

Prolegomena [= Preliminaries] to any Future Metaphysic that can Present itself as a Science

Immanuel Kant

Copyright © Jonathan Bennett 2017. All rights reserved

[Brackets] enclose editorial explanations. Small ·dots· enclose material that has been added, but can be read as though it were part of the original text. Occasional •bullets, and also indenting of passages that are not quotations, are meant as aids to grasping the structure of a sentence or a thought. Every four-point ellipsis indicates the omission of a brief passage that seems to present more difficulty than it is worth. Longer omissions are reported between brackets in normal-sized type.

First launched: July 2004

Last amended: April 2007

Contents

Introduction	1
Preamble on the special features of all metaphysical knowledge	7
General Problems	12
Main transcendental Problem 1: How is pure mathematics possible?	16
Main transcendental problem 2: How is pure natural science possible?	25
Main transcendental problem 3: How is metaphysics possible in general?	45

Conclusion: Determining the boundaries of pure reason	61
Solution of the general question of the <i>Prolegomena</i>: How is metaphysics possible as a science?	73
Appendix: On what can be done to make metaphysics actual as a science	77

Main transcendental problem 3: How is metaphysics possible in general?

40

I have now provided an explanatory justification for •pure mathematics and for •pure natural science. Neither of them needed this for the sake of its own security and certainty; for pure mathematics is supported by the self-evident truth of its propositions; and pure natural science, although its ultimate sources lie in the understanding, is thoroughly supported and confirmed by experience. (Certain as pure natural science is, it can never equal mathematics in that regard, which is why it can't entirely do without the testimony of experience.) Both these sciences therefore needed my enquiry not for themselves but for another science, namely •metaphysics.

Metaphysics is concerned not only with concepts of nature (which always find their application in experience) but also with *pure concepts of reason*, which never find application in any possible experience. No experience can tell us what's true and what's false involving concepts of reason, or even whether these concepts are objectively real or mere fictions. Yet the part of metaphysics that involves them is what the rest of metaphysics is *for*—and that's why this science ·unlike the other two· needs an explanatory justification *for its own sake*. The third question now before us concerns the heart of metaphysics, its special feature, namely reason's preoccupation with itself, and its assumption that by brooding over its own •concepts it can come to know about •objects that it supposes to arise immediately out

of those •concepts without help of any kind from experience.⁶

Reason will never be satisfied until it has solved this problem—i.e. answered the question 'How is metaphysics possible?'. Reason won't let pure understanding be used outside the domain of experience; but reason itself is destined to go beyond those confines. Every particular experience is only a part of the whole domain of experience; but *the absolute whole of all possible experience* is not itself an experience, yet it is something that reason has to think about, as a problem. For reason to present this problem to itself, it needs concepts quite different from those of the understanding. The latter are applied only to items given in experience; but the concepts of reason have a use that is *transcendent*: it transcends all actual and possible experience, because it involves thinking about the completeness of all possible experience, i.e. thinking about

the-totality-of-possible-experience

considered as a single unified item. Such an item couldn't itself be given in experience.

Just as •the understanding supplies *categories*, which are needed for experience, so •reason supplies *Ideas*, by which I mean concepts that one must have though their objects can't be given in any possible experience. •Ideas are as inherent in the nature of •reason as •categories are in the nature of the •understanding. Ideas carry with them an illusion that could easily mislead; this illusion is unavoidable, although it can be prevented from actually leading us astray.

⁶ If we can say that a science is actual, at least in the thinking of all men (•subjectively actual•), as soon as we have established that the problems leading to it are ones that are set before everybody by the nature of human reason, . . . then we are bound to say that metaphysics is *subjectively actual* (and necessarily so), which leads us to the legitimate question: How is it (*objectively*) possible?

All illusion consists in taking the subjective ground of judgment to be objective, as though reason in its use of the Ideas were acquiring a special kind of knowledge. Reason falls victim to this, and is guilty of error, when it takes something that merely concerns •reason's own nature and mode of operation and tries to make it refer to •some object in itself. The only safeguard against this temptation is for reason to know itself—to understand what's going on when it uses Ideas in a transcendent, extravagant manner that goes beyond all possible experience.

41

We must distinguish Ideas, which are pure concepts of reason, from the categories or pure concepts of the understanding; the two correspond to two sorts of knowledge that are quite different from one another •in their natures, •in where they come from, and •in how they are used. In the fundamentals of a science that purports to cover all *a priori* knowledge, the distinction between Ideas and categories is crucial. If we don't respect it, metaphysics will be absolutely impossible—or at best a random, bungling attempt to build a house of cards in ignorance of the materials one is using and of what they are good for. If my *Critique of Pure Reason* had done nothing but make this distinction plain for the first time, it would have contributed more to our grasp of metaphysics—what it is, and how it should be conducted—than all the fruitless efforts to do justice to the transcendent problems of pure reason that had ever before been undertaken. Before the *Critique*, no-one had even suspected that reason was quite different from understanding; so everybody ran the two together, mentioning concepts of the understanding and concepts of reason in the same breath, as though they were of the same kind.

42

All instances of pure knowledge by the understanding have this in common: •although they aren't derived from experience, they involve concepts that can have application in experience, and principles that can be confirmed by experience. Transcendent knowledge by reason contrasts with this. The Ideas that it involves can't be applied in experience, nor can its propositions ever be confirmed or refuted by experience. If any errors creep into the employment of reason, they will have to be discovered by pure reason itself •because neither sensibility nor understanding can have anything to do with them. For reason thus to stand guard over itself is very difficult, because the reason that is standing guard is the very faculty that is necessarily prone to intellectual illusions, and we have no firm objectively grounded procedure for avoiding them—only a subjective enquiry into reason itself as a source of Ideas.

43

My chief aim in the *Critique* was not only to distinguish carefully the various sorts of knowledge but also to derive from their common source the concepts belonging to each of them. I did this so that by knowing the origins of these concepts I could settle how they might safely be used; and it also gave me the priceless though unexpected advantage of knowing, *a priori* and in a principled way, that •my list of concepts, and •my classification and •descriptions of them, are complete. Without this, everything in metaphysics is a mere jumble, in which you never know whether you have enough •for your purpose, or whether and where something is still lacking. This advantage is the very essence of pure philosophy, and isn't to be had anywhere else.

I have derived the twelve categories—the four trios of pure concepts of the *understanding* - from a classification of

kinds of judgment that can be made. The concepts of reason are three in number, and they derive from a classification not of *judgments* but of *logical arguments*—specifically, the three kinds of *inferences of reason*. For these pure concepts of reason (the transcendental Ideas) are given—we simply *do* have them—and if one doesn't want to regard them as something like *innate*, the only source that can be found for them is the activity of reason. That activity in its concern with logical form constitutes the logical element of the inferences of reason; but it also involves recognizing judgments of the understanding as involving this or that *a priori* form of judgment, and in this role it yields transcendental concepts of pure reason.

The basic sorts of argument are: categorical, conditional, and disjunctive. [A categorical argument has a first premise of the form '(Subject) is (Predicate)'; a conditional one has a premise of the form 'If P, then Q'; a disjunctive one has one of the form 'Either P or Q'.] Each Idea involves the thought of a kind of *completeness*. So the Ideas—the concepts of pure reason—are as follows.

- Categorical: the Idea of a complete subject (the Idea of *what is substantial*); this is the Idea of an ultimate 'thing which. . .', like Locke's idea of substance in general; this Idea is psychological because the natural home ground of this thought is in application to oneself: I am a thing which. . . .
- Conditional: the Idea of a complete series of conditions—e.g. the thought of all the causes of the present state of the world; this Idea is cosmological.
- Disjunctive: the Idea of a complete reality that somehow encompasses the entire range of what's possible;

this Idea is theological.⁷ [For an explanation of how theology comes into this, see section 55 on page 59. It's not very helpful, but it's all Kant gives us in this work.]

All three give rise to *dialectics*—i.e. to characteristic dangers of intellectual illusion, insoluble problems, lurking contradictions, and the like. But their ways of doing so are different, and so we have—corresponding to the trio

- categorical, •conditional, and •disjunctive
- a three-part division of the dialectics into its •Paralogism, its •Antinomy, and its •Ideal.

Through this way of coming at things we can feel assured that all the claims of pure reason are completely represented, nothing missed, because we have completely surveyed the faculty of reason itself, from which they all take their origin.

44

It should be borne in mind that the Ideas of reason, unlike the categories, don't help us to bring the understanding to bear on experience. In the knowledge of nature by the understanding, the Ideas of reason are entirely dispensable; indeed they are a positive obstacle to what is going on. (They have, however, their own good use, which we'll come to later.) •The psychological Idea of reason brings up the question 'Is the soul a simple substance or not?' The answer to that is of no interest when we are doing empirical psychology. No possible experience could be evidence for either answer to the question. So far as the description and explanation of our mental histories is concerned, the concept of *simple substance* is quite empty. •As for the questions raised by the cosmological Idea—Did the world begin? Will it end?—answers to these can have no role whatever in explaining

⁷ In disjunctive judgments we consider *the whole range of what is possible* as divided in respect to some particular concept. The ontological principle that *every object falls under one or the other out of each contradictory pair of predicates*, which is also the principle of all disjunctive judgments, essentially relies on this thought of the sum of all possibility—which goes with the thought that every possible object is completely determinate, because it falls under just one out of each contradictory pair of predicates. . . .

any event *in* the world. •And as for the theological Idea: there's a correct maxim of natural science that says that we shouldn't try to explain how nature is by appealing to the will of a highest being, because such an explanation would no longer be natural science, but rather an admission that we have reached the end of it. So the proper use of •Ideas of reason must be quite different from the use of the •categories, i.e. the •pure concepts of the understanding through which experience becomes possible.

Now, reason and understanding are related with one another in a certain manner, which brings in some parts of my laborious analytic of the understanding [i.e. the part of the *Critique of Pure Reason* labelled 'Transcendental Analytic']. How? Well, it can't have to do with the business of getting knowledge of nature through experience: the part reason plays in that—in mathematics and natural science—can be perfectly well played without all this subtle examination of the nature and functions of the understanding. So my analytic of the understanding must link with the Ideas of pure reason for a purpose that lies beyond the empirical use of the understanding. •So now we have a dilemma. On the one hand•, I have said that we can't use the understanding outside the realm of experience, as that would be a meaningless activity, with no subject-matter. On the other hand, the nature of reason must conform with the activities of the understanding, contributing to their perfection and not disturbing them.

Here's the solution—the *truth* about what reason has to do with understanding. What pure reason does is to demand that understanding, when it is brought to bear on the complex of experience, shall achieve *completeness* in its operations. This, however, is only a completeness of principles, not of intuitions and objects. •To put the point in simpler terms: The demand for completeness says 'As long as there's something you don't yet understand, keep

working on it'; it doesn't say 'Aim to grasp the whole story all at once: •survey the mind in such a way that you have all its properties on one side and the ultimate subject that bears those properties on the other; •arrive at results about the world's entire past and entire future; •think in a concrete way about God as the explanation of the entire world'. The illusion—which brings the risk of error—comes from the fact that reason, wanting to make its demand for completeness as sharp and graspable as possible, slips into treating it as though it were a demand for knowledge of something—the ultimate subject of mental states, the world's whole past, etc•.

45: Prefatory remark to the dialectic of pure reason

I have shown in sections 33 and 34 that the freedom of the categories from any input from the senses may mislead reason into extending their use, quite beyond all experience, to things •as they are• in themselves •as distinct from things as they appear to us•; though •no such use is legitimate, for the following reason•. Because the categories lack any sensory element that can give them meaning or sense in particular cases, they can represent *anything* in their role as mere logical functions; but there's *nothing* about which they can, unaided, give specific information. The fancy objects •that reason wrongly tries to bring under the categories• are known as 'noumena', or pure beings of the understanding (or better, beings of thought). Examples include

substance—but conceived without permanence in time,

cause—but not conceived as acting in time,

and so on. •In thinking or talking like this• one attaches to these •supposed• objects predicates whose only •legitimate• use is to enable experience to conform to laws; and yet •by leaving *time* out of it• one deprives them of all the conditions

of intuition that have to be satisfied for experience to be possible, and so these concepts lose all meaning again.

There's no risk that the understanding, when left to itself and not given orders from the outside, will so wantonly roam out of its own proper territory into that of mere creatures of thought. But the empirical use of the rules of the understanding is **conditioned**, and reason can't be fully satisfied with that; so it demands a *completion* of this chain of conditions. [Until this point in the *Preliminaries* the concept of *condition* has been used almost entirely in saying things about **(a)** the conditions that our understanding must satisfy, the conditions that our sensibility must satisfy—e.g. in saying that spatiality and temporality are formal conditions of our sensibility. We now, for almost the first time, encounter a different use of the concept: the topic now is not

(a) the general conditions that have to be satisfied if we are to have certain kinds of engagement with the world,

but rather

(b) the conditions that various items *within* the world have to satisfy if they are to exist.

Examples of **(b)**: Any event in the experienced world occurs only because something causes it to do so; any region of space exists only because there is a larger region in which it is embedded; any period of time exists only because there was an earlier period of time leading up to it; any state exists only if there is something that it is the state *of*; and so on. Kant holds that **(b)** these in-the-experienced-world conditioning relationships can be depended on *because* they are put there by **(a)** the general conditions that our sensibility and understanding have to satisfy. So all this is part of a single unified body of doctrine. Still, it is as well to notice how, after so much talk about 'conditions', we are rather suddenly on the presence of the notion of *chains* of conditions. It is just the shift from **(a)** to **(b)**.]

This forces the understanding to leave its proper domain, so that it can do two things: •represent objects of experience in a series that stretches too far for any experience to capture it, and •look completely outside itself for noumena to which it can attach that chain, thus completing the series,

escaping from the conditions of experience, and making its hold complete. So there they are—the transcendental Ideas. •They are in themselves virtuous, though •they can make trouble; but •the trouble can be averted. •They don't try to produce concepts that are in themselves excessive or extravagant; all they aim for—in conformity with the true but hidden goal to which our reason is naturally drawn—is a limitless extension of the empirical use of the categories. •But through an unavoidable •intellectual• illusion they may seduce the understanding into using the categories in a *transcendent* manner, •i.e. in a manner that isn't related to experience. Deceitful as this misuse is, •it is hard to avoid. To keep yourself from it and confine the categories within the bounds of experience, it won't do merely to resolve in advance to be on your guard against doing so. •What you need is scientific instruction •on how to avoid the trouble, and even then it takes hard work.

I. The psychological Ideas

46

People have long since remarked that in all substances the proper subject—

namely, the *substantial* as such, i.e. what remains after all the qualities (as predicates) are set aside—is unknown, and this limit on our knowledge has been the topic of various complaints. But if our understanding is at fault in this matter, it is not for its *inability to know*—to determine by itself—the substance of things, but rather for its *wanting* to know the substance of things, thereby treating a mere Idea as though it were a given object •into whose nature one might enquire. Pure reason demands that for every predicate of a thing we seek its proper subject; but •this subject can't be anything but a •further• predicate, so

reason tells us to find a subject for •it in its turn, and so on, indefinitely (or until we give up). So we are never to regard anything that we arrive at as an ultimate subject: and our understanding can never have the thought of *the substantial* itself, however deeply it penetrates and even if all of nature is unveiled to us. That's because the special characteristic of our understanding is that when it thinks something it does so by representing it through concepts, and thus through mere predicates; so it can never reach the absolute subject—the sheer *thing*, not understood as thing-that-is-F for any predicate F. Hence all the real properties through which we know bodies are mere qualities of them; and that includes impenetrability, which we can only represent to ourselves as the effect of a power whose subject is unknown to us.

Now, it appears as if we *do* confront this ·absolute· subject in our consciousness of *ourselves* (of the thinking subject), and indeed that we have this in an immediate intuition; for all the predicates of inner sense refer to the *I* as a subject, and I can't conceive myself as the predicate of some other subject. So it *seems* that we are given in experience something that completes the process of relating given concepts predicatively to a subject—given it not merely as an idea but as an object, i.e. the absolute subject itself. But this turns out to be a false hope. For the *I* isn't a concept,⁸ but only a designation of the object of inner sense insofar as we know it by no further predicate. So it can't itself be a predicate of any other thing, any more than it can be a determinate concept of an absolute subject; it's only a *relating* of inner phenomena to their unknown subject. Yet this Idea (which does excellent service as a regulative

principle, totally destroying all materialistic explanations of the inner phenomena of the soul) leads through a wholly natural misunderstanding to a highly plausible argument: from •this supposed knowledge of the substantial status of our thinking being, the argument infers •conclusions about *the nature of the soul*—the nature of it that lies right outside the compass of experience. [See the explanation of 'regulative' at the end of section 56 on page 60.]

47

We may call this thinking self (the soul) *substance*, as being the ultimate subject of thinking that can't be further represented as the predicate of something else; but the concept ·of substance, in this use of it·, remains quite empty, with nothing following from it, if it can't be shown to involve *permanence*—which is what makes fruitful the concept of substances that we encounter in experience.

But permanence can never be proved on the basis of the concept of a substance considered as a thing in itself, but only in relation to experience. This is adequately shown by the first Analogy of Experience ·in the *Critique of Pure Reason*·. If that proof doesn't convince you, try for yourself whether you can derive from the concept of *a subject that doesn't exist as the predicate of another thing that its existence is thoroughly permanent* and that it cannot—unaided or through any natural cause—either come into existence or be annihilated. Synthetic *a priori* propositions such as *that* can never be proved ·of things· in themselves, but only in application to things as objects of possible experience.

⁸ If the representation of self-awareness, the *I*, were a concept through which something could be thought, it could be used as a predicate of other things or would contain such predicates in itself. But it is nothing more than the feeling of something existing, without the slightest concept of it; it is only the representation of *that to which all thinking relates*.

48

So if we want to use the concept of •the soul as substance as a basis on which to conclude that •the soul is permanent, we can do this only in relation to possible experience; if we take the soul to be a thing in itself, and look for a conclusion that holds good beyond the bounds of all possible experience, permanence can't be shown. But all our possible experience requires us to be alive; so the only permanence-of-the-soul result we can establish is that the soul is *permanent throughout one's life*; for the death of man is the end of all experience that the soul could have of itself as an object—unless the contrary is proved, but that 'contrary' is supposed to be the conclusion of the argument for the soul's permanence, so it can't appear among the premises. The most we can show, therefore, is that the soul is permanent throughout one's life—a result that nobody will disagree with! What we want is to show that the soul lasts after death, and this we cannot do, for the reason I have given: the necessary tie between the concept of substance and the concept of permanence is created by the principles of possible experience, and so it holds good only within the domain of possible experience.⁹

49

Here's something else that can be proved •as a requirement for the intellectual management• of experience, but can't be shown to hold of things in themselves: *Our outer experience not only does but must correspond to something real outside of ourselves*. That tells us this much: there is something empirical—thus, some phenomenon in space outside us—•the existence of• which can be satisfactorily proved. •That's all it tells us•, for we have no dealings with objects other than those belonging to possible experience; because objects that can't be presented to us in any experience are nothing to us. What is empirically outside me is what appears in *space*. Now space, together with all the phenomena it contains, belongs to the representations whose objective truth is proved by how they are inter-connected according to laws of experience, just as the actuality of my soul (as an object of inner sense) is proved by how the phenomena of inner sense are inter-connected. Accordingly, •it is by outer experience that I am conscious of the actuality of bodies, as external phenomena in space, just as •by inner experience I am conscious of the existence of my soul in time; but I know this soul only as an object of inner sense—knowing it only through appearances that constitute my inner state;

⁹ It is indeed very remarkable that metaphysicians have always glided comfortably over the principle that *substances are permanent* without trying to prove it. No doubt this is because as soon as they started on the concept of substance they found that every possible basis for a proof had deserted them. Common sense, which strongly felt that perceptions couldn't be unified in experience without this presupposition •of the permanence of substance• filled the gap by a postulate. (•It had to *postulate* permanence instead of *proving* it• because it could never derive the necessary permanence of substance from experience itself, for two complementary reasons. (1) We have no way of tracking substances through all their alterations and dissolutions and finding •empirically• that their matter, their *stuff*, is never lessened. (2) The principle in question involves *necessity*, which is a sure sign of its being an *a priori* principle •and thus not knowable through or provable from experience•.) People then optimistically applied this postulate •about all substances• to the concept of soul as a *substance*, and inferred from this that a man's soul must continue in existence after his death (especially because this substance's having no *parts*—which they inferred from the indivisibility of consciousness—guaranteed that it couldn't be destroyed by falling to pieces). If they had found the genuine source of this principle of the permanence of substance—a discovery requiring deeper researches than they were ever inclined to make—they would have seen that the law of the permanence of substances holds good only for the purposes of •intellectually managing• experience; so it applies to things only so far as they are to be known and conjoined with others in experience. It never applies independently of all possible experience, and consequently it cannot hold good of the soul after death.

the nature of the soul in itself—the thing that *has* these phenomena—is unknown to me. So all that *Cartesian idealism* achieves is to distinguish •outer experience from •dreaming; and to distinguish the •conformity to law that is a criterion of the truth of the former from the •irregularity and false illusion of the latter. In •dealing with• both •outer experience and dreaming• Cartesian idealism presupposes space and time as required for the existence of objects; its only question is this:

•Are the objects of the outer senses, which when awake we put in space, actually to be found in space (as the object of inner sense, the soul, is actually to be found in time)?

That amounts to this:

•Does experience carry with it sure criteria to distinguish it from imagination?

Doubts about this are easy to dispose of. We dispose of them in ordinary life every time we investigate how appearances in both space and time are connected according to universal laws of experience: when the representation of outer things agrees thoroughly with those laws, we can't doubt that they constitute truthful experience. So it is very easy to refute *material idealism*, which •questions the existence of bodies, although it doesn't think of them as things in themselves, but• considers appearances *as appearances* and takes account only of how they are connected in experience: it is just as sure an experience that bodies exist outside us (in space) as that I myself exist according to the representation of inner sense (in time); •I put it like that because• the concept of *outside us* means *existing in space*. Compare these two:

•taking 'body' to refer not merely to outer intuition (in space) but to the thing-in-itself that is the basis of this appearance—the thing that the appearance is an appearance *of*•,

and

•taking the 'I' in the proposition 'I am' to refer not merely the object of inner intuition (in time) but to the subject of consciousness—the thing that *has* the consciousness•.

The 'body' thought generates •the question of whether bodies (which are really phenomena of outer sense) exist as bodies in nature *apart from my thoughts*—a question that can be briskly answered in the negative. The 'I' thought generates •the question of whether I myself (an appearance of inner sense, the soul that empirical psychology studies) exist apart from my faculty of representation in time; and this question is on exactly the same footing as the other, and must likewise be answered in the negative.

Everything is decided and certain in this way, once it has been given its true meaning, Formal idealism (which I have also called 'transcendental idealism') actually abolishes material (or Cartesian) idealism. For if space is nothing but a form of my sensibility, then it is—*as a representation in me*—just as actual as I myself am; and the only remaining question concerns the empirical truth of the representations in space. And if on the other hand space and the phenomena in it are something existing outside us, then the actuality of these •alleged• objects outside us can never be proved •in the way it would *have to* be proved, namely• by applying the criteria of experience beyond the domain of our perception.

II. The cosmological Idea

50

The cosmological Idea is the most remarkable product of pure reason in its transcendent use. It has more power than anything else to rouse philosophy from its dogmatic slumber

and to stimulate it to a hard task, namely making a critique of reason itself.

I term this Idea ‘cosmological’ because it never takes its object from anywhere but the world of the senses, having no use for anything that isn’t given to the senses. So in that way it *stays at home*, does not become transcendent, and is to that extent not a mere Idea. (Whereas the psychological Ideas don’t in that sense ‘stay at home’, because merely conceiving the soul as a *simple substance* involves conceiving something—the simple—that can’t be presented to the senses.) Despite that, the cosmological Idea does in its own way go outside the domain of the senses, because it extends the connection of the conditioned with its condition so far that experience never can keep up with it. In this way, then, it is always an Idea, whose object can never be adequately given in any experience.

51

Note first that in this territory of the cosmological Ideas the usefulness of a *system of categories* is so obvious and unmistakable that this alone would show that it is indispensable in the system of pure reason, even if there weren’t several other proofs of that. There are exactly four such transcendent Ideas, and exactly four classes of categories; but the Ideas differ from the categories in that they are concerned only with the *absolute completeness* of the series of the conditions for a given conditioned item. Matching these cosmological Ideas there are exactly four kinds of *dialectical assertions* of pure reason. Their being dialectical shows in this:

Against each of these assertions we can bring its contradictory, on the strength of principles of pure reason that are as plausible as those supporting the original assertion.

No exercise of metaphysical art can fend off this conflict

between the assertion and its contradictory—except the one that compels the philosopher to look into the first sources of pure reason itself. This Antinomy—i.e. this conflict between dialectical assertions and their contradictories—isn’t something I thought up to amuse myself: it comes from the nature of human reason; so it can’t be avoided or brought to an end. The Antinomy contains the following four *theses* together with their *antitheses*:

(1) *Thesis*: The world has a *beginning* in time and space (a limit).

Antithesis: The world is spatially and temporally *infinite*.

(2) *Thesis*: Everything in the world consists of *elements* that are *simple*.

Antithesis: There is nothing simple; everything is *composite*.

(3) *Thesis*: There are in the world causes through *freedom*.

Antithesis: There is no liberty; all is *nature*.

(4) *Thesis*: In the series of the world’s causes there is some *necessary being*.

Antithesis: There is nothing necessary in the world; in that series *everything is contingent*.

52a

We have here the strangest phenomenon of human reason; no other use of reason produces anything like it. If we think of the appearances of the world of the senses as things in themselves (as we often do), and if we take them to be combined through principles that hold universally for things in themselves rather than merely through principles of experience (which we also often do—indeed without my critique we can’t help it!), there arises a conflict that can’t be removed dogmatically—by proving one side and refuting

the other. That's because thesis and antithesis can *both* be shown by equally clear, evident, and compelling proofs (I guarantee that all the proofs *are* correct), so reason is divided against itself—which gladdens the heart of the sceptic but must make the critical philosopher feel ill at ease.

52b

We can blunder in various ways in metaphysics without any fear of being detected in falsehood. For as long as we avoid self-contradiction, which we can always do when we assert synthetic propositions (even if they are wholly fictitious), our only way of being detected in falsehood is through experience. And experience can't refute us when we assert propositions of the sort involved in metaphysics, namely ones in which the concepts that are involved are mere Ideas, instances of which can't be presented to us in experience. For how can we tell from experience whether the world has lasted from eternity or had a beginning, whether matter is infinitely divisible or consists of simple parts? Such concepts can't be instantiated in any experience, however extensive, and consequently neither the positive nor the negative proposition in the antinomy can be empirically discovered to be false. The only way in which reason could unintentionally reveal its secret dialectic, which it falsely offers as positive doctrine, would be for this to happen:

Reason bases an assertion on a universally admitted principle, and infers the exactly opposite assertion, with the greatest correctness of argument, from another principle that is equally accepted.

That's what actually *does* happen in our present case of

the four natural Ideas of reason, from which arise four assertions and four counter-assertions, each validly derived from universally accepted principles, revealing the *dialectical illusion of pure reason* in the use of these principles—an illusion that would otherwise have stayed hidden for ever.

So this is a decisive experiment, which must necessarily reveal to us any error lying hidden in the presuppositions of reason.¹⁰ Contradictory propositions can't both be false unless they both involve some self-contradictory concept. And then they *can* both be false. For example *A square circle is round* is false (it is false that the circle is round, because it is square), and *A square circle is not round* is likewise false (it is false that the circle isn't round, i.e. that it has corners, because it is a circle.) The logical mark of the impossibility of a concept consists precisely in this, that two contradictory propositions involving it are both false, and as no third proposition can be thought between them, *nothing at all* is thought through that concept.

52c

The first two antinomies, which I call 'mathematical' because they are concerned with the addition or division of the homogeneous, are based on such a self-contradictory concept; and that's how it comes about that in each of them both the thesis and antithesis are false. [Re 'homogeneous': see the start of the next section's second paragraph.]

When I speak of objects in time and in space, I am speaking not about •things in themselves (of which I know nothing), but about •things in appearance, i.e. about experience as a particular way of knowing objects—the only way of knowing

¹⁰ I should therefore like the critical reader of the *Critique of Pure Reason* to attend especially to this antinomy of pure reason, because nature itself seems to have arranged it to pull reason up short in its bold claims, and to force it to look into itself. I take responsibility for every proof I have given for the thesis as well as for the antithesis, and thereby promise to show the certainty of the inevitable antinomy of reason. If this curious phenomenon of the 'proofs' of both P and not-P leads you to go back to examine the presupposition on which it is based, you will feel yourself obliged to join me in inquiring more deeply into the ultimate basis of all knowledge of pure reason.

them that has been granted to mankind. When I think of something as being in time or in space, I must not say ·or think·:

In itself it exists in space and in time, independently of these thoughts of mine;

for if I did I would be contradicting myself. That's because space and time, together with appearances in them, are nothing existing in themselves and outside my representations, but are themselves only modes of representation [= 'ways of being represented-to'], and it is plainly contradictory to say that a mere mode of representation exists outside our representation. So objects of the senses exist only in experience; and to attribute to them a self-subsisting existence •apart from experience or •in advance of it amounts to telling ourselves that *experience* is real apart from experience or in advance of it!

Now if I ask about the extent of the world in space and in time, my complete stock of concepts doesn't enable me to call it infinite or to call it finite. For neither state of affairs can be contained in experience: there can't be experience of an *infinite* space, or of an infinite time elapsed; or of the world's being *bounded* by empty space or by empty time before the world began—all these are mere Ideas. This finite or infinite size of the world, ·not being cashable out in terms of experience·, would therefore have to belong to the world itself apart from all experience. But this contradicts the notion of •a world of the senses, which is merely a totality of appearances that exist and are interconnected only in our representation, that is, in experience, since •this world is not an object in itself but a mere mode of representation. From this it follows that the answer 'Yes' to the question 'Is the world finite in space or in time?' is false, and so is the answer 'No'; because the concept of *a world of the senses existing for itself and in itself* is self-contradictory.

The same holds good for the second antinomy, concerning the division of appearances. For these appearances are mere representations, and their parts exist only in the representation, and consequently exist only *in the dividing*, i.e. in a possible experience that presents them; and the dividing can go only as far as this experience goes. If you assume that an appearance, such as that of a body, contains—in itself, in advance of all experience—all the parts that any possible experience can ever reach, what you are doing amounts to this:

Attributing to a mere appearance, which can exist only in experience, an existence preceding experience; or saying that mere representations are *there* before we encounter them through our faculty of representation.

This is self-contradictory, and consequently so is each answer to the misconceived question ·about divisibility·, whether we answer that bodies in themselves consist of infinitely many parts, or that they have a finite number of simple parts.

53

In the first (the mathematical) class of antinomies the falsehood of the assumed proposition consisted in •taking contradictory items (appearance, thing in itself) to be harmoniously compatible within a single concept. In the second (dynamic) group, on the other hand, the falsehood of the assumed proposition consists in •taking a consistent pair of propositions to be mutually contradictory. Thus, in the first class of antinomies the opposed assertions •were both false, while in the second class the two propositions—which are treated as opposed to one another through mere misunderstanding—•may both be true.

Connecting items mathematically through the concept of ·spatial or temporal· size requires that the connected

items be of the same kind; but dynamic connections by no means require that kind of homogeneity. When it comes to extended magnitudes—i.e. stretches of space or of time—all the parts must be homogeneous with one another and with the whole; but in the connection of cause and effect, although •homogeneity may be found there too, it isn't necessary. Or at any rate the concept of causality doesn't require it, because cause-effect has to do with positing something through something else •quite different from it.

If the objects of the world of the senses were taken for things in themselves, and the laws of nature discussed above were taken to be laws of things in themselves, contradiction would be unavoidable. Similarly, if the subject of freedom were taken to be a mere appearance, like other objects, contradiction would be equally unavoidable, for the same predicate taken in the same sense would be at once affirmed and denied of one and the same object. But if *natural necessity* is tied only to appearances, and *freedom* only to things in themselves, there's no contradiction in assuming or allowing both kinds of causality at once, however hard or impossible it may be to make the latter kind (freedom) comprehensible.

In the realm of appearance every effect is an *event*, something that happens in time; so according to the universal law of nature it must be preceded by a cause, some state of

which leads to the event according to a constant law. But the cause's *entering into this state* that gives it its causal power must likewise *take place or happen*; the cause must have *begun to act*, for without that the effect's following from it cannot be conceived. Without such a beginning, the effect, as well as the effectiveness of the cause, would have to have existed always. This yields the result:

- The state of the cause that makes it effective must also have started among appearances, being an *event* (just as the effect is), and so have been caused in its turn, and so on backwards for ever.

Which in turn yields the further result:

- The condition that governs the coming-into-effectiveness of causes must be natural necessity.

If on the other hand, certain causes of appearances have the property of being *free*, then freedom must be a capacity for starting these appearances—these events—*spontaneously*; there's no such event as the cause's *starting to be effective*, and thus no need for anything outside the cause to prod it into starting to be effective. But in that case the cause must have its effectiveness in a manner that doesn't place it in time; so it can't be an appearance, and must be regarded as a thing in itself, with only its effects being appearances.¹¹ If we can without contradiction think of beings of the understanding—choices, decisions, etc.—as

¹¹ The only acceptable use of the Idea of freedom is in thinking about the relation of the *intellectual* (as cause) to *appearance* (as effect)—the relation between what a *person chooses* and how *his body moves*. So the incessant action through which any portion of matter fills its space—acting so as to keep other matter out of that place—though it takes place from an internal principle [here = 'source' or 'driver'], can't be an exercise of freedom. Nor can we find a concept of freedom that is suitable for purely rational beings such as God. For his action, though independent of external determining causes (because it is only his immanent or caused-from-within action that I am talking about), is determined in his eternal reason, that is, in the divine nature which never changes. It is only if something is to *start* by an action, so that the effect occurs in the time-series or in the world of our senses (e.g. the beginning of the world), that the question arises of whether •the effectiveness of the cause must in its turn have been *started*, or whether instead •the cause can initiate an effect without its own effectiveness *beginning*. In the former case the concept of this causality is a concept of *natural necessity*, in the latter, that of *freedom*. From this you will see that in explaining freedom as the faculty of starting an event spontaneously I have exactly hit the notion which is the problem of metaphysics.

exercising such an influence on appearances, then that enables us to have the second part of the following two-part story:

- Natural necessity is what links all causes to all effects when both cause and effect belong in the world of the senses.
- Freedom is possessed by any cause that isn't itself an appearance though it underlies an appearance.

So *nature* and *freedom* can without contradiction be attributed to the very same thing, but in different relations—on one side as an appearance, on the other as a thing in itself. We have in us a faculty that is not merely

- connected with its subjective determining grounds that are the natural causes of its actions, and is in that way the faculty of a being that belongs to appearances,

but is also

- connected to objective grounds (that are only Ideas), being connected to them in that they can determine [here = 'influence'] this faculty—a connection expressed by the word *ought*.

To spell that out a little in more familiar terms: When a person decides how to act on some occasion, the question 'Why did he make that decision?'—a request to know what *determined* him to make it—can have answers of two entirely different kinds. • One kind explains the decision in terms of his prior state of mind, and the psychological laws that led from that to his decision. • The other kind explains the decision in terms of his *reasons* for it, his beliefs about what he *ought* to do. The former kind of answer invokes natural necessity; the second doesn't, because it doesn't explain the decision as an event arising from prior events; it is the second that takes us out of the realm of appearance, and makes room for freedom. This faculty—the one involved

in the second kind of answer—is called *reason*. To the extent that we consider a man exclusively in the light of his reason, viewed as responding to objective judgments about what he *ought* to do, to that extent we aren't viewing him as a being of sense—an inhabitant of the world of our senses—because this freedom or responsiveness to the 'ought' thought is a property of a thing in itself. We can't at all grasp how it is possible—i.e. how the *ought*, which is not an event, not something that has *happened*, can determine the man's activity, becoming the cause of actions which *are* things that *happen*, and whose effect is an appearance in the world of the senses. Still, although we don't grasp *how* this can be, if reason did relate in this 'influencing' way to a person's decisions, that would bring *freedom* into what happens in the world of the senses to the extent that we can see those happenings as influenced by objective grounds (which are themselves Ideas). For reason's effectiveness as a cause wouldn't depend on subjective conditions—i.e. on facts about the person's prior state of mind—and therefore wouldn't depend on events in time or on the law of nature that controls such events. That's because the grounds of reason—the basic 'ought' thoughts—govern actions in a universal way, according to universal principles, without influence from the circumstances of either time or place.

What I am presenting here is meant merely as an example to make things intelligible. It doesn't necessarily belong to our problem—i.e. the question *How is metaphysics in general possible?*—which must be decided from mere concepts, independently of the properties that we meet in the actual world.

Now I can say without contradiction that • all the actions of rational beings, so far as they are appearances, fall under • the necessity of nature; but • those same actions, considered purely in terms of the rational subject and its ability to act

according to mere reason, are free. For what is required for the necessity of nature? Only that every event in the world of the senses come about in accordance with constant laws, thus being related to causes in the domain of appearance; and in this process the underlying *thing in itself* remains out of sight, as does its causality. But I maintain this:

The law of nature still holds, whether or not the rational being causes effects in the world of the senses through reason and thus through freedom. If it *does*, the action is performed according to maxims whose effects in the realm of appearance are always *consistent with* constant laws; if on the other hand it *doesn't*, the action is not merely consistent with but *subject to* the empirical laws of the sensibility, and in this case as in the other the effects hang together according to constant laws.

This conformity to laws is all we *demand for* natural necessity; indeed, it exhausts all that we *know about* natural necessity. But in the former case where the action is caused by reason, reason is the cause of these laws of nature rather than being subject to them, and therefore it is itself free; in the latter case, where the effects follow according to mere natural laws of sensibility with reason having no input, it doesn't follow that reason is in this case *determined by* the sensibility, which indeed it couldn't be, so reason is free in this case too. Freedom, therefore, doesn't get in the way of natural law in the domain of appearance, any more than natural law brings about a breakdown in the freedom of the practical use of reason, which relates to things in themselves as determining grounds.

Thus practical freedom—the freedom in which reason has causal force according to objectively determining grounds—is rescued, without doing the slightest harm to natural necessity in relation to the very same effects, as appearances.

These remarks explain what I said earlier about transcendental freedom and its compatibility with natural necessity (with a single subject taken in two different ways). For

whenever a being acts from objective causes regarded as determining grounds of reason, the start of its action is a *first beginning*,

although

the same action is in the series of appearances only a *subordinate beginning*, which must be preceded and determined by a state of the cause, which in turn is determined by another immediately preceding it and so on backwards.

In this way we can have the thought, for rational beings and quite generally for any beings whose causality is determined in them as things in themselves, of a being's ability to begin a series of states from within itself, without falling into conflict with the laws of nature. For the relation the action has to objective grounds of reason isn't a temporal one; in this case what determines the causality does not precede the action in time, because determining grounds such as reason provides don't involve objects of sense such as causes in the domain of appearance, but rather determining causes as things in themselves, which don't exist in time. And so the action can without inconsistency be seen (with regard to the causality of reason) as a first beginning and as free, and (with regard to the series of appearances) as a merely subordinate beginning and as subject to natural necessity.

The *fourth* Antinomy is solved in the same way as is reason's conflict with itself in the *third*. The propositions at issue in the fourth—that there is a necessary being, and that there is not—are perfectly reconcilable provided we distinguish *the cause IN the domain of appearance* from *the cause OF the domain of appearance* (with the latter thought of as a thing in itself). Then one proposition says:

Nowhere in the world of the senses is there a cause (according to similar laws of causality) whose existence is absolutely necessary;

the other says:

This world is nevertheless connected with a necessary being as its cause (but of another kind and according to another law).

The 'incompatibility' of these propositions entirely rests on the mistake of extending what is valid merely of appearances to things in themselves, and in general running the two together in one concept.

54

This then is how I *present* and how I *solve* the entire antinomy in which reason finds itself caught when it applies its principles to the world of the senses. The mere presentation would contribute a lot to the knowledge of human reason, even if the solution hadn't yet fully satisfied you—which it well might not, because you have to combat a natural illusion that has only recently been exposed to you and that you have previously always regarded as genuine. For there's no escaping from this: so long as you take the objects of the world of the senses to be things in themselves, and not the mere appearances which is what they really are, you haven't any chance of avoiding this conflict of reason with itself; so you *must* re-examine my deduction of all our *a priori* knowledge and the tests that I put it through, in order to come to a decision on the question. This is all I require at present; for if in carrying this out you take your thoughts deeply enough into the nature of pure reason, you will familiarize yourself with the concepts through which alone the solution of the conflict of reason is possible. Unless that happens, I can't expect complete assent even from the most attentive reader.

III. The theological Idea

55

I have discussed the •psychological Idea(s) and the •cosmological Idea(s). Now:

The •third transcendental Idea is the *ideal* of pure reason. The use of reason for which it provides material is the most important of all; though if it is pursued in a merely •theory-building or •speculative manner, that makes it transcendent—•theorizing outside the domain of possible experience—which in turn makes it dialectical. With the psychological and cosmological Ideas, reason starts with experience, and goes wrong by taking a grandiose view of its grounds and trying to achieve, where it can, the absolutely complete series of grounds. Not so with the third, theological Idea. Here reason totally breaks with experience; and—starting from mere concepts of what would constitute the absolute completeness of a thing in general, and thus bringing in the Idea of •a most perfect primal being—it works down from there to secure the possibility and therefore the actuality of •all other beings. And so the mere presupposition of a being that is conceived not *in* the series of experience but *for the purposes of* experience—for the sake of comprehending its connection, order, and unity—that is, the *Idea*, is distinguished from the concept of the understanding more easily in this case than in the others. Hence we can easily expose the dialectical illusion that arises from our taking the subjective conditions of our thinking to be objective conditions of objects themselves, and taking an hypothesis necessary for the satisfaction of our reason to be an objectively established truth. I have nothing more to say here about the pretensions of transcendental theology, because my remarks about them in the Critique are easily grasped, clear, and decisive.

56: General remark on the transcendental Ideas

The objects given to us by experience are in many respects incomprehensible, and the law of nature leads us to many questions about them which, when carried beyond a certain point (though still in conformity with those laws), cannot be answered. For example: why do material things attract one another? But if we go right outside nature, or stay within it but in thinking about how it is interconnected go beyond all possible experience and so enter the realm of mere Ideas, then we can't say that our subject-matter is incomprehensible and that the nature of things confronts us with insoluble problems. For in this case we aren't dealing with nature—or, to put it more generally, we aren't dealing with given objects—but with concepts that have their origin solely in our reason, and with mere creations of thought; and all the problems that arise from our concepts of them must be soluble, because of course reason can and must give a full account of its own process.¹² As the psychological, cosmological, and theological Ideas are nothing but pure concepts of reason that can't be applied to anything found in any experience, the questions that reason confronts us with regarding them don't come from objects but from mere maxims that our reason lays down for its own satisfaction. It must be possible for them all, as a group, to be satisfactorily answered, which is done by showing that they are principles for bringing our use of the understanding into thorough harmony, completeness, and synthetic unity, so that they do in that way hold good for experience—but for experience as a whole. But although an absolute whole of experience

is impossible, the Idea of a totality of knowledge according to principles is needed if our knowledge is to have a special kind of unity, the unity of a system. Without that, our knowledge is nothing but piece-work, and can't be used for the highest end (which is always the establishment of a general system of all ends). I am talking here not only about practical or moral ends, but also about the highest end of the speculative use of reason.

The transcendental Ideas thus express reason's special role, namely as setting a standard for systematic unity in the use of the understanding. But if the following happens:

We see this unity in our way of knowing as attached to the object of knowledge; we take something that is merely regulative to be constitutive; and we persuade ourselves that by means of these Ideas we can extend our knowledge far beyond all possible experience (and thus in a transcendent manner),

this is a mere misunderstanding in our estimate of the proper role of our reason and of its principles, and it is a dialectic that confuses the empirical use of reason and also sets reason against itself. What makes it a misunderstanding is the fact that really reason serves merely to bring experience as near as possible to completeness *within itself*, i.e. to stop its progress from being limited by anything that can't belong to experience. [A regulative principle is one that guides, advises, or even commands—such as 'Never accept that you have found an event that didn't have a cause'. A constitutive principle is one that gives information, has factual content, purports to tell truths about reality. Kant holds, for example, that 'Every event has a cause' is acceptable as regulative, but not as constitutive.]

¹² Herr Platner in his *Aphorisms* acutely says: 'If reason is a criterion, there can't be a concept that human reason cannot comprehend. Incomprehensibility comes up only with what is actual. . . .' So it only sounds paradoxical and is not really strange to say that although much in nature is beyond our comprehension (e.g. biological reproduction), if we rise still higher and go right out beyond nature everything will be comprehensible again. For then we leave behind the *objects* which can be given us, and occupy ourselves merely with *Ideas*; and here we can easily grasp the law that reason, through them, prescribes to the understanding for its use in experience, because that law is reason's own product.

Conclusion

Determining the boundaries of pure reason

57

After the clearest arguments, which I have provided, it would be absurd for us to hope to know more of any object than belongs to the possible experience of it, or lay claim to the slightest knowledge of anything not taken to be an object of possible experience—knowledge that would tell us what the thing is like in itself. For how could we learn such facts, given that time, space, and the categories—and even more all the concepts drawn from empirical intuition or *perception* in the world of the senses—don't and can't have any use other than to make experience possible, and that even the pure categories are meaningless if they are removed from this relation to perception?

But on the other hand it would be even more absurd if we rejected things in themselves, or declared that •our experience is the only possible way of knowing things, •our intuition of them in space and in time the only possible intuition, •our concept-using understanding the pattern for every possible understanding—all of which would amount to taking the principles of the possibility of experience to be universal conditions of things in themselves.

My principles, which limit the use of reason to possible experience, could in that way become *transcendent*, and the limits of our reason might pass themselves off as limits of the possibility of things in themselves (Hume's *Dialogues Concerning Natural Religion* illustrate this process), if a careful critique didn't *both* watch over the bounds of our reason. . . .and set a limit to its pretensions. Scepticism originally arose from metaphysics and its lawless dialectic. Want-

ing to favour the experiential use of reason, it started out by declaring that whatever transcends this use is worthless and deceitful; but little by little, as the awareness sank in that the *a priori* principles used *in* experience lead (surreptitiously, and seemingly just as legitimately) *further than* experience extends, there came to be doubts even about experience. There's no danger in this •error•, for healthy common sense will doubtless always assert its rights •regarding experience•. But a certain confusion arose in science, which can't work out how far reason is to be trusted—and why just this far and no further?—and the only way to clear up this confusion and forestall any future relapses is through a formal, principled fixing of the boundary of the use of our reason.

It is true: we can't rise above all possible experience and form a definite concept of what things in themselves may be. Nevertheless, we aren't free to stop enquiring into them; for experience never satisfies reason fully; rather, in answering •our• questions it points further and further back, leaving us still hungering for their complete solution. You can see this in the dialectic of pure reason, the solid subjective ground for which consists in just this hunger for completeness. As regards the nature of our soul: having a clear awareness of oneself as a subject, and having become convinced that psychological phenomena can't be explained *materialistically*, •who can refrain from asking what the soul really is? And if no concept of *experience* suffices for the purpose, •who can refrain from accounting for it by a concept of *reason* (the concept of *simple immaterial being*), even though we are totally unable to show its objective reality? As regards

all the cosmological questions about the duration and size of the world, and about freedom versus natural necessity, •who can be satisfied with mere empirical answers to these, when every answer given on empirical principles gives rise to a fresh question, which then requires an answer in its turn, and in this way clearly shows that reason can't be satisfied by explanations relying on facts about how things go in the empirical world? Finally, •who doesn't see, from the utter contingency and dependence of everything he thinks and assumes using mere principles of experience, the impossibility of stopping with those principles? And •who doesn't feel himself compelled, despite all the prohibitions against losing himself in transcendent Ideas, to seek peace and contentment (beyond all the concepts that he can vindicate by experience) in the concept of a being the Idea of which can't be seen to be possible, but which can't be refuted either because it relates to a mere being of the understanding, without which Idea reason would remain forever dissatisfied?

Where extended things are concerned, *boundaries* always presuppose a space existing outside a certain definite place, and enclosing it; *limits* don't require anything like that, but are mere negations, indicating of some quantity that it isn't absolutely complete. But our reason sees around itself a *space for knowledge of things in themselves*, so to speak, though it can never have definite concepts of them and is limited to appearances only.

As long as the knowledge of reason is all of one kind—for example, reasoning within number-theory, within geometry, within natural science, or the like—definite boundaries to it are inconceivable. In mathematics and in natural science human reason recognizes *limits*, that is, recognizes that its inner progress will never be complete; but it doesn't recognize *boundaries*, i.e. doesn't recognize that outside

it there's something it can't ever reach. In mathematics there's no end to the enlargement of our insight or to the new discoveries that may be made; similarly in natural science, there's no end to the discovery of new properties of nature, of new forces and laws, through continued experience and unification of it by reason. So these sciences are never complete, which means that at any time they have limits. But these limits should not be misunderstood—i.e. should not be thought of as boundaries—for mathematics bears only on *appearances*, and so it has no dealings with anything that can't be an object of sensible intuition, such as the concepts of metaphysics and of morals, which means that it has no dealings with anything that could be a boundary for it. Mathematics can never lead to such things, and has no need for them. So there is a continual progress and approach *towards* completion in these sciences, towards the point or line, so to speak, of contact with completeness. The inwardness of things doesn't show up in the domain of appearance, though the Idea of it can serve as the ultimate ground of explanation of appearances; and natural science will never reveal it to us. But it isn't needed for scientific explanations. Indeed, even if such ultimate grounds of explanation were to be offered from other sources (for instance, if angels told us about them), natural science still ought to reject them and not use them to advance its explanations. For those must be based only on what can belong to experience as an object of sense and be connected with our actual perceptions in accordance with empirical laws.

But metaphysics, in the dialectical attempts of pure reason (which we don't undertake arbitrarily or wantonly, being driven to them by the nature of reason itself), leads us to *boundaries*. And the transcendental Ideas, just because we can't evade them and can never realize them in the sense

of *encountering* an instance of one of them, serve not only to tell us that the pure use of reason has boundaries, but also to show us where they are. That's the purpose and function of this disposition of our reason, which has given birth to metaphysics as its favourite child; and this child, like every other in the world, is a product not of blind chance but of an original seed that is wisely organized for great purposes. For metaphysics, perhaps more than any other science, has in its main outlines been placed in us by nature itself, and can't be viewed as the outcome of an arbitrary choice or of an accidental enlargement of our thoughts in the progress of experience—from which indeed it is wholly separate.

Concepts and laws of the understanding suffice for the empirical use of reason, that is, for the use of it within the world of the senses; but they don't satisfy *reason itself*, because it faces an infinite sequence of questions with no hope of ever completely answering them. The transcendental Ideas, which have that completion as their aim, are such problems of reason. Now reason sees clearly that the world of the senses can't contain this completion, neither (therefore) can all the concepts that serve only for understanding the world of the senses—space and time, and the ones I have presented under the label 'pure concepts of the understanding'. The world of the senses is nothing but a chain of appearances connected according to universal laws; so it has no existence for itself, isn't really the thing in itself, and consequently must stand in a relation to something other than itself, namely to what contains the grounds of this experience—to beings that can be known not merely as phenomena but as things in themselves. It is only in the knowledge of *these* that reason can hope to satisfy its demand for completeness in the advance from the conditioned to its conditions.

In sections 33–4 above I indicated the limits of reason with regard to all knowledge of mere creations of thought. [The

word 'limits'—*Schranken*—doesn't occur in those two sections.] Now, since the transcendental Ideas have made it necessary for us to approach them, and thus have led us to the spot where occupied space meets the void, so to speak—i.e. where experience touches that of which we can know nothing, namely noumena—we can settle what the boundaries are of pure reason. For in all *boundaries* there's something positive:

for example, a surface is the boundary of corporeal space, and *is itself a space*; a line *is a space* that is the boundary of a surface; a point is the boundary of a line but yet *is always a place* in space,

whereas *limits* contain mere negations. The limits pointed out in sections 33 and 34 are still not enough [meaning 'not enough to satisfy us'?] once we have discovered that there is still something beyond them (though we can never know what it is in itself). For the question now arises: How does our reason conduct itself in this connection of what we know with what we don't know and never shall? There is here an actual connection of the known with something completely unknown (which will always remain so); and even if the unknown isn't going to become the least bit known (and there's no hope that it will), the concept of this connection must still be capable of being identified and brought into clarity.

So we ought to have the thought of an immaterial being, a world of understanding, and a supreme being (all mere noumena), because it is only in these items—as things in themselves—that reason finds completion and satisfaction, which it can never hope for in deriving appearances from grounds that are homogeneous with them and therefore demand to be grounded in their turn. Another reason why we ought to have those thoughts is that appearances really do bring in something distinct from

themselves (and totally unlike them), in that appearances always presuppose an object in itself ·of which the appearance is an appearance·, and thus they suggest its existence whether or not we can know more of it.

Now, we can never *know* these beings of understanding as they are in themselves, i.e. determinately, but still we have to *assume* them in •relation to the world of the senses and connect them with that world by means of reason; so we shall at least be able to think •*this connection* by means of such concepts as express their relation to the world of the senses. ·This *relational* approach to noumena is the best we can do·. For if we think a being of the understanding •through nothing but pure concepts of the understanding, we really think nothing definite, and consequently our concept has no significance; and if we think it •through properties borrowed from the world of the senses, it is no longer a being of understanding but is thought as one of the phenomena and belongs to the world of the senses. I shall illustrate this with the notion of the supreme being.

The •deistic concept—i.e. the •thin concept of a (not necessarily personal) supreme being—is a wholly pure concept of reason; but all it represents is *a thing containing all realities*. It can't pick out any one reality—thereby saying something in detail about the supreme being—because to do so it would have to use an example taken from the world of the senses, and in that case I (·as a user of the concept in question·) would after all be dealing only with an object of the senses, not something of a radically different sort that can never be an object of the senses. Here's an example:

Suppose that I attribute *understanding* to the supreme being. My only concept of understanding is one that fits understandings like mine—one that has to get its intuitions ·passively· from the senses, and that occupies itself ·actively· in bringing those

intuitions under rules of the unity of consciousness. If I applied that concept to the supreme being, I would be saying that the raw materials of the supreme being's thought come from the realm of appearance; but it was the inadequacy of appearances ·to meet the demands of reason· that forced me to beyond them to the concept of a being that doesn't •depend on appearances and isn't •identified or characterized through them. ·To credit the supreme being with understanding, therefore, I need a concept of understanding from which the notion of getting-data-from-the-senses has been purged·. But if I separate *understanding* from *sensibility* to obtain a pure understanding ·that the supreme being might have·, then nothing remains but *the mere form of thinking without intuitions*; and form alone doesn't enable me to know anything definite, and so it doesn't enable me to point my thought at the supreme being as an object. ·On the one hand, then, I mustn't suppose that the supreme being •thinks about sensible intuitions; on the other, I mustn't suppose that the supreme being •thinks without having intuitions to think about. So·: for my purpose ·of attributing understanding to the supreme being· I would have to conceive another kind of understanding, such as would ·actively· intuit its objects ·*itself*, instead of passively having intuitions of them brought to it by sensibility·. But I haven't the least notion of such an understanding, because human understanding is conceptual, its only way of knowing is through general concepts, ·and it has no ability to present itself with intuitive data·.

I shall run into exactly the same trouble if I attribute a *will* to the supreme being; for I have this concept only by drawing it from my internal experience; I experience will in myself as

based on facts of the form *I shall not be satisfied unless I get object x*, which means that my will is grounded in sensibility, through which desired objects are presented to me; and that dependence on sensibility is absolutely incompatible with the pure concept of the supreme being.

Hume's objections to deism are weak, and affect only the arguments and not the thesis of deism itself. But as regards *theism*, which is supposed to come from adding certain *content* to deism's merely transcendent and thus empty concept of the supreme being, his objections are very strong; indeed they are irrefutable as arguments against certain forms of theism, including all the usual ones. Hume always insists that •by the mere concept of an *original being*, to which we apply only ontological predicates ('eternal', 'omnipresent', 'omnipotent'), we don't think anything definite, and that •other properties must be added if we are to have a concept of a definite, concrete thing. This isn't a trivial requirement, Hume holds. For example, he says that it isn't enough to say *It is a cause*; but we must explain what kind of causality it has—for example, whether it is exercised through understanding and will—and *that* is the point at which his attack begins on his real topic, theism; up to there he had been attacking only the arguments for deism, which isn't a notably dangerous thing to do. All his dangerous arguments refer to anthropomorphism [from Greek meaning 'man-shaped-ism'; in theology anthropomorphism is the view that God is like man]. Hume holds this to be inseparable from theism, and to make it internally self-contradictory; and if anthropomorphism were left out of the theological story, theism would drop out with it, and nothing would remain but deism. We can't make anything out of deism: it is worthless, and can't serve as a foundation for religion or morals. If this anthropomorphism were really unavoidable, no proofs whatever of the existence of a supreme being, even if they were all granted, could give us a detailed concept of this being without involving us in contradictions.

When we connect the command *to avoid all transcendent judgments of pure reason* with the apparently conflicting command *to proceed to concepts that lie beyond the domain*

of immanent (empirical) use, we become aware that the two commands can subsist together, but only right *on the boundary of all permitted use of reason*—for this boundary belongs equally to the domain of experience and to that of the creations of thought [= Ideas]. And through that awareness we also learn how these Ideas, remarkable as they are, serve merely for marking the boundaries of human reason. On the one hand they give warning •not to go on extending our empirical knowledge with no thought of boundaries, as though nothing but sheer *world* remained for us to know, and yet on the other hand •not to overstep the bounds of experience and want to make judgments about things beyond them, as things in themselves.

But we stop *at* this boundary if we limit our judgment merely to how •the world may relate to •a being whose very concept lies beyond the reach of any knowledge we are capable of within the world. For we don't then attribute to the supreme being *in itself* any of the properties through which we represent objects of experience, and so we avoid *dogmatic* anthropomorphism; but we attribute those properties to the supreme being's *relation to the world*, thus allowing ourselves a *symbolic* anthropomorphism, which in fact concerns only language and not the object itself.

When I say that we are compelled to view the world *as if* it were the work of a supreme understanding and will, I actually say nothing more than that a watch, a ship, a regiment, are related to the watchmaker, the shipbuilder, the commanding officer in the same way that the sensible world (or everything that underlies this complex of appearances) is related to the unknown; and in saying this I don't claim to know the unknown as it is •in itself, but only as it is •for me or •in relation to the world of which I am a part.

58

Such knowledge is knowledge *by analogy*. This doesn't involve (as the word 'analogy' is commonly thought to do)

- an imperfect similarity of two things, but rather
- a perfect similarity of relations between the members of two quite dissimilar pairs of things.¹³

By means of this analogy we are left with a concept of the supreme being that is detailed enough *for us*, though we have omitted from it everything that could characterize it *absolutely* or *in itself*; for we characterize only its relation to the world and thus to ourselves, and that is all we need. Hume's attacks on those who want to determine this concept absolutely, taking the materials for doing so from themselves and the world, don't affect my position; he can't object against me that if we give up the objective anthropomorphism of the concept of the supreme being we have nothing left.

Hume in the person of Philo in his *Dialogues* grants to Cleanthes as a necessary hypothesis the *deistic* concept of the original being, in which this being is thought through

nothing but the ontological predicates of 'substance', of 'cause', etc. Two comments on this:

(1) We **must** think the original being in this way; because there's no other way to think it, and if we don't have *any* thought of the original being reason can't have any satisfaction in the world of the senses, where it is driven by mere conditions that are in their turn conditioned, thus driving us back along a never-starting sequence of causes. If the use of reason in relation to all possible experience is to be pushed to the highest point while remaining in complete harmony with itself, the only possible way to do this is to assume a highest reason as a cause of all the connections in the world. Such a principle must be thoroughly advantageous to reason, and can't hurt it anywhere in its natural use.

(2) We **can properly** think the original being in this way, because those predicates—'substance', 'cause' etc.—are mere categories, which yield a concept of the original being that isn't determinate and for just that reason isn't limited

¹³ Thus, there is an analogy between the legal relation of human acts and the mechanical relation of motive powers. I can never do something to someone else without giving him a right to do the same to me in the same circumstances; just as no body can act on another through its motive power without thereby causing the other to react equally against it. *Right* and *motive power* are quite dissimilar things, yet in their relation there is complete similarity. By means of such an analogy I can give a relational concept of things that are absolutely unknown to me. For example, as

a = the promotion of the happiness of children

is related to

b = parental love,

so

c = the happiness of the human species

is related to

x = the unknown in God, which we call love.

Not because (x) God's love has the least similarity to (b) any human inclination; but because we can suppose its relation to the world to be similar to a relation that some things in the world have to some others. But the relational concept in this case is a mere category, namely the concept of cause, having nothing to do with sensibility.

to any conditions of sensibility. In thinking of the original being in this way we don't fall into anthropomorphism, which transfers predicates from the world of the senses to a being quite distinct from that world. We aren't transferring reason as a property *to the original being in itself*, but only *to its relation* to the world of the senses; and so anthropomorphism is entirely avoided. For all we are considering here is the *cause* of something that is perceived everywhere in the world, namely the world's rational form; and insofar as the supreme being contains the ground of this rational form of the world, reason is to be attributed to it. But in saying that the supreme being has reason, we are speaking analogically, expressing only the *relation* that the unknown supreme cause has to the world, and we do this so as to see everything in the world as being in the highest degree *reasonable*. This procedure doesn't involve us in treating reason as an attribute through which we can conceive *God*; what we do, rather, is to conceive *the world* in the way that's needed if we are to tackle it with the greatest possible principled use of reason. In this way we admit that the supreme being in itself is quite inscrutable and isn't even *conceivable* in any determinate way, and that keeps us from two errors that we might otherwise make. Roughly and briefly, they are the errors of trying to explain God in terms of the world, and trying to explain the world in terms of God. A little more fully, one of the errors is that of

making a transcendent use of our concept of reason as an efficient cause (by means of the will), trying to describe the nature of God in terms of properties that are only borrowings from human nature—thereby

losing ourselves in gross and extravagant notions; and the other error consists in allowing our contemplation of the world to be flooded with supernatural patterns of explanation, led by the transfer to God of our notions of human reason—thereby deflecting this contemplation from its proper role, which is to study mere nature through human reason, not rashly to derive nature's appearances from a supreme reason.

The best way to put it, given our weak concepts, is this: we should conceive the world AS IF its existence and its inner nature came from a supreme reason.

In thinking of it in this way, we achieve two things. We recognize what the world itself is like, without wanting to determine what its cause is like in itself. And we see the ground of what the world is like (the ground of the world's rational form) in the *relation* of the supreme cause to the world, not finding the world sufficient by itself for that purpose.¹⁴

Returning now to what Philo granted to Cleanthes: We are perfectly free to predicate of this original being a *causality through reason* in respect of the world, thus moving on to theism; and this doesn't oblige us to attribute this kind of causally powerful reason to the original being itself, as a property attached to it.

Thus the difficulties that seem to stand in the way of theism disappear. We achieve this by joining to Hume's principle:

Don't push the use of reason dogmatically beyond the domain of all possible experience

¹⁴ I shall say: the causality of the highest cause relates to the world in the same way that human reason relates to its artifacts. That leaves the nature of the supreme cause itself still unknown to me: I only compare its effect (the order of the world) which I know, and the conformity of this to reason, with the effects of human reason, which I also know; and hence I call the supreme cause 'reason', without thereby attributing to it what I understand by 'reason' as applied to man, or assigning to it any property of anything else that I know.

this other principle, which he quite overlooked:

Don't consider the domain of experience as something which in the eyes of our reason sets its own boundaries.

The critique of reason here indicates the true middle way between •the dogmatism against which Hume fought and •the scepticism that he wanted to introduce to oppose it. It isn't the usual kind of 'middle way', which one is advised to pick out for oneself as it were mechanically (a little of the one, a little of the other), making nobody any the wiser. Rather, it is a middle way that can be delineated exactly, according to principles.

59

At the beginning of this note I used the metaphor of a *boundary* in order to fix the limits on the proper use of reason. [This 'note' is this Conclusion. Its title includes the term 'boundaries', but the concept isn't discussed until pages 62–3.] The world of the senses contains mere appearances, which aren't things in themselves; but the understanding must assume things in themselves (noumena), because it recognizes the objects of experience as appearances ·and understands that they must be appearances *of* something·. Our reason covers both •appearances and •things in themselves, and the question arises: How does reason go about setting boundaries to the understanding with respect to both these domains? Experience, which contains all that belongs to the world of the senses, doesn't set bounds for itself; it proceeds in every case from some conditioned item to some other ·item that is its condition, and is also a· conditioned object; ·and nothing in this procedure requires it ever to come to a halt·. The boundary of experience must lie •wholly outside it, and that is •the domain of pure beings of the understanding. But when it comes to *finding out* what these beings are like, this

domain is for us an empty space; and when we are dealing with concepts whose instances we can identify and study, we can't move out of the domain of possible experience. But a boundary is itself something positive, which belongs as much to •what lies inside it as to •the space lying outside the given totality; so reason partakes in real positive knowledge when it stretches out to this boundary. Reason doesn't try to go beyond the boundary, because out there it is confronted by an empty space in which it can conceive •forms of things but can't conceive •things themselves. Still, even when it adopts this stance towards the boundary, just in *setting* the boundary reason has *knowledge*. In this knowledge it isn't confined within the world of the senses, but it doesn't stray outside it either; rather, as befits the knowledge of a boundary, it focuses on the *relation between* what lies outside the boundary and what's contained inside it.

Natural theology is such a concept at the boundary of human reason, ·because at that boundary· reason finds itself compelled to look out further towards the Idea of a •supreme being (and, for moral purposes towards the Idea of a world that can be thought but not experienced). It doesn't do this so as to find out anything about this •mere creation of the understanding lying *outside* the world of the senses; its purpose is rather to employ principles of the greatest possible (theoretical as well as practical) unity to guide its conduct *within* the world of the senses—a purpose that is served by relating these principles to an independent reason, as the cause of all the connections ·found in the world of the senses·. The aim isn't to merely *invent* a being ·of reason. Invention isn't in question here·, because beyond the world of the senses there must *be* something that can be thought only by the pure understanding. Reason's aim is to *characterize* this being, though of course only by analogy.

And so we are left with our original proposition, which is

the upshot of the whole critique:

Reason, through all its *a priori* principles, never teaches us about anything except objects of possible experience, and about these it teaches us nothing more than can be known in experience.

But this limitation on what reason can do doesn't prevent it from leading us to the objective *boundary* of experience, i.e. to the *relation* to something that is the ultimate ground of all objects of experience without itself being one of them. Still, reason doesn't teach us anything about what this 'something' is like *in itself*—only about how it *relates* to reason's own complete and utterly high-minded use in the domain of possible experience. But this is all the usefulness we can reasonably want ·reason to have·, and we have cause to be satisfied with it.

60

So I have fully exhibited metaphysics as something we can *do*, showing it as an automatic upshot of *the natural tendency of human reason*, and showing what our essential goal is when we do metaphysics. But we have found that ·things can go wrong in this pursuit·: this *wholly natural* use of such a tendency of our reason, if it isn't reined in and given limits (which can come only from a scientific critique), entangles us in transcendent *dialectical* inferences ·leading to conclusions· of which some are illusory and others are even in conflict with one another; and this fallacious metaphysics is not a help but an obstacle to the advancement of our knowledge of nature. So it is worth our while to investigate the *natural goals* towards which we can steer this liking that our reason has for transcendent concepts, ·and this will counteract the mishaps mentioned above·, because everything that is natural must be originally aimed at some useful purpose.

Such an enquiry is risky, and I admit that what I can say about it is only conjecture, like every speculation about nature's original purposes. But this is permissible, just this once, because I am enquiring not into •the objective validity of metaphysical judgments but into •our natural tendency to *make* such judgments, so that the enquiry belongs not to the system of metaphysics but rather to the study of mankind.

The transcendental Ideas, taken all together, form the real problem of natural pure reason, a problem that compels reason to quit the mere observation of nature, to go beyond all possible experience, and in so doing to bring into existence this thing (whether it is knowledge or sophistry!) called metaphysics. When I consider all of these Ideas, I think I see that

the aim of this natural tendency—this metaphysics—is to free our thinking from the fetters of experience and from the limits of the mere observation of nature, taking this freedom at least far enough to open up to us a domain containing only objects for the pure understanding, which no sensibility can reach.

This is not so that we can speculatively occupy ourselves with this domain (for we can find no ground to stand on while we do that), but so that we can think of moral principles as at least *possible*. ·The connection between their being possible and the domain opened up by metaphysics is as follows·. Reason absolutely requires that moral principles be universal; and they can't achieve universality unless they can fix their expectations and hopes on the domain of transcendental Ideas ·because in the domain of experience strict universality is never to be found·.

Now I find that the •*psychological* Idea, little as it shows me of the nature of the human soul ·thought of as something· elevated above all concepts of experience, does plainly enough show the inadequacy of these concepts, and in that

way steers me away from a **materialist** theory of mind—a theory that’s unfit to explain anything in nature, as well as cramping the use of reason in moral thinking. The •*cosmological* Ideas serve similarly to keep us from **naturalism**, which asserts that nature is sufficient unto itself; they do this through •bringing home to us• the obvious fact that even if we had all possible knowledge of nature, reason’s legitimate demands wouldn’t be satisfied. Finally there is the •*theological* Idea, •whose service to us is as follows•. All natural necessity in the sensible world is conditioned, because it always involves something’s being necessitated by something else that is also conditioned; and thus *unconditional* necessity is to be looked for only in a cause that is different from the world of the senses. And the causality of this cause can’t be yet another example of natural necessity, for if it were it could never make comprehensible the existence of the contingent (as its consequent). So the theological Idea, •which is the Idea of a non-natural cause of everything contingent•, is something that reason uses to free itself from **fatalism**, and to arrive at the concept of a cause possessing freedom, or of a highest intelligence. (This frees us from both versions of fatalism: **(1)** blind natural necessity in the system of nature itself, without a first principle, and **(2)** blind causality of a first principle of nature.) Thus the transcendental Ideas serve, if not to instruct us positively, at least to put a stop to the impudent assertions of *materialism*, of *naturalism*, and of *fatalism*—assertions that restrict the domain of reason—thereby making room for the moral Ideas to operate outside the domain of speculation. This, I should

think, goes some way towards explaining reason’s natural tendency •to engage with Ideas•, which I mentioned earlier.

The facts about the •practical or moral usefulness that a purely •speculative science can have don’t lie within the province of the science itself; so they can be seen simply as a scholium [= ‘explanatory note, marginal comment’] which, like all scholia, is not a part of the science itself. Still, this material surely lies within the boundaries of philosophy, especially of philosophy drawn from the well of pure reason—a part of philosophy in which reason’s •speculative use in metaphysics must necessarily be all of a piece with its •practical use in morals. Hence the unavoidable dialectic of pure reason, considered as something occurring in metaphysics as a natural tendency, deserves to be explained not only as •an illusion that needs to be cleared away but also, if possible, as •an upshot of• *something put in place by nature* for a purpose—though this task lies outside the job-description of metaphysics proper, and so can’t rightly be assigned to it.

The solutions of the questions put forward in the *Critique* at A647–68 = B675–96 should be regarded as a second scholium—this time one that’s more closely related to the content of metaphysics.¹⁵ For that part of the *Critique* presents certain principles of reason that characterize *a priori* •the order of nature or rather •the understanding which is to seek nature’s laws through experience. They seem to have propositional content and not merely to be rules for how the understanding should be employed, and to be law-giving with regard to experience, although they spring from mere reason, which cannot like the understanding be considered as a

¹⁵ Throughout the *Critique* I stuck to my resolve not to neglect anything, however deeply hidden, that could enable the inquiry into the nature of pure reason to be completed. Afterwards it is up to each person to decide how far to take his researches, once he has been shown what remains to be done. This •attitude to further research• can reasonably be expected from someone who has made it his business to survey the whole field, so as to leave it to others for future cultivation and for whatever subdividing of it they think fit. The scholia belong to this •part of the total project•; because of their dryness they can hardly be recommended to amateurs, and hence they are presented only for experts.

principle of possible experience. This agreement between principles issued by reason and what is found in experience may rest on this:

Nature isn't attached to appearances (or to the sensibility through which appearances come) in themselves, but is to be found only in the relation of sensibility to the understanding. And just as nature takes us upwards from sensibility to understanding, so the theoretical use of the understanding takes us further upwards to reason. A thoroughgoing unity in the use of the understanding for the sake of a systematically unified possible experience can be had only if the understanding is related to reason. And so, putting the two together, experience is indirectly subject to

the legislation of reason.

The question of whether this is where the agreement comes from may be pondered by those who want to trace the nature of reason even beyond its use in metaphysics, into the general principles for making general natural history systematic. In the *Critique* I represented this task as important, but I didn't try there to carry it out.

And thus I conclude the analytic solution of my own chief question: *How is metaphysics in general possible?* by starting with the actual doing of metaphysics (or at least with the consequences of that) and climbing from there to the grounds of its possibility. [See the explanation of 'analytic', in this sense of the word, on page 6.]