Must Science Make Cosmological Assumptions if it is to be Rational?

Nicholas Maxwell, Dept. Philosophy, LSE, London WC2A 2AE.

Published in *The Philosophy of Science: Proceedings of the IPS Spring Conference, 1997* edited by Thomas Kelly, IPS, 1997.

1 Outline

The answer is: Yes. Cosmological speculation about the ultimate nature of the universe, being necessary for science to be possible at all, must be regarded as a part of scientific knowledge itself, however epistemologically unsound it may be in other respects. The best such speculation available is that the universe is comprehensible in some way or other and, more specifically, in the light of the immense apparent success of modern natural science, that it is physically comprehensible. But both these speculations may be false; in order to take this possibility into account, we need to adopt an hierarchy of increasingly contentless cosmological conjectures until we arrive at the conjecture that the universe is such that it is possible for us to acquire some knowledge of something, a conjecture which we are justified in accepting as knowledge since doing so cannot harm the pursuit of knowledge in any circumstances whatsoever. As a result of adopting such an hierarchy of increasingly contentless cosmological conjectures in this way, we maximize our chances of adopting conjectures that promote the growth of knowledge, and minimize our chances of taking some cosmological assumption for granted that is false and impedes the growth of knowledge. The hope is that as we increase our knowledge about the world we improve (lower level) cosmological assumptions implicit in our methods, and thus in turn improve our methods. As a result of improving our knowledge we improve our knowledge about how to improve knowledge. Science adapts its own nature to what it learns about the nature of the universe, thus increasing its capacity to make progress in knowledge about the world.

This *aim-oriented empiricist* conception of science solves outstanding problems in the philosophy of science such as the problems of induction, simplicity and verisimilitude.

2 What Does it Mean to Assert that the Universe is Comprehensible?

The thesis that the universe is comprehensible is interpreted here to assert: the universe is such that there is *something* (God, tribe of gods, cosmic goal, pattern of physical law, cosmic programme or whatever), which exists everywhere in an unchanging form and which, in some sense, determines or is responsible for everything that changes (all change and diversity in the world in principle being explicable and understandable in terms of the underlying unchanging *something*).

If the *something* that determines all change is what corresponds out there in the world to a unified pattern of physical law, then the universe is physically comprehensible. The universe is physically comprehensible, in other words, if and only if some as-yet-to-be-discovered unified physical "theory of everything" is true.

As a simple example of a physically comprehensible universe, consider a universe that consists only of the classical electromagnetic field in empty space (there being no charged particles to create, or be acted on, by the field). The way in which changes in the electric field, **E**, produce changes in the magnetic field, **B**, and vice versa, are given by the following equations:

$$(1)\ \nabla\cdot\mathbf{E} = 0 \qquad (3)\ \nabla\times\mathbf{E} = -\frac{\partial\mathbf{B}}{c\partial t}$$

$$(2)\ \nabla\cdot\mathbf{B} = 0 \qquad (4)\ \nabla\times\mathbf{B} = -\frac{\partial\mathbf{E}}{c\partial t}$$

For those not familiar with the notation of the differential calculus of vector fields, a few words of explanation. **E** and **B** are vector fields; each assigns a vector to each space-time point which varies in a continuous way through space and time. Such a vector field may be thought of as assigning a tiny arrow to each space-time point which varies in length and direction with changes in space and time. The length of the arrow, at any spatial point and instant, represents the strength of the field at that space-time point, and the direction represents the direction of the field at that point. A vector field might represent flowing water, in which case the arrow at each space-time point would specify the velocity of the water at that point and instant.

$\nabla\cdot\mathbf{E}$ means $\partial E_x/\partial x + \partial E_y/\partial y + \partial E_z/\partial z$, where E_x, E_y and E_z are the x, y and z components of the vector **E**. **E**, and its x, y and z components, change, in general, with respect to changing positions at any given time, and with respect to changing time. $\partial E_x/\partial x$, a *partial* derivative, tells us how E_x changes with respect to a change in the direction of the x axis, all other changes being kept constant. $\nabla\cdot\mathbf{E}$ thus gives us the sum of the rate of change of **E** in space with respect to the three spatial directions, x, y and z. $\nabla\cdot\mathbf{E}$ is a measure of the extent to which there is net "flow" of electric field out of, or into, a region of space (all the arrows pointing away from, or towards, some common point). Postulates (1) and (2) thus tell us that there are no sources (or sinks) for the electric and magnetic fields, no electrically or magnetically charged particles. (If **B** represented the flow of water, then $\nabla\cdot\mathbf{B} = 0$ would express the fact that nowhere does water flow into, or flow away from, the given quantity of water (by means of some fixed pipe or drain, for example).)

$\nabla\times\mathbf{E}$, itself a vector, means:
$(\partial E_z/\partial y - \partial E_y/\partial z)\mathbf{i} + (\partial E_x/\partial z - \partial E_z/\partial x)\mathbf{j} + (\partial E_y/\partial x - \partial E_x/\partial y)\mathbf{k}$, where **i**, **j** and **k** are unit vectors that point in the positive directions of the x, y and z coordinates. $\nabla\times\mathbf{E}$ is a measure of the extent to which the "flow" of **E** (the spatial array of arrows that represent **E**) has a circular motion or pattern - so that if **E** represented the flow of water, $\nabla\times\mathbf{E} \neq 0$ would express the fact that there are whirlpools. Postulate (3) specifies the manner in which the circular flow of **E** is related to the change of **B** with time; and postulate (4) specifies the way in which the circular flow of **B** is related to the change of **E** with time.

Ordinarily (1) to (4) would be thought of as factual assertions. Here, they are to be reinterpreted as *analytic* statements, which tell us what it is for something to be the electromagnetic field. The empirical import of the theory is contained in a fifth postulate:
(5) The classical electromagnetic field, as characterized by (1) to (4), exists everywhere.
(1) to (4) are to be regarded as specifying a dispositional or necessitating physical property: if it exists everywhere then, necessarily, that which changes, values of the electric and magnetic fields, change in accordance with (1) to (4). *If* this necessitating property exists everywhere

at an instant (on some space-like hyperplane) then it, together with the instantaneous values of the electric and magnetic fields everywhere, determine all subsequent states of the field. In this sense the necessitating property (of being the classical electromagnetic field), specified by (1) to (4), itself does not change but determines how varying values of **E** and **B** do change.

Any fundamental physical theory that specifies how postulated fundamental physical entities evolve and interact can be interpreted in this "essentialistic" way. Instead of basic laws being interpreted as specifying regularities observed by the entities, they can be reinterpreted as specifying necessitating *properties* that the entities possess which determine (necessarily) how the entities evolve and interact. To say that the universe is physically comprehensible (or equivalently, that *physicalism* is true) is to say that there exists an unchanging *something* (capable of being specified by a unified, essentialistically interpreted physical theory) which exists everywhere and which determines (perhaps probabilistically) the way in which everything that changes does change.[1]

3 Standard Empiricism

There is an obvious objection to my claim that science, in order to be rational, must assume the cosmological thesis of physicalism. It is that things are all the other way round. Science must not make any such assumption if it is to be rational. Physicalism is a metaphysical thesis that is incompatible with current theoretical knowledge: hence it cannot conceivably be a part of current scientific knowledge.

My reply is that this objection depends on the adoption of a widely held conception of science - which I shall call *standard empiricism* - which, as it happens, is untenable. The moment standard empiricism is rejected, and a more reasonable conception of science is adopted, it becomes clear that the above objection is invalid. It becomes clear that physicalism *is* a part of current theoretical knowledge in physics.

Standard Empiricism (SE) is the doctrine that in science no substantial thesis about the world can be accepted as a permanent part of scientific knowledge *independent of the evidence*, and certainly not *in violation of the evidence*. In so far as factors other than evidence are appealed to in assessing the acceptability of theories - factors such as the simplicity or explanatory capacity of a theory - this must be done in such a way that no assumption about the nature of the world is permanently upheld, explicitly or implicitly, in science, as a part of knowledge, entirely independently of evidence. It may be that, for a time, choice of laws and theories in science is biased in the direction of some "paradigm" or "hard core", upheld in science in a quasi *a priori* fashion, in the sort of way described by Thomas Kuhn[2] or Imre Lakatos;[3] in the end, however, according to SE, such paradigms or hard cores must be assessed in terms of the empirical success or failure of the research programmes that they support. Paradigms or hard cores cannot be upheld permanently as items of scientific knowledge, independently of empirical considerations. Both Kuhn and Lakatos subscribe to SE in that they hold that in the end what ought to decide what is accepted and rejected in science is *evidence*, empirical success and failure.

We can, in fact, distinguish two versions of SE which differ on just the question of whether, in the end, empirical considerations ought *alone* to determine choice of theory in science, or whether simplicity considerations are important and legitimate *in addition* to empirical considerations. Let us call the first view *bare* SE and the second *dressed* SE.

4 Refutation of Bare Standard Empiricism.

The basic objection to bare SE is that theoretical knowledge in science cannot possibly be established by an appeal to evidence alone. Any law or theory in physical science is a universal statement; it applies, potentially, to infinitely many occurrences. But we can only hope to verify such a law or theory for finitely many occurrences; we must always remain infinitely far away from verifying the law or theory itself. We can have no grounds, so it seems, for holding that theories that have been successful in the past will continue to be successful in the future, especially when applied to new sorts of phenomena.

Bare SE is untenable, in short, because it creates and fails to solve the problem of induction.

There are at least two parts to the problem of induction, the *methodological* part and the *justificational* part. In order to solve the methodological problem it is required, only, to specify the methods in terms of which science selects theories in the light of evidence; in order to solve the justificational problem it is necessary, in addition, to justify the claim that theories, so selected, constitute knowledge (in some acceptable sense of "knowledge"). The methodological problem is clearly very much weaker than the justificational problem; a quite elementary requirement a conception of science must satisfy in order to be even half way adequate is that it solves the methodological problem. Bare SE is so poor a doctrine that it fails even to solve the methodological part of the problem of induction.

In order to appreciate the force of the argument against all versions of SE, dressed as well as bare, it is important to understand the various ways in which the problem of induction (methodological and justificational) can be elaborated and intensified.

We need to appreciate, to begin with, that all the observational and experimental evidence that we can ever possess in support of any physical theory must always be highly restricted and atypical, verifying at most an almost vanishingly small, non-random sample of the predictive content of the theory. Theories such as Newtonian theory (NT) or quantum theory (QT) that apply at all times and places can only be verified by means of phenomena that have occurred on earth during the past three hundred years or so. We cannot *now* verify these theories as far as the future is concerned, or as far as the present or increasingly distant past is concerned for increasingly spatially distant events. Ordinarily we suppose that our telescopes enable us to observe distant stars and galaxies undergoing motion and other changes that happened long ago. But this is only the case if our present optical theories that we use to interpret these observations are more or less *true* when applied to distant places in the past; and it is of course just these applications of these theories that we cannot verify observationally *without assuming as true the very theories we seek to verify* (since we cannot dispense with these theories unless we return to the past). Granted that we do not rely on unverifiable predictions of this type, we are confined to our observations of events in our immediate environment. But if our current theoretical knowledge is even remotely correct, this restricts us to a minute, and highly atypical portion of the history of the universe (see diagram 1). If we possessed a finite number of verifications of a theory scattered at random throughout the history of the universe, the situation might not be so bad (even if we were infinitely far away from verifying *all* the predictions of the theory). This is not our situation at all. All our observations are restricted to a tiny and highly atypical pin-prick of space-time. Furthermore, within this pin-prick, any physical theory can only be verified to some limited degree of accuracy, infinitely far away from verifying *precisely* the predictions of the theory.

But the situation is even worse than this would suggest. For physical theories like NT or QT do not merely make predictions about what actually occurs; they contain an infinite wealth of counterfactual predictions - predictions about states of affairs that could, but do not, exist. NT predicts the path a rocket would pursue were it to be fired at such and such a time and place with such and such a velocity, even though it is not fired at all. The capacity to make counterfactual predictions of this type is what distinguishes genuinely *law-like* statements from so-called accidental generalizations.

A genuine physical theory, in short, makes predictions about an infinity of *possible* physical states of affairs, almost all of which will never occur in reality. Furthermore, it is vital that accepted physical theories make correct predictions for *possible* (or counterfactual) phenomena as well as for *actual* phenomena. This is required if physical theories are to serve their function when used for practical, technological purposes. Whenever we employ physical theory in connection with technology we have created, we apply the theory to circumstances which would not have arisen but for our intervention. We need the theory to be correct for a range of *possible* phenomena we consider during the process of design and planning but which never occur *in actuality*.

It follows that even if a physical theory could be precisely verified for all actual phenomena throughout the entire history of the universe, we would still be infinitely far away from verifying the theory itself, for we would not have verified its predictions for an infinity of *possible* phenomena. Our universe is but one of an infinity of *possible* universes, infinitely many of which would differ from ours (if they existed) in only minor ways (such as that a pin on a table somewhere in the universe is half an inch from its actual position). It is vital that we take such possibilities, such possible universes, seriously in life, in considering alternative lines of action. In pondering alternative lines of actions, we require our theories to be as correct in such (slightly different) possible universes as in the actual universe. But we can only conceivably verify any theory for just *one* of these infinitely many possible universes (and we cannot even remotely do that). The prospects of verifying a physical theory *empirically* seem utterly hopeless.

As David Hume well knew,[4] it does not help to argue merely that theories can be verified if we accept a principle asserting that nature is uniform, for at once the question arises as to how we can justifiably accept this principle. We cannot argue that it is verified by experience, for this would require that we assume the very principle we seek to verify. Nor does it help, as Hume again understood,[5] to argue that we should forego *conclusive verification* and accept instead *probable verification*: the above considerations apply with equal force to any such probabilistic version of SE.

Bayesianism tries to get round Hume's arguments by interpreting probabilities of scientific hypotheses as being entirely *subjective* in character, recording no more than the subjective confidence, or lack of it, of the individual scientist in this or that theory. But there are two major difficulties. First, we may doubt that Bayesianism really can show how diverse prior probabilities, which different scientists give initially to an untested scientific hypothesis, gradually converge on some common posterior probability, as a result of subsequent empirical testing. Second, and much more seriously, Bayesianism cannot hope to generate rational grounds for preferring one (unrefuted) scientific theory to another, just because it restricts itself to considering subjective probabilities only. According to Bayesianism, there is in science nothing more than the different *prejudices* of individual scientists, partly clashing and partly in accord with one another, prejudices which are modified somewhat by the results of empirical testing. Why should the mere prejudices of professional scientists be taken any more seriously than the prejudices of (consistent) madmen? Every time we cross a bridge, take medicine or fly in an aeroplane we entrust our lives to the prejudices of scientists; and yet Bayesianism fails completely to explain why it is rational to do this, why the prejudices of scientists deserve to be taken any more seriously than those of lunatics. In so far as the judgements of experienced scientists *do* deserve to be taken seriously, there must, surely, be some objective reason for this; Bayesianism cannot begin to provide any such reason - a point admitted by Bayesians themselves when they concede that their doctrine does not solve the problem of induction.[6]

It does not help, either, to forego verification altogether and accept, with Popper, that there is at most empirical falsification of theory in science.[7] For, given any theory, however massively empirically successful, the above considerations in effect show that there will always be infinitely many rival theories which agree precisely with the given theory for all available evidence but which disagree for some as yet unobserved phenomena. All we need to do in order to construct such "aberrant" versions of an accepted empirically successful theory, T (such as NT in the last century, let us say, or QT or general relativity today), is to modify T *for as yet unobserved phenomena* in any way we please. In this way we can always construct endlessly many of the following five kinds of aberrant rival theories to T. These aberrant rivals to T would never be taken seriously for a moment in scientific practice; but bare SE, whether verificationist or falsificationist (i.e. Popperian), can provide no grounds for their wholesale exclusion from science. Indeed, the following aberrant theories become increasingly scientifically acceptable, as we move down the list, from the standpoint of bare SE, a devastating indictment of the position.

(1) The new aberrant theory, T_1, agrees with T for all observed phenomena, disagreeing only for phenomena within some as yet unobserved space-time region. Example: If T is NT, then in order to create the aberrant theory, NT_1, the Newtonian postulate "$F = Gm_1m_2/d^2$ at all times and places" needs only be modified to read: "$F = Gm_1m_2/d^2$ for times before 12 pm, 31st December 1999 and $F = Gm_1m_2/d^3$ for times at or after 0.0 am, 1st January 2000". Until the 1st January 2000, the aberrant theory fits all the available evidence just as well as NT does, although after that date it yields drastically different predictions. Doubtless NT_1 will be decisively refuted when the year 2000 dawns, but there will always remain infinitely many further aberrant theories of this type to be refuted.

(2) The aberrant theory, T_2, agrees with T for all observed *kinds* of phenomena (wherever and whenever they may occur) but disagrees for some *kind* of phenomenon never observed. We saw, above, that almost all the possible states of affairs to which a theory such as NT refers will never occur, thus remaining unobserved even if all of space and time is observed. Thus we can easily specify endlessly many different *kinds* of physical states of affairs to which the theory applies that will never have been observed, and may never *be* observed, but *could* be observed. For example: "Six spheres, within such and such size and accuracy, made of solid gold, to such and such degree of purity, move without colliding under gravitation in a vacuum, of such and such degree, within such and such a spatial region." If the spheres must have a mass greater than one million tons (and permitted departures from precision are minute), it is quite likely that no technologically advanced civilization anywhere, in the entire history of the cosmos, will create the six-gold-sphere system specified in the way indicated. One particular example, then, of this second type of aberrant theory is: "For all systems that are not a six-gold-sphere system, $F = Gm_1m_2/d^2$ holds, but for all six-gold-sphere systems, $F = Gm_1m_2/d_3$ holds". There is an infinity of such aberrant versions of NT which, though violently at odds with NT, will nevertheless never clash with NT for anything that actually occurs, anywhere in the history of the cosmos.

(3) The aberrant theory, T_3, agrees with T for all observed kinds of phenomena but disagrees for some particular kind of phenomenon observed many times *except for some irrelevant,*

bizarre detail. Let us suppose that the phenomenon in question is Galileo's experiment: balls roll with (nearly) constant acceleration down inclined planes. This confirms NT. The unobserved phenomenon is the following variant of Galileo's experiment: 5 lb. of gold dust is sprinkled around the inclined plane on the laboratory floor. Instead of Newton's law of gravitation, NT_3 asserts: "For phenomena except this particular kind of phenomenon, $F = Gm_1m_2/d^2$, but for this modified-Galileo kind of phenomenon, $F = Gm_1m_2/d^3$". Any kind of experiment (or phenomenon), however often observed to confirm some theory T, can always be modified in this sort of bizarre, irrelevant way, in endlessly many different ways, so as to ensure that *precisely* this kind of (modified) phenomenon has never been observed, and no doubt never will be observed. This third kind of aberrant theory, in other words, is a special case of the second kind of aberrant theory - one that is especially easy to test experimentally.

(4) Taking T_3 as our starting point, we may add some independently testable and confirmed postulate, h, to T_3 to form $T_3 + h = T_4$, the fourth kind of aberrant variant of T. Thus to NT_3 we might add: "Copper expands when heated (between such and such a temperature range, in such and such conditions)", in this way forming a new theory, NT_4. Here, NT_4 has greater empirical content than NT_3, the excess content being in part corroborated empirically.

(5) So far it has been assumed that T yields nothing but correct predictions. This assumption is unrealistic. All actual physical theories, however empirically successful, have their limitations and problems. For even the most empirically successful physical theory, a more realistic assessment of overall empirical support would be something like the following: the predictions of T are accurate for phenomena A, more or less accurate for phenomena B, not forthcoming for phenomena C (because here the equations cannot be solved), apparently incorrect for phenomena D (perhaps because the derivations from T are incorrect, or the experiments have not been done properly), and not forthcoming for phenomena E because beyond the scope of T (see diagram 2). There are now infinitely many aberrant variants of T which assert: everything occurs as T predicts in A; in B phenomena occur well within experimental error (and this is specified piecemeal); in C the observed phenomena occur (and this is specified piecemeal); and so too for D and E. All theories of this type are vastly more empirically successful than T, in that they successfully predict everything that T does; they are not refuted where T is (in B and D); and they make accurate predictions where T yields no predictions (in C and E).[8]

It is important to appreciate that empirically successful aberrant theories, of types (1) to (5), arise all the time, even when firmly established physical laws and theories are applied in utterly standard circumstances to standard phenomena, in order to design or manufacture a transistor radio, let us say, a bridge or an aeroplane. No one bridge is exactly like any other, in all details, especially if we take the environment of the bridge into account. We ordinarily assume, of course, that the differences are irrelevant from the standpoint of the trustworthiness of physical laws that determine the strength of steel and so on. The above considerations show that it will always be possible, however, to formulate rival (aberrant) laws, just as well verified empirically as ordinary, accepted laws, if not more so, which postulate that, in the precise circumstances of such and such a bridge, unprecedented phenomena occur: steel will become soft as butter, and the whole edifice will collapse.

There are *infinitely* many aberrant rivals to T in each of (1) to (5) above, each aberrant rival to T being just as empirically successful as T, at the very least. Indeed, each one of the infinitely many aberrant rivals to T in (4), and in (5) is actually *empirically more successful* than T. We can of course set out to refute experimentally these aberrant theories, but as there are infinitely many of them, there will always remain infinitely many unrefuted. How, then, can it be justifiable to accept any physical law or theory as constituting knowledge, even conjectural knowledge, on the basis of empirical success and failure *alone*? All versions of bare SE fail hopelessly to solve the problem. Indeed, all versions of bare SE fail merely to specify *methods* which lead to the exclusion of infinitely many empirically successful aberrant theories from scientific consideration: bare SE fails to solve even the methodological part of the problem of induction.[9]

5 Refutation of Dressed Standard Empiricism

According to dressed SE, two kinds of criteria govern choice of theory in science, criteria having to do with empirical success, and criteria having to do with the simplicity, unity or explanatoriness of theories. In order to be accepted as a part of scientific knowledge a theory must satisfy both kinds of criteria sufficiently well. The above empirically successful aberrant theories, of types (1) to (5), may satisfy empirical criteria sufficiently well for acceptance, but they fail dismally to satisfy non-empirical criteria of simplicity, unity or explanatoriness. It is for this reason that such theories fail to be accepted as a part of knowledge. Such theories, indeed, are not even formulated as candidates for consideration; merely in order to qualify as a potential theory, a testable conjecture must meet requirements of simplicity or unity.

There can be no doubt that dressed SE does better justice to scientific practice than bare SE does. It comes much closer to solving at least the methodological part of the problem of induction, insofar as it specifies methods which exclude empirically successful aberrant theories, of types (1) to (5), on the grounds that these theories lack simplicity, unity or explanatoriness.[10] But it does so at the cost of violating the basic idea of SE. Persistently to prefer simple, unified, explanatory theories to *ad hoc* theories, of types (1) to (5) above, that are just as empirically successful if not more so, *is* to make a persistent, substantial assumption about the nature of the universe, independently of or even against the evidence, to the effect that the phenomena themselves are, at least approximately, simple, unified, non-aberrant, amenable to explanation (see diagram 3). This is equivalent to *rejecting* SE. Dressed SE is, in other words, inconsistent: it both demands that no permanent assumption be made about the nature of the universe independent of the evidence and also demands, by implication, that just such an assumption must be made.

Suppose that, in physics, only theories that postulate the existence of atoms that interact by contact are ever considered. Field theories, however empirically successful they might be empirically, are rejected out of hand just because they are not atomic theories. In this case, surely, we would all declare, without any hesitation, that physics is just assuming that the world is made up out of atoms entirely independently of the evidence, or even in violation of the evidence. Precisely the same point arises, I claim, for a physics which persistently chooses *simple*, *unified* theories in preference to empirically more successful complex, disunified, aberrant theories. To do this *is* to assume that the world is, at least approximately, simple or unified independently of, or despite, the evidence. It is to violate SE.

Those who defend versions of dressed SE try to have their cake and eat it. That is, they try to develop a conception of theoretical simplicity or unity which, when invoked in science, simultaneously (a) suffices to exclude empirically successful aberrant theories of types (1) to (5) above, and (b) does not commit science to any persistent, substantial assumption about the world. But this cannot be done. In excluding empirically successful aberrant theories, of types (1) to (5), on the grounds that they fail to comply with a methodological principle of simplicity, we *are* making a substantial assumption about the world, and are thus violating SE. If the methodological principle of simplicity is weakened sufficiently so that adopting it in science does not commit science to any assumption about the world, it becomes too weak to exclude aberrant theories of types (1) to (5), and the bare SE nightmare of science being drowned in an infinity of rival theories is upon us.

Dressed SE is, in other words, caught in a trap. The more a notion of simplicity is developed whose use in science is compatible with SE, the more useless such a notion becomes from the standpoint of ruling out empirically successful aberrant theories, of types (1) to (5). On the other hand, the more a notion of simplicity is developed which does effectively rule out empirically successful aberrant theories, the more its use in science becomes incompatible with the basic idea of SE.

Being inconsistent, it is not surprising that dressed SE fails to solve the justificational problem of induction. Ignoring the inconsistency, the inability of dressed SE to solve the problem can be seen as follows. In order to solve the problem it is necessary to justify science persistently rejecting infinitely many empirically successful aberrant theories; but this in turn requires justifying an assumption about the nature of the universe that is implicit in rejecting aberrant theories. But dressed SE provides no such justification; indeed, it *cannot*, because it does not, and cannot, recognize that a persistent assumption about the universe is being made in persistently rejecting aberrant theories (and one cannot justify what one cannot acknowledge).

Dressed SE cannot, in the end, even solve the methodological part of the problem of induction. In order to specify methods that exclude all aberrant theories on the grounds that they lack "simplicity"[11], it is necessary to specify precisely what "simplicity" is.[12] There is an obvious, immediate response: a theory is "simple" if and only if it is "non-aberrant", that is, if the theory's *content* is non-aberrant, i.e. what the theory asserts about the world is non-aberrant. We have, in other words, that a theory is "simple" if and only if it makes a non-aberrant assertion about the universe. But dressed SE cannot adopt this solution to the problem. To do so would be to make absolutely explicit that science, in preferring "simple" theories, is assuming that the universe itself is "simple" (i.e. non-aberrant). Those who seek to solve the problem of what "simplicity" is, within the confines of dressed SE, are obliged to develop a conception of "simplicity" which is such that persistently preferring "simple" theories, in this special sense, does *not* implicate science in accepting some assumption about the world to the effect that the world itself is "simple". Attempts to perform this impossible task are, not surprisingly, unsuccessful! They tend to move the reference of "simplicity" so that it no longer applies to the *content* of theories - and so too to the world - but instead applies to the theories themselves, considered as objects in their own right, to the form or structure of the theories. I have in mind attempts to solve the problem of simplicity proposed, for example, by Goodman, Friedman, Kitcher and Watkins.[13]

It is easy to see that all such attempts must fail. Consider a theory, T, which makes a non-aberrant assertion about the world and thus satisfies the above "obvious" requirement of simplicity. Holding the non-aberrant *content* of T constant, we can endlessly reformulate "T" itself in potentially infinitely many different ways, employing different terminology, "languages", concepts, axioms. Some of these will have immensely complex forms or structures, immensely complex formal, axiomatic or inferential structures, however exactly we propose to conceive of such things. Hence no conception of "simplicity", which refers to the form or structure of theories (a la Goodman, Friedman and Co.) can hope to do justice to what is required for the solution to the methodological problem of induction, namely "simplicity" as "non-aberrant *content*" of theories.

Richard Feynman[14] has provided the following amusing illustration of the difficulties one gets into if one tries to characterize "simplicity" in terms of form or structure. Consider an appallingly aberrant universe governed by 10^{10} quite different, distinct laws. Even in such a universe, the true "theory of everything" can be expressed in the dazzlingly simple, unified form: A = 0. Suppose the 10^{10} distinct laws of the universe are: (1) F = ma; (2) $F = Gm_1m_2/d^2$; and so on. Let
$A_1 = (F - ma)^2$, $A_2 = (F - Gm_1m_2/d^2)^2$, and so on for all 10^{10} distinct laws. Let $A = \sum a_r$. The "theory of everything" of this universe can now be formulated as: A = 0. (This is true if and only if each $A_r = 0$.)

To sum up. Bare SE is an honest doctrine. It is at least consistent. But it fails hopelessly as a conception of science in that it fails to exclude infinitely many aberrant theories from science, thus failing to solve even the methodological problem of induction. Dressed SE, by contrast, is a dishonest doctrine; it claims that science makes no permanent assumption about the world but then commits science to preferring simple theories; if this is to be effective it must carry the implication that science does assume that the universe itself is simple. Once this is honestly conceded, it becomes clear that dressed SE is inconsistent. If this inconsistency is going to be obscured, then the assumption that the universe is non-aberrant, implicit in the methods of science, must be repressed, denied. This ensures that the justificational problem of induction cannot be solved; and it ensures that even the methodological problem cannot be solved, since the key notion of "simplicity", central to the methodology, cannot be adequately explicated.

Thus both bare and dressed SE are untenable. Science does, and must, make a substantial cosmological assumption, independently of empirical considerations, if there is to be any theoretical scientific knowledge. The question arises: What ought this cosmological assumption to be? And how is it to be justified?

In answering these questions, two conflicting desiderata need to be taken into account. On the one hand we want to make an assumption that is as contentless as possible, as unlikely to be false as possible, as near to being a requirement for any knowledge at all to be possible. On the other hand we want to make an assumption that is as heuristically and methodologically fruitful as possible, that holds out the greatest hope, if true, for real progress in theoretical scientific knowledge. These desiderata are likely to conflict because the more heuristically and methodologically fruitful a metaphysical assumption is potentially, so the more likely it is to make a quite precise, contentful assertion about the universe, and so the more likely it is to be false. (I am assuming here that, other things being equal, content is inversely related to likelihood of being true.)

6 Aim-Oriented Empiricism

The way to do the best possible justice to *both* ostensibly conflicting desiderata is to make, not just *one* assumption, but rather a number of assumptions, a hierarchy of increasingly attenuated cosmological assumptions, until we arrive at assumptions which are such that doubting them cannot help the growth of knowledge, whatever the nature of the universe may be. At the top of the hierarchy, there are two cosmological assumptions that are so attenuated, so relatively contentless, that there is a good chance of them being true; they are such, I shall argue, that it can never serve the goal of acquiring knowledge about the world to reject them, whatever the world may be like. The top assumption, indeed, is such that its truth is a necessary requirement for us to have any factual knowledge about anything. As we descend the hierarchy of assumptions we adopt, at each level, that cosmological assumption which, if true, holds out the greatest hope for the growth of knowledge, and which seems best to support the growth of knowledge. Corresponding to each cosmological assumption at each level there is a methodological rule which asserts: give preference to those theories which, taken together, best accord with the cosmological assumption. If currently adopted cosmological assumptions, and associated methods, fail to support the growth of knowledge, or fail to do so as apparently successfully as rival assumptions and methods, then assumptions and associated methods are changed, at whatever level appears to be required below the top two assumptions, which are permanently accepted. In this way we give ourselves the best hope of making progress, of acquiring authentic knowledge, while at the same time minimizing the chances of being taken up the garden path, or being stuck in a cul de sac. The hope is that as we increase our knowledge about the world we improve the cosmological assumptions implicit in our methods, and thus in turn improve our methods. As a result of improving our knowledge we improve our knowledge about how to improve knowledge. Science adapts its own nature to what it learns about the nature of the universe, thus increasing its capacity to make progress in knowledge about the world. This capacity of science to improve knowledge-about-how-to-improve-knowledge in the light of improving knowledge is, I wish to claim, of the essence of scientific rationality; it helps to explain the relatively recent explosive growth of scientific knowledge.

According to the conception of physics here being advocated which I call *aim-oriented empiricism* (AOE), we need to display knowledge at the following ten levels (see diagram 4).

Level 1: P_1. Empirical data (low level observational and experimental laws).

Level 2: P_2. Accepted fundamental physical theories (such as general relativity (GR) and QT).

Level 3: P_3. Best available more or less precise version of physicalism which is, at present, I suggest, a doctrine that may be called *Lagrangianism*. According to Lagrangianism, the universe is such that all phenomena evolve in accordance with Hamilton's principle of least action, formulated in terms of some unified Lagrangian (or Lagrangian density), L. We require, here, that L is not the sum of two or more distinct Lagrangians, with distinct physical interpretations and symmetries, for example one for the electroweak force, one for the strong force, and one for gravitation, as at present; L must have a single physical interpretation, and its symmetries must have an appropriate group structure. We require, in addition, that current quantum field theories and GR emerge when appropriate limits are taken.

Level 4: P_4. *Physical Comprehensibility or Physicalism* (already indicated in section 2 above). The universe is such that there is an impersonal, unchanging, knowable *something*, U, that exists everywhere, and determines (deterministically or probabilistically), how that which varies, V, does vary, from instant to instant. (Given U, and given the value of V at an instant throughout the universe, all subsequent values of V are determined uniquely, given determinism, or determined probabilistically, given probabilism.) In other words, the universe is such that some as-yet-to-be-formulated, unified theory of everything is true (specifying the nature of U and V), which might, but need not be, formulatable in terms of a Lagrangian and Hamilton's principle.

Level 5: P_5. *Comprehensibility* (already indicated in section 2 above). The universe is such that there is a knowable *something* (God, tribe of gods, cosmic goal, unified pattern of physical law, cosmic programme or whatever) which, in some sense, determines or is responsible for everything that changes (all change and diversity in the world in principle being explicable and understandable in terms of the underlying unchanging *something*).

Level 6: P_6. *Near Comprehensibility*. The universe is sufficiently approximately comprehensible for the assumption of perfect comprehensibility to be more fruitful than any comparable assumption from the standpoint of improving knowledge.

Level $_7$: P_7. *Rough Comprehensibility*. The universe is such that some assumption of partial comprehensibility is fruitful from the standpoint of improving knowledge.

Level $_8$: P_8. *Meta-Knowability*. The universe is such that there is some discoverable[15] assumption that can be made about the nature of the universe which aids, and does not hinder, the growth of knowledge.

Level 9: P_9. *Epistemological Non-Maliciousness*. The universe is such that it does not exhibit comprehensibility, meta-knowability, or even mere partial knowability more generally, in our immediate environment only. However radically different phenomena may be elsewhere in the universe, the general nature of all such phenomena is such that it can in principle be discovered by us by developing knowledge acquired in our immediate environment. If aberrant phenomena occur, their occurrence is discoverable by us in our immediate environment.

Level 10: P_{10}. *Partial Knowability*. We possess some factual knowledge of our immediate environment, and some capacity to improve this knowledge, sufficient at least to make partially successful action in the world possible; the universe is such that the possession and acquisition of such knowledge is possible.

A few words of clarification concerning the principles at levels 3 to 10. They are to be understood (as I have indicated) in such a way that P_r implies P_{r+1} for r = 3,...9, but not *vice versa*. P_3 has the most content, and is therefore the most likely to be false, while P_{10} has the least content, and is thus the least likely to be false. It is more than likely that P_3 *is* false, progress in theoretical physics requiring that a revised version of this "level 3" principle be accepted (recent developments in quantum gravity, having to do with "duality", suggesting that this may well be the case). It is less likely that P_4 is false, less likely that progress in theoretical physics will require this "level 4" principle be revised, although this is still a possibility. And as we ascend, from r = 3 to r = 8, the corresponding principles become increasingly contentless, and increasingly unlikely to require revision, although the possibility always exists. The cosmological theses, P_3, ... P_8, have the form that they do have in part because of the way that natural science in general, and theoretical physics in particular, have developed during the last four hundred (or two thousand) years. A radically different history, with a radically different outcome, would result in a different set of principles up to some value of r less than 9, the more radically different so the greater the value of r.

P_3 requires, perhaps, a few words of explication. All fundamental, dynamical theories accepted so far in physics, from Newtonian theory, (NT), classical electrodynamics, to general relativity, (GR), non-relativistic quantum theory, (QT), quantum electrodynamics, quantum electroweak-dynamics, quantum chromodynamics, and the so-called standard model, can be formulated in terms of a Lagrangian and Hamilton's principle of least action. In the case of NT, this takes the following form. Given any system, we can specify its kinetic energy, KE (energy of motion), and its potential energy, PE (energy of position due to forces), at each instant. This enables us to define the Lagrangian, L, equal at each instant to KE - PE. Hamilton's principle states that, given two instants, t_1 and t_2, the system evolves in such a way that the sum of instantaneous values of KE - PE, for times between t_1 and t_2, is a minimum value (or, more accurately, a stationary value, so that it is unaffected to first order by infinitesimal variations in the way the system evolves). From the Lagrangian for NT (a function of the positions and momenta of particles) and Hamilton's principle of least action, we can derive NT in the form familiar from elementary textbooks.

It is this way of formulating NT, in terms of a Lagrangian, L, and Hamilton's principle, that can be generalized to apply to all accepted fundamental theories in physics. Thus P_3 asserts that the universe is such that a true, unified theory of everything, T, can be formulated, T being such that it can be given a Lagrangian formulation in the way indicated.[16]

P_4 asserts, a little more modestly, that the universe is such that some kind of true, unified theory of everything, T, can be formulated, T not necessarily being such that it can be given a Lagrangian formulation. P_5 asserts, more modestly still, that the universe is comprehensible *in some way or other*, but not necessarily physically comprehensible. P_6 asserts, even more modestly, that the universe is sufficiently nearly comprehensible for the assumption that it is perfectly comprehensible to be more fruitful from the standpoint of the growth of knowledge than any comparable rival assumption, relative to our existing knowledge. P_7 asserts, more modestly still, that the universe is such that some assumption of partial comprehensibility is more fruitful than any rival, comparable assumption. It might be the case, for example, that the universe is such that there are *three* fundamental forces, theoretical revolutions involving the development of theories that progressively specify the nature of these three forces more and more precisely. In this case, the assumption that there are three distinct forces would be more helpful than that there is just *one* fundamental force (required if the universe is to be perfectly comprehensible physically). Alternatively, it might be the case that the universe is such that, progress in theoretical physics requires there to be a series of theoretical revolutions, there being, after each revolution, one more force: in this case, the assumption that the universe is such that the number of distinct forces goes up by one after each revolution would be more helpful for the growth of knowledge than the assumption that there is just one fundamental force. P_8, even more modestly, asserts merely that the universe is such that existing methods for improving knowledge can be improved. These methods might involve consulting oracles, prophets or dreams; they need not involve developing explanatory theories and testing them against experience. P_9 asserts, still more modestly, that the universe is such that local knowledge can be developed so that it applies non-locally;[17] and P_{10} asserts, even more modestly, that the universe is such that some factual knowledge of our immediate environment exists and can be acquired.

Given these metaphysical principles about the nature of the universe, we can also consider infinitely many further such principles, $\{Q_3\}$,...$\{Q_8\}$, such that, for r = 3,...8:-

(i) Q_r implies P_{r+1} but is not implied by P_{r+1};

(ii) Q_r is incompatible with P_r.

In view of (i) and (ii), we can regard each Q_r as a possible rival to P_r. Solving the justificational problem of induction involves justifying preference for P_r over rivals $\{Q_r\}$.

7 Grounds for Accepting Aim-Oriented Empiricism

AOE is to be accepted and SE is to be rejected because AOE is able to solve the following six problems, none of which can be solved granted SE.

(1) The problem of excluding empirically successful aberrant theories from science.

(2) The problem of what simplicity is.

(3) The problem of induction (i.e the problem of justifying persistent preference for simple theories in science).

(4) The problem of verisimilitude.

(5) The problem of the discovery of fundamental physical theories.

(6) Wigner's problem of the unreasonable effectiveness of mathematics in the natural

sciences.[18]

Space permits consideration of only the first three problems; for the remaining three the reader is referred to my discussion of these problems elsewhere.[19]

That AOE solves the problem of excluding empirically successful aberrant theories from science is obvious. Physicalism implies that no aberrant theory can be true. But that AOE solves the next problem on the list is perhaps not quite so obvious.

8 The Solution to the Problem of Simplicity

"Simplicity"[20] in the present context apparently means the simplicity of the *form* of a law or theory - the extent to which the *functions*, the *equations*, of the theory are simple. But it also means the extent to which a theory is non-*ad hoc*, or *explanatory*, or *elegant*, or *unified*, or *conceptually coherent*, or possessing what Einstein called *inner perfection* or, in other contexts, *beauty, comprehensibility* or *intelligibility*.

In judging some theories to be "simple" and others to be "complex", physicists may mean only that some theories are such that it is an easy matter to solve their equations and extract their empirical predictions, whereas other theories have equations that are fiendishly difficult, if not impossible, to solve. This highly pragmatic meaning of simplicity is, of course, of immense importance in physics - especially in less fundamental, more phenomenological parts of physics, where the aim is primarily the instrumentalist one of predicting phenomena as easily and accurately as possible. There is, however, no particular reason why simplicity, in this pragmatic sense, should be an indication of truth. Here our concern is only with simplicity insofar as this *is* (or is taken to be) an indication of truth.

The problem of what simplicity is breaks up into the following subordinate problems.

(i) The terminological problem: As we have seen, whether a theory is simple or complex appears to depend on how the theory is formulated, the terminology or concepts used to formulate it. But how can such a terminology-dependent notion of simplicity have any significant methodological or epistemological role in science? What determines the "correct" terminology, in terms of which theories are to be formulated so that their simplicity may be appraised? How can there possibly be any such thing as the "correct" terminology? If there is not, does not the whole notion of simplicity of theories collapse? On the one hand, the simplicity or complexity of a theory must, it seems, depend on the terminology used to formulate it, but on the other hand, this cannot, it seems, be the case if the simplicity is to be significant as an indication of truth.

(ii) The problem of changing notions of simplicity. As science develops, what simplicity means changes. What it meant to Newton is different from what it would have meant to a nineteenth century physicist, which is different again from what it would mean to a late 20th century physicist. How can justice be done to the changing nature of simplicity (and to variability from one discipline to another)?

(iii) The problem of the multi-faceted nature of simplicity.

"Simple" is the generic term that philosophers of science tend to use for a whole family of notions that scientists appeal to in assessing the non-empirical merits of theories, as I have indicated above. An acceptable theory of simplicity ought to pick out just one concept as fundamental, but at the same time do justice to the role that the other concepts appear to have in assessing theories in physics.

(iv) The problem of ambiguity. An indication of the complexity of the notion of simplicity in physics is given by the fact that one theory may be, in an obvious sense, much more complex than another, and yet, at the same time be, in a much more important sense, much simpler. The classic case of this ambiguity of simplicity is provided by a comparison of Newton's and Einstein's theories of gravity. In one obvious sense, Newton's theory is much simpler than Einstein's; in another sense, Einstein's theory is the simpler. An adequate theory of simplicity must resolve this puzzling state of affairs.

(v) The problem of doing justice to the intuition of physicists. Physicists are by no means unanimous in their judgements concerning the simplicity of theories, but there is a considerable level of agreement. An acceptable theory of simplicity must do justice to such agreed intuitions.

(vi) The problem of improving on the intuitions of physicists. An acceptable theory of simplicity ought to be able to improve on the intuitions of physicists, if it provides a genuine clarification of the nature of simplicity.

The first of these problems, the terminological problem, is by far the most serious. It has the form of a paradox. The degree of simplicity of a theory both must, and cannot possibly, depend on terminology. AOE solves this paradox as follows.

According to AOE, *the* non-empirical requirement that the totality of fundamental theories, T_n, in physics must satisfy in order to be ultimately acceptable is that T_n is a precise version of the vague, level 4 thesis of physicalism. The key notion is thus *unity* - unity of the *content* of the totality of fundamental dynamical theory. Given two rival total theories, T_n and T_{n+1}, T_{n+1} is simpler than T_n if and only if T_{n+1} exemplifies physicalism better than T_n does. In other words, if T_n is more disunified than T_{n+1} (in one or more of the eight different ways discussed below) then T_n is less simple.

This account of simplicity can be extended to individual theories in two different ways. An individual theory, T*, is "simpler" than a rival, T**, if T_m+T* exemplifies physicalism better than T_m+T** does, where T_m is the conjunction of all other current, accepted, fundamental theories. We can also, however, treat an individual theory as if it is a "theory of everything", ignoring all phenomena which lie outside the domain of the theory. Given two rival individual theories, T_1 and T_2, we can regard them as rival "theories of everything" and consider their relative simplicity, i.e. unity, i.e. their success, when so regarded, of being precise versions of physicalism.

Furthermore, this account can be straightforwardly extended to do justice to the point that notions of simplicity evolve with evolving knowledge. Theoretical physics does not just, in practice, presuppose physicalism; at any given time it presupposes some more precise version of physicalism, some blueprint, B_n, which is almost certainly false and will need to be changed so that it subsequently becomes B_{n+1}, which in turn becomes B_{n+2}, and so on. B_n might be the blueprint that the world is made up of small, rigid, spherical corpuscles with mass that interact only by contact, and B_{n+1} might be the Boscovichean blueprint that the world consists of point-particles with mass that interact by means of a rigid, spherically symmetrical field of force that varies with distance from each point-particle. Other blueprints important in the history of physics are: the aether blueprint of 19th century physics, the aether being an elastic medium filling space which transmits gravitational and electromagnetic forces and light, and in which matter is embedded; the unified field/particle blueprint, charged particles being both sources of the field and acted on by the field; the unified self-interacting field, particles and matter being merely intense regions of the field; the geometrical blueprint, particles and forces being nothing more than topological or geometrical features of space-time; the quantum field blueprint of modern physics, particles being excitations of the field; the Lagrangian blueprint, discussed above; the superstring blueprint, according to which particles have the form of minute strings embedded in space-time of ten or twentysix dimensions, those in excess of four being curled up into a minute size.

In accepting a blueprint B, we accept that the fundamental physical entities and force(s) are as specified by B; we accept a set of invariance or symmetry principles, specific to B, related to the geometry of space-time, and the general dynamical/geometrical character of the fundamental physical entity (or entities), postulated by B.

Thus, the Boscovich blueprint may be so understood that it asserts that fundamental physical entities - point-particles with mass - are all of one type (symmetric with respect to particle exchange), rigid throughout all motions, and rotationally symmetric. The time *evolution* of any physical system is invariant with respect to translations in space and time, changes of orientation, and changes in fixed velocity with respect to some inertial reference frame. By contrast, the field/particle blueprint, as understood here, associated with classical electrodynamics - postulates the existence of two distinct kinds of fundamental entities, point-particles and fields of force; it asserts that force-fields are non-rigid (changes in the field travelling at some definite, finite velocity). This means that spherical-symmetry will be restricted to the case when a charged particle is motionless in a spatial region within which the field is otherwise zero. Again, whereas the Boscovich blueprint may be taken to imply Galilean invariance, the field/particle blueprint may be taken to imply Lorentz invariance.

A level 2 theory, T, may clash with physicalism and yet exemplify physicalism to some degree, in that it is disunified to some degree in one or more of the following eight ways of being disunified.

(1) T has a different *content* in the N different space-time regions, $R_1, \ldots R_N$,

(2) T postulates, in an arbitrary fashion, N distinct, unique, spatially localized objects, each with its own distinct, unique dynamic properties.

(3) T postulates that, for distinct ranges of physical variables, such as mass or relative velocity, in distinct regions, $R_1, \ldots R_N$ of the space of all possible phenomena, distinct dynamical laws obtain.

(4) T postulates N different kinds of physical entities,[21] differing with respect to some dynamic property, such as value of mass or charge, but otherwise interacting by means of the same force.

(5) As in (4) except that the distinct kinds of entities interact by means of distinct forces. If there are N distinct kinds of entities, and each kind interacts with all other kinds in distinct ways, then there will be N^2 distinct interactions, forces, and component dynamic theories.

(6) Consider a theory, T, that postulates N distinct entities (e.g. particles or fields), but these N entities can be regarded as arising because T exhibits some symmetry. If the symmetry group, G, is not a direct product of subgroups, we can declare that T is fully unified; if G is a direct product of subgroups, T lacks full unity; and if the N entities are such that they cannot be regarded as arising as a result of some symmetry of T, with some group structure G, then T is disunified.

The way in which relativistic classical electromagnetism unifies the electric and magnetic fields is an example of this kind of unity. Given the electric field, then the magnetic field must be adjoined to it if the theory is to exhibit the symmetry of Lorentz invariance. Again, the way in which chromodynamics brings unity to the eight gluons, and to quarks that differ with respect to colour charge, postulated by the theory, provides another example of this kind of unity. The diverse gluons and colour charged quarks of the theory are required to exist if the theory is to have its distinctive locally gauge invariant character, in this case the symmetry group being SU(3). The electroweak theory of Salam and Weinberg is an example of partial unity of this type, in that, in this case, the symmetry group, corresponding to the locally gauge invariant character of the theory, is SU(2) X U(1) - a group that is a direct product of subgroups. The theory only partially unifies the diverse quanta of the associated fields, the photon of electromagnetism and the vector bosons of the weak force.[22]

(7) According to GR, Newton's force of gravitation is merely an aspect of the curvature of space-time. As a result of a change in our ideas about the nature of space-time, so that its geometric properties become dynamic, a physical force disappears, or becomes unified with space-time. This suggests the following requirement for unity: space-time on the one hand, and physical particles and forces on the other, must be unified into a single self-interacting entity, U. If T postulates space-time and physical "particles and forces" as two fundamentally distinct kinds of entities, then T is not unified in this respect.

(8) If (apparent) disunity has emerged as a result of a series of cosmic spontaneous symmetry-breaking events, there being manifest unify before these occurred, then the relevant theory, T, is unified. If current (apparent) disunity has not emerged from unity in this way, as a result of spontaneous symmetry-breaking, then the relevant theory, T, is disunified.

We have here, then, eight *different* ways in which the totality of fundamental physical theory can exemplify physicalism to some degree N.

Analogously, T may clash with a blueprint, B, and yet exemplify B to some degree, in that it postulates B-type entities, forces and symmetries, but at the same time violates, to some degree, and in one or more ways, the specific kind of unity postulated by B. The ways in which T violates B may differ in some respects from the eight ways in which T may violate physicalism. If B is the Boscovichean blueprint, only the first five of the eight unity/disunity distinctions just specified are relevant. In this case, B does not postulate anything like a force exhibiting local gauge invariance (6); it does not unify space-time and matter (7); and it does not postulate spontaneously broken symmetries (8). Consequently, even though T fails to be unified in ways (6), (7) and (8), this does not mean that T lacks B-type unity. If, on the other hand, the point-particles postulated by T have force-fields that lack rigidity and spherical symmetry, this would constitute a violation of B-type unity or simplicity, even though this does not, as such, violate physicalism.

Given two distinct blueprints, B_n and B_{n+1}, which postulate somewhat different kinds of entities, forces and symmetries (even though there is some overlap), we have two distinct notions of simplicity, B_n-simplicity and B_{n+1}-simplicity. A theory, T, may have a high degree of B_n-simplicity, and a low degree of B_{n+1}-simplicity.

Blueprints can themselves be assessed with respect to simplicity, with respect, that is, to how well they exemplify physicalism.

The simplicity of level 2 theories can, in short, be assessed in two distinct ways, in terms of what may be called P-simplicity and B-simplicity (degree of exemplifying physicalism and some blueprint, B, respectively).[23] The P-simplicity of a theory, T, assesses how successfully T realizes physicalism, and remains fixed as long as physicalism does not change its meaning. The B-simplicity of T assesses how well T realizes the best available overall blueprint for physics; B-simplicity evolves with evolving blueprints. Furthermore, that blueprints evolve with evolving knowledge is, according to AOE, essential to the rationality of science, a vital, necessary component of scientific progress (granted that we are ignorant of what version of physicalism is true). There is thus, according to AOE, no mystery about evolving notions of simplicity; that the notion of simplicity should evolve is essential to rationality, a vital component of progress. Simplicity criteria, associated with level 3 blueprints, do not merely *change*; they can *improve*. We learn more about the precise way in which Nature is simple or unified as science progresses.

This, in barest outline, is the aim-oriented empiricist solution to the problem of what simplicity *is*.[24]

8 Content and Form

Why does this proposed AOE solution to the problem of what simplicity is succeed where all SE attempts at solving the problem fail? The decisive point to appreciate is that, according to AOE, in assessing the relative simplicity of two theories, T_1 and T_2, what matters is the *content* of the two theories, not their *form*. It is what theories *assert* about the world that must accord, as far as possible, with physicalism, with the thesis that a unified *something* runs through all phenomena. Thus questions of formulation, axiomatic structure, etc., are essentially irrelevant when it comes to assessing the simplicity of theories in a methodologically significant sense. The fact that a theory may seem simple when formulated in one way, highly complicated or *ad hoc* when formulated in another way - a fact that defeated SE attempts at solving the problem - has, according to the AOE view, no bearing whatsoever on the simplicity or unity of the theory in an epistemologically and methodologically significant sense, which has to do exclusively with *what the theory asserts about the world* (which remains constant throughout mere terminological reformulations). What matters, in short, is the simplicity or unity, not of the *theory itself*, but of what the theory *asserts to be the case*. A perfectly simple or comprehensible possible universe may be depicted by a theory that is formulated in a horribly complex fashion; and *vice versa*, a horribly complicated or incomprehensible universe may be depicted by a theory formulated in a beautifully simple way.

So far I have indicated how AOE solves the first two problems concerning what simplicity is: the terminological problem, and the problem of changing, or evolving, conceptions of simplicity. What about the remaining four problems indicated above? I take these in turn.
(iii) The multi-faceted problem. AOE is quite clear: the key notion behind the generic term "simplicity" is unity or explanatoriness. These two notions are connected as follows. The more *unified* a dynamical theory is, other things being equal, so the more *explanatory* it is. To explain, in this sense, is, ideally, to show that apparently diverse phenomena are really just different versions of the *one* kind of phenomenon, differing only with respect to initial conditions but otherwise evolving in accordance with the same force. Thus NT explains the diverse phenomena it predicts by revealing that these phenomena all evolve in accordance with Newtonian gravitation. As long as the totality of physical theory is disunified, explanation is inadequate; the explanatory task of physics is only at an end when all physical phenomena have been shown to be just *one* kind of phenomenon, all differences being differences of initial conditions of the *one* kind of entity or stuff.[25]

Other terms banded about - simplicity, symmetry, elegance, beauty, comprehensibility, etc. - all devolve, more or less straightforwardly, from the central notion of unity-throughout-diversity, or explanatoriness. Thus the requirement that a theory satisfies symmetry principles is related to unity in the ways indicated in (6) and (8) above, on pages 38-39.

It may be asked: Does simplicity (in the non-generic sense) play a role in distinguishing between physically comprehensible and incomprehensible universes? If it does, it takes second place to considerations of unity. This point is best discussed in connection with the fourth problem.

(iv) The problem of ambiguity. General relativity (GR) is, in a quite straightforward sense, a much more complicated theory than Newton's theory of gravitation (NT). NT determines the gravitation field by means of *one* equation, whereas GR requires a system of *six* equations. Furthermore, NT is a linear theory, in the sense that, as one adds more massive bodies to a system of bodies, the gravitational forces due to the new bodies merely add on to the forces already present. GR, on the other hand, is non-linear: the gravitational field interacts with itself. (The gravitational field itself contains energy, which induces curvature into space-time, and thus has gravitational effects.) Finally the equations of GR are vastly more difficult to solve than those of NT; GR is much more complex than NT in terms of the pragmatic notion of simplicity indicated above.

GR has, however, much greater unity than NT. According to NT, gravitation is a force that exists as something entirely distinct from, and in addition to, space and time; according to GR, gravitation is nothing more than the variable curvature of space-time. The field equations of GR specify how the presence of mass, or energy more generally, causes space-time to curve. According to GR, bodies "interacting gravitationally" do not, in a sense, interact at all; all bodies move along the nearest thing to straight lines in curved space-time, namely curved paths called geodesics, that constitute, roughly, the shortest distance between points (the four-dimensional analogue of great circles on the earth's surface). Ordinarily one would think of the earth's motion round the sun as constituting a spiral in four dimensional space-time; according to GR, the mass of the sun causes space-time near the sun to be curved in such a way that the path executed by the earth is a geodesic, the shortest distance between any two space-time points occupied by the centre of the earth at successive times.

GR unifies by eliminating gravitation as a force distinct from space-time; space-time has a variable curvature, as a result of the presence of matter or energy, and this variable curvature affects what paths constitute geodesics, and thus what paths bodies pursue; gravitation, as a force, vanishes. As a result, GR does not need an analogue to Newton's second law F = ma; all that is required is a generalization of Newton's first law: every body continues in its state of rest or uniform motion in a straight line, except in so far as a force is imposed upon it. ("Uniform motion in a straight line (in Euclidean space)" needs to be generalized to become "geodesic in Riemannian space-time".)

Despite its greater complexity, GR exemplifies physicalism better than NT because of its greater unity. And there is a further, crucial point. Given the basic unifying idea of GR, namely that gravitation is nothing more than a consequence of the variable curvature of space-time, the equations of GR are just about the simplest that are possible. The complexities of GR are not fortuitous; they are inevitable, granted the fundamental unifying idea of GR.

From this discussion of NT and GR we can draw the following general conclusion. Given two theories, T_1 and T_2, if T_2 has greater unity than T_1 then, other things being equal, T_2 is the better theory from the standpoint of non-empirical considerations, even if T_2 is much more complex than T_1. This will be the case, especially, if the greater complexity of T_2 is an inevitable consequence of its greater unity. Simplicity considerations may have a role to play, on the other hand, if there are two theories, T_1 and T_2, that are unified equally, in the same way, so that, at a certain level, T_1 and T_2 have a common blueprint, but T_1 is much simpler than T_2. In this case, T_1 is a better theory than T_2 on non-empirical grounds. A universe that exemplifies unity in a way that is highly complex in comparison with other possible universes (other possible dynamic structures) unified in the same sort of way, is, we may argue, not fully comprehensible. To this extent, comprehensibility requires simplicity.

(We have here a ninth way of drawing the distinction between unity and disunity, to be added to the eight indicated above.)

The extent to which simplicity considerations, of this limited type, ultimately play a role in what it means to say that the universe is comprehensible depends, to some extent, on the character of the true theory of everything, T. Given T, there are, we may assume, any number of rival theories,
$T_1 \ldots T_n$, that exemplify physicalism just as well as T does, as far as the eight requirements for unity indicated above are concerned. It is conceivable that a level 3 blueprint, B, can be specified which is such that T, together with a proper subset of $T_1 \ldots T_n$, are all equally well B-unified, as far as the eight requirements for unity are concerned, but *one* of these B-unified theories is much simpler than the others. In this case we could declare that, for unity, we require that the simplest of these theories is true.

(v) The problem of doing justice to the intuition of physicists. I shall restrict myself to considering just five points (five items of data, as it were, that any theory of simplicity ought to be able to account for). First, physicists are generally at a loss to say what simplicity is, or how it is to be justified. Second, despite this, much of the time most theoretical physicists are in broad agreement in their judgements concerning the non-empirical simplicity requirements that theories must satisfy to be accepted, at least to the extent of agreeing about how to distinguish non-aberrant from aberrant theories (although of course they do not use this terminology).[26] But third, in addition to this, non-empirical simplicity criteria intuitively accepted by physicists tend to change over time. Fourth, during theoretical revolutions there are often spectacular, irreconcilable disagreements.[27] Rationality tends to break down during revolutions, as graphically described by Kuhn.[28] But fifth, despite all this, intuitive ideas concerning simplicity, at least since Newton, have enabled physics to meet with incredible (apparent) success.

According to the account of simplicity being advocated here, the more nearly the totality of fundamental dynamical theory exemplifies physicalism, so the greater is its degree of simplicity. In practice physics accepts physicalism, even though this may be denied by physicists (because it clashes with the official doctrine of SE). This view accounts for the above five points as follows.

The failure of physicists to say what simplicity is, or how it should be justified, is due to the fact that most physicists accept some version of dressed SE; within this framework no adequate account of the role of simplicity in physics can be given, as we have seen. The general, more or less implicit acceptance of physicalism in practice means that there is, in practice, at any given time, broad agreement concerning judgements of simplicity. (Physicists may merely require that any acceptable theory must be such that it can be given some more or less specific kind of formulation: but this in practice is equivalent to demanding that any theory accord with some blueprint corresponding to the concepts, the language of the formulation.)

According to AOE, even if at level 4 there is no change of ideas, at level 3 it is entirely to be expected that there will be changes over time. (It would be astonishing if, at level 3, the correct guess was made at the outset.) The historical record reveals just such an evolution of blueprint ideas, from the corpuscular hypothesis, via the Boscovichean blueprint, the classical field blueprint, the empty space-time blueprint (with variable geometry and topology), the quantum field blueprint, Lagrangianism, to the superstring blueprint. Thus, over time, judgements concerning simplicity both do, and ought to, evolve with evolving level 3 blueprint ideas. During theoretical revolutions, it is above all level 3 blueprint ideas that change. During such revolutions, some physicists will hold on to the old, familiar blueprint, while others will embrace the new one. This means that physicists will assess the competing theories in terms of somewhat different conceptions of simplicity, related to the different, competing blueprints. General agreement about simplicity considerations will, in these circumstances, break down. Arguments for and against the competing theories will be circular, and rationality will tend to break down in just the way described so graphically by Kuhn. Finally, the success of physics is due, in large part (a) to the acceptance in practice of physicalism (or some fruitful special case such as the corpuscular hypothesis or Boscovicheanism), and (b) to the fact that physicalism is either true or, if false, "nearly true" in the sense that local phenomena occur as if physicalism is true to a high degree of approximation.

(vi) It deserves to be noted that AOE does not merely account for basic facts about physicists' intuitions; it clarifies and improves on those intuitions. Once AOE is generally accepted by the physics community, the breakdown of rationality during theoretical revolutions, noted by Kuhn, will no longer occur. If the revolution is a change from theory T_1 and blueprint B_1 to theory and blueprint T_2 and B_2, an agreed framework will exist for the non-empirical assessment not only of T_1 and T_2, but of B_1 and B_2 as well. Kuhn argues that the breakdown of rationality during revolutions is due to the fact that, ultimately, only empirical considerations are rational in science. During a revolution, empirical considerations are inconclusive; the new theory, T_2, will not have had time to prove its empirical mettle (etc.). Thus rational assessment of rival theories must be highly inconclusive. Insofar as physicists appeal to rival paradigms (as Kuhn calls them), B_1 and B_2, the arguments are circular, and thus irrational (persuading only those who already believe). Accept AOE, and this situation changes. Rational considerations do exist for the (tentative) assessment of the relative merits of B_1 and B_2; we are justified in assessing how adequately they exemplify physicalism. This means, in turn, that we can judge rationally whether we are justified in assessing T_2 (or T_1) in terms of B_2. Such judgements, though rational, will be fallible even if physicalism is true: acceptance of AOE thus makes clear that dogmatism, at the level of paradigms, or level 3 blueprints, is wholly inappropriate. This in itself promotes rationality in physics.

10 The Problem of Induction

AOE belongs to a well known approach to the problem of induction, the approach, namely, of appealing to a principle of the uniformity of nature.[29] Theses at levels 3 to 9 are all uniformity principles (and so is the thesis of weak unity). Even the thesis of partial knowability, at level 10, may be regarded as a highly restricted, qualified uniformity principle. There are, however, on the face of it, three decisive objections to the idea that any such approach can solve the justificational problem of induction.

(1) Any attempt to solve the problem in this way must rest on a hopelessly circular argument. The success of science is justified by an appeal to some principle of the uniformity of Nature; this principle is then in turn justified by an appeal to the success of science. As Bas van Fraassen has put it "From Gravesande's axiom of the uniformity of nature in 1717 to Russell's postulates of knowledge in 1948, this has been a mug's game".[30]

(2) Even if, by some miracle, we knew that Nature is uniform in the sense that the basic laws are invariant in space and time, this still would not suffice to solve the problem of induction. Given any empirically successful theory, T, invariant in space and time, there will always be infinitely many rival theories which will fit all the available data just as well as T does, and

which are also invariant in space and time.

(3) We cannot even argue that the principle of uniformity, indicated in (2), must be accepted because only if the principle is true is it possible for us to acquire knowledge at all. One can imagine all sorts of possible universes in which knowledge can be acquired even though the uniformity principle, as indicated above, is false.

These objections may well be decisive against some traditional attempts to solve the problem of induction by appealing to a principle of the uniformity of Nature, but they are harmless when directed against AOE.

What differentiates earlier "uniformity" views from AOE is that whereas the earlier views appeal to just *one* (possibly composite[31]) principle of uniformity, strong AOE appeals to *eight* distinct uniformity principles upheld at *eight* distinct levels, these principles becoming progressively more and more contentless as we ascend from level 3 to level 10. This difference is decisive as far as the above three objections are concerned.

Reply to (1): It is obviously fallacious to justify the uniformity of Nature by an appeal to the success of science, and then justify the success of science by an appeal to the uniformity of Nature. Any view which appeals to just *one* (possibly composite) uniformity principle becomes fallacious in this way the moment it appeals to the success of science. The only hope of a valid solution to the problem along these lines is to justify accepting the specified uniformity principle on the grounds that there is no alternative: if the principle is *false*, all hope of acquiring knowledge disappears, and thus we risk nothing in assuming the principle to be true. Just this kind of justification is given by AOE for principles accepted at levels 10 and 9 - a kind of justification which makes no appeal to the success of science, and thus entirely avoids the above fallacy. In addition, however, according to AOE, we need to choose between rival, much more specific, contentful uniformity principles in such a way that we choose those that seem to be the most fruitful from the standpoint of promoting the growth of empirical knowledge. Choice of principles, at levels 3 to 7 at least, *is* influenced by the (apparent) success of science, or the (apparent) success of research programmes within science. But this does *not* mean that AOE commits the above fallacy of circularity: principles at levels 9 and 10 are justified without any appeal to the success of science at all. Just because AOE appeals to eight principles, graded with respect to content, it becomes possible to give different justifications for these principles at different levels, something that is not possible if an appeal is made to only *one* uniformity principle.

Reply to (2): As a result of specifying eight uniformity principles, graded with respect to content, AOE is able to uphold, at level 3 or 4, uniformity principles much stronger than the principle that laws should be uniform in space and time, sufficiently strong indeed to pick out, at any given stage in the development of physics, that small group of fundamental dynamical theories that do the best justice (a) to the evidence and (b) to the best available level 3 or level 4 principle.

Reply to (3): Traditional "uniformity" views that appeal to just *one* uniformity principle have the impossible task of formulating a principle which is simultaneously (i) sufficiently *strong* to exclude empirically successful aberrant theories and (ii) sufficiently *weak* to be open to being justified along the lines that it is impossible to acquire knowledge if the principle is false. AOE, as a result of specifying eight principles, graded with respect to content, is not required to perform this impossible task. At levels 9 and 10 uniformity principles are accepted that *are* sufficiently weak to be justified along the lines that it is impossible to acquire knowledge if they are false; at levels 3 and 4 principles are adopted that are sufficiently strong to exclude empirically successful aberrant theories. These latter principles are not such that they must be true if any advance of knowledge is to be possible; circumstances are conceivable in which these strong principles ought to be revised in the interests of further acquisition of knowledge. Indeed, at level 3, such revisions have occurred a number of times during the development of modern physics.

In outline, then, the proposed solution to the problem of induction amounts to this. We are justified in accepting the cosmological assumptions that the universe is partially knowable and epistemologically non-maliciousness, at levels 10 and 9, because we have nothing to lose; it cannot harm the pursuit of knowledge to accept these assumptions in any circumstances whatsoever. The same cannot be said for the level 8 assumption of meta-knowability: if we accept this assumption and it is false, we will be led fruitlessly to search for improved methods for the improvement of knowledge. On the other hand, granted that it *is* possible for us to acquire knowledge (as level 10 and 9 theses assert), it is not unreasonable to suppose that existing methods for the improvement of knowledge can be improved. Unless we try to discover improved methods for the improvement of knowledge, we are unlikely to discover such methods, even if meta-knowability (relative to our existing knowledge) is true. And if new methods rapidly generate new knowledge that satisfies existing criteria for knowledge, and survive our most ferociously critical attempts at refutation, then the thesis of meta-knowability deserves to be taken very seriously indeed. The immense apparent success of science fulfils these conditions, and indicates that meta-knowability deserves to be adopted as a part of our conjectural knowledge.

Granted level 1 evidence and the level 3 thesis of Lagrangianism, current fundamental physical theories, the standard model and general relativity, deserve to be accepted. But why should Lagrangianism be accepted? Granted the evidence and the level 4 thesis of physicalism, there is no other available cosmological conjecture, at an equivalent level of generality, that has appeared to be as fruitful for the generation of (level 2) theoretical knowledge as Lagrangianism. But why should physicalism be accepted? Granted the evidence and the level 5 thesis of comprehensibility, there is no other conjecture, at an equivalent level of generality, which has appeared to be so fruitful for the generation of (level 2) theoretical knowledge as physicalism. Why should the comprehensibility thesis be accepted? Granted acceptance of the level 6 thesis of near comprehensibility, it is all but tautological that the level 5 thesis of perfect comprehensibility should be accepted. But why should near comprehensibility be accepted? Because, granted the level 7 thesis of rough comprehensibility, no thesis other than near comprehensibility, at a comparable level of generality, has appeared to be as fruitful for the generation of level 2 knowledge. And why should rough comprehensibility be accepted? Because, granted the level 8 thesis of meta-knowability, no other thesis, at a comparable level of generality, has appeared to be as fruitful for the generation of level 2 knowledge.

11 Is Aim-oriented Empiricism Necessary for the Solution of the Problems of Simplicity and Induction?

I have argued that *if* AOE is accepted, problems of simplicity and induction can be solved. But is it *necessary* to accept AOE in order to solve these problems? Might it not be sufficient to accept a much weaker assumption than Lagrangianism at level 3, or physicalism at level 4? Might it not be sufficient to assume that the universe is such that it behaves as if non-aberrant to a sufficient degree of approximation as far as phenomena are concerned that we encounter in practice, for all empirically successful theories that are *aberrant* somewhere within this region of possible phenomena to be *false*? If we are justified in accepting this assumption of *local approximate non-aberrance* (as we may call it) then, so we may argue, we are justified in excluding all empirically successful locally aberrant physical theories from consideration, which would suffice to solve the problem of induction. But the assumption of local approximate non-aberrance is much weaker than physicalism: it is thus more rational to accept it rather than physicalism. Hence AOE is not required to solve the problem of induction.

I have two replies to this objection. First, local approximate non-aberrance does not suffice to solve the two basic problems of simplicity: for that physicalism *is* required. Second, it is wrong to assume that it is rational to accept the weakest possible cosmological assumption that suffices to exclude empirically successful aberrant theories. As I emphasized above, two considerations govern choice of cosmological assumptions built into the methods of science: we require that such assumptions are (a) as unlikely as possible to require revision, and (b) as fruitful as possible for the growth of empirical knowledge. Assumptions at levels 9 and 10 satisfy (a); assumptions at levels 3 to 7 satisfy (b). From the standpoint of fruitfulness, actual and potential, physicalism is, it seems, vastly superior to the thesis of local approximate non-aberrance, given the historical record of physics, and it is on that ground that we are justified in accepting it as a part of (conjectural) scientific knowledge.

Notes

1. For further details concerning this essentialistic approach to physics, and the hypothetical existence of necessary connections between successive events, see: N. Maxwell, Can there be Necessary Connections between Successive Events?, *The British Journal for the Philosophy of Science 19*, 1968, pp. 1-25; Induction and Scientific Realism: Einstein versus van Fraassen. Part Two: Aim-Oriented Empiricism and Scientific Essentialism, *The British Journal for the Philosophy of Science 44*, 1993, pp. 81-101.
2. T.S. Kuhn, *The Structure of Scientific Revolutions*, 1962, University of Chicago Press.
3. I. Lakatos, Falsificationism and the Methodology of Research Programmes, in I. Lakatos and A. Musgrave, eds., *Criticism and the Growth of Knowledge*, 1970, Cambridge University Press, pp. 91-195.
4. D. Hume, *A Treatise on Human Nature*, vol. 1, 1959, Dent, Section XII.
5. Ibid.
6. For a detailed critical examination of Bayesianism see J. Earman, *Bayes or Bust?*, 1992, MIT Press.
7. K.R. Popper, *The Logic of Scientific Discovery*, 1959, Hutchinson; *Conjectures and Refutations*, 1963, Routledge and Kegan Paul, Ch. 1.
8. The idea of formulating the problem of induction in terms of empirically successful aberrant theories was introduced by me in 1974; see my The Rationality of Scientific Discovery, Part I, *Philosophy of Science 41*, 1974, pp. 123-153. This generalizes and improves on Goodman's way of formulating the problem, in terms of "aberrant" predicates"bleen" and "grue": see N. Goodman, *Fact, Fiction and Forecast*, 1954, Athlone Press.
9. For earlier formulations of this argument see N. Maxwell: A Critique of Popper's Views of Scientific Method, *Philosophy of Science 39*, 1972, pp. 131-52; The Rationality of Scientific Discovery, Parts I and II, *Philosophy of Science 41*, 1974, pp. 123-53 and 247-295; Induction,

Simplicity and Scientific Progress, *Scientia 114*, 1979, pp. 629-53; *From Knowledge to Wisdom: A Revolution in the Aims and Methods of Science*, Basil Blackwell, 1984, ch. 9; Induction and Scientific Realism: Einstein versus van Fraassen, Parts I and II, *The British Journal for the Philosophy of Science 44*, 1993, pp. 61-79 and 81-191.

10. In the end it fails even here, as we shall see.

11. "Simplicity" is being used here as a blanket term, to cover a number of other terms used in this context, such as "explanatoriness", "non-*ad hocness*", "non-aberrance", "unity", "symmetry", "beauty", "elegance", "conceptual coherence", "inner perfection". Whether there is just one relevant notion here, or two or more notions, I leave open, at the present stage of the argument. The common idea lurking behind the diverse terms is that there are non-empirical criteria for the selection of theories in science, which have to do with the form or nature of theories (or of what they assert), rather than their empirical success or failure.

12. Granted dressed SE, the methodological problem of induction becomes the problem of specifying what "simplicity", in a methodologically significant sense, *is*. Even defenders of dressed SE themselves tend to forget that the problem of specifying what simplicity is *is* the methodological problem of induction. In the literature it sometimes seems that the problem of what simplicity is is one problem, and the problem of induction is another more or less distantly related problem, whereas the former is in fact a basic part of the latter. The tendency to divide up problems in this way is due to specialization; it has the disastrous consequence that solutions to philosophical problems, which require the integration of problems from diverse fields, become all the more difficult to discover. (For the cure for the disease of specialism, see N. Maxwell, Science, Reason, Knowledge and Wisdom: A Critique of Specialism, *Inquiry 23*, 1980, pp. 19-81.)

13. N. Goodman, *Problems and Projects*, 1972, Bobbs-Merrill; Michael Friedman, Explanation and Scientific Understanding, *Journal of Philosophy 71*, 1974, pp.5-19; Philip Kitcher, Explanatory Unification, *Philosophy of Science 48*, 1981, pp. 507-531; John Watkins, *Science and Scepticism*, 1984, Princeton University Press, pp. 206-213.

14. R. Feynman, R. Leighton and M. Sands, *The Feynman Lectures on Physics*, vol. II, 1964, Addison-Wesley, pp. 25-10 - 25-11.

15. The notion of "discoverable" is problematic. As I am using the term, no aberrant thesis about the universe is discoverable if the aberrant phenomena, postulated by the thesis, lie beyond our experience.

16. For accounts of Lagrangian formulations of classical and quantum mechanical theories see: R. Feynman, R. Leighton and M. Sands, *The Feynman Lectures*, vol. II, ch. 19; F. Mandl and G. Shaw, *Quantum Field Theory*, 1984, John Wiley, ch. 2; H. Goldstein, *Classical Mechanics*, 1980, John Wiley.

17. P_9 is a kind of "principle of the uniformity of nature". P_9 is, however, intended to be very much weaker than uniformity principles as these are usually formulated and understood. It does not assert that all phenomena are governed by the same laws everywhere, since the possibility of (some) aberrant phenomena is conceded. Instead, P_9 asserts that if aberrant phenomena occur anywhere they occur in our immediate environment. P_9 does not even assert that approximately lawful phenomena occur everywhere, but merely that whatever it is that makes our immediate environment partially knowable extends throughout the universe. We might live in a partially knowable world even though no laws strictly obtain, as the notion of law is understood in natural

science.

18. E. Wigner, The Unreasonable Effectiveness of Mathematics in the Natural Sciences, chapter 17 of his *Symmetries and Reflections,* 1967, MIT Press.

19. N. Maxwell, The Rationality of Scientific Discovery, Parts I & II, *Philosophy of Science 41,* 1974, pp. 123-53 & 247-95; *From Knowledge to Wisdom: A Revolution in the Aims and Methods of Science,* 1984, Basil Blackwell, Ch. 9; Induction and Scientific Realism: Einstein versus van Fraassen. Parts I, II & III, *The British Journal for the Philosophy of Science 44,* 1993, pp. 61-79, 81-101 & 275-305; *The Comprehensibility of the Universe* (forthcoming).

20. Tradition has forced me to use the word "simplicity" in two different senses. On the one hand, in discussing "the problems of simplicity" (as it is traditionally called), I use the word as the generic term, to stand for all the terms that may be used in the present context: "unity", "explanatoriness", etc. On the other hand "simplicity" may be used to refer to just one (hypothetical) aspect of theories in addition to the other aspects, such as unity, explanatoriness, etc. In this second sense, "simplicity" really does mean simplicity. I hope this ambiguity of usage is not too confusing.

21. Counting entities is rendered a little less ambiguous if a system of M particles is counted as (a somewhat peculiar) field. This means that M particles all of the same kind (i.e. with the same dynamic properties) is counted as *one* entity. In the text I continue to adopt the convention that M particles all the same dynamically represents one *kind* of entity, rather than one entity.

22. For accounts of the gauge group structure of quantum field theories see: K. Moriyasu, *An Elementary Primer For Gauge Theory*, 1983, World Scientific; I.J.R. Aitchison and A.J.G. Hey, *Gauge Theories in Particle Physics,* 1982, Adam Hilger, Part III; D. Griffiths, *Introduction to Elementary Particles*, 1987, John Wiley, ch. 11. For introductory accounts of group theory as it arises in theoretical physics see: C.J. Isham, *Lectures on Groups and Vector Spaces for Physicists*, 1989, World Scientific; or H.F. Jones, *Groups, Representations and Physics*, 1990, Adam Hilger.

23. "Theory", throughout this discussion, means the conjunction of fundamental dynamical theories required to cover, in principle, all known phenomena, or the conjunction of laws of some domain of phenomena if no theory of the domain exists.

24. For a more detailed discussion see my forthcoming *The Comprehensibility of the Universe*.

25. Strictly speaking, two non-empirical requirements are involved in the explanatory character of a theory. The explanatory character of a theory becomes all the greater the more (i) unified it is, and (ii) the greater its empirical content. This second requirement is related to the goal of theoretical physics of discovering unity throughout *all* the diverse physical phenomena that there are; the goal of theoretical physics, we might say, is to discover dynamic unity running through all possible diversity.

26. If this were not the case, there would be no generally accepted theories in physics. Given any empirically successful non-aberrant theory, there are infinitely many aberrant rivals that are just as empirically successful. Without general agreement that these rivals fail to satisfy non-empirical simplicity considerations, there would be no basis for excluding them from consideration. (Aberrant theories may be accepted because it is not noticed how aberrant they are: I have, for many years, argued that orthodox quantum theory is unacceptably aberrant, in a widely unnoticed way, in that it is an aberrant addition of quantum postulates and classical postulates for a treatment of measurement: see N. Maxwell, *Am. J. Phys.* ***40*** 1431-1435 (1972)

1431; *Am. J. Phys.* ***41*** 1022-1025 (1973); *Found. Phys.* ***6*** 275-292 & 661-676 (1976); *Found. Phys.* ***12*** 607-631 (1982); *Brit. J. Phil. Sci.* ***39*** 1-50 (1988); in *Bell's Theorem and the Foundations of Modern Physics* (eds. van der Merwe, A., Selleri, F. & Tarozzi, G.) 362-370 (World Scientific, Singapore, 1993); *Phys. Lett.* ***A 187***, 351-355 (1994); in *Fundamental Problems in Quantum Physics* (eds. Ferrero, M. & van der Merwe, A.) 205-214 (Kluwer Academic, 1995).)

27. See, for example, Kuhn's discussion of such disagreements in his *The Structure of Scientific Revolutions*, 1962, Chicago University Press, especially chs. VII-XII.

28. See especially ch.IX of Kuhn's *The Structure of Scientific Revolutions*.

29. For earlier expositions of this approach to the problem of induction see: N. Maxwell, A Critique of Popper's Views on Scientific Method, *Philosophy of Science 39*, 1972, pp. 131-52; The Rationality of Scientific Discovery, *Philosophy of Science 41*, 1974, pp. 123-53 & 247-95; *What's Wrong With Science?*, 1976, Bran's Head Books, Frome; Articulating the aims of science, *Nature 265*, 1977, p. 2; Induction, Simplicity and Scientific Progress, *Scientia 114*, 1979, pp. 629-53; *From Knowledge to Wisdom*, 1984, Basil Blackwell, ch. 9; Induction and Scientific Realism, *British Journal for the Philosophy of Science 44*, 1993, pp. 61-79, 81-101 & 275-305. My proposed solution to the problem of induction has some features in common with views advocated by Mill, Keynes, Russell, Popper, Reichenbach, and Lakatos: see J.S. Mill, *A System of Logic*, book 3, ch. 3, in J.M. Robinson (ed), *Collected Works of John Stuart Mill*, 1973-4, Toronto University Press, Toronto; J.M. Keynes, *A Treatise on Probability*, 1921, Macmillan, London; B. Russell, *Human Knowledge: Its Scope and Limits*, 1948, Allen and Unwin, London, Part VI; K.Popper, *The Logic of Scientific Discovery*, 1959, Hutchinson, London; H. Reichenbach, *Experience and Prediction*, 1938, University Press, Chicago; I. Lakatos, Falsification and the Methodology of Scientific Research Programmes, in I. Lakatos and A. Musgrave (eds), *Criticism and the Growth of knowledge*, 1970, Cambridge University Press, Cambridge, pp. 91-195. The basic difference between any of these views and the one I argue for here is that, as AOE requires, I stress the need to consider an hierarchy of increasingly attenuated cosmological conjectures implicit in the methods of science, low-level conjectures requiring revision as knowledge advances which in turn leads to revision of associated methods. This difference is decisive from the standpoint of solving the problem of induction.

30. B. van Fraassen, Empiricism in the Philosophy of Science, in P.M. Churchland and C.A. Hooker, (eds.) *Images of Science*, 1985, University of Chicago Press, pp. 259-260.

31. Russell argues that *five* postulates are "required to validate scientific method": see B. Russell, *Human Knowledge*, 1948, Allen and Unwin, p. 506. These postulates are not, however, ordered with respect to content and implication in the way specified by AOE: they are all on the same level and may, therefore, be treated as five components of one composite postulate.