# Introducing a

# non-physicalist conception of panpsychism

[Dissertation for an MA in Philosophy; autumn 2009]

🙞 🙝 🙟 🙜

Introduction 1

1. Physicalism 2

2. Existence 15

3. Idealist panpsychism 30

4. Idealist panpsychism and the physics of our universe 38

5. Physicalism versus idealist panpsychism 43

6. Conclusions 45

References 46

Table 2.1 Definitions of existence 16

Figure 3.1 A dodecahedron 31

Figure 3.2 A percept 34

# Introduction

The aim of this dissertation is to provide an elementary example of a panpsychic universe. It is one in which – in contrast to physicalism – the experiential (defined in terms of the possession of a qualitative, subjective, structured percept) is the foundation of existence. In this panpsychism, physics is secondary because it is arrived at by collating the experiences of a collection of experiential entities as they perceive one another. Before presenting this example I need to do some preparatory work.

Section 1 discusses physicalism, the currently-dominant metaphysical viewpoint. It is largely devoted to a detailed critique of physicalism as characterised by Jeffrey Poland in his book “*Physicalism: the philosophical foundations*” (1994), but other physicalist viewpoints are discussed.

Section 2 follows on from this. I give detailed arguments that the physicalists’ conceptions of physical existence are ambiguous and inadequate, and that we need to have some clear definition of actual existence to distinguish this concept from mere mathematical existence. I conclude by *defining* actual existence in terms of experiential existence: the property of having a qualitative, subjective, mathematically structured percept of other entities that actually exist. This definition clearly states what it is for an object to be instantiated. Although it is counterintuitive – because it entails panpsychism – the definition is not in conflict with either what we know from science, or with our everyday experience.

Section 3 first introduces some contrasting ideas of Max Tegmark, who believes that we and the world we live in are nothing but an (incredibly complex) formal mathematical system: “instantiation” is an empty concept for him. He also proposes that all formal mathematical objects such as a dodecahedron have actual existence in a Platonic realm. Most of section 3 is devoted to my dissimilar panpsychic universe with a dodecahedral physics, which might conceivably *actually* (rather than just mathematically) exist. This universe is extremely elementary, and a good deal would need to be added – not least a time dimension – before we arrive at a universe in any way resembling our own. (This is why I say “Introducing” in the title of this dissertation.)

Section 4 very briefly sketches how the strange properties of quantum mechanics might begin to explain the no-sign and combination problems well-known to panpsychists. The no-sign problem asks, “Why is it that physicists see no sign of mind in the lawful behaviour of elementary particles?” The combination problem asks, “How do the simple minds of elementary particles combine to give a more-or-less unitary sophisticated mind in the brain of a human being?” The section goes on to show how the conception of panpsychism proposed here might be testable – at least in principle.

Section 5 contrasts physicalism with the panpsychism proposed here, and Section 6 concludes the dissertation.

# 1. Physicalism

Jeffrey Poland’s (1994) book is uniquely valuable in being entirely devoted to a study of the foundations of the metaphysical project of physicalism. Here I will summarise some aspects of physicalism as stated by Poland, contrast this with a more reductive conception of physicalism proposed by Papineau (in Stone & Wolff, 2000), and then provide my own critique of Poland. (In this section, references given as page numbers only are to Poland (1994).)

### Formulation of physicalism

Poland’s chapters 1 and 4 are devoted to the formulation of physicalism. His chapter 1 discusses three core ideas and values of the physicalist project:

First, with respect to ontology, the physicalist holds that everything that exists or occurs in nature is *ontologically dependent upon* the physical domain: that is, nothing can exist or occur in the absence of physical objects, attributes, and events.

(Emphasis in the original, pp. 14-5)

[Second, …] physical truths and facts provide the conditions [or base] for all *objective* truths and facts. This has been expressed in recent years by such claims as that the physical facts determine all the facts, that there is no difference between individuals without there being a physical difference between them, and that the physical truth determines all truth.

(Emphasis in the original, p. 19)

Here truths and facts are understood to be “objective” when “they exist independently of human activity, interest, knowledge, etc” (p. 20).

The third [key idea of physicalism] is that of a unified explanatory system in which different branches of knowledge are organised hierarchically with physics at the foundation, and in which the generalisations and phenomena studied at each level in the hierarchy are explainable in terms of generalisations and phenomena at lower levels. In such a hierarchy, physics is seen as the ultimate repository of mechanisms, processes, states, etc. that can be appealed to in an explanation of why some phenomenon occurs or why some regularity holds. […] *all such appeals* are ultimately grounded in phenomena and regularities at the physical level.

(Emphasis in the original, p. 21)

His chapter 4 presents physicalism in terms of a number of theses. Among them are:

(T0) All objects and attributes that are (or can be) instantiated must be instantiated in regions of space-time.

(p. 190)

(T2) For any possible complete distribution of the physical facts, there is exactly one possible complete distribution of the non-physical facts.

(p. 203)

His thesis (T3) (p.205) is somewhat technical, and will not be stated here, but it has the consequence that “A sentence *s*, formulable in some language Li, has a legitimate claim to objective status only if its truth value is fixed once the physical truths are fixed…” (pp. 205-6).

### Earlier formulations

Poland’s chapter 2 (pp. 10-108) amounts to over a quarter of his book, and is a detailed critique of earlier formulations of physicalism, which he finds to be inadequate. My discussion of this chapter will be brief, both because of lack of space, but also because I am largely in agreement with him. Here I will only summarise what is needed to understand Poland’s own views, and to develop my own.

He defines *classical reductionism* in terms of two theses: “Every term in the languages of the special sciences is definable in the language of physics” and “Every law in the special sciences is derivable from the laws of physics and the definitions” (p. 46). Poland gives psychology, biology and economics as examples of the special sciences (p. 64). Classical reductionism is based on the false view that science consists “of a single, total theory in one language and as embodying a set of basic principles that are sufficient for all theoretical purposes. [Physics alone constitutes] the total explanatory system of science” (p. 11).

Poland (contra classical reductionists and opponents of physicalism) denies that physicalism implies either that all facts are physical facts, or that all facts can ultimately be expressed in the language of physics (pp. 60-65). Examples of non-physical facts are: “The sky looks blue”, “She ate because she was hungry”, “The economy is in the doldrums”, “genocide is bad”, “Judges have authority to try cases” and so on (based on Poland, p. 344).

Poland is opposed to classical reductionism because it amounts to eliminativism about such non-physical facts (p. 11). He criticises the eliminative materialism of the Churchlands (pp. 221-222, 295-296). Identity theories such as attribute identity theories in which “the attributes of individuals in the domains of the special sciences are identical to physical attributes” (p. 66) also eliminate experience (pp. 66-70). These theories are eliminative in effect, even if their proponents disagree. Poland’s overall conclusion is that classical reductionism is “disastrous” (p.65).

Poland also considers non-reductive theories of which there are many and various kinds. Several are based on the idea of supervenience. This can be defined in several ways and Poland discusses definitions of global, local and weak supervenience (pp. 78-104). One possible definition of global supervenience in this context is that a minimal physical duplicate of our world is a duplicate of our world in all respects. In other words, the totality of all physical facts in our world determines all the facts. The “minimal” requirement is needed to eliminate extra, epiphenomenal (and hence dubious) objects such as ectoplasm. Poland, rightly in my view, rejects supervenience as being too weak and to abstract to form an explanation (p. 105). Supervenience of any kind can only describe how non-physical entities and attributes covary with the underlying physical dynamics. Description of covariation does not amount to explanation (p.105).

### Definition of physics

Poland’s chapter 3 is an attempt to characterise what it is to be “physical.” Without an adequate characterisation the physicalist project is clearly vacuous. “Strategically, the problem for the physicalist is that, unless there is some antecedently specifiable principle for specifying the physical bases, physicalism cannot be formulated in a significant way” (p. 119).

Poland gives a lengthy discussion of attempts to provide such bases, and objections to them (pp. 109-23). He sums up by listing thirteen principles that should be followed in defining the physical bases (pp.123-4). Among other things, these principles require that the physical bases: should be as narrow as possible; should be in terms of physics as distinct from other sciences; should allow for the development of physics away from current physics; and be free of vagueness. He goes on to give what might be regarded as his partial definition of physics:

[P]hysics is the branch of science concerned with identifying a basic class of objects and attributes and a class of principles that are sufficient for an account of space-time and of the composition, dynamics, and interactions of all occupants of space-time. The crucial features of these classes are that they are minimal with respect to the descriptive and explanatory purposes they serve, that the magnitudes are defined for all regions of space-time, and that each occupant of space-time satisfies the principles governing these magnitudes. […] The minimality condition in effect keeps physics from being gratuitously identified with all of natural science or with physics plus chemistry.

(p. 124)

He goes on to say that physics is a “universal” science in the sense that it “needs to be understood in terms of kinds of questions that physical theories typically attempt to answer”, namely:

* What are the fundamental constituents of *all* occupants of space-time?
* What are the fundamental processes that underlie *all* causation and *all* interaction between such occupants?
* What parameters are relevant to the dynamic unfolding of *all* systems in space-time and hence to *all* change?
* What is the nature of space-time itself […]?

(Emphasis in the original, p. 125)

### Vertical explanation

Poland’s essential idea is:

[A]ll phenomena are to have explanations in terms of physically-based explanations. A system of knowledge structured by such vertical explanations will be a physically-based explanatory system in the sense that all explanations, *of whatever sort*, will only make appeals to objects and attributes that are ultimately grounded in the physical domain, via a chain of vertical explanations.

(Emphasis in the original, p. 207)

He presents this idea as a thesis:

(T4) All instantiations of non-physical objects and attributes are vertically explainable in terms of physical or physically-based objects and attributes.

(p. 208)

He goes on to say that vertical explanations provide answers to questions of the following form:

1. In virtue of what lower-level attributes did the instantiation of such-and-such higher-level attributes occur?
2. How did such-and-such lower-level attributes constitute an instantiation of such-and-such higher-level attribute?

[Likewise for objects]

(p. 208)

### Realisation theories

Poland introduces the central idea of a *realisation theory*:

The idea is that of a theory, associated with a given attribute N, that abstractly characterises the kinds of attributes that are sufficient for N and *that shows how such attributes can combine to actually constitute N in specific cases.* […] *In all such cases an appeal is made to a theory of how the instantiation of certain relevant physically-based attributes suffices for the instantiation of the target non-physical attribute.*

(Emphasis added, p. 210)

Realisation theories are not themselves explanations, but are used to generate vertical explanations:

For the physicalist, given a realisation theory[[1]](#footnote-1) for the attribute N, and given a distribution of physically-based attributes, vertical explanations can be generated to account for the specific instantiations of N, if they occur. Unlike the vertical explanations they help to construct, a realisation theory for an attribute is a generic, abstract account of how an attribute can be constituted.

(p. 210)

Poland says that realisation theories are “essential for effectively answering questions (1) and (2) above” and thus central to the physicalist project, (pp. 210-1). Without endorsing functionalism, Poland uses it as an example of a realisation theory that attempts to account for mental phenomena (pp. 211-3).

### Poland on identity, emergence and realism

Poland rejects identity theories – sometimes thought of as a variety of physicalism. Even if identity theories work in certain limited cases (which Poland doubts) they do not work generally (p.198). Rather than *identity*, he asserts that physicalism is to be understood in terms of *realisation*. “Certainly physicalists want to say that the actual world is a physical world, but not in a way that denies the existence of non-physical objects and attributes, and not in a way that identifies such entities with physical entities,” (p. 309). Poland gives as examples of non-physical objects chairs, social institutions, works of art, qualitative mental states, and values, among others (pp. 196, 344).

Regarding emergence, Poland states of objects:

From the ontological perspective, physicalists have always been concerned to rule out ‘ontologically emergent’ phenomena of any sort (for example, spirits). Thus the physicalist holds the view that the ontology of physical theory is exhaustive of all that exists in the sense that there are no objects that are neither basic physical objects nor objects ‘built up’ out of, or ‘realised by’, basic physical objects. […]: all objects that exist do so in virtue of physical objects. There are no ghosts!

(p. 15)

Poland goes on to give a more complex argument against the existence of emergent *attributes* as opposed to entities (pp. 16-17). However he does allow for there to be emergent physical *laws* that only come into play when physical configurations of a certain complexity are realised (say within brains). He describes such emergent laws as “innocuous” (pp.116-117).

Poland’s discusses realism at some length. He says:

[P]hysicalism does imply a form of scientific realism. An instrumentalist or fictionalist account of scientific theories in general and of physics in particular does not square with the physicalist’s concern about ontological matters […]. The basis for all objective fact and truth and of all entities and influences must be real, not just a convenient fiction and not something about whose existence we need have no beliefs.

(p. 355)

However, Poland is neutral between what he calls metaphysical realism and non-realism. He characterises the difference between these concepts as follows:

In the case of the [metaphysical] realist there is only one actual world, whereas in the case of the [metaphysical] non-realist there are many. But in all cases, what is an objective fact of the matter is internally pinned down by the physical basis operative in a given actual world.

(p. 326)

Moreover:

[P]hysicalism is indeed a doctrine that is developed in a parochial manner: it is concerned with one world at a time, and in the present case, it is our own. Thus other possible worlds with a physics different from the physics of this world do not fall within the scope of physicalist theses as framed by us.

(p. 150)

Poland thus unsatisfactorily defines metaphysical non-realism/realism according to whether or not alien universes – causally disconnected from ours and with their own distinct physics – exist. This makes a decision between metaphysical realism versus non-realism a matter of faith. A metaphysical non-realist according to Poland’s definition is not being sceptical or ‘non-realist’ about the existence of anything in our universe. In the next section I will explain how Poland seems to have arrived at this mistaken definition.

Poland claims that all scientifically realistic interpretations of science “can be trumped by a determined opponent of metaphysical realism. For example the existence of physical entities can be one or another sort of ‘dependent’ existence, be it dependence upon mind or theory or language” (p. 356). But this last sentence has nothing to do with metaphysical non-realism as Poland defines it, because it is clearly intended to speak of the character of physical entities *in this world*. What Poland calls “a determined opponent of metaphysical realism” is in fact a determined opponent of physicalism.

Poland affirms that most physicalists are metaphysical realists (p. 184), and states that “Within a thoroughgoing metaphysical realist framework, the evidential equivalence of incompatible theories does not spell disaster for physicalism, since the independent world [in which we live] will determine which (if any) of the evidentially equivalent theories is true” (p. 154). Throughout the dissertation I will assume this thoroughgoing metaphysical realism, which Poland characterises here in terms of the facts of our world.

### Critique of Poland

The bulk of my critique of Poland’s physicalism falls under two major categories: a critique of his concept of realisation theories; and a critique of the adequacy of his concept of ‘the physical.’ The latter critique will apply to physicalists generally. Some criticisms fall outside of these categories.

* Poland’s concept of metaphysical realism is mistaken

This has already been shown above, but I wish to explain how the mistake has arisen. Poland imagines the following situation: “What if there is no fact of the matter as to which of the [incompatible] evidentially equivalent theories is *the true physics*? (i.e. what if no realist scenario is viable?)”, (Emphasis in the original, p. 154). He goes on to say:

My view is that, in such a situation, it is possible to develop our total theory of nature and our physicalism on the basis of either or both of these competing theories. Within some non-realist philosophical settings, two incompatible theories that are each evidentially correct might be viewed as, for example, ‘being true in different worlds’. [He cites Goodman (1978).] And, in such a philosophical setting, physicalism is a doctrine ‘internal’ to each system of knowledge and internal to each world. Thus ‘in each world’ or ‘relative to each theory’ the [physical] bases will be determinate […]

(p. 155)

This whole discussion is predicated on the assumption that no realist scenario is viable and that non-realism is true. When Poland speaks of a “world” in this context, this is more accurately understood as shorthand for “a particular domain or portion of our experience relative to which a particular theory is evidentially correct.” So in this setting a “world” is a portion of *our* universe – the non-realist world of our experiences, taken as a whole. In particular, a “world” as defined here is not an alien universe, causally disconnected from our own.

For a non-realist, it is reasonable to say that in some sense there might be several incompatible, true theories of physics, each applying to a “world,” where “world” refers to a particular domain of our experience. For the realist physicalist, however, there is only one physics that is true of our world, where “our world” is understood in the more usual sense of the external universe in which we live. (A realist physicalist can be neutral as to whether alien universes, causally disconnected from ours, exist.)

The two usages of the word “world” in the first two sentences of the previous paragraph are entirely different: the non-realists’ “particular domain or portion of our experience” is nothing like a “world or universe” as understood by realists. Poland combines these two sentences, while conflating these two wholly distinct usages of the word “world”, to arrive at his mistaken definition of the difference between metaphysical realism and non-realism (p. 326, quoted at the top of page 7 above).

* The definition of a realisation theory is to a certain extent question-begging

The first part of Poland’s definition of a realisation theory (p. 210, quoted on page 5 above) is that it “abstractly characterises the kinds of attributes that are sufficient for N.” The second part of the definition, which I have italicised above, asserts that a realisation theory “shows how” and is a “theory of how” N is constituted/instantiated. This second part therefore asserts that a realisation theory amounts to an explanation, but it does not state the manner in which this explanation is achieved. But, of itself, an *abstract characterisation* of the kinds of attributes A that are sufficient for N does not amount to an *explanation* of why N occurs whenever A occur. This will become clearer when I discuss the particular example of functionalism below.

* There are no examples of successful realisation theories that are not cases of classical reduction

Poland promotes realisation theories (pp. 210-216) as a broader alternative to what he agrees is the failed concept of classical reduction. He discusses two examples: transparency and functionalism. With regard to transparency, this, insofar as it has been explained, has been explained by classical reduction – in terms of the physical structure of a piece of matter that allows light an unobstructed spacetime path through it. Functionalism with regard to mind is a realisation theory that is not merely a case of classical reductionism. I hope to show that functionalism fails as a satisfactory realisation theory after giving Poland’s account of it:

For example, functionalist accounts of mental states delineate the pattern of causal relations which, if instantiated in some physical or non-physical system, suffice for the instantiation of a given mental state, M. The realisation theory in such a case consists of a characterisation of the required pattern of causal relations [while being neutral as to how these relations are implemented.] And the realisation theory effectively clarifies what it is that makes it the case that the specified pattern of causal relations, rather than some other, constitutes M. Then, for a given instantiation of the mental state in question, the *physicalist* vertical explanation would consist in citing a set of instantiated physically-based causal relations and showing, by appeal to the functionalist realisation theory, why those relations suffice for the realisation of the mental state.

(Emphasis in the original, p. 211)

He goes on to say that a realisation theory for an attribute N, “while specifying what it takes to instantiate N, does not normally make any reference to physical attributes. Rather, a realisation theory specifies something like the essence of an attribute, although I don’t wish to hang too much on the use of that term” (p. 211). He sums up: “According to the functionalist, the essence of a mental state, M, is a certain pattern of causal relations. Such patterns can be exhibited in a physically-based world of the sort envisioned by the physicalist […]” (pp. 211-212).

It is here that the physicalist/functionalist realisation theory, and indeed functionalism in general may be called into question. The mental state M is something whose essence is, at least in part, given to us when we experience it. M might be the qualitative experience of a toothache. But the functionalist realisation theory for M “specifies something like the essence” of it in terms of its pattern, P, of causal relations: certain patterns of neural firing and chemical processes in the brain, causally related to increased heart rate, sweating, screams and complaints of pain, withdrawal, running away, and so on.

If the proponent of such a realisation theory claims that mental states M are in essence no more than these causal relations P, then he or she is denying that qualitative, subjective feelings of pain are real. The same applies if these causal relations P constitute M. In this case the qualitative, subjective feeling of pain has been “explained away” to be replaced by P. On the other hand, if the realisation theorist does not make the claim that the essence of M is given by P, (or that M is constituted by P), then he or she has not explained M, but has merely given an abstract characterisation of those causal circumstances P in which M appears.

The success of functionalism as an explanation is thus at least questionable; and even Poland himself states that he does not endorse functionalism (footnote, p. 211), and further calls the theory “sadly incomplete and defective” (p. 61). As Chalmers notes (emphasis in the original), “When it comes to conscious experience, this sort of explanation fails. […] *Why is the performance of these* [cognitive and behavioural] *functions accompanied by experience?*” (1995, p. 203).

Poland’s only examples of realisation theories are transparency, which insofar as it succeeds does so because it is an example of classical reduction, and physicalist functionalism, which fails as an adequate explanation. He gives no widely endorsed example of a non-trivial physicalist realisation theory for any branch of science or of human knowledge. This calls into question whether successful realisation theories that are not just classical reductions exist. Related to this is:

* In the absence of a classical reduction, realisation theories do not amount to explanations. At best they describe what happens to be true in our universe

Some non-physical theories might turn out to be translatable directly or indirectly into physics. It is plausible that chemistry might one day be fully reducible to physics, and it is at least reasonable that biology (ignoring mind) might one day be reducible to chemistry. If this is so, then a world having an identical physics to ours would necessarily have the same chemistry and biology. When non-physical sciences can be translated directly into physics, then a vertical explanation really does explain.

Moreover, today we have partial vertical explanations between biology, chemistry any physics, but these partial explanations only do the work of explanation insofar as they are classical reductions, as in the case of transparency above. For instance the role of blood in biology is made clearer by understanding that it contains haemoglobin, and that this molecule interacts with oxygen according to certain physical principles, and so on.

However, Poland admits the possibility that some facts about the world might not be translatable (even indirectly) into the language of physics (p. 65). In this situation – precisely the situation where classical reductions fail and Poland makes use of his concept of a realisation theory – it is difficult to see how any realisation theory could provide a vertical *explanation*. Facts about subjective, qualitative experiences provide a plausible example of such a situation:

Let S be the statement “A brain in a particular biological state X is, by virtue of that state, having a particular subjective, qualitative experience Y loosely describable as ‘the taste of cherry brandy’ because SS(XX, YY).” Here SS is a sub-statement involving biological and other objective facts XX, and experiential facts YY. By hypothesis, Y and YY are not translatable (even indirectly) into the language of physics. S is part of a realisation theory which asserts how (in one particular instance) a phenomenological fact Y is realised in biology as X.

Now, for Y to be explained in terms of X there must be some sort of linguistic relationship between Y and X. This is because an *explanation* is by its nature a linguistic means of intellectually ‘grasping’ Y in terms of X. Without any (at least partial) translation or identity or other comprehensible relationship between X and Y, there can be no ‘grip’ and thus no explanation. For this reason statement S, despite containing the phrases “by virtue of” and “because” cannot form part of an explanation of Y in terms of X.

* Unpredictable emergent laws are cannot form the basis of an explanation

Poland asserts that there may be non-physical law or laws L “that are not derivable, and hence not predictable, on the basis of knowledge of lower-level laws,” moreover they might “‘kick in’ at certain stages of the development of the universe. Realisation theories and vertical explanations based upon them will enable us to understand why certain phenomena [P] occur […] only when certain kinds of structure have evolved. Neither of these sorts of emergent phenomena falsify physicalist theses of course” (p. 223).

Law L is an essential part of the realisation theory which generates the vertical explanation that supposedly explains P. The explanation would be some argument, based in part on the premise that law L holds, that arrives at the conclusion that P occur. But why should anyone accept the premise that law L holds, especially given that it is neither derivable nor predictable from lower-level laws? The explanation fails because the argument on which it is based cannot be shown to be sound. It clearly cannot even be shown that the premise that L holds is at all plausible.

* Given Poland’s assumptions, and given the failure of realisation theories, Poland fails to explain the coming-into-being of non-physical entities, and he is wrong about identity

Poland believes that entities such as qualia, social institutions, works of art and values are ontologically real, but are not in themselves physical. That they are non-physical shows that he is not an identity theorist. His claim that these novel entities are “non-emergent” is based on the assumption that all non-physical realities have an explanation as being “realised by” basic physical objects and exist “in virtue of” those objects. These explanations are given by way of realisation theories.

However, this is not the way the word “emergence” is usually used. Early in the history of the universe its ontology did not contain these non-physical entities, whereas nowadays it does. This fact is sufficient for us to say that, in the usual sense of the word, these novel entities “emerged”. (Poland is accepting explicable emergence, but rejecting inexplicable emergence.) Here I use the neutral term “coming-into-being” to mean “emergence in the usual sense”. Poland accepts the coming-into-being of non-physical entities such as qualia and social institutions. But until he gives a successful realisation theory for these entities, Poland has failed to account for their coming-into-being over the course of the history of the universe.

Poland fails to respond adequately to what David Papineau calls the *causal argument* for a strongly reductionist version of physicalism. It goes: P1. The physical world is causally closed; P2. All mental occurrences have physical effects; P3. There is no causal over-determination. Conclusion: Mental occurrences must be identical with physical occurrences (Stone & Wolff, 2000, p. 180).

The argument is valid but Poland rejects the conclusion. Which of these premises would Poland reject? Poland explicitly accepts P1: he defines physics as involving the study of “the fundamental processes that underlie *all* causation” of all occupants of space-time (Emphasis original, p. 125).[[2]](#footnote-2) He also explicitly accepts P3: “[I]f it is a high-level regularity that non-physical attribute A is the cause of non-physical attribute B, then an instance of this regularity is explained by citing the physical attributes C and D, which *realise* A and B respectively on that occasion, and by explaining that C causes D. [Footnote:] Among other things, C and D constitute the causal powers of A and B on this occasion” (emphasis original, p. 219).

So Poland must reject P2. If fact, he should logically be driven to say *a fortiori* that “There is no mental causation independent of the underlying physical causation.” Papineau claims (I believe correctly) that rejection of P2 amounts to either epiphenomenalism or to pre-established harmony (Stone & Wolff, 2000, p. 181). Both of these positions are extremely uncomfortable to hold, and Poland says nothing in favour of them. Rather, Poland’s talk of causation at the physical level “constituting” causation at the mental level is simply wrong. If A is a scent of honey, B a desire to eat the honey, C and D are the underlying neurological states, and c the causal chain of physical events from A to B; and if all of these things are ontologically real and distinct, then there is no causation from A to B. One can say that A is followed by B, but this is a regularity, not causation.

Poland seems to be taking a position similar to John Searle’s *biological naturalism* which Searle presents in the form of four theses that are summarised here (Searle, 2004, pp. 113-114):

1. (a) Conscious states are real phenomena. We cannot do an *eliminative reduction* of consciousness, showing it to be an illusion. (b) Nor can we do an *ontological reduction* (p.119) of consciousness showing it to be nothing but its neurobiological basis, because this would leave out the subjective, qualitative and intentional features of conscious states.
2. Conscious states are entirely caused by lower level neurobiological processes in the brain. They have no causal powers of their own, independent of the neurobiology. Conscious states are thus *causally reducible* to neurobiological processes.
3. Conscious states are realised in the brain as high-level features of the brain system. Individual neurons are not conscious, but portions of the brain system composed of neurons are conscious.
4. Because conscious states are real features of the world, they function causally. My conscious thirst causes me to drink water for example.

Searle’s position has been critiqued extensively. There are obvious tensions between 1(b) and 3 (fudging the conclusion of Papineau’s argument), and between 2 and 4 (fudging premise P2 of Papineau’s argument), for example. I will not discuss Searle’s theory further here as I critiqued it extensively in an earlier essay. Poland has failed to rebut Papineau’s causal argument that concludes by identifying the mental with the physical.

(My own rejection of Papineau’s argument is sketched as follows: we know hardly anything of the physical world except its abstract mathematical structure; but toothaches are real and concrete; and all we know of alleged ‘physical causation’ is experienced lawful regularity. We should reject P1 for these and other reasons. Suppose we nevertheless accept the conclusion that the mental and physical are identical. In this case we should identify the (abstract and little-known) physical to be in truth the (concrete and immediately-given) mental, rather than the other way around. I will flesh out this argument throughout the rest of the dissertation.)

* Is the term ‘physical’ epistemological or ontological?

What does Poland mean by the claim that some object or attribute is ‘physical’? Is it something characterised empirically – as something that we know by means of the human discipline of physics, or is ‘the physical’ something metaphysical and ontologically fundamental – the ‘physics of the world’ in which we live?

There is a difference between these two concepts. Poland admits that there might be physical entities that we could not in principle detect. For example, the Planck length is about 1.6 × 10-33 cm. Quantum general relativity “predicts violent fluctuations in the [spacetime] geometry at distances of the order of the Planck length,” (Misner *et al.*, 1973, p. 10). Stephen Hawking states, “To probe to shorter distances would require particles of such high energy that they would be inside black holes,” (2001, p. 176). Most physicalists would want to be realists about spacetime, but these facts imply that spacetime itself can never – even in principle – be measured at the finest scales.

* The ontological physical basis of the universe cannot be identified unambiguously from a study of the language of physical theory alone.

Poland believes that we can, at least provisionally, read off the ontological physical bases – what actually exists in the universe – from the current theory of physics. He says that the final stage of doing this is “identification of an ontological structure that provides an intended interpretation of the language of physical theory” (p. 148).

But this is not the case, and was never the case even in the era of classical physics. Let us suppose we are living in the nineteenth century, and are attempting to discover the ontological physical basis in the manner suggested by Poland. Newtonian physics, to the limits of our ability to perform measurements, seems to be a plausible candidate for the true physics of the world. There are at least two possible ontologies for Newtonian physics:

1. Pieces of matter act directly upon one another at a distance.
2. Gravitational fields actually exist. Each piece of matter has an associated gravitational field that extends throughout the whole of space. Each piece of matter responds locally to all of the other gravitational fields where it is situated.

In A. gravitational fields do not actually exist – they are nothing more than mathematical tools that we may choose to use in our calculations; whereas in B. gravitational fields actually exist and do some real causal work in the world. Even if we were indeed living in a Newtonian universe, no possible experiment could distinguish these ontologies. Similar ontological ambiguities are inevitable for all theories of physics.

* Poland, like most physicalists, does not adequately define physical existence
* Physicalism is not the only position one can take that is compatible with science

These last two criticisms require considerable discussion and I will examine them over the next two sections.

# 2. Existence

This section attempts to answer the question, “What do we *mean* when we say that our universe or an object within it *actually exists*, and how does actual existence go beyond or differ from mere mathematical existence?”

The concept of existence is one that you and I have everyday, commonplace notions about. You know that this piece of paper you are reading from exists, for example, and that Spiderman does not. Everyone is competent in speaking about existence, but it is a different matter to have a sufficiently clear understanding as to what existence means precisely. Can we define existence? What do scientists mean when they claim that quarks exist?

Moreover, what do philosophers mean when they claim that, say, mountains exist? Physicalist philosophers take the existence of objects as being basic, and so have tended not to define it. Poland characterises existence in terms of “space-time and of the composition, dynamics, and interactions of all occupants of space-time” (1994, p. 124) – for a fuller quotation see page 4 above. In this section I hope to complete my argument that Poland’s conception provides an inadequate foundation.

In response to Anselm’s ontological argument for the existence of God the (near) idealist Kant, rightly countered that existence is not a predicate:

I should have hoped to put an end to these idle and fruitless disputations in a direct manner, by an accurate determination of the concept of existence, had I not found that the illusion which is caused by the confusion of a logical with a real predicate […] is almost beyond correction. […]

‘*Being*’ is obviously not a real predicate […]

(Kant, 1781/7, A599-600/B627-628)

But Kant did not say what existence positively amounts to. On the contrary he claimed that the noumenon (the cause of our experiences) is, in any positive sense, completely unknowable (Savile, 2005, pp. 108-109; Kant, 1781/7, A249/B307).

The phrase “this book *is real*” is used in this section as a synonym for “this book *actually exists*.” Moreover, when the context is clear, I will sometimes abbreviate the latter phrase to “this book *exists*.” It is this concept of the concrete existence of our universe and of the objects within it that I am trying to pin down and capture in what follows.

### Four definitions of existence

We begin our analysis by examining four concepts or definitions of existence. These are: 1.*Experiential existence* of minds that possess subjective, qualitative, perceptual states. (We are beings with experiential existence, but there are other, much simpler examples). 2. *Empirical existence* of external objects that can be inferred by collating the percepts of experiential beings. 3. *Material existence* of entities obeying physical laws without reference to experiential beings. 4. *Mathematical existence*, which is merely a formal description or mathematical model that is logically consistent.

We will ask how each of these concepts of existence measures up to our everyday intuitions about actual existence. None of these definitions considered individually will be adequate, but several, taken together will turn out to provide a good candidate to capture the target notion of actual existence. The four concepts of existence are summarised in Table 2.1:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | ***Definition of existence*** | ***Examples*** | ***Claim to existence*** | ***Actually exists?*** |
| **1.** | **Experiential existence** | *Minds such as:*  *Human mind,*  *Earthworm’s mind* | Has experiences | ½🗸 |
| **2.** | **Empirical existence** | *Rock, Water, Quark,*  *Human as animal* | Can be experienced  as obeying precise, mathematical, physical laws | ½🗸 |
| **3.** | **Material existence** | *Rock, Water, Quark,*  *Human as animal,*  *Early universe* | Obeys precise, mathematical, physical laws independently of any concept of experience | 🗴 |
| **4.** | **Mathematical existence** | *Arithmetic,*  *Algebra,*  *Dodecahedron* | Logically consistent formal structure | 🗴 |

### Table 2.1: Definitions of existence

**1. Experiential existence**

**Definition:** to *exist* *experientially* is to exist as a mind that possesses qualia or qualitative experiences. Anything that has experiential existence will be called an *experiential being*.

Here, following Searle (1992, pp. 94-95), Strawson (2006, pp. 5-7), Poland (1994, p213), and Chalmers (1996, pp. 6-11) among others, I am taking a realistic view of qualia such as red, pain, and of all qualitative experiences. A *qualitative experience* is a set of qualia structured into a single percept (the qualitative experience of seeing a table for instance, without presupposing the table’s existence). According to this definition it is not necessary for a mind with experiential existence to have any cognitive abilities that enable it to reflect, to know that it exists, or even for it to have any thoughts at all.

Talk of a “being” raises a final point. In our language there is a “self” or “being” that “has” a qualitative experience and therefore appears to be distinct from it. If fact, the concept of a qualitative experience is incoherent without some minimal concept of a perceiver, for what is a qualitative experience with no perceiver? See Strawson (2006, p. 26 footnote 46). To talk of a qualitative experience is to talk of a being that has this experience, but, for the moment, nothing more is assumed to be known about this being. In particular, I am not assuming Cartesian substance dualism.

**2. Empirical existence**

When several human beings look at a particular rock they will have closely related percepts of it. They can discuss these percepts among themselves. It might be about as big as their fists and shaped like a potato for instance. This agreement is why they can recognise it with reasonable certainty as being a real object, and each can be confident that it is not merely a hallucination that he or she is having.

Science is based on essentially the same intersubjective process of collating human experiences. Early science was concerned with entities that were immediately observable, the mechanics of everyday objects, the flow of water, the motions of the stars, and so on. A major advance came when mathematical laws were found that connected these observations. The observed regularities were now more formal, and could be tested. We can now postulate the (empirical) existence of other entities based on intersubjective regularities in the percepts of experiential beings:

**Definition:** to *empirically exist* is to have the power to cause (directly or indirectly in any manner) systematic, intersubjectively consistent regularities in the percepts of experiential beings.

An example of an entity with empirical existence is the rock described above. The definition of empirical existence is framed in such a way that it is a characteristic of the putative entity itself, and not of the observers. Empirical existence makes a claim about the rock – it is causing experiences – and this is more than a claim about “what is going on in the minds of the observers.” A mirage in the desert is due to the empirical existence of sand, sunlight and hot air. It is not the empirical existence of water, even though the victims of the mirage may believe this.

Another example is the Moon. People looking in a particular direction into the night sky at a particular date and time have very similar experiences of seeing a shiny crescent. The rigorously-comparable character of our experiences leads us to conclude with confidence that there is some external cause of them which we call “the Moon.” Examples of intersubjective regularities are: We can study the Moon through telescopes and map its craters. We have a scientific theory describing mathematically how the Moon orbits the Earth, and all our observations agree with this theory. When we do not see the Moon in the expected location, there is always a clear explanation: “It’s cloudy tonight.”

As I have so far explained it, this definition is anthropocentric. But it is perfectly feasible for alien intelligences, either in our own universe, or in a different universe, to have their own science, and hence their own concept of empirical existence. There can even be empirical existence giving rise to sensation in the absence of sophisticated thought. Consider a group of earthworms passing a jagged rock. The rock, by virtue of its empirical existence, in particular related to its specific shape, will produce a consistent pattern of sensations in the simple minds of the worms.

**3. Material existence**

Cosmology and physics accurately describe the universe as it was long before the arrival of life, and it seems that the arrival of life (and hence, on the majority view, the arrival of experience) is a contingent fact about our universe. This has given rise to the need for a distinct concept of *material existence* or of *material entities*.

**Tentative definition:** *Material entities* are spatiotemporal configurations of matter-energy that interact with other material entities according to the mathematical laws of completed physics. Material entities all have *material existence*. Material existence amounts to actual existence, and material entities do not have to be observed in order to exist.

A water molecule is a candidate example of an entity that has material existence. We can safely assume that rain fell and rivers flowed into oceans long before the advent of life on Earth. Current physics seems to give us clear-enough examples of what material existence is – quarks, photons, electrons, spacetime itself, electromagnetic fields, and so on, all have material existence, and completed physics would provide us with a perfect catalogue of material entities, their properties and laws of interaction (perhaps constructed from p-dimensional membrane-like ultimate entities in a spacetime with extra dimensions; or perhaps something different (Hawking, 2001, chapter 7)).

Poland describes the ontological bases (1994, pp. 131-133) of physicalism. He summarises:

[T]he physical ontological base is completed by a characterisation of the class of all possible total [spacetime] distributions of physical objects and attributes […] subject to the constraint that the laws of physics are satisfied. […] One of these distributions [W] will correspond to the current total distribution of such objects and attributes; the others will be alternative total distributions that define the nature and limits of what is physically possible in this world.

(1994, p. 132)

It turns out that, despite its apparent clarity, the concept of material existence is problematic. W gives the current distribution of material entities in our world. Other distributions give alternative (possible but non-realised) distributions of merely-putative material entities. The question is: what distinguishes our actual universe W from its unrealised alternatives? Here are some possibilities:

* Material entities in W are instantiated whereas putative material entities in alternative, merely-possible worlds are not instantiated.

But what does the term “instantiated” mean? In the absence of a definition of this word we have not clarified the concept of material existence.

* Our universe W possesses a real spacetime in which material objects are instantiated (and this is what “instantiation” means), whereas “spacetime” in merely-possible alternative worlds is no more than an abstraction.

But now we have the problem of defining what it is for a spacetime to be real as opposed to being a mere abstraction. This is very similar to the problem of defining instantiation as above; the more so since (in contrast to Newtonian mechanics) general relativity treats spacetime itself as a material object playing a dynamic role in the temporal evolution of a physical system. Misner *et al.* for example speak of general relativity as “geometrodynamics” and explain “the effect of geometry on matter, and the reaction of matter back on geometry” (1973, pp. 3, 43). This is most spectacularly observed when gravitational waves in the spacetime geometry are generated by massive, rapidly orbiting binary stars (1973, chapters 35-37).

The problem of what it means for spacetime to be instantiated is essentially the same as the problem of what it means for a material entity to be instantiated.

* Our universe W possesses real causation whereas a merely possible world does not.

Here the problem is in defining “real causation” as opposed to abstract causation. Take the situation in our world where a cue ball strikes a red ball. Putting aside Humean doubts, the cue ball really does cause the red ball to move. But real causation depends in part upon the two balls being real. (In a mathematical model – a merely possible universe – a cue ball might strike a red ball “causing” it to move, but here the balls, movement and causation would all be no more than abstractions.) If we do not know what it is for the balls to be real, then we do not know what real causation is.

We could try to anchor the concept of real causation by linking it to those instances of causation giving rise to human experiences. (In this dissertation I am not entertaining sceptical doubts about this latter type of causation. Nor am I doubting that science can go some way to providing the links. When I am watching snooker there are indeed balls in the external spatiotemporal world in front of me. Light from these balls enters my eyes and eventually gives rise to my experience.)

John Stuart Mill made this influential suggestion. He defined material objects as being “the permanent possibilities of sensation;” (Berkeley, 1996, p. xxxv). But tying material existence to causation and then tying causation to human experience makes material existence depend essentially on human experience. Anchoring causation to reality in this way breaks the final clause in our tentative definition: “A material entity exists independently of being observed.” Material existence then collapses into empirical existence. This is undesirable to the physicalist for two reasons: (1) the physicalist intends to have a concept of existence that is independent of human existence; (2) it makes human experience more basic than material existence, and this is the opposite of what the physicalist desires for his or her project.

Having a precise definition of material existence is crucial to the project of physicalism because, for the physicalist, the totality of things that materially exist in our universe constitutes the ontological base of everything that exists. The bulk of the rest of this section will be devoted to showing that the concept of material existence cannot be modified in such a way to give it a clear and precise definition. If this is the case then physicalism cannot be based on a secure foundation.

**Mathematical existence**

There is no consensus as to the character of mathematical existence, but there have been two major approaches to it that might be of interest to physicalists. (The names of these approaches are my own):

The first, *Platonic realism*, whose advocates include Plato, Roger Penrose, Martin Gardner, and Kurt Gödel, asserts that mathematical objects have an existence in a special abstract realm, which exists independently of the existence of the universe (Penrose, 1989, pp. xii, 146-148). Plato, in his parable of the cave, argued that the physical world of our experience is merely the imperfect flickering shadow of the perfect mathematical world.

The second, *pragmatic realism*, asserts that mathematical entities are merely abstractions from observed regularities in our own particular universe. The Christian philosopher Bishop George Berkeley took this pragmatic view (1710, section 12), as did the thoroughgoingly secular populariser of mathematics Lancelot Hogben. The latter wrote, as part of a polemic against Plato:

This supremacy of the head is very flattering to intellectuals who have no practical problems to occupy them. […] An educational system which is based on Plato’s teaching is apt to trust the teaching of mathematics to people who put the head before the stomach […]. Naturally this repels healthy people for whom symbols are merely the tools of organised social experience, and attracts those who use symbols to escape from our shadow world in which men battle for the little truth they can secure into a “real” world in which truth seems to be self-evident.

(Hogben, 1936, p. 25)

When we use mathematics to describe “all possible worlds,” we are implicitly taking the Platonic realism approach, and are supposing that mathematical existence is ontologically prior to (and somehow deeper than) the existence of our universe. This is because we are using this discipline to constrain which other worlds might possibly exist.

### About these definitions of existence

In these four definitions, I have proceeded from what is known most directly and immediately, and each concept of existence depends upon the previous one:

Mind other minds empirical existence material existence

This section distinguishes between empirical existence and material existence, which are often confused to the great detriment of our understanding. I insist upon this distinction because the two concepts are clearly different. Mind is essential for empirical existence, but is unnecessary for material existence. Physicalists generally conflate empirical existence and material existence, calling both “physical existence.” Physicalists need some sort of concept of material existence because they wish, among other things, to explain how mind arrives in a materially existing, initially insentient universe.

We each of us as individuals only know of the existence of our universe indirectly, by way of such experiences of ours that are coherent, lawful and intelligible. These experiences, as experiences, come to us directly and immediately. We know nothing about the universe except by having experiences, and so these should be considered first.

We initially know of objects in the world by observation. Objects make themselves known to us by appearing as regular structures in our experiences: objects cause these experiences. We can confirm that percepts are veridical and not systematic illusions by conferring with other people (minds like our own). By means of this intersubjective process of observation by a community of observers, we come to recognise the empirical existence of objects in the world. It is even the case that we know of the empirical existence of our own bodies in this way.

Our confidence in the empirical existence of objects depends on other minds. This is why empirical existence is placed after other minds in the diagram above. Arthur Eddington expresses this:

It is true that I have a strong impression of an external world apart from any communication with other conscious beings. But apart from such communication I should have no reason to trust the impression. [He gives vivid dreams as an example.] So long as we have to deal with one consciousness alone, the hypothesis that there is an external world responsible for part of what appears in it is an idle one. All that can be asserted of this external world is a mere duplication of the knowledge that can be much more confidently asserted of the world appearing in consciousness. The hypothesis only becomes useful when it is the means of bringing together the worlds of many consciousnesses occupying different view-points.

(Eddington, 1928, p. 284)

By way of theories of lesser or greater sophistication, we come to some knowledge of objects outside of the range of our direct observation and experience: the Moon when it is behind clouds, the existence of electrons, and so on. But can this theoretical knowledge be developed into a concept of what it means for unobserved entities to exist, as is required by physicalists? I will argue below that we cannot arrive at a satisfactory concept of material existence.

Do any of these four concepts of existence amount to actual existence? My opinions, based on the following reasons, are roughly summarised in the final column of Table 2.1 above.

1. Descartes wanted base philosophy on a foundation of absolute certainty. His procedure was to discard any notion of which there was the slightest question. Whatever remained, he reasoned, must be a certain truth. He argued that his perceptions might be entirely illusory in the sense that it was at least conceivable that the external world might not exist at all. He knew however that he thought, and so he concluded that he existed as a thinking being: *cogito ergo sum*.

*Cogito ergo sum* remains a certain truth even if we reject absolute certainty as the appropriate standard by which to judge metaphysical questions. Similarly, the conditional assertion that “*If X has experiential existence then X actually exists*” remains true whether we accept or reject Descartes’ criterion of absolute doubt. (Mathematical circles provide examples of Xs without experiential existence.)

We know absolutely that we actually exist as minds, but we can be extremely confident that we are not merely disembodied minds, existing in isolation, experiencing the illusion of a non-existent universe. We can be extremely confident that our existence as minds is not all that our actual existence amounts to. For this latter reason, there is a half-tick in the top row of Table 2.1.

2. The definition of empirical existence accurately describes what laypeople and scientists do in coming to the firm conclusion that something “actually exists” (or just “exists”). The Moon provides an explicit example. Each person has a similar experience, perhaps of a bright crescent. These experiences are in such concordance with one another and with scientific knowledge, that we are justly confident in asserting the Moon’s actual existence. But there must be more to the Moon’s actual existence than its empirical existence. We certainly want to say that the Moon actually exists when clouds obscure the night sky, and that it existed in the distant past, before the arrival of life on this planet. These arguments explain the half-tick in the second row of Table 2.1. (Physicalists would want to say that the Moon continues to materially exist even when humans are not observing it, but I will argue below that no adequate concept of material existence can be given).

4. A logically consistent formal structure is sufficient for mathematical existence, but mathematical existence does not of itself amount to actual existence. For example, Hawking and Ellis describe several exact solutions to Einstein’s field equations, and all of these validly exist as mathematical objects (Hawking and Ellis, 1973, chapter 5). These mathematical objects are very different from one another in terms of the physical predictions they make, and so it is impossible that all of them are even approximately consistent with what we know about our own universe. For example, part of the De Sitter spacetime is a model for the now discredited Bondi, Gold and Hoyle steady state theory of the universe, (Hawking and Ellis, 1973, p. 126). We have no reason to believe that there are other universes that instantiate all of these mathematical objects, even approximately. The same applies to any mathematical object whatsoever. This is my reason for placing a cross in the final row of Table 2.1. (But compare Tegmark (2003), to be discussed later.)

This completes the explanation for the contents of Table 2.1, except for the cross in the third row, and this I turn to now.

### Physicalism and material existence defined

We need to examine this viewpoint of physicalism, the currently dominant philosophical position about the universe. Some of its assumptions are:

* Completed physics (even though we might never arrive at it) gives a complete, objective (observer-independent) description of the ontology of the cosmos.
* The ultimate entities of the universe (electrons for example) can be characterised *completely* in terms of completed physics. Ultimates have no intrinsic, hidden properties. For example, according to physicalists these ultimate constituents are not sentient.
* The same applies to simple entities such as rocks, water, flashes of lightning, and volcanoes.
* The universe, in its initial years, existed entirely without consciousness (construed broadly as any experiential quality).
* Consciousness only exists in small pockets of the universe where life has evolved to a certain degree of complexity.
* Consciousness requires complex structures of matter, such as water, carbon and other chemicals, and can be explained in terms of these complex material structures.

The project of physicalism is to first give an account of material existence, and then to explain consciousness in terms of complex configurations of matter.

A volcano is an example of an entity that has material existence. We can safely assume that lava flowed from volcanoes that spewed sulphurous fumes into the air long before the advent of life on Earth. Most scientists and philosophers would agree that if one had a complete description of the volcano in terms of completed physics, then that is all there is to the volcano.

According to physicalists, in the early years of the universe the entirety of the matter within it was insentient. Today, however, we know that some material objects such as an apple or a mountain can cause some complex systems of matter, namely us, to have experiences. Trivially, matter must have some potential to cause experiences. But from the point of view of physicalism, the fact that matter potentially has empirical existence is useless as an explanation for consciousness. If I asked a physicalist “How does matter cause consciousness?” and she were to reply “Matter has the potential to cause consciousness” then I would object “This is not an explanation but merely a restatement of the question I wish you to answer.” Physicalists need a concept of material existence that is independent of the concept of “potential for consciousness.”

On page 18 above I gave a tentative definition of material existence. This definition could be made somewhat more precise, and details might have to change as physics advanced. Nonetheless, it does capture the essence of material entities and of material existence, as these concepts seem to be intuitively understood by physicalists. I say “seem to be intuitively understood” because many physicalists don’t give any account of material existence before attempting to explain consciousness in terms of complex configurations of matter.

My tentative definition is inadequate because it is circular. Poland’s definition of material existence, which he calls the “physical ontological base” (1994, p. 132), is quoted on page 18 above. It is unsatisfactory because it rests on an unanalysed concept of spacetime: either spacetime is itself material, in which case the definition is circular; or it is unexplained in which case the definiendum is essentially an abstraction, as will become clear. I hope to show below that no satisfactory account of material existence can be given.

### The early universe

Some scientific theories such as evolution and cosmology deal with eras before human beings existed. Do physicalism and the above definition of material existence succeed in giving such theories a sound ontological foundation?

Physicalists maintain that consciousness does not appear in the cosmos until the time when life has developed to a certain degree of complexity. Let **UT** be our universe up until the moment before the first glimmer of consciousness appears.

There is neither experiential existence nor any empirical existence within **UT**. There is only material existence. The early universe has a mathematical model UT. The question arises, “If physicalism is true, how does **UT** differ from this mathematical object UT?” When you reflect on it, the answer is “Not at all.” Physicalists might claim that the matter in **UT** actually exists whereas the matter in UT merely has mathematical existence, but the physicalism gives no clue as to what the difference between mathematical existence and actual existence is.

If physicalism is true, our early universe **UT** is *conceptually* *indistinguishable* from the mathematical object UT. This conclusion is highly counterintuitive, and most physicalists would deny it, and so it needs to be explained.

First consider our early universe, **UT**, (here I will use the convention that everything with material existence will be shown in bold):

1. Let P be the set of mathematical laws of completed physics, which might be deterministic or statistical. P mathematically models the evolution in **time** **t** of **UT**. Within **UT** itself is the **initial state** of the **universe**, defined by the part of **UT** where **t = 0**.
2. P gives a formal mathematical definition of what it is to be a spatiotemporal configuration of matter-energy [a material entity] within **UT**. For example P tells us what **water** is, and that it is made of **H2O**. Examining the mathematical structure of a small portion of **UT**, the laws P might tell us that here **rain** is falling on a **planet** for example. This is so despite the fact that there are no observers within **UT**. **Water**, **H2O**, **rain**, and **planet** all have material existence.
3. P also gives a formal mathematical definition of what it is to be an ultimate within **UT**. An **electron** and a **photon** are examples of ultimates within **UT**. Examining the mathematical structure of a small portion of **UT**, the laws P might tell us that here an **electron** is absorbing a **photon** for example.

The mathematical model UT of the early universe has very similar properties. (In this second list, naturalistic-sounding words and phrases such as “time,” “H2O,” “planet,” “photon,” “a spatiotemporal configuration of matter-energy,” and so on, are merely labels for specific substructures in the mathematical model):

1. Let P be the set of mathematical laws of completed physics, which might be deterministic or statistical. P mathematically models the evolution in time t of UT. Within UT itself is the initial state of the universe, defined by the part of UT where t = 0.
2. P gives a formal mathematical definition of what it is to be a spatiotemporal configuration of matter-energy within UT. For example P tells us what water is, and that it is made of H2O. Examining the mathematical structure of a small portion of UT, the laws P might tell us that here rain is falling on a planet for example. This is so despite the fact that there are no observers within UT. Water, H2O, rain, and planet all have formal existence.
3. P also gives a formal mathematical definition of what it is to be an ultimate within UT. An electron and a photon are examples of ultimates within UT. Examining the mathematical structure of a small portion of UT, the laws P might tell us that here an electron is absorbing a photon for example.

*These 3-point lists are important because they characterise all the statements that can be made about* **UT** *and* UT*, expressed in a mixture of natural language and mathematical terms.* In the first list, I used the convention that everything with material existence was shown in bold. Corresponding items in the second list, which merely have formal mathematical existence, are shown in regular font. The lists are identical, except that some words in the first list are in bold, whereas the same words in the second list are in regular font.

##### Theorem

If physicalism is true, then our early universe **UT** is conceptually indistinguishable from the mathematical object UT

##### Proof

For every true statement about actual structure in **UT** there is a corresponding true statement about the formal structure within UT, and vice versa.

For example, if in **UT**, “materially existing rain is falling on a materially existing planet,” then in UT, “formally existing rain is falling on a formally existing planet.” We can reason in the opposite direction: if we know a fact about UT, then we know the corresponding fact in **UT**.

Other examples are that in **UT**, “A material electron absorbs a material photon,” whereas in UT, “A formal electron absorbs a formal photon.” In **UT**, “actual time actually passes”, whereas in UT, “formal time formally passes.”

How is material existence distinguishable from (a specific, corresponding type of) formal existence? The answer is “It is not.” The physical laws P are stated in purely mathematical terms, and the definition of material objects (point 2 of the first list) only makes use of the formal mathematical structure of the early universe **UT**. So *this definition is purely mathematical*, and turns out to be exactly the same as the definition of specific, formal substructures within the mathematical model UT (point 2 of the second list).

The same argument applies to actual ultimates (point 3 of the first list). They are defined solely in terms of the mathematical laws P and the mathematical structure of **UT**, and this definition is identical to the definition of formal ultimates (point 3 of the second list).

Similarly, the early universe itself, its initial state, and the flow of time are defined solely in formal, mathematical terms, and these definitions are identical to the corresponding formal concepts of the mathematical model.

We have shown that **UT** and UT are conceptually identical: For every statement that can be made about **UT** there is an identical statement that can be made about UT and vice-versa. While the word “existence” in the first case is intended to be “material existence” and in the second case “mathematical existence,” physicalists do not explain what it is that distinguishes these two concepts of existence.

If **UT** and UT are conceptually identical, then they must be identical, for to suppose otherwise is to suppose that there is an ineffable, occult distinction between them. Such distinctions are forbidden to rational thinkers.

(Note that, as already argued on page 23 above, from the point of view of physicalism, to make the contrasting statements that “**Matter** in **UT** has the potential to give rise to consciousness”, whereas “Matter in UT does not have the potential to give rise to consciousness” is to make an occult distinction that is useless as an explanation of why consciousness arises from **UT** and not from UT. Adding these contrasting statements to the respective lists above does not amount to making a genuine conceptual distinction between **UT** and UT.)

**QED**

Why does the above theorem work? Can the situation be remedied by defining material existence in terms of different, abstract physical laws? I believe that the answer is fairly clearly “No.” Advances in physics might mean that we would have to redefine material existence, perhaps in terms of string theory say, but the argument would go through in exactly the same way. Physics works by means of observation, experiment, discovering mathematical laws, and constructing mathematical models. Once you have abstracted away the experiential beings (the physicists), all that you are left with is the mathematical model. This is shown in the first 3-point list, where all of the supposed definitions of “material entities,” “ultimates,” and so on were made in purely mathematical terms. (To the admittedly controversial extent that the quantum character of our universe prevents the human observer being abstracted away from it, physicalism is false because the universe does not exist independently of us. This is discussed on page 40 below.)

### Consequences for physicalism

We are being realists about qualia such as the taste of double chocolate mouse, or the first glimmer of sensitivity experienced by a primitive creature. According to physicalists, the universe is wholly insentient up until a certain time **T** when it first appears. Some physicalist theories focus on this problem, and these might be called theories of primeval emergence. However, all physicalist theories, if they are to be considered comprehensive, must explain primeval emergence.

If physicalism is true then the emergence of qualia in the early universe is incomprehensible. Insofar as the physicalist project has been developed to date, the early universe is conceptually indistinguishable from a mathematical object. This implies that for any explanation that contemporary physicalists might supply of how consciousness arose in our universe, there would be a corresponding explanation of how consciousness arises in its mathematical model. This is absurd, for it is inconceivable that the slightest glimmer of a qualitative experience, such as a tickle, could emerge from any mathematical object, no matter how complex. (Max Tegmark (2003) bites this gigantic bullet, as will be discussed later.)

I believe that the above arguments make the case that, until physicalists can explain what the existence of physical objects amounts to when these objects are unobserved (and the prospects doing this do not appear good), then no physicalist theory will be able to explain the primeval emergence of qualia in the early universe. The failure to give any account of what it is for unobserved matter to exist is, in my opinion, a sign that the metaphysical project of physicalism is in severe difficulties. All of the above justifies the cross in the third row of Table 2.1.

Physicalism remains popular, despite its fatal lack of foundations. The reasons for this seem to be: First, many physicalists are overconfident in their intuitions about matter. Because of this, they do not give any account of what it is for matter to exist, in preparation for their attempts to explain consciousness in terms of complex configurations of matter. Second, there is the widespread misconception that physicalism is a necessary part of our current scientific understanding, rather than being a set of disputable, metaphysical assumptions. Six physicalist assumptions were listed on page 23 above. In the theory I will develop here, all six assumptions are false. Most physicalists incorrectly believe that the first assumption is a necessary truth.

### Introducing idealist panpsychism

This final subsection introduces a particular metaphysical theory, *idealist panpsychism*, which will be elaborated in subsequent sections.

Table 2.1 summarises the four concepts of existence:

* If we know that an entity has experiential existence, that is to say, if we know that it exists as a mind, then we know that it actually exists. But there is more to actual existence than experiential existence because we know (beyond reasonable doubt) that all minds are embodied in some way.
* If we know that an entity has empirical existence, then we know that this entity actually exists, but we know that there is more to an entity than its empirical existence because we know for instance that the Moon continues to exist when it is behind the clouds.
* Concepts of material existence are inadequate. If we assert that there is no more to material existence than obeying the laws of physics, then material existence is merely another name for mathematical existence, and this is insufficient for actual existence. If, as many physicalists do, we extend the concept of material existence by stating that there must also be some potential for causing consciousness, then this extended definition is useless in explaining the primeval emergence of consciousness.
* Mathematical existence does not amount to actual existence.

We still need some rational concept of actual existence whose domain includes the very early universe and its contents, which must undoubtedly have existed in some way. The inadequate concept of material existence needs to be replaced. I am now ready to give the core definitions of idealist panpsychism.

**Definition:** an *experiential entity* is a unitary entity with both experiential existence and empirical existence.

**Definition:** To *actually exist*, (or just *exist* for short) is precisely to be an experiential entity, or to be composed of experiential entities.

This latter definition is explicitly metaphysical. It aims to capture the concept of actual existence, and is therefore supposed to apply to all universes.

Experiential existence and empirical existence invariably occur together in any entity, and it is this unitary co-occurrence that constitutes the entity’s existence. In our universe, or indeed in any universe, *to be an experiential entity* is both to possess qualia or qualitative experiences, and also to cause systematic intersubjective regularities in the percepts of other experiential entities.

In brief, an experiential entity is an experiencer that can be experienced by other experiencers. Everything that exists in any universe is – by the very definition of existence – composed of experiential entities. (Anomalies such as Descartes’ solipsistic mind are logically conceivable, but do not count as true universes.)

The theory given by these definitions is idealist, because every individual thing that exists is, in essence, a (perhaps exceedingly simple) mind. It is also panpsychist for exactly the same reason, when we note that existent things include the ultimate, elementary entities of completed physics.

Earlier in the section there was the fear that, if the Moon’s existence amounted to no more than empirical existence, then the Moon would cease to exist when clouds obscured the sky. This fear was reinforced when the alternative concept of material existence proved inadequate. According to idealist panpsychism, however, the Moon’s actual existence is a combination of empirical and experiential existence. The countless particles that constitute the Moon are in fact experiential entities that are always experiencing and being experienced by each other. The Moon therefore continues to exist even when clouds obscure the sky, and it used to exist before the coming of life on Earth. Idealist panpsychism is thus thoroughgoingly *realistic* about objects in the universe – despite its idealism. Indeed, this panpsychism’s realism is better grounded than that of physicalism, because, unlike the latter, it gives an explicit and adequate definition of what it is for objects to exist.

Experience is fundamental, and was present everywhere in the universe from the beginning of time. The elementary particles of physics are in truth extremely unsophisticated experiential entities. They can combine in hierarchies to form complex, organic experiential entities such as ourselves. Or they can combine in far more stereotypical ways to form experiential entities whose behaviour is so dull and predictable that they *appear* to be insentient.

Do these definitions capture well enough our intuitive ideas of actual existence, despite the counterintuitive consequences of both idealism and panpsychism? I would argue *Yes*: They capture the notion of the existence of ourselves as human beings with our unitary existence as both minds and as bodies. They capture the notion of the empirical existence of an item such as this piece of paper when it is part of our immediate experience. They also capture the notion of the paper’s continued existence when it is shut away in a locked room. The task of idealist panpsychism is not to explain experience, because experience is fundamental. Instead, the task is to explain how the physics of the world arises from the interactions of experiential entities, including how mind and body can exist as a unity.

In an argument, very much condensed here, Eddington makes a similar case:

*Actuality*. “Knowableness to mind” is moreover a property which differentiates the actual world of our experience from imaginary worlds in which the same general laws of nature are supposed to hold true. [… … In an imagined world] unreal stars emit unreal light which falls on unreal retinas and ultimately reaches unreal brains. […] Is the brain disturbance translated into consciousness? That will test whether the brain is real or unreal. This property, which is evidently not definable with respect to any of the laws of Nature, we describe as “actuality” […]

(Eddington, 1928, pp. 265-266)

# 3. Idealist panpsychism

The section opens with a brief description of Max Tegmark’s unusual physicalism, giving his example of a dodecahedral universe, and stating why I believe his attempt fails. Max Tegmark’s dodecahedron can be adapted to define a conceivable idealist panpsychist universe. The structure of this putative universe is given both in terms of its physics, and in terms of the percepts of its “inhabitants.” The universe is trivial in its simplicity, but its importance lies in the completeness with which it is described.

### Tegmark’s physicalism

Tegmark is a strong reductionist, believing that everything can be reduced to physics, and that the physics of the world is essentially pure mathematics. For him mathematical existence and physical existence are the same concept. We and the world we live in are nothing but an (incredibly complex) formal mathematical system. There is no intrinsic character to any physical object, and there is no extra fact that distinguishes our instantiated world from its mathematical model: “instantiation” is thus an empty concept for him (Tegmark, 2003, section IVB; Hut *et al.*, 2006, section IIA). His reasons for adopting this position are: the extraordinary success of physical theories that describe the world in mathematical terms (Hut *et al.*, 2006, pp. 2-3); and the lack of any evidence for intrinsic, non-mathematical properties in the external world (*op. cit.*, p. 12).

Tegmark’s conception of consciousness is that there are (mathematically defined) self-aware substructures in a universe (2003, p. 14). Tegmark goes further and proposes that it might be the case that all formal mathematical objects have actual (he says “physical”) existence in a Platonic realm. A mathematical dodecahedron, for example, has no more or no less actual existence than the world in which we live:

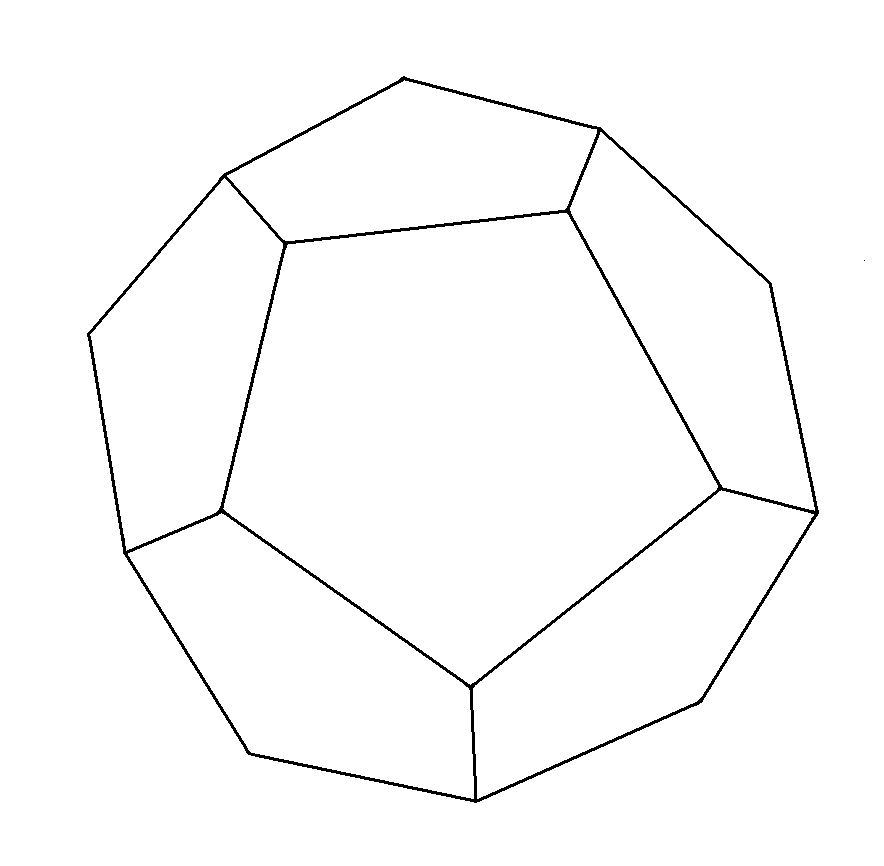
Now suppose that our physical world really is a mathematical structure, and that you are a self-aware substructure within it. In other words, this particular mathematical structure enjoys not only mathematical existence, but physical existence as well. What about [all other mathematical structures]? Do they too enjoy physical existence? If not, there would be a fundamental, unexplained ontological asymmetry built into the very heart of reality, splitting mathematical structures into two classes: those with and without physical existence. As a way out of this philosophical conundrum, I have suggested […] that complete mathematical democracy holds: that mathematical and physical existence are equivalent, so that *all* mathematical structures exist physically as well.

(2003, p. 14)

Tegmark thus proposes an amazingly wide concept of existence.

### A Tegmarkian dodecahedral universe

One elementary example of such a Tegmarkian universe is a dodecahedron, Figure 3.1. The physics of this universe amounts to no more than its geometry. The dodecahedron is composed of an (uncountably) infinite set of points {a, b, c …}, together with a particular distance relationship d(a, b) between any pair of these points. Nothing else needs to be specified. For instance, angles between points can be calculated using the cosine rule.



### Figure 3.1: A dodecahedron

### Critique of Tegmark

Criticisms of Tegmark come under two main headings: plausibility and consciousness.

Several authors criticised Tegmark on the grounds of plausibility, and he has given interesting rebuttals to this accusation (Hut *et al*., 2006). I do not agree with Tegmark’s critics on this point because plausibility is subjective and depends upon what one’s (at least current) foundational beliefs are. Tegmark’s ultra-Platonism is a logical consequence of his being unable, given his physicalist assumptions, to arrive at a substantive concept of actual existence or instantiation. As was shown in the previous section, any physicalist who claims that physical entities have no intrinsic character and are fully characterised by mathematical laws will be driven into a position very close to Tegmark’s. The burden on such physicalists is to explain how they would avoid this.

Tegmark says little about consciousness: “I believe that consciousness is the way information feels when being processed. […] I am what I am and will continue to enjoy the way I subjectively feel regardless what the underlying explanation turns out to be” (Hut *et al.*, 2006, p. 3). This suggests he is a qualia realist, but he is of necessity a functionalist of sorts, and I have already criticised this position.

Further criticisms can be made, however. How is information “processed” in a mathematical object? Tegmark defines time formally, but processing seems to require a real flow of time: how does he explain our perceived flow of time from youth to old age? How is there any unity in our conscious experience? The state of our brain at any instant must be a vastly complicated mathematical substructure with many components. Nothing binds these components into a single experience. Another difficulty (Searle, 1992, pp. 214-222) is that mathematical objects are purely syntactical and that syntax has no causal powers.

I suspect that Tegmark would bite a couple of bullets and assert that the flow of time is an illusion and that there is no causation in the universe. But on unity he has severe difficulties: It seems impossible to explain the unity of our experiences; and to deny unity as utterly as his theory requires amounts to denying consciousness altogether (James, 1983, pp. 266-268).

We are now ready to contrast Tegmark’s theory with idealist panpsychism.

### Idealist panpsychism

Here is a semi-formal definition:

**Definition:** *idealist panpsychism* is the doctrine that the universe is composed of hierarchies of experiential entities, and of nothing else. These experiential entities can both perceive one another and be perceived.

An *experiential entity*, (which I will occasionally call an *experiential being*), has already been defined as an entity with both experiential existence and empirical existence. That is to say, an experiential entity possesses qualitative experiences, and it causes systematic intersubjective regularities in the percepts of other experiential entities. Experiential entities are the fundamental entities of idealist panpsychism. All other concepts arise from this basic concept. So the concept of “matter,” for example, arises from collating the experiences of these experiential entities; and it is not the case that consciousness or experiential entities arise from matter.

**Definition:** A *theory of physics* is a set of procedures and mathematical rules that collates the experiences and percepts of experiential entities.

The definition of physics given here is thoroughgoingly intersubjective but it accurately reflects what physicists actually do in arriving at their theories. Physicists observe the world, devise theories, test them with experiments, repeat experiments to confirm the results of scientists in other laboratories, publish their results in peer-reviewed journals, and so on. Our current, imperfect theories of physics come under this definition. We also hope that there exists a *completed physics* for our universe that completely describes its workings in terms of mathematical laws, and correctly predicts the experiences of all experiential entities (importantly ourselves) whenever any specific procedures are carried out. In particular the completed physics correctly predicts the outcome of all possible experiments. A completed physics could still exist for our universe, even if it was forever beyond our human ability to discover it.

**Definition:** A *universe* is (roughly) a maximal set of causally-connected experiential entities, together with a completed physics that describes the laws of behaviour or interrelationships of these entities. The set cannot be extended without breaking the laws of physics. Moreover, these entities are experientially and causally isolated from any other experiential entities.

This definition as it stands is approximate because it discusses time and causation in insufficient depth.

### A panpsychicdodecahedral universe

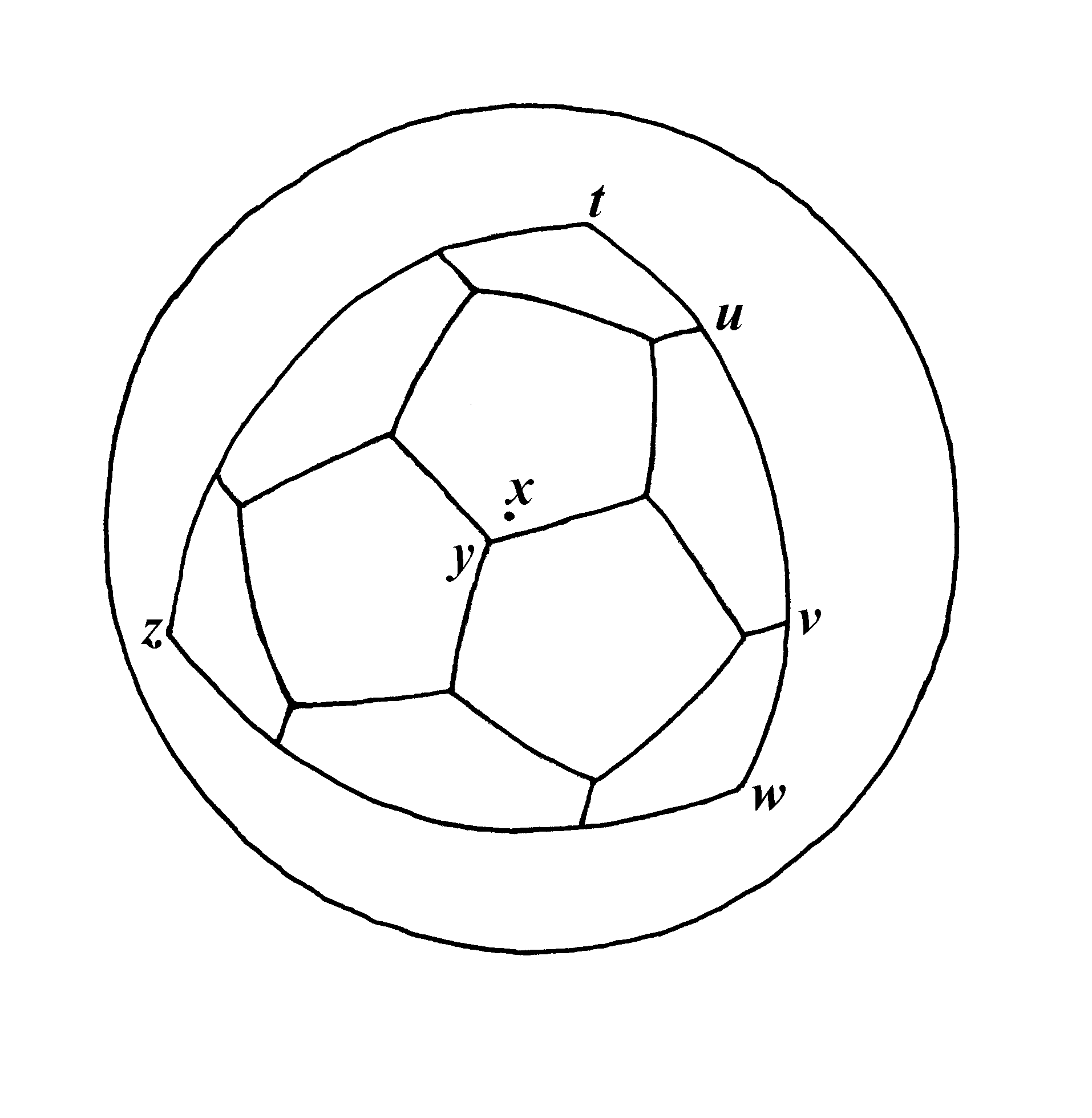
I now wish to give a complete description of a conceivable actual universe in terms of the concepts of idealist panpsychism. This is more than merely a model of a universe, defined solely in terms of abstractions. Rather, it is a conceivable universe, because it is defined in terms of (putative) experiential entities and their percepts. The description of the universe is complete in the sense that both the relationships between these entities (the physics of the universe), and the percepts of each of these entities will be given. The physics of this universe will turn out to be identical to Max Tegmark’s mathematical dodecahedron, but there similarities end.

The universe comprises a set of *experiential entities* {**a**, **b**, **c**… }, each of which can experience the others. There is also a *distance relationship* **d**(**x**, **y**) between any pair of entities from this set. The distance relationship is such that, for any pair (**x**, **y**) of them, entity **x** experiences a sensation of intensity proportional to e‑**d**(**x**, **y**) in the direction of **y**, and vice versa, regardless of any intervening entities. That is to say, the intensity of the percept declines exponentially with distance. There is a one-to-one correspondence between the set of entities {**a**, **b**, **c**… } and the set of points {a, b, c… } of Max Tegmark’s dodecahedron such that **d**(**a**, **b**) always equals d(a, b) for any **a** and **b**, and so the entities are laid out in the form of a dodecahedron.

The above description is not quite in the spirit of idealist panpsychism because it derives the percepts of the experiential entities from the physics of the world (the geometry of the dodecahedron), whereas we wish to derive the physics from the experiential entities and their percepts.

Let us consider, for a general experiential entity, **x**, the percept **P(x)** that belongs to it. In this particular example, the percept is in the form of a sphere, where each point on the sphere represents a given direction r from **x**, and each point is also marked by the intensity of sensation in that direction. The nature of this sensation is left unspecified and unknowable, but we might imagine it as being brightness. The brightness in a given direction increases with the thickness of entities in that direction. More precisely, the intensity of sensation in given direction r is given by ∫U e-r dr, where U is the subset of the ray in direction r in which experiential entities are located.

The entity at the centre of the dodecahedron will have a symmetrical percept in which the perceived faces of the dodecahedron brighten slightly towards the edges, and where the vertices are brightest of all. An entity that lies close to one of the faces of the dodecahedron will “see” this face filling a larger part of its perceptual field, and as being fainter than the opposite, more distant faces. More formally, each entity’s perceptual field has a particular mathematical structure, given by the above exponential decline rule. For each entity, there is a particular, calculable total intensity of perception in each direction.



**P(x)**

### Figure 3.2: A percept or, more precisely, the mathematical structure abstracted from the percept of an experiential entity

I have tried to sketch the percept of one such entity in Figure 3.2. The sphere **P(x)** represents the percept or mind of experiential entity **x**, which happens to lie at one of the vertices of the dodecahedron. We can regard the centre of the sphere, labelled x, as being **x**’s point of view. The only part of the percept with intensity greater than zero lies within the spherical triangle twz, in which **x** is “looking within or towards” the dodecahedron. The arc tuvw is **x**’s percept of one of the three faces adjacent to it, namely the face with corners tuvwx. This face is experienced as an arc because it is “seen” edge-on. Points u and v are the points of most intense sensation along arc tuvw. Point y on the sphere is **x**’s percept in the direction towards the opposite vertex y of the dodecahedron. It is the point of most intense sensation of all, because **x** is “looking through” and experiencing the entire diameter of the dodecahedron.

The dodecahedral universe is actual or real in the same strong sense that our own universe is real. It is not merely a mathematical abstraction. Each entity **x** in the universe has an experiential existence. There is “something like it is to be” that entity. Although entity **x** cannot think, it does have a percept, **P(x)**, and, moreover this perception is veridical, in that it accurately reflects **x**’s position in the dodecahedron. Modelling Descartes’ argument, we can say on behalf of such an entity, “It perceives, therefore it is.”

The physics of this universe is given by the mathematical model of the dodecahedron. Although there are no physicists in this universe, there are a multitude of observers. A particular entity **a** is perceived by all of the other entities in the dodecahedron. From the viewpoint each of these others, **a** is just a perceptual speck of a certain magnitude in a certain direction. These percepts can be collated to place **a** at a certain point in the dodecahedron. As viewed by the others, **a** is just a point. Just as in our universe, **a**’s percept is private, and the other entities have no access to it.

### Strict idealist panpsychism

In the above account, I have introduced the physics of the universe early in the discussion. But it is one of the principles of idealist panpsychism that experiential entities, their percepts and interrelationships constitute the fundamental facts the universe, and that the physics of the universe is a secondary formal description of these fundamental facts.

What should I do in order present this example in strict accordance with this principle? The argument would have to be somewhat as follows:

This particular idealist panpsychist universe is a given set of experiential entities, {**a**, **b**, **c**… }, together with their corresponding percepts, {**P(a)**, **P(b)**, **P(c)**… }. These percepts are supposed to have some experiential quality, but we do not know what this is. These facts are all that is given about this particular universe.

There is no concept of time in this example universe, and the minds of these experiential entities comprise only their individual percepts. They have no powers of cognition and so they cannot reflect upon their observations; nor do they have any means of communication. Such entities cannot develop a theory of physics for themselves. Nonetheless, from the information above, we can collate their percepts on their behalf to work out a well-defined physics (which in this very elementary example, reduces to geometry).

Taking a sample of percepts, we can plot their structure as spheres of intensity, perhaps using a computer with a graphical display. We would be able to see for ourselves the dodecahedral character of these percepts. There would be only one percept from the entire set which was symmetrical: Twenty symmetrically placed points on the spherical percept would have identical and greatest intensity, and the arcs joining these points would be more intense than nearby locations. (The computer image would resemble a football.) We might speculate that this percept belonged to the entity at the centre of a dodecahedron, where the twenty points of maximal intensity were the vertices of the dodecahedron, and the arcs were the edges.

Given this symmetrical percept alone, there many other candidate shapes for the universe. Most candidates would be rather like dodecahedra, but with convex or concave rather than planar faces. However, we could examine the other members of the set of percepts and find among them exactly twenty in which the intensity of the percept was zero except within a spherical triangle, as sketched in Figure 3.2. We could speculate that these twenty percepts belonged to the entities at the vertices of a dodecahedron. By examining these percepts, as already discussed, we might conclude that perceived arc tuvw might represent the planar face tuvwx of a dodecahedron as perceived edge-on. Using similar evidence we would be strengthened in our belief that the universe was in the shape of a dodecahedron, we could state this belief as a formal hypothesis.

We may then, by examining the detailed pattern of intensities of the percepts of the experiential entities, particularly the one in the centre and those forming the vertices, attempt to discover the rule governing the distance apart of these entities. Hopefully, we would quickly arrive at a proposed law of physics where distance is defined in terms of intensity of perception:

**Proposed law of physics (definition of distance):** if, for any pair of entities **x**, **y**, entity **x** experiences a sensation of infinitesimal intensity e‑d dr, in the direction r, caused by **y**, then the *distance* **d**(**x**, **y**) between **x** and **y** is defined to be (the exponent) d.

Using the complete set of entities and their percepts we could check that this law was consistent with all of the entities forming a dodecahedron, thus confirming the hypothesis. We could safely conclude that this physics is true of this world because it accurately and consistently reflects the percepts of all the entities in the world.

Thus, at least for this elementary example, we can derive the physics from the elementary entities and their percepts. It is much more difficult to try to do the same for our universe. We initially have only a very partial knowledge of the percepts of only a few of the most complex entities - ourselves. We would have to guess at laws determining what the percepts of the ultimate constituents of physics might be, and how these ultimates might combine lawfully into composite experiential entities. The geometry of the universe and the laws of physics are much more complicated. There is a concept of time in which the entities interact in complex ways, and with only partial information about each other. Nonetheless, this example demonstrates in principle how idealist panpsychism can work.

### Mind, body, and world

The dodecahedral example also illustrates the principles by which mind, body, and the physics of the world are interrelated.

The idealist panpsychic dodecahedral world has a mathematical model. This consists of a dodecahedron, together with a sphere associated with each point of the dodecahedron, and with a positive or zero number (representing intensity) associated with each point on the surface of each sphere. Does this entire mathematical model constitute the physics of this world? No. The physics of the world is just the dodecahedron, because each experiential entity within the world has a percept of a dodecahedron, and of nothing else.

Each entity **x** in this world has a percept **P(x)**, and, in this example, this percept constitutes the entirety of the mind of **x**. There are still, however, some analogies between minds in this simple world, and minds in our world:

* Each mind in the simple world is hidden, and it cannot be observed by other entities within the world using the methods of physics.
* Each mind in the simple world is associated with its physical body. In the simple world this body is perceived by the other entities in the world as no more than a point on the dodecahedron. In our world, each (human) mind is associated with a physical body that is located in space in a very small region, say of the order of one cubic metre.
* Minds in the simple world are perspectival, and the perspective they have on the world is closely associated with the positions of their physical bodies in the world.

For example, when I have the experience of seeing the interior of Blackwell’s coffee shop in Oxford, I see my hands in front of me, and my reflection in the mirror. My companions also experience seeing my body in this same location. My mind thus has a perspectival view of the universe that is closely linked with the position of my body in the universe.

Because the dodecahedral example is so elementary, it throws light on the way in which mind and body are interrelated in our universe. For each experiential entity, **x**, considered as a whole, we can clearly see the relationship between it and its various aspects: (1) its location as perceived by others within the dodecahedron; (2) its qualitative percept, **P(x)**, of the dodecahedron; and (3) its centre of perspective, which is the central point of the sphere **P(x)**. We can also see the distinction between (4) *mind* in the abstract, where sphere **P( )** is the arena in which the intensity at each point could in principle take on any non-negative value; and (2 again) the *contents of mind*, which is a particular percept, **P(x)**, giving particular values of intensity to each point on the surface of the sphere.

In a more complex example, where the universe possessed a concept of time, there could also be additional aspects to **x**: (5) its memory of past experiences; and (6) its choices for future action based on its memory and its percept.

We can also see from the example that because the percept is structured, it contains implicit information, even though the primitive entity cannot think. Thus there might be implicit information that the intensity in one direction was twice the intensity in a second direction; that the angle between these directions has a particular size; and so on. If there were no such implicit information, then the structured percept might as well not exist.

It is wrong to think of the idealist panpsychic universe as comprising “material” points arranged in the form of a dodecahedron, with perceptions tacked on as an afterthought. Rather, it comprises entities, whose minds constitute the essence of what they are. These experiential entities are interrelated in a coherent way, definable in terms of their perceptions, and it is this intersubjectivity that gives rise to a mathematically describable physics. Physics thus arises from sentience, and not the other way round. With this idealist panpsychic universe it is coherent to ask whether or not it actually exists, because, despite its much greater simplicity, the character of its existence is the same as that of our own universe.

### Our universe

The dodecahedral universe is very much a “toy” example. My own belief is that our own universe is an idealist panpsychist world, but the theory would require considerable development in order to show this. There are many conceivable worlds conforming to the principles of idealist panpsychism. The perceptual space of entity **x** might not be in the form of a sphere, but instead have some different shape or structure. The rules for determining the particular percept could have any degree of complexity. Again, the physics of the world could be far more complicated, for example including the concepts of time, causality, and laws of motion. Time also allows for entities with memories and choices of actions. The panpsychism presented here needs to be developed to account for all the complexities of our world, perhaps by regarding time as a process along the lines suggested by Whitehead (1978).

# 4. Idealist panpsychism and the physics of our universe

Previous sections have argued for idealist panpsychism: the thesis that the fundamental level of reality is hidden from even completed physics, and consists entirely of experiential entities that can organise themselves into hierarchical systems. Even the most basic level of completed physics (so-called “fundamental” physics) could not deal directly with the ultimate constituents of reality. It could merely deal with the structure of the appearances of these constituents, which are in essence minds.

This section asks: what do we expect the properties and characteristics of completed physics to be if these arguments are sound? It then goes on to describe in general terms quantum physics, the closest approach we have to completed, fundamental physics at the present epoch, and shows that this has the properties and characteristics as predicted under the hypothesis of idealist panpsychism.

### Predictions of idealist panpsychism

Idealist panpsychism makes six predictions about the properties and characteristics of fundamental physics. The first four apply to panpsychism in general:

1. *The behaviour of an elementary entity depends upon the detailed configuration of all of the other entities in its environment*

This is because panpsychism treats the fundamental units of nature as being experiential in character. They can perceive their environment, and act according to the whole of this percept.

1. *Fundamental physics is information-theoretical in character*

This is because experiential entities have incomplete percepts of one another, and act on the information or knowledge contained in these percepts.

1. *Elementary entities can amalgamate to form indecomposable compound entities*

We presume that a complex mind such as the human mind is built from successive hierarchies of experiential entities, beginning with fundamental experiential entities that correspond to the ultimates of physics. Let us call the latter *experiential ultimates*. The problem of explaining how these hierarchies are built is called the *combination problem* of panpsychism, (Skrbina, 2005, pp. 145-149; James, 1983, pp. 160-164). Human minds have a recognisably coherent psychic unity which, although imperfect, is clearly more unified than it could ever be if the human brain were nothing more than a disorganised heap of independently acting experiential ultimates. Minds must therefore be able to amalgamate into compound minds. This has to take place at all levels, from the level of experiential ultimates to the level of the human brain. The proposal here is that the psychic unity of each compound mind manifests itself as indecomposable behavioural unity when this mind is observed as a physical object (or system) by other experiential entities.

1. *Fundamental physics will be found to be inextricably bound up with consciousness*

Panpsychism clearly predicts that this will be the case. Whether or not this is actually true is controversial, as we shall see.

The final two points are specific to idealist panpsychism. (Different theories of panpsychism will have different things to say about ontology and mental causation. Some say little at all.)

1. *Fundamental physics will have difficulties in describing a coherent ontology*

We can ask of a theory: What do the equations tell us about what is real and what is truly happening? Idealist panpsychism is unusual in that it posits that the ontological base (the structured percept) is invisible to completed physics. Even completed physics is no more than the most fundamental level of appearance, as discussed earlier. Worldviews that do not recognise this fact will be in difficulty if they try to fix ontology at the level of (current or even completed) physics.

1. *In any given environment, elementary entities show an irreducible spontaneity of behaviour*

This last point does not follow from what has so far been proposed. It is conceivable to have a deterministic version of idealist panpsychism in which the percept of each entity at any given time determines precisely what is going to happen next. However, such determinism is implausible as it contradicts the everyday view we have of ourselves as persons having some limited freedom, able to a certain extent to choose our own goals in life.

### The character of quantum theory

Idealist panpsychism makes six general predictions about the character of fundamental physics. It is easy to check that classical theories, for example Newtonian mechanics, do not match these predictions at all well. But does idealist panpsychism correctly predict the characteristics of quantum theory – the fundamental physics of the universe insofar as we know it at the present day?

**Point 1:** *The behaviour of an elementary entity depends upon the detailed configuration of all of the other entities in its environment*

This is true. If, for example, a particle is launched into any spatially distributed environment of apparatus and detectors, then the probability that it will arrive at any particular detector depends upon the entire environment (Mattuck, 1992, pp. 28-36). Richard Feynman, by no means a panpsychist, is lead into making the metaphor: “So light doesn’t *really* travel only in a straight line; it ‘smells’ the neighbouring paths around it.” (1990, p. 54).

**Point 2:** *Fundamental physics is information-theoretical in character*

This is uncontroversial, and quantum computing and quantum cryptography are currently active areas for research. See for example Bouwmeester *et al.* (2000).

**Point 3:** *Elementary entities can amalgamate to form indecomposable compound entities*

Experimental tests of quantum theory have confirmed this fact. The term for this indecomposable amalgamation is *entanglement*, (Bouwmeester *et al.*, 2000, pp. 7-11).

**Point 4:** *Fundamental physics will be found to be inextricably bound up with consciousness*

This is controversial. The Copenhagen interpretation of quantum mechanics certainly takes the view that consciousness is essential. Rudolph Peierls says that the wave function tells us about our knowledge of a system and that the so-called “collapse of the wave function” occurs when a *conscious* (his emphasis) observer makes an observation thus changing the observer’s state of knowledge (Davies & Brown, 1986, p. 73). Peierls emphasises that the theory is expressed entirely in terms of knowledge and says nothing about “reality,” a word whose meaning is unclear in this context. Peierls does not believe that an inanimate object such as a computer or photographic plate could collapse the wave function (*op. cit*., p. 74). John Wheeler understands the Copenhagen interpretation somewhat differently. He believes that the blackening of a grain of a photographic plate is a real event – an irreversible amplification of a quantum event. However, the information stored on the plate is not *used* until it has been observed by a community of observers (*op. cit.*, p. 63). This concept of used information is closely related to the definition of *empirical existence* given here earlier.

Eugene Wigner tried to give an ontology to quantum mechanics by arguing that the first consciousness to observe the system caused the collapse of the wave function as a real physical event. This theory has been criticised on the grounds that it is not clear what constitutes a conscious system (Baggott, 1992, pp. 187-188).

Most other interpretations have been devised in an attempt to avoid a causal role for consciousness at the level of physics. Hugh Everett introduced the “relative state” or “many universes” interpretation in which the universe branched whenever an observation was made (1957); John Taylor is an advocate of the “statistical interpretation” in which the theory can only be meaningfully applied to ensembles of identically prepared quantum systems (Davies & Brown, 1986, pp. 113, 115); and David Bohm has proposed a deterministic, non-local, hidden-variable theory (Bohm, 1980, chapter 4). These attempts to devise alternative interpretations of quantum mechanics all eliminate a causal role for consciousness in the world. This effectively eliminates consciousness itself, even if this is not their proponents’ intention. Such alternatives to the Copenhagen interpretation have been criticised by several physicists. See for example Henry Stapp’s critique (2007, chapter 10).

**Point 5:** *Fundamental physics will have difficulties in describing a coherent ontology*

This is a historical fact about quantum mechanics. The Copenhagen interpretation is unassailable, but only describes an epistemology. Alternative attempts at interpreting quantum mechanics have tried to give an ontology, but all of these are problematic. There is no doubt that if a clear, unproblematic ontology were to be presented then physicists worldwide would vigorously acclaim both it and its authors. This has not yet happened.

**Point 6:***In any given environment, elementary entities show an irreducible spontaneity of behaviour*

This is an *empirical* *fact* about quantum mechanics. No matter how tightly we constrain the experimental setup, and no matter how tightly we try to fix the prior quantum state of an elementary entity, the final, observed state of that entity can only ever be determined probabilistically (Feynman, 1990, p. 19).

While it is true that David Bohm’s model of quantum theory is deterministic, this determinism is hidden from observers (1980, chapter 4). Bohm himself did not think that his theory was more than a stopgap to deeper understanding (*op. cit.*, chapter 7). In my opinion, the excellent empirical evidence for point 6 carries more weight than Bohm’s proposal of an essentially unobservable determinism.

We see that idealist panpsychism successfully predicts all six of these characteristics of quantum mechanics, whereas physicalism predicts none of them. Contrary to what some critics assert, elementary entities do give some signs of possessing experience.

### Testability

To what extent is idealist panpsychism testable? Tegmark says:

To avoid this conclusion that mathematical and physical existence are equivalent, one would need to argue that our universe is somehow made of stuff perfectly described by a mathematical structure, but which also had other properties not described by it. However, this […] implies either that it is isomorphic to a more complicated mathematical structure or that it is not mathematical at all. The latter would make Karl Popper turn in his grave, since all those additional bells and whistles that make the universe non-mathematical by definition have no observable effects whatsoever.

(Hut *et al.*, 2006, p. 12).

Idealist panpsychism is of this latter type (but having subtle observable effects) and so it is incumbent upon me to explain how and to what extent it could be tested. Testing would be according to the following steps:

1. Guess the mathematical structure of the percepts of experiential entities that correspond to physical ultimates
2. Guess the statistical laws that govern the wilful spatiotemporal behaviour of these entities based on their percepts
3. Check that guesses 1 and 2 correctly predict the physical properties and dynamics of physical ultimates
4. Entangled particles behave as a unified entity. Make a guess as to what the percept of a pair of entangled entities is, given the percepts of each member of the pair
5. Guess the statistical laws that govern the wilful spatiotemporal behaviour of this unified entity
6. Check that guesses 4 and 5 correctly predict the physical properties and dynamics of the entangled pair
7. Ditto for entanglements of a larger number of entities
8. Get human reports of their experiences and willings
9. Match entanglements in subjects’ brains to their reports

What are the prospects for success? Given that physics is able to describe the properties and dynamics of physical ultimates in fairly tractable mathematical terms, steps 1 to 3 seem to be within the bounds of plausibility. Suppose steps 1 to 3 were successful. Someone not convinced by the arguments of this dissertation would probably regard the mathematical percepts as, at best, useful calculating devices, and would want to use Occam’s razor to cut them off from the ontology of the world. On the other hand, someone convinced by my arguments would regard the achievement of steps 1 to 3 as weak confirming evidence. Steps 4 to 6 are likewise plausibly achievable and would provide additional support for those who already have independent reasons to favour panpsychism. Step 7 is equally straightforward provided the number of entangled particles is very small.

Achieving step 7 to the extent needed to account for the complexity of the human brain is far beyond what is currently humanly achievable, and may never be achieved. Likewise psychologists know of many pitfalls in accepting human reports, and of their many inaccuracies and limitations. Matching these together might prove impossible.

At first sight this is at best an extremely weak degree of testability, but we must compare it with the testability of rival metaphysical positions. If one takes qualitative experiences to be part of the ontology of the cosmos, then physicalism is an untestable metaphysical position because, as described earlier, realisation theories fail to deliver any vertical explanations.

Tegmark’s unusual mathematical multiverse ontology also has a small degree of testability but, as he rightly says himself, the problems of doing so are “horrendous” (2003, p. 16). Testing it involves: listing all possible mathematical objects; devising a ‘suitable’ statistical measure over these mathematical objects; devising a formal mathematical definition of what it is for an entity to be conscious; Finding out which universes thus defined contain consciousness; Calculating if our universe is typical of those universes containing consciousness. If our universe is in some sense ‘typical’ of all possible universes containing consciousness, then this counts as evidence supporting Tegmark’s ontology (2003, pp. 16-17).

In my opinion the prospects for testing idealist panpsychism, though weak, are much more promising, or at least no less promising, than the prospects for testing Tegmark’s ontology. Other physicalist positions are not testable at all.

# 5. Physicalism versus idealist panpsychism

Physicalism and idealist panpsychism differ with regard to their concepts of concrete as opposed to abstract existence. I have given these two contrasting concepts – of the concrete existence of objects in the world – different names:

* In physicalism, *physical existence* is defined in terms of material existence and (at least potential) empirical existence, with the former being more fundamental.
* In idealist panpsychism, *actual existence* is defined in terms of experiential existence and empirical existence, with the former being more fundamental.

Questions can be raised about both theories and I will list some here, giving the best answers I can.

### A question about physicalism

***Without explicit definitions of “instantiation”, “spacetime” or “causation”, how could the physicalist explain the way in which our world is not a mathematical object?***

*Answer: Give an adequate explanation (difficult), or bite the bullet like Tegmark and assert that our world is precisely a mathematical object.* The second option drives one towards a quasi-eliminativist position with regard to qualitative experience.

*Alternative answer: Accept that the potential for causing consciousness is a brute fact that distinguishes concrete from mathematical existence.* In this case one must also accept our consciousness as a brute fact, forever beyond explanation.

### Questions about idealist panpsychism

***How can percepts exist without being supported by matter?***

*Answer: They are basic, and do not need any support*. Supporting percepts with matter is nonsense because it is supporting the concrete with a will-o’-the-wisp: no adequate definition of material existence has been given. To suppose that matter as understood by physicalists is somehow concrete is to fall into the fallacy of *misplaced concreteness* (Whitehead, 1978, p. 18).

***How can the percepts of experiential entities exist if they are not in spacetime?***

*Answer: Perfectly well. Asking this begs the question of physicalism.* In idealist panpsychism an entity’s location in the world is given by its percept, (and an intelligent entity could verify this). Other entities perceive the given entity and all agree that it is situated at this same particular location. This applies both to simple entities, as in the dodecahedron example, and to more complex entities, as in the Blackwell’s example. Idealist panpsychism agrees with the empirical fact that you cannot locate my qualitative experiences by examining my brain (Leibniz, 1714, ¶17). (Contrast this with Poland’s thesis (T0) given at the bottom of page 2 above.)

***Is idealist panpsychism a form of dualism?***

*Answer: No.* Experimental entities are essentially ‘mental’ in the broadest sense, and the physics of a universe is ontologically dependent upon its experiential entities. In the dodecahedral example the physics of this universe (the dodecahedron itself) is *instantiated* precisely in the sense that it is composed of mutually-perceiving experiential entities in the manner described above. There is nothing more to physical existence than this, and so there is no physical ‘substance’.

***How can an experiential entity be basic if it is so complicated as to have a qualitative, structured percept and a certain wilful freedom of action?***

*Answer: Structure is no barrier to being fundamental.* An electron is considered to be fundamental despite possessing mass, charge, location, spin, momentum, and so on.

***Aren’t the percepts of experiential entities just mathematical objects, so that idealist panpsychism is also no more than mathematics?***

*Answer: No. Percepts are qualitative through and through.* They are also structured, and we can abstract this structure to get a mathematical model of a percept, as in Figure 3.2.

***Doesn’t an idealist panpsychist run into similar problems of explaining physics in terms of experience, just as a physicalist does in explaining experience in terms of physics?***

*Answer: No.* The essential reason why physicalists cannot explain experience is the sui generis character of qualitative experience, relative to the facts of physics. For the idealist panpsychist however, percepts have a pre-existing structure, as in Figure 3.2. Mathematical structure, which gives rise to physics in the manner already explained, is *abstracted* from the percept. Thus there is nothing sui generis.

***Doesn’t the definition of actual existence run into similar problems of circularity as material existence?***

*Answer: No.* The definition of actual existence is grounded for each of us in *cogito ergo sum*.

***Isn’t the theory is in a very primitive state?***

*Answer: This is true.*

# 6. Conclusions

I gave a critique of the metaphysical project of physicalism. My major finding is that realisation theories fail to provide the vertical explanations necessary for the project to succeed. Furthermore, the physicalists’ conceptions of physical existence are inadequate, and we need to have some definition of concrete existence that can distinguish this concept from mere mathematical existence. I concluded by *defining* actual existence in terms of experiential existence: the property of having a qualitative, subjective, mathematically-structured percept of other entities that actually exist.

Some ideas of Max Tegmark were described by way of introduction to a discussion of my alternative, panpsychic universe with a dodecahedral physics, which might conceivably *actually* (rather than just mathematically) exist. This universe is very elementary, and much would need to be added – not least a time dimension – before we arrive at a universe in any way resembling our own.

If panpsychism is true then six bizarre features of quantum physics are not unexpected. Physicalism does not predict these features. I conclude that fundamental physics provides some limited evidence favouring panpsychism. Idealist panpsychism has a weak degree of testability, but this compares favourably with the testability of Tegmark’s proposal, and also with the zero testability of other physicalist accounts. Contrasting physicalism with idealist panpsychism I found that the latter fared better when examined by asking questions about these theories.

My overall conclusion is that, despite huge efforts by its proponents, the overarching project of physicalism is bound to fail. The idealist panpsychism introduced here is one example (doubtless among others) of promising, under-explored, non-physicalist avenues for understanding the ontology of the world.

# References

*If a URL is given in addition to the reference, then page numbers given in the text above refer to the document as it appears on the internet.*

Baggott, J. (1992), *The meaning of quantum theory* (Oxford: Oxford University Press).

Berkeley, G. (1710/1996), *Principles of Human Knowledge* and *Three Dialogues* (Oxford: Oxford University Press).

Bohm, D. (1980), *Wholeness and the implicate order* (London: Routledge).

Bouwmeester, D., Ekert, A. & Zeilinger, A. (eds, 2000), *The Physics of Quantum Information* (Berlin: Springer).

Chalmers, D. (1995), ‘Facing up to the problem of consciousness’ *Journal of Consciousness Studies*, **2** (3), pp. 200-219.

Chalmers, D. (1996), *The conscious mind – In search of a fundamental theory* (Oxford: Oxford University Press).

Davies, P. & Brown, J. (eds, 1986), *The ghost in the atom* (Cambridge: Cambridge University Press).

Eddington, A. (1928), *The nature of the physical world* (Cambridge: Cambridge University Press).

Everett, H. (1957), ‘'Relative state' formulation of quantum mechanics’ *Review of Modern Physics*, **29**, pp. 454-462.

Feynman, R. (1990), *QED: the strange theory of light and matter* (London: Penguin).

Goodman, N. (1978), *Ways of worldmaking* (Indianapolis: Hackett).

Hawking, S. & Ellis, G. (1973), *The large scale structure of space-time* (Cambridge: Cambridge University Press).

Hawking, S. (2001), *The universe in a nutshell* (London: Bantam).

Hogben, L. (1936), *Mathematics for the million* (London: George Allen & Unwin).

Hut, P., Alford, M. & Tegmark, M. (2006), ‘On math, matter and mind’ *Foundations of Physics*, **36**, pp. 765-794.

http://arxiv.org/pdf/physics/0510188

James, W. (1890/1983), *The principles of psychology* (Cambridge MA: Harvard University Press).

Kant, I. (A1781/B1787/1998), *Critique of pure reason* (Cambridge: Cambridge University Press).

Leibniz, G. (1714), *Monadology*. English translation taken from http://www.marxists.org/reference/subject/philosophy/

works/ge/leibniz.htm

Mattuck, R. (1992), *A guide to Feynman diagrams in the many-body problem* (New York: Dover).

Misner, C., Thorne, K. & Wheeler, J. (1973), *Gravitation* (San Francisco: Freeman).

Penrose, R. (1989), *The emperor’s new mind* (Oxford: Oxford University Press).

Poland, J. (1994), *Physicalism: the philosophical foundations* (Oxford: Clarendon Press).

Savile, A. (2005), *Kant’s critique of pure reason – an orientation to a central theme* (Oxford: Blackwell).

Searle, J. (1992), *The rediscovery of the mind* (Cambridge MA: MIT Press).

Searle, J. (2004), *Mind – a brief introduction* (Oxford: Oxford University Press).

Skrbina, D. (2005), *Panpsychism in the West* (Cambridge, MA: MIT Press).

Stapp, H. (2007), *Mindful universe – quantum mechanics and the participating observer* (Berlin: Springer).

Stone, M. & Wolff, J. (eds, 2000), *The proper ambition of science* (London: Routledge).

Strawson, G. (2006), ‘Realistic Monism: why physicalism entails panpsychism’ *Journal of Consciousness Studies*, **13** (10-11), pp. 3-31.

Tegmark, M. (2003), ‘Parallel universes’ in Barrow, J., Davies, P. & Harper, C. (eds), *Science and ultimate reality: from quantum to cosmos*, honouring John Wheeler’s 90th birthday (Cambridge: Cambridge University Press).

http://space.mit.edu/home/tegmark/multiverse.pdf

Whitehead, A. (1929/1978), *Process and reality – an essay in cosmology* (New York: Free Press).

1. Poland actually abbreviates “a realisation theory” to “an RT” here. Similarly he frequently abbreviates “vertical explanation” to “VE”. In quoting Poland I will expand such abbreviations without comment throughout. [↑](#footnote-ref-1)
2. Poland contradicts himself on this point in one place only. He mentions approvingly, but without giving an example, that physicalists can accept top-down influence (p. 34, footnote). [↑](#footnote-ref-2)