there are multiple worlds which serve equally well as domains of our experience. To give a simple example, suppose the observed universe for a certain class of intelligent beings looks like $\mathbb{Z}^3$: the beings can move one spatial unit per unit time, and the direction of the spatial unit moved can be up or down or to either side. At each time, some of the elements of the spatial lattice at that time (isomorphic to $\mathbb{Z}^2$) are labelled with states of consciousness, and these labels characterize our beings. Through scientific study, the beings arrive at the theory that their universe is characterized by a function on $\mathbb{Z}^3$ (this function giving the position and state of all objects at all times). Is this theory correct? Let us assume that everything in the experience of our beings accords perfectly with this model of their universe, so that in a strong scientific sense it is correct. Nevertheless, it would also be perfectly consistent with the experience of these beings for their universe to contain points outside their scientific model. Suppose one of our beings claims that there are spatial points infinitely distant from him (and so also all the other beings, who we may reasonably suppose are at any time contained in some finite sublattice of $\mathbb{Z}^2$). Others may decry this theory as unscientific because it cannot be substantiated in a finite amount of time, and perhaps they are correct (this depends on precisely what we mean by a 'scientific' hypothesis, a point to which we shall return later in section 4.2). Does this necessarily mean that the bold suggestion of our not-entirely-empirically-minded being is incorrect? Under the assumptions of mathematical metaphysics, it certainly does not. For example, we may take a nonstandard model $\mathbb{Z}^*$ of $\mathbb{Z}$, which necessarily contains Z-chains, and think of the universe of our beings as a function on $(\mathbb{Z}^*)^3$ rather than $\mathbb{Z}^3$. All the elements of these Z-chains are points infinitely distant from the beings, just as our bold metaphysician-being claimed to exist. Do they really exist in the world of our beings? Assuming (as seems eminently reasonable) that our beings theory of their universe is first-order, they have no means available to them for distinguishing between these two universes (assuming any method they use must provide an answer within a finite amount of time). If instead we allow our beings an infinite amount of time to test their hypotheses, we find that this unscientific theory of our metaphysician-being is in fact empirically testable: A being need only walk in some direction for an infinite amount of time, and after an infinite amount of time has passed, he will find himself at a spatial point infinitely distant from his starting point. One might want to deal with these difficulties by requiring as a point of methodology that a physical model should be as small as is consistent with the observations. However, let us suppose that our beings can only observe objects within some fixed finite distance of their position. Then at each time there is some finite sublattice of $\mathbb{Z}^2$ which suffices for a description (via a function) of everything observable to the beings at that time. The universe of the beings could then be taken to be the time-indexed union of these finite sublattices, and we should obtain a physical model smaller than the $\mathbb{Z}^3$ model and still sufficient to describe everything observable by the beings. However, this is intuitively less elegant than the $\mathbb{Z}^3$ model, so one might like to avoid a scientific methodology which requires the use of such small models. What I should like to argue is that the precise model is indeterminate: A being exists simultaneously in all universes which are consistent with his experience.

In a sense what we have examined above–a metaphysician who claims there is more in the world than science can establish–was the easy case. What if our metaphysician makes claims which appear to contradict science? For example, let us now think of the world with which we are familiar, and suppose both Peter and John have the experience of sitting in a room together and engaging in a philosophical discussion. Peter claims that in fact they are not in the same room, but rather in a swamp separated by one hundred meters. He says their experience of the room is merely a construction of their minds, hiding from them the reality of their situation. Is Peter correct? It is apparent that the situation Peter describes is consistent with all their experiences and any experiment they could possibly perform, and so by our criterion for which worlds a consciousness is embedded in (anything consistent with the experiences of that consciousness), we see that Peter is in fact correct, insofar as his and Johns consciousnesses are indeed embedded in worlds where they are in a swamp separated by one hundred meters and their minds generate an artificial experience of sitting in a room. However, he is also incorrect insofar as there are other worlds where they actually are sitting in a room, and yet other worlds where they are in fact in a desert, etc. Thus we see that our position has resulted in a peculiar pluralism about reality. But surely some classes of possible worlds are more likely than others; intuitively what Peter and John are actually experiencing is their discussion in a single room together, and the swamp world, though possible, is highly unlikely. To formalize these thoughts, we must ask what is meant by likely. The standard mathematical theory of this notion is that of probability, and I believe this will suffice (as it did in the discussion of nondeterministic physical processes in the preceding subsection). What we want to say is that for any consciousness there is a probability measure on possible worlds consistent with the experience of that consciousness, and properties of these worlds which are true with high probability (especially probability 1) will be considered by the consciousness as likely to hold in reality. From whence do we obtain this probability measure? Since this is associated with conscious experience of reality, I suspect and will assume in the sequel that this probability measure can be obtained from a consciousness. That is, given a consciousness, one can define a probability measure associated with that consciousness such that the possible worlds experienced by the consciousness tend to be those with high probability. As was discussed in the nondeterminism objection in the preceding subsection, it should be noted that some events with low probability should be expected to occur (in fact it is common that some event with probability zero must occur), but a consciousness would be surprised if it could successfully predict that an outcome with low probability will occur as the outcome of a specific event. This surprise is unformalized and somehow tied to the probability measure associated with the consciousness, but in this paper I have nothing further to add. Exploring the connection between mind and probability further is intruiging from any philsophical perspective, and mathematical metaphysics offers an interesting new perspective to approach it from.

One apparent difficulty with the account of minds and experiences we have been giving is the existence of dreams. We appear to have direct experience of something unreal in dreams, but since dream worlds are conceivable, they are possible and hence exist according to mathematical metaphysics. Thus are experiences in dreams just as real as experiences in what we intuitively consider to be reality? This is true in the same sense that the swamp world from the preceding paragraph was just as real as the world perceived by Peter and John. However, perhaps we can use the language of probability to be more precise. It is clear that the dream state is an altered state of consciousness, and so we should not be surprised to find that the probability measure associated with a consciousness is related to experience in a different way during a state of dreaming rather than a wakeful state. Our experience indicates that objects in the physical world are stable and can be manipulated at will only by physical means, but objects in a dream are ephemeral and can be manipulated not only by what is perceived in the dream as physical means, but also by direct mental manipulation (for example, it is common for one to be able to fly in a dream by simple mental exertion, which is obviously contradictory to our experience of the physical world). It seems reasonable to say that the probability that we stay in the real world throughout our lives, with dreams being merely patterns of interaction of neurons in our brains, is much higher than the probability that we somehow travel between the real world and various dream worlds each time we dream and wake, even though the latter is possible and hence does occur. Therefore we conclude that the former is what is occurring in reality.

(ii) Alien worlds.

We can imagine a world so completely alien from the physical world of our experience that it seems absurd to regard it as a possible world (for example the structure $(\omega_1, <)$). Thus it would appear that mathematical metaphysics permits too broad a range of possible worlds. However, we can also view this as simply an artifact of the precise definition of 'possible world' which we have chosen (namely taking this term to be synonymous with 'mathematical structure'). If a more restrictive definition of 'possible world' is more convenient for some purposes (e.g. physical world, or world consistent with my experience thus far), then certainly mathematical metaphysics is open to such an alternative definition. The main difficulty is making such a more restrictive definition precise.

2.3 Objections concerning minds

Before getting to the objections, let us examine more explicitly what a mind should be, given mathematical metaphysics. We saw in the section on fictitious worlds that just as the structure of the natural numbers is simultaneously embedded in every structure which contains an object and a successor function, so a mind (there called a consciousness) is simultaneously embedded in every structure which contains that mind as a substructure. Thus a mind does not inhabit a single world, but rather participates simultaneously in a multitude of worlds: every world in which it is logically possible for that mind to participate. It is important to note here that for a mind M to participate in a world w, the world w must contain a structure isomorphic to M; it is not sufficient for M to be an object in w. To see why, note that we may take a collection of countably many minds and define a successor function to obtain a model of the natural numbers. However, none of the minds participate in this world; they are merely the structural placeholders for numbers, which are not conscious in any ordinary sense. Also, when a mind is instantiated in many worlds, one might ask what it means for different instantiations to be the same mind. The answer is that two structures instantiate the same mind iff they are isomorphic as minds (a notion which cannot be made precise without a precise understanding of what it means to be a mind, but this need not concern us yet).

This view of minds clearly results in a problem for our intuition, because the world of our conscious experience appears to be just one world, not every world we could possibly participate in. For example, it is consistent with my experience that I am a crustacean at the bottom of an ocean being deceived by Descartes demon. Thus on mathematical metaphysics, that is my situation in some possible worlds. However, we intuitively say that such a state of affairs is extraordinarily improbable. On what grounds do we make that judgment? I cannot ultimately give better than intuitive grounds, but I can explain a mechanism by which we can perhaps better understand our experience of apparently just one world. As mentioned before, I do not know how to give a mathematically precise definition of a mind, but let us suppose that we have such a definition, and that through this we find naturally associated to each mind a probability measure such that the experience of the mind accords with the measure: If a class of worlds has small measure, the mind rarely finds its experience aligning with these worlds, while if a class of worlds has large measure, the mind regularly finds its experience aligning with these worlds. Thus such a mind generally does not find itself surprised by accurate predictions of low-probability events, as we outlined in the section on nondeterminism in section 2.1.

In a sense a mind participates 'more fully' in classes of worlds to which it assigns high probability than those to which it assigns low probability, though it is simultaneously instantiated in all worlds which are in any way consistent with its experience, including those of very low probability. If a mind were to follow a possible future in which it only experiences events it considers to have low probability, it would be thoroughly and consistently surprised. However, note that the probability measure associated to a mind is allowed to change with time, so if the world I am accustomed to were to vanish and I found myself decidedly in a swamp world, my probability measure would quickly adapt to assign high probability to the regular happenings of the swamp world.

Some explanatory remarks are in order. First, I am assuming that minds are essentially temporally extended entities, while states of consciousness are not. Thus we may find that in some world w there is an instant of time t where the structure of a state of consciousness c which is part of a mind M can be found, but not so for any later instant of time. In this scenario, M participates in w through c at time t but subsequent to that ceases to participate in w (since no state of consciousness c which is part of M can be found in w after time t), so that I should say its experience ceases to align with w after time t. One might point out here that if the states of consciousness are simply two symbols, these would be found in any world containing at least two objects, thus causing difficulty for the above discussion. Here I should say that minds do not have experiences instantaneously, but rather over short spans of time. Thus though the symbols can be found in nearly any world, their temporal sequence across a short span of time can be found in many fewer worlds. We can therefore reasonably assume that the states of consciousness encode information about the states immediately preceding and following (just because two symbols may suffice does not mean this is the representation we must use). This is similar to the mathematical notion of a germ.

Now let us note that a single state of consciousness may simultaneously participate in many minds. Suppose a mind M finds itself in state c at time t. What future states will that mind enter? It will enter all future states which are consistent with the state of consciousness c having occurred at time t, and with the definition of a mind (which we have assumed known). Intuitively, there are very many paths the mind could take into the future, all consistent with its being a mind and having had state c at time t. We are now in a position to say with greater precision what I mean by a mind rarely finding its experience to align with classes of worlds of small measure, and regularly finding its experience to align with worlds of large measure. The set of possible futures of a mind is also equipped with a probability measure (which should be determined by the measures associated to the mind at each time t), and with high probability the mind finds itself on a path from a class with large measure, while with low probability it finds itself on a path from a class with small measure, where again probability is interpreted in the intrinsically mental sense of determining expected experience. It is also important to note that if a mind finds itself in state c at time t, its past is not entirely determined. It is intuitively obvious that there are many possible ways a mind could have come to that state at that time, and so all of these occur. However, the memories of the mind are clearly related to a probability measure on past worldlines leading to the present in which the mind finds itself, and so we have an explanation for the apparent determinateness of the past.

Since minds are not bound to any particular world (but rather simultaneously embedded in many worlds), and are constantly branching in both the future and past directions, we have a neat way to explain what is meant by an action that a mind could have performed. If a mind is in state c at time t and there is some future path where it performs the action A, than M can perform action A, in a weak sense. If furthermore the probability of the class of futures where it does in fact perform A is reasonably large, then M can perform action A in a stronger sense. If there is some past path to a state c out of which there exists a future path along which M performs A, then M could have performed A in a weak sense, while if both of these paths have reasonably high probability, then M could have performed A in a stronger sense. One might ask whether at any moment there are any actions a mind could not have performed. The answer depends on exactly how we define 'mind' and 'action,' (if we define them so any mind performed every action in some possible past, then the answer is no), but what we can say definitively is that a mind M could not have performed an action A if and only if the assertion 'M performed A' is inconsistent given the definitions of M, the 'performance' relation, and A.

Perhaps a few technical remarks about measures on sets of worlds are in order. The collection of all possible worlds is very large, larger than any proper class in any particular model of Goedel-Bernays set theory. Thus one might worry about what kinds of probability measure make sense on such a space. However, the class of worlds which are potential physical models of the world we experience appears to be much smaller. In fact, assuming any physical system can be modelled by a Borel function on a manifold, we have that there are only $2^{\aleph_0}$ many physically possible worlds. Thus we may reasonably assume that the probability measure of a physical mind is concentrated on these worlds, in the sense that any set of worlds disjoint from this set of physically possible worlds has measure zero.

Having reviewed and expounded upon the view of minds which mathematical metaphysics naturally leads to, we examine some specific objections.

(i) The Matrix objection

If physical objects are mathematical, why can we not manipulate them with our minds in the same way that we manipulate abstract mathematical objects, thus resulting in a world in which reality-manipulating actions like those in The Matrix are possible? First, one must note that under mathematical metaphysics, matrix-like actions are indeed possible. However, it appears that the probability of finding oneself in a world where these actions occur is small. Since we can intuitively manipulate abstract mathematical objects, how can this be so? The answer lies in the distinction between holding an object and holding a thought about an object: Our minds have the ability to think about mathematical objects in a rather less constrained way than they have the ability to manipulate physical objects. This is because the (mathematical) relation of manipulation is mediated through our physical bodies, while the relation of thinking about is mediated through our minds as physically instantiated in our brains. Note that the physical body and brain are still mathematical objects on this account; what distinguishes them as physical is their presence in a physical structure, namely our universe. That the universe is physical should follow from the definition of 'physical,' though unfortunately a precise characterization of which structures should count as physical appears to be out of reach for the time being.

We can take the matrix objection further. Suppose you are holding a spoon. You can conceive of the spoon bending purely in response to your mental exertion, and therefore in some possible worlds it does. Suppose you are about to exert your mind in just this way, an exertion which we may say occurs at time s. The current time is some $t < s$. At time t, it is possible that at time s the spoon will indeed bend in response to your mental exertion, and thus the event that it does indeed bend is in your future at time t. Would it not then be appropriate to say that the spoon will bend in response to your mental exertion at time t? Yes, but this is only in the sense that you are in a swamp world. Just as consciousness induces a probability measure on possible present worlds, so it induces a probability measure on possible future trajectories, as we saw in the introduction to this section. At time t there are future trajectories where the spoon bends at time s, but our experience indicates that this is highly unlikely, so we conclude that the measure of the set of such trajectories is small.

(ii) Boltzmann brains

Assuming a physical universe rather similar to our own, it appears that Boltzmann brains are by far the statistically most likely instantiations of consciousness. Here I understand Boltzmann brains to be brains which randomly assemble from high-entropy surroundings, complete with detailed memories of a non-existent past, and disappear back into the high-entropy background after just enough time to have some thoughts. Thus we should expect that there are far more Boltzmann brains than brains which evolved from an even lower entropy past, and hence should expect that we are Boltzmann brains. It seems to me that this is in a sense correct; a random brain which instantiated a small interval of my thoughts should be a Boltzmann brain with high probability. However, the experiences of Boltzmann brains presumably vary wildly and are short-lived on average, so the stream of events which is deemed most likely by the probability measure associated with my consciousness is more stable. An evolution from a low-entropy past singularity certainly seems like a good candidate for a stable past and future (following the second law of thermodynamics), and indeed this is what is predicted by modern physics. Thus although Boltzmann brains are far more common than brains which exist in a non-equilibrium universe (such as our own), we should not be surprised to find that we do not have the experience of Boltzmann brains.

(iii) Minds can make choices, while mathematical objects cannot.

To deal with this objection we simply need to understand what a choice means, just as the objections that mathematical objects are atemporal and acausal were met by understanding what time and causality mean. A mind M can be said to make a choice at time t iff there are at least two states it may assume at that time which will lead to different futures. Assuming as in mathematial metaphysics that minds can be perfectly modelled as mathematical objects, it is clear that mathematical objects can make choices; all that is needed is for the mathematical object to model a mind and satisfy the aforementioned property.

(iv) Minds can manipulate mathematical objects, while mathematical objects cannot manipulate anything.

Now we need to understand what manipulating a mathematical object means. A natural answer would be that a mind manipulates a mathematical object if it forms a substructure of itself which instantiates some properties of the mathematical object, and then evolves this substructure until extracting some knowledge (or at least belief) about the structure. Since the mind can by assumption of mathematical metaphysics be perfectly modelled as a mathematical object, we see that mathematical objects can indeed manipulate other mathematical objects.

(v) Minds can have direct knowledge of themselves, while they can only have knowledge of mathematical objects through mathematical proof.

The problem here is equivocation between two different meanings of 'knowledge.' The first is a mental sense roughly corresponding to our intuitive notion of knowledge; the second is the precise mathematical sense of mathematical proof. In the strict mathematical proof sense, a mind does not know anything about itself unless it has a perfect mathematical model of itself and proofs of those facts which it knows about itself. Similarly, our intuition provides a guide to knowledge concerning mathematical objects in much the same way it provides a guide to knowledge of ourselves. There is no incompatibility between the two clauses of this objection when we understand 'knowledge' in these two distinct senses.

Before moving on, I note quickly that Nagel [4] finds it difficult to see how a complete physical (let alone mathematical) description of a bat could fully capture what it is like to be a bat. This is not exactly an objection concerning minds, and the best way I can see to deal with it is to ask what sort of entity 'what it is like to be a bat' is; perhaps if we knew this, we could capture it mathematically. Otherwise, I see no reason to worry about such entities as 'what it is like to be a bat;' perhaps these are nothing more than linguistic utterances which trigger responses in our brains but do not refer to an object.

3 What Mathematical Metaphysics Does

Having dealt with some intuitive objections to the position of the mathematical metaphysician, we now examine what positive philosophical benefits can be derived from it. We shall look at some of what mathematical metaphysics has to offer in the areas of epistemology (3.1), philosophy of mathematics and science (3.2,3.3), philosophical logic (3.4,3.5), semantics of natural language (3.6), physics (3.7), and general metaphysics (3.8) by examining a specific way that the mode of thought of the mathematical metaphysician clarifies some perennial issues.

3.1 Precise epistemology

The quest for a precise characterization of what it means to know has a long history in philosophy and has proven exceedingly difficult without incurring major compromises. Mathematical metaphysics does not provide a perfect solution, but at least provides a very natural candidate for the semantics of knowledge: An agent a knows a proposition φ if and only if the agent possesses a proof of φ. This entails that humans know very little indeed about the physical world (and furthermore very little can be known at present as we are not in possession of a perfect mathematical model for the physical world), a contention that was briefly examined in the last paragraph concerning the first objection in section 2.1. However, this definition does have some very nice consequences from the point of view of epistemic logic. In particular, if we interpret $K_a(\varphi)$ to mean that agent a can form a proof of φ using knowledge he already possesses, then the following rules of epistemic logic hold:

(K) $K_a(\varphi) \wedge K_a(\varphi \to \psi) \vdash K_a(\psi)$

(Truth) $K_a(\varphi) \vdash \varphi$

(Positive Introspection) $K_a(\varphi) \vdash K_a(K_a(\varphi))$

Note that negative introspection, the assertion that $\neg K_a(\varphi) \vdash K_a(\neg K_a(\varphi))$, does not necessarily hold, because in particular for constructive logics there may be no proof possible that a proposition lacks a proof. Axiom (K) follows easily since we regard $K_a(\psi)$ to mean not that a knows ψ immediately, but rather can conclude that ψ is true upon reflection (by giving a proof). The (Truth) axiom formalizes a fundamental intuition that if we know something, it is true. Given this, the (Positive Introspection) axiom can cause difficulty for accounts of knowledge which allow it to be derived from sources which are not completely reliable, as then it is possible that an agent a has justification for belief in φ from an unreliable source, and that φ happens to be true and a knows φ, but $K_a(K_a(\varphi))$ is false because agent a does not know that he knows φ because he does not know φ is true due to the unreliable source of justification, and hence because of his knowledge of (K) and the truth axiom cannot justify belief in $K_a(K_a(\varphi))$. We shall not be detained here by a discussion of how traditional and contemporary accounts of knowledge deal with this issue, but note that it is intuitively very plausible that the only completely reliable source of knowledge is mathematical proof, in which case knowledge as mathematical proof is the only explication of knowledge which avoids unreliable sources.

3.2 Precise connection between mathematics and physics

Mathematical metaphysics avoids the mysterious Platonic bridge between the mathematical and the physical by simply identifying physical objects with certain mathematical objects. Thus Wigner's question [6] about why mathematics is relevant to the physical world is easily answered: The mathematics which constitutes the physical world is certainly relevant to it. One might still worry that since we do not have a perfect mathematical model of the physical world, the mathematical tools we think are relevant to the physical world may not actually be. It is true that, for instance, spacetime could turn out to be discrete, thus invalidating the assumption that it is a Riemannian manifold; but certainly even if spacetime is not literally a Riemannian manifold there is some sense in which it is well-approximated by such a manifold. Assuming such approximations are relevant to the physical world, the mathematical tools used by physicists today should (mostly) still be relevant to the physical world even if it is wildly different than what the use of these mathematical tools assumes. To strengthen our intuition that these approximations are indeed relevant to the physical world, we may consider classical Newtonian mechanics, which is an approximation to both quantum mechanics and general relativity and yet immensely useful for making a wide range of predictions concerning the physical world. Surely there is a well-defined sense in which Newtonian mechanics is mathematically associated to a perfect mathematical model of the physical universe as a relevant approximation.

Another possible concern is that if all structures satisfying a mathematical definition of what it means to be physical really are physical, then why are not all the perfect mathematical models studied by applied mathematicians physical in our universe? This conundrum appears to rest on a dual meaning of the word 'physical:' I have been referring both to physical worlds in general (mathematical structures which fit the definition of physical structures), and 'the' physical world, namely our universe. It should be noted again that which world is 'the' physical world depends on the context of a discourse. Beings from a physical world completely different from ours (except that they still employ the English language to express their thoughts) would consider their world to be the physical world, while those of us in our physical universe would disagree. Perfect mathematical models of physical worlds other than our own appear to us to be purely abstract rather than concrete simply because we live in a certain physical world and define 'concrete' with reference to that world. In other worlds, our world would seem purely abstract.

One final note is that since our minds are simultaneously embedded into all worlds consistent with their experience, many worlds which we would normally consider purely abstract are 'real' in the sense that our minds are instantiated in them. Of course we can try to define 'the' real world as the world with highest probability, but such a world may fail to exist. What is intuitively clear is that the coherence of our experience suggests we experience primarily the embedding of our minds into a single structure.

3.3 Precise meaning of structure and isomorphism

Philosophers of science, particularly structuralists, often discuss the structure of physical systems and their correspondence to mathematical models via isomorphisms, but unfortunately the precise meanings of these terms in the context of physical objects are rarely made clear. Mathematical metaphysics makes available a simple and compelling solution to this problem: The structure of a physical system is simply its mathematical structure (the collection of its components together with all their relations), and isomorphism has the precise meaning which it is given by mathematicians. Again mathematical metaphysics provides a neat way to avoid the mysterious Platonic bridge between the mathematical and physical realms by simply collapsing the physical into the mathematical.

3.4 Precise meaning of alethic modality

Since mathematical metaphysics entails the existence of an object fitting any consistent description, it allows us to conclude the existence of a Lewis-like assemblage of possible worlds, which then makes available a natural interpretation of necessity as truth in every possible world and possibility as falsehood in every possible world. However, this notion is not interesting if we allow every structure as a possible world (if the empty strucuture is allowed, everything is possible and by duality nothing is necessary; while if the empty structure is excluded but a sufficient variety of other structures are allowed, only logical truths are necessary). Consequently necessity and possibility should be interpreted in a restricted domain of structures, say physical structures, or plausible histories and futures of the physical universe we perceive. One might then reasonably ask what it is that makes a modality 'alethic;' because of the natural interpretation of necessity as truth in all possible worlds and possibility as truth in some possible world, it seems natural (assuming Kripke semantics for modal logic) to call a modality alethic if and only if its accessability relation is complete (connects every pair of worlds).

3.5 Precise meaning of probability

Pinning down exactly what it means for a particular event to have a certain probability is a long-standing philsophical problem which appears to still be far from solution. Mathematical metaphysics does not offer a complete solution, but does offer a framework for further work which appears quite promising. This is described in the answer to the objection from section 2.1 that mathematical objects cannot be nondeterministic, unlike physical objects, and further pursued in the introduction to section 3, but we briefly review it here. The concept of probability makes sense for a probability space, the points of which are possible worlds. An agent does not necessarily know what properties hold of the possible world w which he is in, but would be surprised if he could correctly predict that w has some property which holds with low probability. The meaning of 'surprised' here should be made precise and does not correspond exactly to common usage, but it should be reasonably intuitively clear what is going on. Thus because the probability of an event (property of worlds) is linked to an occurrence in the agent's mind (surprise), we find that the probability measure is linked to the agent's mind and governs what sorts of experiences are unsurprising. To put things another way, we should expect to find ourselves in a reasonably generic world among those we could possibly inhabit, and genericity is determined by a probability measure associated to our minds (and hence indirectly experiences). Certainly this was not very precise, and it leaves many questions open, but it at least opens an interesting direction for future inquiry.

3.6 Explication of soft concepts

Philosophers consistently find it very important to be precise about their use of terms, but for concepts like 'beauty,' 'chair,' or 'green,' it appears to be impossible or at least impractical to be completely precise. Such concepts I call soft concepts. Mathematical metaphysics offers no help with the difficulty of making soft concepts precise, but at least gives reason to believe this can indeed be done, and perhaps some idea of what that would look like.

Let us start with the concept of 'green.' To simplify our discussion we shall restrict our attention to which range of wavelengths of light should be considered green. It is clear that not just any range is acceptable (certainly nothing which is obviously blue should count as green, for example), and that the precise boundaries of the range of green wavelengths are difficult to specify. For an object of a colour near one of these boundaries, some people may say it is green, others may say it is not, and still others may suspend judgment. Thus it seems we have a two-parameter family of candidate definitions of green, where one parameter is the left endpoint of the range and the other its right endpoint (we could increase the number of parameters by deciding whether to include or exclude each endpoint, but will ignore that issue for simplicity). These parameters themselves are restricted to a certain range, and we assume this range is fixed (otherwise we could iterate the process and give ranges for the endpoints of the parameter ranges, etc.). Each choice of a pair of parameters gives a possible definition of green. Which one is the 'right one' depends on our definition of 'right,' and it is certainly intuitively plausible that any of these definitions is the right definition in some contexts. We have not specified a range of wavelengths until choosing a pair of parameters, but it is notable that the whole ensemble of possible definitions of green forms a single mathematical object and can be seen as giving precise meaning to the phrase 'the soft concept "green."'

The concept of 'chair' is somewhat more complicated since it is not easily reducible to a few parameters as in the case of colours, and perhaps has something to do with the soft concept of purpose (must a chair have the purpose of being sat upon?). The concept of beauty is yet more complicated since people's judgment varies more wildly over time than for the other two concepts we have considered, and it is conceivable that two people would disagree in every instance about what is beautiful. However, it is still very intuitively plausible that a family of possible definitions can be distinguished for each of these soft concepts, and that this family will have a structure which we may identify with the soft concept.

3.7 Quantum Physics

The famous measurement problem of quantum physics is the question of how and why wavefunctions collapse when observed. For example, if a photon passes through a beamsplitter, its wavefunction takes both paths and can even interfere with itself, but as soon as the location of the photon is measured the wavefunction collapses into a reasonably well-defined location (bound by Heisenberg uncertainty, of course). Mathematical metaphysics provides an intriguing way of thinking about this: Since minds are simultaneously embedded in all possible worlds consistent with their experience, perhaps the photon appears to have taken multiple paths until its location is measured because following either path is fully consistent with our experience until we make an attempt to measure the location of the photon. Certainly there is more work to be done here, such as determining why wavefunctions can interfere with each other on such an account, but this does look philosophically promising.

Another interesting consequence of mathematical metaphysics which is related to quantum physics is the fact that mathematical metaphysics implies the existence of an Everett-style many-worlds ensemble. Since the many-worlds interpretation of quantum mechanics is widely regarded as solving the quantum measurement problem, we have found a major problem in physics whose solution is provided by mathematical metaphysics. Also, the discussion of the preceding paragraph may still shed light on why wavefunctions collapse at the times they do.

3.8 Ontological maximalism

As mentioned near the end of section I, the class of all mathematical objects is extremely broad; perhaps the broadest imaginable class of objects. Thus the ontology of mathematical metaphysics is not conservative at all, but is rather maximal in a sense (it includes as many mathematical objects as can possibly exist). Since many philosophers strive to do with as few ontological commitments as possible, we shall do well to examine how a position of ontological maximalism might be justified.

The justification that I offer is that ontological maximalism is very intuitive: If it is consistent for an object to exist, why should it not? What harm could it do? This is related to the attitude of set theorists that cardinal numbers as large as consistent should exist, and to the intuition that there is no reason to avoid passing to a conservative extension of a given theory when convenient, even one with new ontological commitments (as in the passage from Zermelo-Fraenkel set theory to von Neuman-Bernays-Goedel set theory). Because of this intuition, one might wonder why philosophers have such a strong tendency toward avoiding ontological commitments. My opinion is that they seek to protect their arguments from attack by reducing their assumptions and ontological commitments as much as possible. This is certainly a valuable pursuit, just as it is valuable in mathematics to prove a theorem from a minimum of assumptions, but it really is not necessary once we have accepted the position of mathematical metaphysics. Because philosophers naturally seek to make their arguments as independent of particular metaphysical assumptions as possible, we should expect them to continue avoiding ontological commitment if necessary, but this does not reduce the intuitive appeal of ontological maximalism.

4 How Mathematical Metaphysics Might Be Used

In this section we explore ways of thinking in the framework of mathematical metaphysics which would plausibly be seen as relevant to the everyday lives of people who are not philosophers.

4.1 Universal immortality

A surprising consequence of the view of minds taken in the defense of mathematical metaphysics above is that all minds are immortal in the sense that there exist times arbitrarily far into the mind's future. One way to see this is to note that there is no contradiction which results from a mind continuing to exist arbitrarily far into the future, and so this certainly occurs in some possible worlds. Because a mind only has experience in worlds where it exists, this shows that minds always have an immortal experience (the possible futures in which a mind dies will not be followed past the point of death).

One might worry that the assumption that it is consistent for any mind to be immortal is too strong, especially in the context of an argument that all minds are immortal. This can be partially remedied by the following argument using weaker assumptions:

(0) Time has the structure of an ordered commutative monoid without largest element (whose operation is thought of as addition of time intervals).

(i) If a mind M is alive at a time t, then there exists $\varepsilon > 0$ such that M is alive at time $t + \varepsilon$.

(ii) If a mind M is alive at every time in a bounded set T, then M is alive at a time s such that $s \geq t$ for every $t \in T$.

Premises (i) and (ii) can be collapsed into

(ii') If a mind M is alive at every time in a bounded set T, then M is alive at a time s such that $s > t$ for every $t \in T$,

but perhaps (i) and (ii) are separately more plausible than (ii'). Premise (0) asserts some minimal and highly intuitive structure for time. Premise (i) would perhaps be unintuitive in a single world, but it does seem very plausible that in the infinite ensemble of possible futures of M there exists one where M lives a tiny bit longer. The last assumption is a kind of closure assumption which roughly says a mind cannot suddenly disappear, and certainly has the weakest intuitive justification of the three premises. From these assumptions, we have that for every mind M, if M is alive at time t, then for every time s there exists a time $s' \geq s$ such that M is alive at time s', which intuitively means that M is immortal.

To see this, assume M is a mind which is alive at time t, and let $s \geq t$. By transfinite induction, construct a sequences (t_α), (ε_α) such that for every ordinal α, M is alive at time t_α, $\varepsilon_\alpha > 0$, and the following conditions hold:

(a) $t_0 = t$

(b) $t_(\alpha + 1) = t_\alpha + \varepsilon_\alpha$

(c) For limit β, if $(t_\alpha \colon \alpha < \beta)$ is bounded choose t_β such that $t_\beta \geq t_\alpha$ for every $\alpha < \beta$; otherwise halt the construction.

This is possible because M is alive at t_0 by assumption, ε_α as in (b) exists by premise (ii), and t_β as in (c) exists by premise (iii). If the collection of times is not a proper class in the set-theoretic sense, this construction must eventually halt (certainly before stage κ where κ is the cardinality of the set of times). If the set of times is a proper class, this construction need not halt at any ordinal stage, but this problem can be solved by embedding the entire model of set theory (say $\mathcal{N}$) in which we are working into a larger structure (call it $\mathcal{N}'$) which contains the set of all ordinals in $\mathcal{N}$. $\mathcal{N}'$ then witnesses that the construction must halt. Supposing the construction halts at stage β, we have that $(t_\alpha \colon \alpha < \beta)$ is unbounded, and so for every time s there exists a time $s' \geq s$ such that M is alive at time s', as was to be proved.

Any of the assumptions of the argument just given can certainly be challenged, and they appear in order of plausibility (at least according to the author's assessment). In particular, one could imagine a mind M which ceases to be alive at a time $s > 0$ (so M does not exist at any time $s' \geq s$) but is alive at every time $0 < t < s$; such a mind can satisfy premise (i) but cannot satisfy premise (ii) (or else we should conclude the contradiction that M is in fact alive at time s). However, the above argument from premises (0)–(ii) at least strengthens the earlier claim that it is consistent for any mind to continue existing indefinitely.

Note that in the above argument we assumed nothing about how one can tell whether two states of consciousness belong to the same mind, or whether a mind can cease to exist for an interval of time and then come back into existence. Intuitively I should think that two states of consciousness belong to the same mind iff there is a 'consciously continuous' path of states between them; this terminology is merely suggestive, and I have no idea how to make it precise. Note that this leaves open the possibility that the same state of consciousness can belong to different minds; indeed we should expect this given the highly plausible conclusion that a mind may posses multiple possible futures. Furthermore, a consciously continuous path of states need not necessarily be continuous in time, and so the possibility that a mind may cease to exist for a period of time and then return is also left open. A somewhat intuitively plausible case where this may occur is a deep coma.

One might easily think that the exact same arguments as those used to conclude that all minds are immortal can be used to show that any object continues to exist indefinitely. Since we have used the term 'object' to refer to an element of a mathematical structure, it certainly is possible for an object x to cease to exist in a temporal structure (in the sense that there exists a time t such that x does not exist at any time $t' \geq t$). However, there is a sense in which the argument does still work when we view the future of an object in time as indeterminate: for any element x of the world at time t in a temporal structure $\mathscr{N}$ there exists a temporal structure $\mathscr{N}'$ which agrees with $\mathscr{N}$ up to t and is such that x is an element of the world at time t' in $\mathscr{N}'$ for every $t' \geq t$. This intuitively means that an object which exists at some time t will continue to exist indefinitely in some possible futures.

4.2 Scientific truth and methodology

First let us note that scientific truth is a soft concept in the sense of section III.6, and so admits various interpretations. The term 'scientific knowledge' as it is used in practice has a time-indexed definition which is totally impractical to make precise, but we shall see that mathematical metaphysics makes feasible a modification of the verificationist criterion for scientific truth.

Since observers are simultaneously embedded in many possible physical universes, one might wonder how we can gain scientific knowledge of the real world. First, we should note that 'real world' is a soft concept whose precise definition we should not expect to be universally agreed upon, though we should certainly expect scientific generalities to hold on a subset of the possible worlds in which we are embedded that has high probability (this is part of the soft concept of scientific truth). This certainly opens the door to the possibility of multiple incompatible physical theories which describe reality equally well, and the fact that just this situation appears to hold with quantum mechanics and general relativity gives some plausibility to the assertion that the real world is underdetermined. However, it certainly is consistent with mathematical metaphysics that there is a unique perfect model of the physics of our universe.

How are we to go about answering questions concerning our universe? One way is to seek uncertain knowledge (a soft concept as discussed in sections 3.1,3.6), and that is what is pursued in practice. However, in order to really answer a question we should obtain certain knowledge about it. This is generally regarded as impossible or completely impractical, but surprisingly on mathematical metaphysics there is a simple way to certain knowledge concerning physical objects. All one needs to do is test a hypothesis in every conceivable way, which can certainly take an infinite amount of time, but this is available by the argument for immortality from the preceding paragraph. Because of this method of obtaining certain knowledge, a verificationist interpretation of scientific truth is perfectly reasonable given mathematical metaphysics, provided we are willing to wait infinitely long for verification. Using this methodology, one can obtain every empirical fact, and thus prove everything which can be known empirically about physical objects. Since non-empirical information about physical objects is simply an artifact of the embedding into a mathematical structure, this in fact yields everything which can be known about a physical object. Perhaps these musings are not useful for practical scientific methodology, but this is at least an interesting thought.

4.3 Framework for examining theological arguments

It is a fun and interesting exercise to see how some standard arguments for the existence of God fare under mathematical metaphysics.

(i) Ontological argument

There are many different versions of this argument, but a basic modern one due to Plantinga (a modal ontological argument) is as follows:

(a) If God exists, he exists necessarily.

(b) It is possible that God exists.

Conclusion God exists, necessarily.

From premise (b) and the general logical assumption that a nonempty possible world exists, we have that there exists a world w where God exists. Consequently by (a) God exists necessarily and hence in every possible world.

That there is something wrong is clear, because this entails God exists in the empty world (or if we exclude the empty world, that God is a natural number, etc.). The problem seems to be with premise (a): On mathematical metaphysics, nothing exists necessarily (from the point of view of all structures being regarded as possible worlds), and so (a) is impossible. Thus if (a) is a true assumption about God then God does not exist. It is certainly possible for this argument to succeed by restricting attention to a subset of the possible worlds, but in that case the subset was of course necessarily chosen so all its members contain God, which trivializes the argument.

(ii) Leibniz cosmological argument

(a) Anything contingent has a cause

(b) There are no cycles or infinite regressions of causes

(c) Contingent things exist

Conclusion There exists a necessary cause.

Premise (b) could be suspect depending on one's notion of causality, and both (a) and (b) are certainly suspect in structures lacking any notion of causality (or perhaps even time). However, in addition to the intuition that everything exists contingently because nothing exists in the empty structure, we also have that everything exists necessarily in a broader sense because all structures are fixed and unchanging; they and their components cannot be created or destroyed (from a perspective external to the structures). This latter notion of necessary existence seems to me most appropriate for analyzing the Leibniz cosmological argument, in which case (c) is false but the conclusion is true (all causes are necessary). Certainly my judgment of how to interpret necessary existence in this context can be challenged, but the above response to the Leibniz cosmological argument does appear to be worth investigating. A final remark which will be pursued further when we examine the next argument is that the necessary cause which the Leibniz cosmological argument need not have any connection with God.

(iii) Kalam cosmological argument

(a) Anything which begins to exist has a cause

(b) The universe began to exist

Conclusion The universe has a cause.

This is perhaps the best-studied argument for the existence of God in recent Christian philosophy. Let us begin by examining the first premise. Whether this is true depends on the definition of cause, which unfortunately is far from clear. For example, do virtual particles have causes? One could define the notion of cause such that the first premise is true, but it is not clear whether such a definition can be made to correspond to intuition. Let us grant the first premise for now. In order to understand what we have granted and to begin examining the second premise, we must understand what it means for something to begin to exist. A straightforward definition would be that x begins to exist at time t iff x exists at time t and for every time $s < t$, x does not exist at time s. However, this does not appear to make sense in the case of the universe, because in order for there to be time, the universe must exist, and so there is no time before the universe existed. Thus the second premise seems highly suspect. Even if we grant the whole argument, this does not at all lead to the conclusion that the cause of the universe must be God. We could define it to be God (keeping in mind that there could be many causes, and so we may end up with many gods), but then we are simply using an unfamiliar meaning of the word God. However, assuming the intuitive notion of God is coherent, it is possible that this being exists and created the universe, thereby causing it to exist. Consequently this did happen in some pasts of our universe (rendering any consistent account of creation true in a weak sense), but it is not at all clear that the probability of such a past should be high (or even nonzero).

(iv) Teleological argument

It is clearly possible for life to come into being without a designer (after all, one possibility for the history of our universe is that it popped into existence five minutes ago, completely uncaused), and so a designer is not necessary for life. One might still ask whether it is most probable that life was designed; this seems to me a task for scientific argument, and I have nothing further to add here.

(v) Moral argument

The argument is that adherence to a system of morality cannot be normative without a being which determines what is moral. On mathematical metaphysics, any system of morality can be defined (say by saying which actions in all possible worlds are moral according to this system, and perhaps as a soft concept), and no being is needed to determine this definition. As to why a particular definition used is normative, our minds have moral sensibilities as part of their structure, and so our personal morality is determined from our minds, and morality in a society is determined by the interactions of many minds, etc. This does imply that morality is relative (what is considered moral by one mind may not be by another), but such relativity does not necessarily imply loss of normativity; it merely implies that normativity is also relative, which is certainly plausible (it is obligatory to remove one's shoes before entering a mosque but not before entering a church).

4.4 Religious pluralism

On mathematical metaphysics, minds are simultaneously embedded into all worlds which contain their structure, and these quite clearly include worlds where the central premises of various religions are true. Thus one can meaningfully say that the religions are true, and this could perhaps be given as an argument for practicing them. However, one must keep in mind that in the same sense we are in a swamp world; probabilities do have bearing on our experience. Thus one might ask what is the probability of various religions. I should say for most religions it is significantly lower than that of worlds which developed with histories considered scientifically plausible, simply because most religions posit events which would be considered highly unlikely if one were to claim that they happened today. Thus perhaps we can reject most religions on the same grounds that we reject the hypothesis that we are living in a swamp world. What is this rejection? It is not rejection of the veracity of religions, because the religions are true in the sense made precise above. It is also not rejection of the practice of religions, which is largely a cultural phenomenon. Rather it is rejection of the scientific truth of most religions, where scientific truth means roughly truth in a mathematical model with structure such that the probability that our universe is such a structure is high (see discussion in section 4.2). It should be noted that it is possible for a religion to make no claims which are scientifically unlikely, in which case our argument in this paragraph has no bearing on deciding whether it is true in the 'real world,' a soft concept which we nevertheless have a good intuitive grasp of.

One might well object that this is simply an intuitive judgment against religion, with no more force than the judgment of the religiously inclined that their religion is true with high probability. This objection is valid; however, it seems to me much more difficult to argue that any particular religion is true with high probability than to note that events which would today be considered to have very low probability are assumed by most religions to have occurred, and so by the intuitive principle that the past should be similar to the present, conclude that the probability of most religions is low.

However, a believer may certainly find his religious experience sufficiently compelling to warrant belief, and in fact mathematical metaphysics can be used to provide support for such a belief, especially if it does not entail any specific physical predictions. The line of reasoning is the same as that in the Matrix objection from section 2.3: Assuming the believer's religious beliefs are consistent with his experience (if not he had best revise them), there are possible worlds in which he is embedded and such that the claims of his religion are true. If he regards these worlds as the most relevant, there is a well-defined sense in which he may say his religion is true in the 'real' world (which intuitively must certainly be relevant). In fact, it is conceivable that a religious person could strengthen his faith by concluding using mathematical metaphysics that it is true, which requires only the very weak assumption that it is consistent (having already made the strong assumption that mathematical metaphysics is true), and then supporting the apparently very lenient claim that it is relevant to himself (and perhaps others).

4.5 Relativity of Reality and Action

On mathematical metaphysics, everything consistent with my current experience actually does occur (in some of the worlds into which my mind is embedded), and so if I do something and then immediately claim I did not do it, my claim is in a sense true (since there are possible pasts where I did not in fact do it, and everyones memory is failing in such a way that it appears I did). Thus when we say what happens, we generally do not mean this simpliciter, but rather what happens with high probability. Why should high probability, rather than truth simpliciter, be what is most important to us when evaluating everyday situations? It appears that this is because what happens with high probability is what matters for our experience; what happens with high probability is what actually happens (with rare exceptions).

5 Conclusion

We have seen that mathematical metaphysics is a natural alternative to nominalism for eliminating the mysterious Platonic bridge between the mathematical and the physical. We saw in section 1 that this position can be stated very precisely, in section 2 that it meets intuitive objections brought against it (it should be clear from the discussion in that section how a wide array of similar objections can be met), in section 3 that it has sufficient advantages to be worth studying, and in section 4 that this position has consequences which could well be relevant to society and the everyday lives of non-philosophers. Both Platonism and nominalism have been studied extensively for centuries, while mathematical metaphysics has almost no adherents and nearly all are from this century (Tegmark [5] comes to mind), so one should expect to find the position of mathematical metaphysics far less developed than the Platonist and nominalist positions. The fact that mathematical metaphysics provides such a novel perspective on many traditional philosophical problems is reason in itself to study it alongside the more traditional metaphysical stances on mathematical objects, and if it does turn out that mathematical metaphysics provides a much more robust metaphysics than Platonism and nominalism, perhaps philosophers will eventually recognize this and the position will become more common.

References

[1] Roman Frigg and Ioannis Votsis. Everything you always wanted to know about structural realism but were afraid to ask. *European journal for philosophy of science*, 1(2):227–276, 2011.

[2] Kenneth Sylvan Guthrie and David R Fideler. *The Pythagorean sourcebook and library: an anthology of ancient writings which relate to Pythagoras and Pythagorean philosophy.* Red Wheel/Weiser, 1987.

[3] Charles McCarty. Completeness and incompleteness for intuitionistic logic. *J. Symbolic Logic*, 73(4):1315–1327, 2008.

[4] Thomas Nagel. What is it like to be a bat? *The philosophical review*, pages 435–450, 1974.

[5] M. Tegmark. The Mathematical Universe. *Foundations of Physics*, 38:101–150, February 2008.

[6] Eugene P Wigner. The unreasonable effectiveness of mathematics in the natural sciences. richard courant lecture in mathematical sciences delivered at new york university, may 11, 1959. *Communications on pure and applied mathematics*, 13(1):1–14, 1960.