

Critique of Pure Reason the Dialectic

Immanuel Kant

Copyright ©2010–2015 All rights reserved. Jonathan Bennett

[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. Each four-point ellipsis indicates the omission of a brief passage that seems to present more difficulty than it is worth. Longer omissions will be reported between square brackets in normal-sized type. This version follows (B) the second edition of the *Critique*, though it also includes the (A) first-edition version of the Paralogisms of Pure Reason. Undecorated marginal numerals refer to page-numbers in B; ones with an 'A' in front refer to A, and are given only for passages that don't also occur in B. The likes of ..356 in the margin mean that B356 (or whatever) started during the immediately preceding passage that has been omitted or only described between square brackets. These marginal numerals can help you to connect this version with other translations, with the original German, and with references in the secondary literature. Cross-references to other parts of this work include the word 'page(s)', and refer to numbers at the top-right corner of each page.—The Transcendental logic divides into the Transcendental analytic, which started on page 45, and the Transcendental dialectic, which starts here.

First launched: January 2008

Contents

Introduction	155
1. Transcendental illusion	155
2. Pure reason as the seat of transcendental illusion	157
Book 1: The concepts of pure reason	163
1. The ideas in general	163
2. The transcendental ideas	166
3. System of the transcendental ideas	171
Book 2: The dialectical inferences of pure reason	174
Chapter I, The paralogisms of pure reason (first edition)	175
First paralogism: Substantiality	178
Second paralogism: Simplicity	179
Third paralogism: Personhood	182
Fourth paralogism: ideality (in regard to outer relation)	184
The paralogisms of pure reason (second edition)	197
Chapter 2: The antinomy of pure reason	206
1. System of cosmological ideas	207
2. Antithetic of pure reason	211
First antinomy	213
Second antinomy	216
Third antinomy	219
Fourth antinomy	222
3. What's at stake for reason in these conflicts	225
4. The transcendental problems of pure reason, considered as downright <i>having</i> to be soluble	230
5. A sceptical look at the cosmological questions raised by the four transcendental ideas	233
6. Transcendental idealism as the key to sorting out the cosmological dialectic	235
7. Critical solution of reason's cosmological conflict with itself	238
8. Applying the regulative principle of pure reason to the cosmological ideas	242
9. Putting the regulative principle of reason to work empirically, in connection with the cosmological ideas	244

Chapter 3: The ideal of pure reason	263
1. The ideal in general	263
2. The transcendental ideal	264
3. Speculative reason's arguments for the existence of a supreme being	269
4. There can't be a successful ontological argument for the existence of God	272
5. There can't be a successful cosmological argument for the existence of God	276
6. There can't be a successful physico-theological argument for the existence of God	283
7. Critique of all theology based on speculative principles of reason	287
Appendix to the transcendental dialectic	292
1. The regulative use of the ideas of pure reason	292
2. The final purpose of the natural dialectic of human reason	302

Introduction

1. Transcendental illusion

..349 [Kant is about to warn us not to think that a 'logic of illusion' is a 'doctrine of probability'. The warning looks more apt in German than it does in English. The word standardly translated as 'illusion' is **Schein**, cognate to the verb *scheinen* = 'seem'. And the German for 'probability' is *Wahrscheinlichkeit* = 'true-seemingness'.] I have already characterized dialectic in general as a *logic of illusion* [see page 45]. I should head off right away two possible misunderstandings of that. **(1)** I don't mean that it's a doctrine of *probability*. For probability is truth. It's admittedly truth that is known on insufficient grounds, so that the knowledge of it is imperfect; but that doesn't mean that it is *deceptive*; so probability theory belongs not here in the dialectic but in the analytic part of logic. [A reminder: in this version, 'know' translates *erkennen*, which doesn't imply anything of the sort 'known for sure' or 'known through overwhelming evidence'; see the note on pages 2-3.] **(2)** It's even more wrong to identify *illusion* with *appearance*. The essence of *illusion* is that it leads to *error*, so the concept of illusion belongs only in contexts where 'true' and 'false' are in play. Now, there's no work for true/false to do in connection with •intuitions, as distinct from •judgments that are made on the basis of intuitions. (That's why it is right to say that the senses don't err—not because they always judge correctly but because they don't judge at all!) The domain of operation of the concepts of truth/falsity/ error/illusion is that of the *judgment*—that is, the relation of the object to our understanding. A representation of the senses never involves error, because it never makes any judgment whatsoever; and there is no error, either, in any item of knowledge that completely accords with the laws of understanding—a natural force can't deviate from its own laws unless it is

influenced from outside itself. . . . But the senses and the understanding are our only sources of knowledge; so error must be brought about by the understanding's being influenced from outside itself—specifically, by the unobserved influence of sensibility on the understanding. What happens is that the subjective grounds of the judgment join forces with the objective grounds, making the objective grounds deviate from their true function.¹ Analogously, a moving body would keep moving for ever along the same straight line if nothing interfered with it, but it swerves away from that line if another force acts on it in a different direction. To distinguish the proper action of the understanding from the external force that is mixed in with it, we have to regard its erroneous judgment as the diagonal between two forces—forces that push the judgment in different directions that enclose an angle (so to speak)—and to break this composite action down into the simple actions of the understanding and of the sensibility. In the case of pure *a priori* judgments this task is performed by transcendental reflection, through which, as I have already shown [see the explanation on page 144], every representation is given its place in the corresponding faculty of knowledge, so that the understanding's influence can be distinguished from that of sensibility.

I'm not concerned here with empirical illusions (e.g. optical illusions) that occur in the empirical use of rules of understanding that are otherwise correct, and through which the faculty of judgment is misled by the influence of imagination. My topic is *transcendental illusion*, which exerts its influence on principles that aren't even meant for use in experience

¹ When sensibility is subordinated to understanding, as providing the understanding with something to work on, it is the source of real items of knowledge. But when that same sensibility influences how the understanding works and dictates the judgments that it makes, it's the basis of error.

(if they were, we would at least have a criterion of their correctness). Transcendental illusion defies all the warnings of criticism and sweeps us out beyond the empirical use of categories and fobs us off with a merely deceptive extension of *pure understanding*. I shall label as *immanent* any principles that are usable only within the limits of possible experience, and I'll label as *transcendent* any principles that profess to go beyond these limits. ['Immanent' comes from Latin meaning 'remain inside', and 'transcendent' comes from Latin meaning 'climb over'.] Don't confuse *transcendent* with *transcendental*—they are not equivalent terms. On the one hand we have:

- the *transcendental* use or misuse of the categories, which is merely an error of the faculty of judgment when it isn't properly reined in by criticism, so that it doesn't pay enough heed to the boundary of the territory in which (the *only* territory in which) pure understanding is allowed to run free.

353 I said 'use or misuse', but in fact it is always a misuse. The principles of pure understanding, which I have expounded in the *Analytic of Principles* [pages 89–154], are for empirical use only, and not for transcendental use, i.e. use extending beyond the limits of experience. In contrast to that we have

- transcendent* principles—actual principles that encourage us to tear down all those boundary-fences, step across them, and claim an entirely new domain that recognises no limits of demarcation.

If my critique can succeed in exposing the illusion in these alleged principles, then the principles that are for merely empirical use can be set off against the others by being called *immanent* principles of pure understanding.

- Logical illusion—the illusion that a formally invalid argument is valid—is a mere imitation of the form of reason, and it happens only when we don't attend carefully enough to the

•relevant logical rule. When we look more carefully at the given case, the illusion—*click!*—vanishes. •Transcendental illusion, on the other hand, persists even after it has been detected and clearly revealed as invalid by transcendental criticism; an example of this is the illusion in the proposition: *The world must have a beginning in time*. [Kant will return to this on page 213. •In this version, 'criticism' = 'critique'. A note on page 12 explains the choice of which word to use in a given context.] The cause of this kind of illusion is the fact that certain basic •rules and maxims that tell us how to use our reason (subjectively regarded as a faculty of human knowledge) have all the appearance of being •objective principles. •That is, rules that really mean things of the form 'When engaged in cognitive activities, do X' have the appearance of being of the form 'The real world is Y'. It's not just that they are really subjective yet appear to be objective; but also they are really advice or commands, yet they appear to be informative propositions. Some ways of connecting our concepts are advantageous to the understanding, and are •in that sense •*subjectively necessary for us*, and we see these as *objectively necessary for the world*, i.e. as statements about what things must be like in themselves. This is an *illusion* that can no more be prevented than. . . an astronomer can prevent the moon from appearing larger as it rises, even though he knows that it isn't really.

354

So the transcendental dialectic will content itself with exposing the illusion of transcendent judgments, while also keeping us from being deceived by it. It can't make the illusion actually disappear (as logical illusion does), because what we have here is a •*natural* and inevitable illusion, trading on subjective principles that it foists on us as objective; whereas logical dialectic in exposing deceptive inferences has to deal merely with a failure to follow the rules—i.e. with an illusion •*artificially* created by something imitating

a valid inference. So there we have it: there's a natural and unavoidable dialectic of pure reason. It's not something that a bungler might get tangled in through ignorance, or something that a sophist has contrived so as to confuse thinking people. Rather, it is inseparable from human reason; even after its deceptiveness has been exposed, it will go on playing tricks with reason, continually tricking it into momentary aberrations that have to be corrected over and over again.

2. Pure reason as the seat of transcendental illusion

A. Reason in general

All our knowledge starts with the •senses, moves up from there to the •understanding, and ends with •reason—our highest faculty—so that it can work up the materials provided by intuition, bringing them under the highest unity of thought. As I set myself to explain this highest cognitive faculty, I find myself in some difficulty. Like the understanding, reason can be used **(1)** in a merely *formal* (i.e. logical) manner, in which it abstracts from all content of knowledge. But reason also has **(2)** a *real* use, because it contains within itself the source of certain concepts and principles that it doesn't borrow either from the senses or from the understanding. For a long time now logicians have defined our ability to use reason in the **(1)** formal way as our ability or faculty for making mediate inferences. .i.e. for drawing conclusions from two or more premises;. . .but this doesn't throw any light on the **(2)** other use of reason, in which reason itself gives birth to concepts.

[Kant's next two sentences are hard to understand unless one knows what is to come later. The present version of this paragraph is *much*

longer than the original, but adds nothing to what he meant.]

We are faced, then, with a division of reason into **(1)** reason-used-logically and **(2)** reason-used-transcendentally, and now we have to hunt for a higher concept of reason that has **(1)** and **(2)** as special cases of it—i.e. the more general concept of *reason* that is an ingredient in both of the more specific concepts of **(1)** and **(2)**. (This is the 'difficulty' to which I have referred.) To characterize the higher or most general concept of reason, we need to assemble a table or chart setting out all the concepts that fall under it; and a clue to doing that is provided by what we found with the faculty of understanding. That too has both logical and transcendental uses, and it turned out that the logical uses provide the key to the whole story of the understanding, including its transcendental part. Just think back or look back to the way we moved from •the table of judgment-kinds [page 49] to •the table of categories [page 52] and from that to •the table of principles of pure understanding [page 99]. Well, I shall show that a disciplined account of the basic logical ways in which reason can be used will point to an over-all account of the nature and shape of reason as a whole—the genealogical tree of the concepts of reason. In doing this I'll be taking reason to be the faculty of *principles*. (This is in contrast to the understanding, which I have been treating as the faculty of *rules*. I have sometimes spoken of 'principles of the understanding'; I'll explain that shortly.)

[We are about to encounter talk about syllogisms. Any argument with the form of this:

- (1)** Some bullies are cowards,
- (2)** All cowards are depressed, therefore
- (3)** Some bullies are depressed

is a syllogism. Its major premise is **(2)**, because it contains the predicate of the conclusion ('depressed'); and **(1)** is the minor premise.]

The term 'principle' is ambiguous. It is often used to stand for any item of knowledge that can be *used as a principle*,

even if its origin doesn't qualify it as *being* a principle. Any universal proposition, even one derived from experience through induction, can serve as the major premise in a syllogism, which involves its being 'used as a principle' in the weak sense of being used as a basis from which to infer something else, but that doesn't mean that it is a principle properly so-called. Mathematical axioms such as 'There can be only one straight line between two points' are instances of universal *a priori* knowledge, and therefore relate as 'principles' to all their instances. (And because they are known *a priori* they are *more* like 'principles' strictly so-called than are empirically established propositions such as 'All cowards are depressed'.) But this doesn't entitle us to say that this property of straight lines is something that we know from principles. In fact we know it only through pure intuition [see page 30].

[Kant now offers an obscure paragraph whose main point is that because any 'All. . .' proposition can be the major premise in a syllogism, and so can in that way be *used as* a principle, therefore the *a priori* propositions associated with the understanding 'can be called "principles" in the sense that they can be used as principles'. Then:]

But if we consider those basic propositions of pure understanding in the light not of how they can be used but rather of what their source is, then we can see that they are *nothing like* principles in the strict sense of propositions expressing knowledge based on concepts. Our ability to have the *a priori* knowledge they express comes from two sources, neither of which consists in *concepts*, namely:

- pure intuition (for the mathematical propositions) or
- conditions that have to be satisfied for any experience to be possible (for the others).

Consider a proposition of the second type: *Every event has a cause*. This can't be inferred merely from the concept of

event; on the contrary, it's a basic proposition that points to the conditions that enable us to have a determinate *event* concept in the first place. [In this passage and some later ones, 'event' translates a German phrase meaning 'thing that happens'. That's what events are—things that happen.]

Thus, the understanding can't supply us with any *synthetic items of knowledge derived from concepts*; and those are the only things that I call 'principles' *period*; though any universal proposition—and especially any that can be known *a priori*—can be described as 'relating to such-and-such in a principle-like way'. 358

It has long been wished. . . that instead of the endless complexity of the laws of the land we could find their *principles*, because that's our only hope of 'simplifying' the law. There's nothing problematic about the thought that there are such principles, because the laws we are considering here are only constraints that we have imposed on our freedom, which means that if they are harmonised and simplified under very general legal principles, the latter are directed to something that is entirely our own work—something that we have generated out of our own legal concepts. Contrast that with the thought that objects in themselves, the very nature of things, stand under principles, and are determined according to mere concepts. That is problematic: it seems impossible, or at least quite contrary to common sense. Well, we'll look into that in due course. My present point is just that

- knowledge derived from principles, strictly and properly so-called

is something quite different from

- knowledge obtained merely through the understanding,

though the latter can be principle-like in being more basic than some other knowledge. . . . 359

Just as understanding can be seen as the faculty that uses rules to unify appearances, reason can be seen as the faculty that uses principles to unify the understanding's rules. Thus, reason never applies directly to experience or to any object. What it applies to is the understanding: its role is to give an *a priori* unity by means of concepts to the understanding's complex web of items of knowledge. This 'unity of reason', as we may call it, is nothing like the unity that the understanding can create by bringing appearances under its concepts.

That's the best I can do to explain the general concept of *the faculty of reason*—or the best I can do without using examples. *They* will be provided later on.

B. *The logical use of reason*

We distinguish what is immediately known (e.g. *A figure bounded by three straight lines has three angles*) from what is only inferred (e.g. *The sum of those three angles equals two right angles*). We're constantly in need of inferences, and eventually we get used to inferring—so much so that we stop being aware of the difference between immediate and inferred knowledge, and often treat as being immediately perceived what has really only been inferred. . . . In every process of reasoning there is

- 360 **(1)** a fundamental proposition (the premise), and
 (2) another proposition (the conclusion that is drawn from **(1)**), and finally
 the inference (logical sequence) by which the truth **(2)** is inseparably connected with the truth of **(1)**.

If **(2)** is contained in **(1)** in such a way that it can be derived from **(1)** without the mediation of a third proposition, the inference is called *immediate*. But if **(2)** can't be reached from the item of knowledge contained in **(1)** until another

judgment is added, then the inference of **(2)** from **(1)** is non-immediate. I call immediate inferences *inferences of the understanding*, and non-immediate ones *inferences of reason*. The proposition 'All men are mortal' contains the propositions 'Some men are mortal', 'Some mortal beings are men', and 'Nothing that isn't mortal is a man'—so these are all immediate conclusions from it, each drawable in an inference of the understanding. On the other hand, the proposition **(1)** 'All men are mortal' doesn't contain **(2)** 'All learned beings are mortal' (it doesn't involve any use of the concept of *learned being*), so **(2)** can only be inferred from **(1)** only with the help of a mediating judgment.

[Warning: Kant's use here of 'mediating judgment' (*Zwischenurteil*, between-judgment) is misleading. If he were using it properly, he would be talking about the case where to get from P to R you have to get from P to Q and then from Q to R. (That would make the need for a mediating judgment a subjective matter: although I can't see that $P \Rightarrow R$ except by bringing in Q, you are smart enough to see that $P \Rightarrow R$ without getting help from Q.) Anyway, that is *not* what Kant means when he speaks of mediating judgments. His real topic here is simply cases where R doesn't follow from P alone but does follow from P together with Q. His announced theory really is that what is logically special about reason is that it is used in inferring conclusions from *pairs* of premises. That isn't a load-bearing part of what's important in the Dialectic, but it figures in some of Kant's preliminary moves, so we need to get straight about it.—Something else that needs to be understood: the standard German word for 'syllogism' is *Vernunftschluss* = 'inference of reason'. Some of what Kant says about such inferences really does fit syllogisms and only syllogisms (e.g. the technical term 'major premise'), but much of the time he is talking more broadly about inferences-from-pairs-of-premises. From now on this version will usually translate *Vernunftschluss* by 'inference of reason'. In the following paragraph, the schematic S-M-P example, which isn't Kant's, is tied to one very simple and basic kind of syllogism, narrowly so-called. But you'll soon see that his topic is broader than that.]

In every inference of reason I first think a *rule* (the major premise) through the *understanding*:

All M are P.

Secondly, I bring a known item ·S· under the condition ·M· of the rule by means of *judgment* (the minor premise):

All S are M.

Finally, I determine [here = 'establish some fact about'] the known item ·S· by applying the predicate ·P· of the rule:

361

All S are P,

arriving at this—the conclusion—a *a priori* through *reason*. There are different kinds of inference of reason—three of them, in fact, corresponding to the three kinds of judgments. They are:

- (1) categorical,
- (2) hypothetical,
- (3) disjunctive.

How do we decide which category a given inference of reason belongs to? By looking at the form of its major premise, i.e. at *how* that premise relates the two items that it involves. [The rest of this paragraph is an addition to what Kant wrote, but it consists only of borrowings from things he will say later.] In a (1) categorical inference of reason like the one semi-illustrated above, it is the subject-predicate relation expressed in the proposition that *all M are P*. In a (2) hypothetical one, of the form:

If P then Q,
P,
therefore Q,

it is the ground-consequent relation expressed in the proposition that *if P then Q*. And in a (3) disjunctive inference, of the form

R or S,
Not-R,
therefore S,

it is the parts-of-a-logical-division relation expressed in the proposition *R or S*.

In most cases the judgment that forms the conclusion is set as a problem—to see whether it follows from judgments already given, ones through which a quite different object is thought. [The element M in the categorical case, P in the hypothetical case, and R in the disjunctive case is 'quite different from' anything in the conclusion.] I look in the understanding to see how this conclusion is situated there; I'm trying to discover whether it stands under certain conditions according to a universal rule. If I find such a condition 'embodied in a rule', and if the conclusion relates to it in the right way, then the conclusion is deduced from the rule—which is also valid for other objects of knowledge. We see here that in inference reason tries to reduce the complex web of knowledge obtained through the understanding to the smallest number of principles (universal conditions) thereby bringing it into the highest possible unity.

C. The pure use of reason

[In this paragraph, Kant speaks of the *Vergleichung* of one proposition with another, standardly translated as 'comparison'. It's hard to avoid that, but what he really means here is 'comparing' not in the sense of *likening* P to Q but only in the sense of *laying them side by side so as to take in their inter-relationship in a single thought*.] Here are two *prima facie* possible accounts of the basic status of reason: 362

- (1) Reason can be considered all on its own; it is an independent source of concepts and judgments that come from it alone and give it a relation to objects.
- (2) Reason is a merely subordinate faculty, whose role is to impose a certain 'logical' form on given items of knowledge. It is through reason that things known by means of the understanding are determinately related

to one another, with lower items of knowledge being brought under higher ones, . . . this being done by comparing them.

Which of these is right? That's the preliminary question we are now facing up to (you'll see in a moment why I call it 'preliminary'). The answer is **(2)** rather than **(1)**. Reason is perhaps not 'merely subordinate', in that it *demands* that the multiplicity of rules of the understanding be unified by principles of reason. (In doing this work,

- reason makes the output of the understanding hang together in a thoroughly connected whole, by bringing it under principles,

just as

- the understanding connects up the various outputs of intuition, by bringing them under concepts.)

But a principle of reason doesn't prescribe any law *for objects*; it doesn't contain anything that is needed as a basis for knowing objects or knowing anything about them (that being what enables the understanding to prescribe laws for objects). Reason is merely a subjective law for the orderly management—the *housekeeping*—of our stock of understanding-outputs, . . . aiming at the greatest possible economy in our use of them. It doesn't entitle us to demand that *the objects* have a uniformity that will make things easier for our understanding and increase its reach; so we can't ascribe any objective validity to the maxim in which reason demands that the output of the understanding be unified. With the preliminary question thus answered, we now come to the big question that will be with us for a long time. In a word, the question is: Does reason in itself—i.e. does *pure* reason—contain *a priori* synthetic principles and rules, and what might such principles consist in?

If pure reason is capable of a transcendental principle through which it yields synthetic knowledge, what will it

be based on? We get sufficient guidance in answering that from two points about the formal and logical procedure of reason.

First, an inference of reason doesn't concern itself with intuitions, aiming to bring them under rules (as the understanding does with its categories). What it deals with are concepts and judgments. Thus, even if pure reason does *somehow* concern itself with objects, what it is *immediately* related to are not objects and the intuition of them, but rather the understanding and its judgments, which *do* deal at first hand with the senses and their intuition for the purpose of establishing facts about their object. The unity of understanding is the unity of a possible experience, but the unity of reason is nothing like that. The proposition *Every event has a cause* contributes to making the unity of experience possible; it's because of this making-experience-possible that the understanding can use its concepts to pull experience together through synthetic propositions like that one. In contrast with that, reason doesn't have the job of making experience possible; and that deprives it of any chance of imposing on experience any such synthetic unity as is imposed by *Every event has a cause*.

Secondly, when reason is put to use logically, it starts with some judgment and tries to find a universal rule of which the judgment is a special case. (The universal rule is the major premise, and the judgment in question is the conclusion.) In doing this, reason is acting on its maxim:

When you have an item of knowledge, find something more general of which it is a special case, or, to say the same thing in more technical terms,

- When you have an item of knowledge, find the condition by which it is conditioned.

Now, the major premise of any inference-of-reason also falls within the scope of reason's seek-the-condition maxim,

which means that reason tells us to look for something still more general from which *it* follows. That involves going from the inference—

P₀, Q, therefore R—

to a prior inference of reason whose conclusion is P₀ and whose major premise is some proposition P₋₁; and from that to a still earlier one whose conclusion is P₋₁, and so on backwards and upwards. All of this happens in accordance with reason's principle—its very own principle—

Given any conditioned item of knowledge obtained through the understanding, find the unconditioned whereby the understanding can be completely unified.

We may want to treat this logical maxim as a principle of pure reason—i.e. to regard it not merely as a command that tells us what to do, but as a statement saying that something is the case—namely that for everything conditioned *there is* a condition. This would involve us in assuming that if something conditioned is given, the whole series of conditions. . . is likewise given, i.e. is contained in the object and its connection. (Notice that if this whole series exists, it is itself *unconditioned*.)

Such a principle of pure reason is obviously *synthetic*; something that's conditioned is analytically related to some condition but not to the unconditioned. And other synthetic propositions must follow from it—propositions of which pure understanding knows nothing, because it deals only with items that are conditioned. [These days we might say that *the unconditioned* doesn't appear on the understanding's radar screen.] If there actually is anything unconditioned, we'll have to pay special attention to all the features of it that distinguish it from everything that is conditioned, and they'll provide the raw material for many synthetic *a priori* propositions.

Any principles arising from this supreme principle of pure reason will be *transcendent* in relation to all appearances, i.e.

there can't be any adequate empirical use for the supreme principle. (So it will be entirely different from all principles of understanding, because their use is wholly *immanent*—they don't transcend experience, because their only theme is the possibility of experience.) Consider the principle:

•The series of conditions extends to the unconditioned.

(This might be offered either as telling us a truth about what there is in the world out there or as making a demand about how we are to behave when we think about the world out there.) Does this principle have objective applicability? What does it imply concerning the empirical use of the understanding? Or is there no such objectively valid *principle* of reason, but only a logical *command* that instructs us to advance towards completeness by working our way up to ever higher conditions, thereby giving our knowledge the greatest possible unity of reason? Might it be the case that this *command* should never have been viewed as a transcendent *principle* of pure reason, and that we went too fast when we postulated that the objects out there in the world actually include an unrestrictedly complete series of conditions? And if that is how things stand, what *other* misunderstandings and delusions may have crept into the inferences of reason that we conduct? . . . Answering these questions will be my task in the transcendental dialectic, which I'm aiming to develop from its deeply concealed sources in human reason. I'll shall divide the Dialectic into: •Book 1 [pages 163–173] on the transcendent *concepts* of pure reason, •Book 2 [pages 174–291] on its transcendent and dialectical *inferences*, and •an Appendix [page 292–315].

366

Book 1:

The concepts of pure reason

In advance of settling whether there can be any concepts derived from pure reason, we know this much: if we can get them, it will be through inferences and not through reflection. Concepts of understanding are also thought *a priori*,
 367 antecedently to experience and for the sake of experience; but all they give is the unity of reflection on appearances that have to belong to a possible empirical consciousness. Those concepts make it possible for us to have knowledge and to settle facts about the objective world. And they don't come to us through inferences; there are indeed two reasons why they couldn't possibly do so. (1) They first provide the material required for making inferences, so we can't do any inferring until we *have* those concepts. (2) There's nothing we could infer them *from*: they aren't preceded by any *a priori* concepts of objects from which they could be inferred. Their objective reality isn't based on anything inferential; its sole basis is the fact that they constitute the intellectual form of all experience; so it must always be possible to show their application in experience.

The label 'concept of reason' tells us from the outset that we're dealing here with something that can't be confined within experience, because it concerns a body of knowledge of which any empirical knowledge is only a part—indeed it may be that the whole of possible experience. . . . is only a part of it. No actual experience is ever completely adequate to it, yet every actual experience belongs to it. Concepts of reason enable us to *conceive*, and concepts of understanding enable us to *understand*. . . . Just as I have labelled the pure concepts of understanding 'categories', so I shall give the concepts of pure reason a new name, calling them

transcendental 'ideas'. I'll now explain and justify this label.

1. The ideas in general

Despite the great wealth of our languages, a thinker often finds himself at a loss for the expression that exactly fits his concept, and this lack prevents him from being really intelligible to others or even to himself. He could coin a new word, but that amounts to claiming to *legislate* for
 369 language—and you can't often get away with that! Before trying that way out—which is always a long shot—we should scout around in a dead learned language, to see if *it* provides both the concept and a suitable word for it. Even if those who first launched the word were a bit careless, so that the use of it was somewhat wobbly even back then, it's always better to latch onto the meaning that distinctively belongs to it (whether or not we're sure that it was originally used in precisely this sense) than to defeat our purpose by making ourselves unintelligible.

When we want to distinguish a certain concept from related ones, and there's just one word whose existing meaning exactly fits it, we would be wise to use that word sparingly—keeping it to its own proper meaning and not also using it, for stylistic variety, as a synonym for other expressions. Otherwise we may stop focussing intently on that word, mixing it up with lots of other words whose meanings are quite different; and if that happens, we'll lose also the
 370 thought that the word expresses and could have preserved.

From the way Plato used the term 'idea' we can see that he meant it to stand for something that not only couldn't be borrowed from the senses but even extends far beyond the concepts of understanding (which Aristotle was busy with)—because nothing that fits it can ever be encountered in experience. Plato held that ideas are archetypes of the

things themselves [= 'models from which things are copied'], unlike the categories, which are merely keys to possible experiences. In his view, ideas issued from highest reason, through which human reason comes to share in them; but our reason is no longer in its original state, and has to strain to recall the old now-obscure ideas, by a process of recollection (which is called philosophy). I'm not going to conduct a textual enquiry into what this great philosopher meant by 'idea'. I merely remark that it isn't at all unusual to find. . . . that we understand an author better than he has understood himself. Not having pinned down his concept exactly enough, an author's intention is sometimes belied by what he has said, or even by what he has thought.

371 Plato knew very well that •our faculty of knowledge feels a need for something much higher than merely spelling out appearances according to a synthetic unity so as to be able to read them as experience; and •that our reason naturally soars to items of knowledge that have to be recognised as having their own reality rather than being mere fictions of the brain, despite the fact that they go far beyond the bounds of experience—so far that no empirical object can ever fit them.

Plato found the chief instances of his ideas in the field of the practical [here = 'moral'], i.e. in what rests on freedom, which is the subject of items of knowledge that are produced only by reason.² If you try to derive the concepts of virtue from experience. . . . you'll turn virtue into something that varies according to time and circumstance, a slippery non-entity that can't be brought under any rule. We're all well

² It's true that he also extended his concept of *idea* so as to cover items of speculative knowledge, provided that they were pure and given completely *a priori*. He even extended it to mathematics, although what that science is about can be found only in possible experience. I can't go with him down that road. . . .

aware that the truth is nothing like that. We know that if someone is held up as a 'perfect example of virtue', we judge this by comparing the person, the alleged 'perfect sample', with the true original that we have in our minds. This original is the *idea of virtue*. Objects of experience can serve as 'approximate' examples of it (showing that proofs that what the concept of reason commands is at least somewhat feasible), but they can't serve as 'perfect' archetypes of virtue. [This uses the word (*Urbild*) that was translated as 'archetype' two paragraphs back; but here, and from now on, Kant thinks of an *Urbild* not as a model *from which other things are copied*, but rather as a model or ideal example *to which we may approximate*.] None of us will ever act in a way that matches up to what is contained in the pure idea of virtue, but that doesn't prove that there's something chimerical about this thought. It's only by means of this idea that we can make any judgment as to moral worth or unworth; so the idea serves as an indispensable basis for any approach to moral perfection—even if we don't get very close because of the obstacles in human nature. . . .

Plato's *Republic* has become proverbial as a striking example of the kind of dreamy perfection that could only exist in an idle thinker's brain, and he has been ridiculed for claiming •that a monarch can't rule well unless he participates in the ideas. [Here as everywhere in this half of the *Critique*, 'idea' is used only as a Platonic or Kantian technical term.] We would, however, be better advised to run with •this thought, and, where the great philosopher leaves us without help, to shed light on it through our own efforts rather than discarding it on the wretched and harmful pretext that it isn't practicable.

A constitution providing for *the greatest possible human freedom* under laws that make *the freedom of each consistent with the freedom of everyone else*

—that is, to put it mildly, a *necessary* idea; it must be made basic not only in the initial design of a constitution but also in all its laws. (I state the idea in terms of •freedom, not of •the greatest happiness, for happiness will take care of itself if freedom is assured.) In drafting a constitution, we must initially abstract from the present obstacles. ·Let’s think a little about the nature of these obstacles to successful government·. Rather than being inevitable upshots of human nature, perhaps they arise rather from something that *could* be remedied, namely the neglect of genuine ideas in making laws. Legislators commonly ·excuse their failures by· appealing to ‘adverse experience’—i.e. to contingent circumstances that thwarted their plans·. Actually, nothing could be more harmful or more unworthy of a philosopher than that. The ‘adverse experience’ wouldn’t have occurred if at the right time those institutions had been set up in accordance with ideas, rather than the ideas being displaced by crude conceptions which, just because they were derived from experience, nullified all ·the legislators’ good intentions. The more legislation and government are brought into harmony with the above idea [i.e. the one indented earlier in this paragraph], the less punishment there would be; so it was quite reasonable for Plato to maintain that in a perfectly structured state no punishments would be needed. It may be that this perfect state won’t ever come into being; but that doesn’t

374 stop the idea from being valid. What it does is to set this *maximum*—‘the greatest possible human freedom’—before us as an archetype, something we can move towards, so as to bring the legal organisation of mankind ever nearer to its greatest possible perfection. How far *can* we go along that line? How big a gap *must* there be between the ·archetypal· idea and what we actually achieve? No-one can answer this, and no-one should try, because this is all about freedom, which can pass beyond any specified limit.

Plato saw clear proofs that ideas have an explanatory role not only •in the moral sphere, where human reason exhibits genuine causality so that ideas are working causes of actions and their outputs, but also •in regard to nature itself. A plant, an animal, the orderly arrangement of the cosmos—presumably therefore the entire natural world—clearly show that they are possible only according to ideas. No individual creature coincides ·exactly· with the idea of what is most perfect in its kind; just as no human being coincides ·exactly· with the idea of humanity, though each of us carries that idea in his soul as the archetype of his actions. Despite this, these ideas are completely determinate unchangeable individuals in the Supreme Understanding ·of God·, and they are the ultimate causes of things.

[Kant offers a guarded expression of approval for Plato’s appeal to ideas outside the moral sphere. Then:] But where Plato’s doctrine renders a very special •service is in connection with the principles of morality, legislation, and religion, where the experience (of the good) is itself made possible only by the ideas—incomplete as their empirical expression must always remain. This •service hasn’t been recognized, because it has been judged in accordance with empirical rules—the very things that Plato’s approach has shown can’t validly be treated as principles ·in our moral thinking·. When we are studying nature, experience supplies the rules and is the source of truth; but when it comes to the moral laws, experience is (alas!) the mother of illusion! It is *very* bad behaviour to derive laws prescribing what *I ought to do* from what *is done*, or to limit laws on that basis. .375

Following out these considerations is what gives philosophy its own special dignity; but just now we must occupy ourselves with a less grand but still worthwhile task, namely levelling the ground and making it firm enough to support these majestic moral edifices. ·Why does it need to be *made* 376

firm?· Because this ground has been honeycombed by subterranean workings that reason, in its confident but fruitless search for hidden treasures, has carried out in all directions. What we have to do now is to get some insight into the transcendental use of pure reason, its principles and ideas, so that we can be in a position •to get the facts about what influence pure reason has and •to make a judgment as to its value. [Kant pleads with the philosophically serious reader to use 'idea' only in its original meaning rather than using it as a label for 'any and every species of representation'. There are plenty of terms for each kind of representation, he says, and he gives a list—a 'chart'—of terms with their definitions.] [•In this version, each bold-type item is the one that re-appears at the next level up. •In this one case *Erkenntnis* is translated as 'cognition', because the generally preferred 'item of knowledge' sounds too peculiar. See note on pages 2–3. •Despite its prominence here, this is the last we hear of 'notion' as a technical term.]

Bottom level:

The genus is 'representation'.

When this is accompanied by consciousness it is **perception**.

Second level:

Perception considered merely as a state of the person is 'sensation'.

Perception considered as perception *of* something is **cognition**.

Third level:

A cognition relating directly to an individual object is an 'intuition'.

A cognition relating indirectly to objects, through features that many objects may share, is a **concept**

Fourth level:

Empirical concepts.

Pure concepts.

Fifth level:

·Pure concepts can be schematised, i.e. amplified by something sensible·.

A pure concept originating solely in the understanding, with no input from sensibility, is a **notion**.

·And so at last we rise to our present topic, which involves 'notion' but seems not to come from any two-part division of notions, namely:

Sixth level·:

A concept that is formed from notions and outruns the possibility of experience is an **idea**.

Anyone who has familiarised himself with these distinctions must wince when he hears the representation of the colour *red* called an 'idea'. It oughtn't even to be called a concept of understanding, a notion.

2. The transcendental ideas

The Transcendental Analytic gave us an example of how the mere logical form of our knowledge can give rise to pure *a priori* concepts which represent objects prior to all experience. (Strictly speaking, rather than •representing objects they •indicate the synthetic unity without which we couldn't have empirical knowledge of objects.) The different forms of judgment. . . .generated categories that direct all our use of understanding in experience. In the same way we can expect that the different forms of inferences of reason. . . .will generate special *a priori* concepts (we can call them 'pure concepts of reason' or 'transcendental ideas') which will determine how understanding is used in dealing with experience as a totality.

The function of reason in its inferences is to give ·greater· universality to items of knowledge. . . . Consider the proposition, 'Caius is mortal'. I could get this from experience

378

by means of the understanding alone, leaving reason out of it. But I am after something more general; I'm looking for a concept (in this case, the concept *man*) that contains the condition under which the predicate. . . of this judgment ('is mortal') is given; and after I have brought the predicate under this condition taken across its whole range ('All men are mortal'), I proceed on that basis to settle on the item of knowledge about my object ('Caius is mortal').

379 Accordingly, in the conclusion of an inference of reason we restrict a predicate to a certain object, having first thought it in the major premise under a given condition taken across the whole range of that condition. This fact about the size of the range is called *universality* or *totality*. . . . So the transcendental concept of reason is nothing but

- the concept of the *totality* of the *conditions* for any given conditioned item.

What makes possible the totality of conditions is the *unconditioned*, and conversely the totality of conditions is always itself unconditioned. [This use of 'conversely' here suggests that Kant meant to say that something involves a totality of conditions *if and only if* it involves something unconditioned. But that isn't what he actually says.] So we can give a general explanation of what a pure concept of reason is—i.e. what an idea is—by saying that it's a concept of something unconditioned, when the concept is thought of as a basis for the synthesis of the conditioned. That means: as a basis for a process of connecting conditioned items with one another; for example,

- discovering a causal chain among certain events would be
- conducting a synthesis of (causally) conditioned items, and similarly with the other relevant relations. There will be exactly as many •pure concepts of reason as there are •kinds of relation that the understanding represents to itself by means of the categories. There are just three of these,

expressed by **(1)** 'S is M', **(2)** 'If P then Q', and **(3)** 'R or S'. So we have to look for three kinds of unconditioned item: **(1)** the *categorical* synthesis in a *subject*; **(2)** the *hypothetical* synthesis of the members of a *series*; **(3)** the *disjunctive* synthesis of the parts in a *system*.

So there are exactly three kinds of inference of reason, each of which moves up through prosyllogisms to the relevant unconditioned item: **(1)** to the subject that is never itself a predicate; **(2)** to the presupposition that doesn't presupposes anything further; **(3)** to an aggregate of the members of the division of a concept such that nothing further is needed to complete the division. So the pure concepts of reason—

- concepts of totality in the synthesis of conditions, i.e. concepts of *going the whole way* in looking for a condition for every conditioned item—

—are necessary at least as *setting us the task* of extending the unity of understanding, where possible, right up to the unconditioned. They are based on the nature of human reason, which is essentially committed to the demand for conditions. It may be that there isn't anything for these transcendental concepts actually to apply to; in which case the only good they do is to direct the understanding in such a way that when it is extended to the uttermost it is completely free of inconsistency.

THE RIGHT WAY TO USE 'ABSOLUTE'

While I'm dealing with 'the totality of conditions' and 'the unconditioned' as equivalent labels for all concepts of reason, I come on another expression (as well as 'idea') that I can't do without but can't safely use, because long-standing misuse has made it ambiguous. The word is 'absolute'. Like just a few others, this word in its original meaning was fitted to a concept that no other word in the language exactly suits. So if the word is lost, or if (same thing) it is used

381 with several different meanings, •the concept itself will be lost too. This is a concept that reason is very busy with, and giving it up would do great harm to all judgments in transcendental philosophy. **(1)** The word ‘absolute’ is now often used merely to indicate that something is true of a thing considered *in itself*, and therefore true of its *inward* nature; in this sense, ‘x is absolutely possible’ means that x is in itself possible—which is the *least* that could be said about it. **(2)** But the word is also sometimes used to indicate that something holds true in all respects, without limitation (e.g. absolute despotism), and in this sense ‘x is absolutely possible would mean that x is *in every relation* (in all respects) *possible*—which is the *most* that can be said of x’s possibility. From here on, though I shall be *discussing* both senses of ‘absolute’, I shall *use* the word only in sense **(2)**, reserving ‘intrinsic’ for sense **(1)**. Sometimes a statement is true in both senses of ‘absolute’: if something is **(1)** intrinsically impossible then it is **(2)** impossible in any relation, and therefore absolutely impossible. But in most cases the two meanings are infinitely far apart: if something is **(1)** in itself possible, we can’t conclude that it is also **(2)** possible in every relation, and thus absolutely possible. We’ll see later on that absolute necessity doesn’t always depend on intrinsic necessity, and therefore shouldn’t be treated as equivalent. If the opposite of something is intrinsically impossible, this opposite is of course impossible in all respects, and the thing itself is therefore absolutely necessary. But we can’t run this inference the other way, arguing that if something is absolutely necessary its opposite is intrinsically impossible, i.e. that the *absolute* necessity of things is an *intrinsic* necessity. . . . The loss of a concept that is of great importance for speculative philosophy must *matter* to you if you are a philosopher. [In Kant’s usage, ‘speculative’ is the opposite of ‘practical’ or ‘moral’; it means ‘having to do with the truth

382

of theories’, and doesn’t carry any of the sense of ‘guesswork’ that the word has today.] I hope, then, that it will matter to you that we should pin down and carefully preserve the word on which the concept depends.

So there it is: I shall use the word ‘absolutely’ in contrast to what holds only comparatively, i.e. in some particular respect; referring to what is valid without restriction in contrast to what is restricted by conditions. [As well as *absolut*, which he is discussing here, Kant often uses *schlechthin*, which means ‘without qualification’. It could often be translated by ‘absolutely’, and in previous translations it often is; but the present version will use ‘absolute(ly)’ only for *absolut*, and translate *schlechthin* by ‘utterly’ or ‘unqualifiedly’ or some such expression. When Kant contrasts **(1)** things that are principle-like in this or that way with **(2)** things that are *schlechthin Prinzipien*, he is translated on page 158 as contrasting **(1)** with ‘principles *period*’. Grossly unhistorical, but it does capture his meaning.]

Now a transcendental concept of reason always aims at absolute totality in the synthesis of conditions, and its only terminus is in what is unqualifiedly unconditioned, i.e. is not conditioned in any respect. For pure reason leaves to the understanding everything that kicks off from the objects of intuition, or rather from the synthesis of such things in the imagination. Reason’s only concern is with absolute totality in the use of the concepts of the understanding; it takes the synthetic unity that is thought in the category and tries to track it up to something unqualifiedly unconditioned. We can call this the *unity of reason* in appearances, and that expressed by the category the *unity of understanding*. Reason isn’t concerned with the understanding considered as containing the ground of possible experience. Why? Because •no experience is unconditioned, so •the concept of the absolute totality of conditions isn’t applicable in any experience, •so •reason has nothing to do or say down at that

level. But reason is concerned with the understanding in another way: it tells the understanding what direction to take towards a certain unity, of which the understanding itself has no concept. *What* unity? It's the unity that would come from uniting all the acts of the understanding, in respect of every object, into an *absolute* whole. The objective use of the pure concepts of •reason is, therefore, always *transcendent*, while that of the pure concepts of •understanding must always be *immanent*, because the only way to use them is in application to possible experience.

[Kant now has a paragraph in which he repeats what he has already said •about the 'transcendental' nature of pure concepts of reason, •about their role as direction-setters for the understanding, and •about the ideas of *practical* reason as having a larger and more direct role in human life than do those of *speculative* reason, the latter being our concern in this book. Then:]

In saying (as we must) that the transcendental concepts of reason are *only ideas*, we aren't taking them to be superfluous and empty. Although they can't latch onto any object, they can in a basic and unnoticed way be useful to the understanding as a canon for its extended and consistent use [re 'canon', see note on page 25]. What this provides for the understanding is not

•more knowledge than it would have by means of its own concepts •unguided by reason•,

but rather

•better and more extensive guidance for the acquiring of knowledge.

386 Not to mention the fact that concepts of reason may enable us to move across from thoughts about •nature to thoughts about •morality. . . . I'll deal with that in a later work. In *this* work our concern is. . . .only with reason in its speculative use—and indeed, more narrowly, with its *transcendental*

speculative use. Let's take a tip from our procedure in the deduction of the categories, by considering the *logical form* of knowledge through reason. . . .

Reason. . . is the faculty of *inferring*, i.e. judging mediately (by bringing the condition of a possible judgment under the condition of a given judgment). The given judgment is the universal rule (major premise)

•All men are mortal•.

What brings the condition of another possible judgment under the condition of the rule is the minor premise

•Caius is a man•.

The judgment which applies the predicate of the rule (•'mortal'•) to the *brought-under* case •of Caius• is the conclusion

•Caius is mortal.

387

The rule says of some predicate that it applies to everything that satisfies a certain condition. That condition (•mortality•) is found to be satisfied in an actual case (•Caius•). What has been asserted to be universally valid under that condition is therefore to be regarded as valid also in the present case, which satisfies that condition. It's easy to see what is happening here: reason is arriving at an item of knowledge through acts of the understanding that constitute a series of conditions. Here is an example, concerning my way of arriving at the proposition that **(3)** *All bodies are alterable*. I start from the proposition that

(1) Everything composite is alterable.

This item of knowledge is quite distant from **(3)**; it doesn't involve the concept of body, though it does involve the condition of that concept, alterability. I then proceed from **(1)** to a proposition that is less remote from **(3)**, and stands under the condition of **(1)**, namely the proposition that

(2) •All• bodies are composite.

From this I finally pass to

(3) •All• bodies are alterable,

which connects the more distant item of knowledge (alterable) with the knowledge actually before me. By this procedure I have arrived at an item of knowledge (a conclusion) by means of a series of conditions (the premises). [In one fiercely compressed sentence, Kant says things that can fairly be spelled out as follows: An inference of reason in which we pass from conditions (in the premises) to something conditioned (in the conclusion) can sometimes be part of a longer series of conditions-to-conditioned inferences, going in either direction.

- In one direction, the longer series takes the conditions in the premises of the original inference and provide conditions of *them*, and then conditions of those conditions, and so on upwards.
- In the other direction, the longer series takes the conditioned item in the conclusion of the original inference and provide items of which *it* is a condition, and then items of which those items in turn are conditions, and so on downwards.

This can't happen with disjunctive inferences of reason, but it can happen with either of the other two forms of inferences of reason—categorical (= subject-predicate) and hypothetical (= if-then). The most natural kind of example (Kant doesn't give any) of the **hypothetical** form of inference takes the use of the hypothetical 'If . . . then' to express facts about what *causes* what. We explain the fact that •Q by putting together our knowledge that •P's being the case would cause Q to be the case and our knowledge that •P. Then we can move upward into the fact that P is caused to be the case by O, which is caused to be the case by N, and so on back up the causal chain; or downwards into the fact that Q causes R to be the case, which causes S to be the case, and so on down the causal chain. Examples of an elongated inference of reason that has the **categorical** form are harder

to provide, or even to describe; they will be returned to [page 172]. The **disjunctive** form doesn't come into this because a disjunction doesn't have a direction.]

..388

But we soon become aware that how the faculty of reason works in the •ascending series of inferences of reason, in which we infer items of knowledge by looking at

•conditions as being conditioned in their turn, is quite different from how it behaves in the •descending series, in which we look at

•conditioned items as being conditions in their turn. In the ascending inference the item of knowledge is given only as conditioned; to arrive at it by means of reason we have to assume that all the members of ·the ascending series·, the series on the side of the conditions, are given—the crucial point being that we have to think of that entire series as *already complete* In the ·descending· series, the one on the side of the conditioned, the one that looks at consequences, our only thought is of a series *in process of coming into existence*, not one already presupposed or given *in its completeness* Thus, if an item of knowledge is viewed as conditioned, reason is forced to regard the series of conditions in the ascending line as completed and as given in its totality. But if the same item of knowledge is viewed as a condition of further items of knowledge that constitute a series of consequences in a descending series, reason doesn't care •how far this downward series extends, or •whether a totality of the series is possible. That's because reason doesn't need any such series in order to draw its conclusion. [Kant's development of this point is expressed rather technically. What it comes down to, expressed here (though not by him) purely in terms of the causal kind of hypothetical inference of reason, is this:] Reason is compelled to regard any present event as the upshot of *all its causes*; without knowing whether that series has a first member (an

389

uncaused cause) or rather stretches back to infinity, with no first member, it has to regard the event as having such a totality of causes in its ancestry; the proposition reporting this one event can't be counted as true unless the entire series of its causes is unconditionally true. (This holds even if it is admitted that we can't possibly grasp a totality of conditions.) Reason *requires* this, by announcing that its knowledge is *a priori* determined as necessary, either •in itself (in which case it needs no grounds) or else •derivatively as a member of a series of grounds—a series which is, taken as a whole, unconditionally true.

3. System of the transcendental ideas

390

Our topic is not *logical* dialectic, which ignores the •content of knowledge and confines itself to exposing the fallacies concealed in the •form of inferences of reason. Rather, it is a *transcendental* dialectic that has to contain, completely *a priori*, the origin of certain items of knowledge derived from pure reason as well as of certain inferred concepts whose objects •can't ever be given empirically and therefore •lie wholly beyond the reach of the faculty of pure understanding. The transcendental use of our knowledge, both in inferences and in judgments, has a natural relationship to its logical use; and this relation has shown us •that there can be only three kinds of dialectical inference of reason, corresponding to the three kinds of inference through which reason can arrive at knowledge by means of principles, and •that in all of these its business is to ascend from the conditioned synthesis to the unconditioned—i.e. from something to which the understanding always remains restricted to something that the understanding can never reach.

[Kant now presents an obscure account of three kinds of relation that can be involved in a representation, from

which he derives a three-part relation-based classification (r-bc) of concepts of pure reason (i.e. transcendental ideas). In the paragraph after this one he will say that this r-bc coincides with the classification he has already presented on the basis of •logical form, namely the division into **(1)** categorical, **(2)** hypothetical, and **(3)** disjunctive. What really matters is not the r-bc itself but rather a classification that Kant supposedly derives from it, namely:]

- (1)** ideas containing the absolute (unconditioned) *unity* of the *thinking subject*,
- (2)** ideas containing the absolute *unity of the series of conditions of appearance*, and
- (3)** ideas containing the absolute *unity of the condition of all objects of thought in general*.

The thinking subject is what **(1)** *psychology* is about, the sum-total of all appearances (the world) is what **(2)** *cosmology* is about, and the thing that contains the highest condition of the possibility of all that can be thought (the Being of all beings) is what **(3)** *theology* is about. Thus, pure reason provides the ideas for **(1)** a transcendental doctrine of the **soul**, a rational psychology, **(2)** a transcendental science of the **world**, a rational cosmology, and **(3)** a transcendental knowledge of **God**, a rational theology. The understanding can't produce even a sketch of any of these sciences—even when it is supported by the highest logical use of reason, i.e. by all possible inferences through which we aim to move from given appearances right up to the most remote members of the empirical synthesis. Each of these sciences is an entirely pure and genuine product of pure reason—or *problem* of pure reason!

392

How, exactly, do the pure concepts of reason come under these three headings? I'll answer that fully in the next chapter, where we'll see that they follow the guiding-thread of the categories. If you are wondering how the categories,

which are concepts of the •understanding, come into this story about our concepts of •reason, I'll point out here that pure reason latches directly not onto *objects* but onto the understanding's *concepts of objects*. •So much for the general point, but what about the details? I shall contend that

- (1) reason, simply by the synthetic use of that very function of which it makes use in categorical inferences of reason, is necessarily brought to the concept of the absolute unity of the *thinking subject*; that
- (2) the logical procedure used in hypothetical inferences of reason leads to the ideal of the utterly unconditioned *in a series* of given conditions, and finally that
- 389 (3) the mere form of the disjunctive inference of reason must necessarily involve the highest concept of reason, that of a *Being of all beings*—a thought that, at first sight, seems utterly paradoxical.

When I complete my account •in the next chapter•, I shall make clear *how* all that can be the case.

Strictly speaking, there can't be an *objective deduction* of these transcendental ideas, like the one I gave for the categories, because the ideas—just because they *are* ideas—don't relate to any object in such a way that they could be (•or, for that matter, fail to be•) true of it. But a subjective derivation of them from the nature of our reason can be given, and in this chapter I have given it.

[We are about to meet three technical terms that have to be understood:

- (1) 'inherence',
- (2) 'dependence',
- (3) 'concurrence'.

In (1) a categorical or subject-predicate proposition, some property is said to *inhere* in a subject—e.g. mortality inheres in Caius. In (2) a hypothetical proposition something is said to *depend* on something else—e.g. the ball's starting move depends on its having been hit. In (3) a disjunctive proposition, two or more possibilities are said to divide the

whole range of possibilities amongst them; rather than some being made subordinate to others, they are all treated as on a level, as somehow going together or concurring. You'll recognize that this is just the same 1-2-3 that we have been dealing with in the past few pages. This note makes a feeble job of relating disjunction to 'concurrence', but the blame for that may lie with Kant.] It's easy to see that what pure reason has in view is the absolute totality of the synthesis *on the side of the conditions* (whether of inherence, of dependence, or of concurrence); it isn't concerned with absolute completeness *on the side of the conditioned*. It's only the former that is needed in order to presuppose the whole series of the conditions and present it *a priori* to the understanding. Given a complete (and unconditioned) condition, we don't need any concept of reason for the continuation of the series: every step in the downward direction from condition to *conditioned*—from 394 conditions to what they are conditions *of*—is taken by the understanding itself. The transcendental ideas, therefore, serve only for going *up* the series of conditions to the unconditioned, i.e. to principles. As regards the intellectual journey *down* from conditions to the conditioned, reason does indeed make a very extensive logical use of the laws of understanding, but it's not a transcendental use. If we form an idea of the absolute totality of a synthesis in a downward series—e.g. an idea of the whole series of all *future* alterations in the world—this is a mental entity that we have *chosen* to create, not something we are *forced* to presuppose by the nature of our reason. . . .

Finally, we also come to realize that the transcendental ideas themselves hang together to form a certain unity, and that it's by means of them that pure reason draws all its items of knowledge together to form a system. The advance from (1) the knowledge of oneself (the soul) to (2) the knowledge of the world, and by means of this to (3) the primordial being, •God•, is so natural that it seems

395 to resemble reason's *logical* advance from premises to a conclusion.³

[The phrase 'primordial being' translates the German *Urwesen*. The prefix *Ur-* is used to convey the idea of something that is the basic source of x, the fundamental origin of all the Fs, or the like. (English has no such resource except in words openly borrowed German—e.g. such English words as 'urkingdom' and 'urtext'.) Some Kant translators use 'original being'; but 'original' doesn't colloquially carry the weight and solemnity of *Ur-*. Thus, 'primordial', here and throughout; with apologies, and this explanation.]

³ Metaphysics has only three ideas as the proper objects of its enquiries: *God*, *freedom*, and *immortality*—so related that the combination of *God* with *freedom* leads inevitably to *immortality*. Any other matters that metaphysics may deal with are merely means of arriving at these ·three· ideas and of establishing their reality. Reason needs the ideas not for the purposes of natural science but in order to pass beyond nature. Insight into them would put the faculty of speculative reason in sole charge of *theology* and *morals*, and, through the union of these two, likewise *religion*—which means that it has sole charge of the highest ends of our existence. [Kant is now going to use 'analytic' and 'synthetic' in a way that was quite standard in his day but is entirely different from the senses he has given these words up to here. In the present sense, 'analytic' and 'synthetic' are labels for two *methods of presentation* of some doctrine. An •analytic presentation starts with things we all know to be true and works its way from those to the theory or doctrine that explains and is supported by them. A •synthetic presentation goes in the opposite direction: it starts with the fundamental theses of the doctrine to be expounded, and works from those to various of their consequences, which could include the things-we-already-know that are the *starting-point* for the analytic format.] In a •systematic presentation of the ideas, the synthetic order would be more suitable; but before we get to that there has to be a •preliminary working-through of the materials, and for that the analytic order—the reverse of the synthetic—is better. It lets us start from what is immediately given us in experience—advancing from the doctrine of the *soul* to the doctrine of the *world* and from that to knowledge of *God*.

Book 2:

The dialectical inferences of pure reason

..396 The object of a purely transcendental idea can be said to be something of which we have no concept, despite the idea's being something that reason is compelled by its own inherent nature to produce. That 'can be said', and it's true: Of an object that satisfies the demands of reason it is indeed impossible for us ever to form a concept of the understanding, i.e. a concept that could be exhibited and intuited in experience. Still, it might be better, and less likely
397 to mislead, if we said instead that although we can't have any knowledge of the object that corresponds to such an idea, we do nevertheless have a *problematic* concept of it.

The transcendental (subjective) reality of the pure concepts of reason depends on our having been led to such ideas by a necessary inference of reason. There will therefore be inferences of reason, having no empirical premises, through which we infer from something we know something else of which we have no concept though an inevitable illusion leads us to regard it as objectively real. Because of the conclusions they come to, these movements of the mind would be better called 'sophistical' [*vernünftelnde*] rather than 'inferences of reason' [*Vernunftschlüsse*; note the similarity of the two words—one might translate the former as 'fooling around with reason'], though their origin gives them some claim to the latter title, since they aren't fictitious and have arisen not by chance but from the very nature of reason. They are sophistries [*Sophistikationen*] not of men but of pure reason itself, and not even the wisest of men can free himself from them. If he works hard at it he may be able to guard himself against actual error; but he'll never be able to free himself from the illusion, which unceasingly mocks and teases him.

So there we have it: there are exactly three kinds of dialectical inferences of reason—just as there are three ideas in which their conclusions result.

(1) I call the first kind of inference of reason the transcendental *paralogism*. In it I conclude from the transcendental concept of the subject, which contains nothing manifold, the absolute unity of this subject itself, though in doing this I have no concept whatsoever of this subject. [Kant will explain this later. *Very* briefly and sketchily, the thought is this: The transcendental concept of *myself* is what's involved in every thought I have of the sort 'I now experience x', 'I now think about y'. It is 'transcendental' in the sense that it isn't the concept of thinking-being-with-such-and-such-characteristics; I can attribute to myself various properties, but when I do that, the transcendental concept is the concept of the *I* that does the attributing, not the *I* to which the properties are attributed. In that sense, then, my transcendental concept of myself doesn't reflect *any* of my complexity, i.e. 'contains nothing manifold'. And I commit a paralogism = invalid-inference of pure reason when I go from that •premise about the total uncomplexity of the transcendental *I* to a •conclusion about my not being in any way complex.]

(2) I shall call the state reason is in when conducting the second kind of sophistical inference the *antinomy* of pure reason. It involves the transcendental concept of the absolute totality of the series of conditions for any given appearance—e.g. the series of *all* the causes of a given event. [Note that whereas 'paralogism' is a label for a certain kind of *inference* that reason conducts, 'antinomy' here is the name of the *state that reason is in* when it conducts a certain kind of inference—a state of **conflictedness**, in which has two conflicting but equally bad ways of looking at something. Kant switches to calling individual pairs of conflicting propositions 'antinomies' = **conflicts** only when he gets to 'Comment on the first antinomy' on page 215.] When I think about my concept of the unconditioned synthetic unity of the series in one of the two ways, I find the concept to be

398

self-contradictory, so I take it in the other way, inferring that *that* is the truth of the matter, though in fact I have no ·acceptable· concept of that either. [A rough, quick example: When I try to think about all-the-causes-of-event-E on the assumption that every one of those causes also had a cause (so that the chain of them had no first member), I get into intellectual trouble; so I rush to the conclusion that some causes were not themselves *caused*, but were rather exercises of *freedom*; and that turns out to be intellectually problematic too.]

(3) Finally, in the third kind of sophistical inference, from •the totality of conditions for thinking of objects as such that I could be confronted with I infer •the absolute synthetic unity of all the conditions for things to be possible. That is, from things that I don't know (because I have merely a transcendental concept of them) I infer a *Being of all beings*, which I know even less through any transcendental concept, and of whose unconditioned necessity I can form no concept whatsoever. I'll label this dialectical inference of reason the *ideal* of pure reason.

Chapter I

The paralogisms of pure reason (1st edition)

399

A •logical paralogism is an inference of reason that is fallacious in form, whatever its content is. It counts as a •transcendental paralogism if there's a transcendental basis for the formal fallacy. A fallacy of this sort is based on the nature of human reason; the illusion it gives rise to can't be avoided, though it may be rendered harmless.

A concept that wasn't included in the general list of transcendental concepts must yet be counted as belonging to that list. I'm talking about the concept (or the judgment, if

you like) 'I think'. It's easy to see that this is the vehicle of all concepts: ·the only way for the concept C to come before me or enter into my scheme of things is for it to be the case that *I think* C·; and that includes transcendental concepts. So *I think* must itself count as transcendental. But it can't have any special label, because all it does is to bring forward, as belonging to consciousness, *any* thought that one has; and that's why its omission from the initial list doesn't mean that the list was defective. Although it's not an empirical concept, it belongs on one side of a certain distinction that can be drawn empirically: the distinction between •myself considered as a thinking being, a soul, an object of inner sense, and •myself as a body, an object of outer sense. ·Obviously, the transcendental *I* belongs on the mental/soul/inner side of that divide·. I label as the 'rational doctrine of the soul' the kind of psychology whose subject-matter is expressed purely through the transcendental concept *I*. It is 'rational'—in the sense of having-to-do-only-with-reason·—because in it I don't try to learn anything about the soul from experience. In the •empirical doctrine of the soul I appeal to experience ·through inner sense·, and get specific detailed information about my soul; but in the rational doctrine of the soul I let all those details go, set aside *all* empirical input, and restrict myself to what I can learn about my soul considered just as something that is present in all thought.

400

So we have here something purporting to be a science built on the single proposition *I think*. How good are the grounds for thinking that there *is* such a science? That's the question we have to address now. You might want to object:

'The proposition *I think*, which expresses the perception of oneself, contains an inner experience. So the ·supposedly· rational doctrine of the soul built on this proposition is never pure—it is always to that extent based on an empirical principle.'

401

[Kant replies, at unhelpful length, that *this* ‘inner perception’ involves no details, doesn’t serve to mark off oneself from other things, and is simply a necessary accompaniment of *all* thought and experience; so that it shouldn’t be regarded as empirical knowledge. Then:] If to this all-purpose representation of self-consciousness we added the slightest object of perception (even if it’s only pleasure or unpleasure), that would immediately transform •rational psychology into •empirical psychology.

Thus, *I think* is rational psychology’s sole text, from which its whole teaching has to be developed. Obviously, if this thought is to be *about* something (myself), it can involve only transcendental predicates of that something, since the slightest empirical predicate would destroy this science’s rational purity, its independence from all experience.

What we have to do here is follow the guidance of the categories, with just one difference. In the transcendental logic I have always taken the categories in the order

•quantity, quality, relation, modality;

and I stand by that ordering considered as an aspect of the theory of categories, but in our present context I have to vary it by adopting the order

•(1) relation, (2) quality, (3) quantity, (4) modality.

That’s because our starting-point here is a given *thing*—*I* as a thinking being—so we must start with the category of *substance* (which is one of the categories of relation). Starting from there, we’ll be going through other classes of categories in reverse order [not strictly true!]. Thus, the topic [= ‘logical geography’] of the rational doctrine of the soul, from which everything else that it contains must be derived, is this:

(1)

The soul is *substance*

(2)

In quality it is *simple*

(3)

Through the different times when it exists, it is *one*, i.e. unity and not plurality

(4)

It relates to *possible* objects in space

All the concepts of pure psychology can be assembled out of these elements, with no other source being called upon. Here is how: 403

- this substance, merely as an object of inner sense, yields the concept of *immateriality*;
- as simple substance, it yields the concept of *incorruptibility* [here = ‘indestructibility’],
- its being a thinking substance that lasts through time yields the concept of *personhood*;
- all three of those combine to yield the concept of *spirituality*; and
- the substance’s relation to objects in space yields the concept of *causal interplay with bodies*, which in turn
- leads us to represent the thinking substance as the source of life in matter, i.e. as soul (*anima*), and as the basis of *animality*. Finally,
- animality, when combined with spirituality, yields the concept of *immortality*.

Out of all this there arise four paralogisms of a transcendental psychology that is wrongly regarded as a body of knowledge about the nature of our thinking being—knowledge that we acquire through pure reason. The only basis we can find for it is the simple, intrinsically empty representation *I*; 404

and this doesn't even qualify as a concept; it's merely a bare consciousness that accompanies all concepts. All that is represented through this *I* or *he* or *it* that thinks is a transcendental subject of thoughts = *x*. [In adding '= *x*' Kant wants to convey that this item is characterless, empty, a sort of place-holder, rather than something with a describable character of its own.] It is known only through the thoughts that are its predicates; apart from them we can't have any concept of it; any attempt that I make to characterize my transcendental *I* will use my representation of it in thoughts of the type: 'I conclude/think/see/believe/suspect/know that I am F'—where the first 'I' is the transcendental one—so that the attempt to describe it must revolve in a perpetual circle. There's no escape from this, because consciousness as such isn't a representation that picks out one object as distinct from others; rather, it is a form of representation in general. . . .

405 It must at first seem strange that something that is a pre-condition for *my* thinking—i.e. something that is merely a property of myself as a thinking subject—also holds for everything that thinks. That is the strangeness of the thesis that we can use a seemingly empirical proposition as the basis for a necessary and universal judgment, namely the judgment that anything thinks must be constituted in the way that the voice of self-consciousness declares that I am constituted. But although it is strange, it is also true, and here is why: It is *a priori* necessary that I attribute to a thing all the properties that are preconditions of my having any thought about them. Now, I can't have the slightest representation of a thinking being through any outer experience; I have to get it through inner self-consciousness; which means that I get my thoughts about thinking beings other than myself by transferring my consciousness to them. [Kant's next sentence is long and hard to follow. Its gist

is this: When I want to think about (for example) *you* as a thinking being, and so 'transfer my consciousness' to you, I am not mentally transferring to you any of my individual qualities. The transferable 'I think' that is involved here isn't what Descartes took it to be (when he argued from it to 'I exist'), namely a perception of an existent thing. And the use I am making of it is merely problematic; i.e. I'm using it only to ask some questions—I want to know what can be inferred from such a simple proposition, whether or not its 'I' stands for something that actually exists. Then:]

If our knowledge-from-pure-reason of thinking beings in general were based on

- more than the *cogito*, i.e. the inevitable, always-present, empty 'I think'
- our observations of how our thoughts come and go, and the natural laws of the thinking self that we derived from these observations,

that would give rise to an **empirical** psychology, a theory about the workings of inner sense. Perhaps it could explain the appearances of inner sense; but it couldn't ever reveal properties that don't in any way belong to possible experience (e.g. properties that something has because it is *simple*), or yield any knowledge of absolutely necessary truths about the nature of thinking beings as such. So it wouldn't be a **rational** psychology. 406 A348

Since the proposition 'I think' (taken problematically) contains the form of every single judgment of the understanding, and accompanies all categories as their vehicle, it is obvious that when we draw conclusions from that proposition we must be using our understanding only in a *transcendental* manner. [Why 'understanding' rather than 'reason'? Presumably because these would be inferences from a single premise, whereas Kant defines 'reason' in terms of inferences from two or more premises.] Since using the understanding in this way keeps

out any admixture of experience, and in the light of what I have already shown, we can't have much optimism about what we are going to achieve in this way. Well, let's keep a critical eye open as we follow this procedure through all the basic concepts of pure psychology.

From here until page 197 the material all comes from (A) the first edition of the *Critique*; the second-edition (B) version begins at page 197.

First paralogism: Substantiality

- If our representation of something x is the *absolute subject* of our judgments, so that x can't be used as determination of something else, x is *substance*.
- I, as a thinking being, am the *absolute subject* of all my possible judgments, and this representation of myself can't be used as predicate of anything else.
- Therefore I, as thinking being (soul), am *substance*.

Critique of the first paralogism of pure psychology

In the analytical part of the Transcendental Logic I showed that pure categories—one of which is the concept of substance—
 A349 •have no objective significance except when they are brought to bear on an intuition, and •are applied to the complex web of intuition as unifiers. In the absence of this web, they are merely forms of a judgment, without content. I can say of any thing that because it is a thing it 'is substance', in the sense that I am distinguishing it from mere predicates and states of things. •And from that I get something like the paralogism:

In all our thought, the *I* is the subject, in which thoughts inhere only as states; and this *I* can't be represented as the state of something else. So everyone *must* regard himself as substance, and regard

•his• thinking as merely properties that he has, states that he is in.

But what *use* am I to make of this concept of a substance? I certainly can't infer from it that I as a thinking being *persist* for myself and don't in any natural manner either *arise* or *perish*. But there's no other use I can make of the concept of the substantiality of myself as a thinking subject; if I can't use it to infer my permanence, I can't use it for anything. [Recall that in the Analytic Kant treated permanence, or never-going-out-of-existence, as the essence of the empirically usable category of substance.]

To see how far we are from being able to deduce permanence from the pure category of substance, consider how we have to proceed when we want to use the concept in an empirically useful way: to do this we must, at the outset, have an object that is *given in experience* as permanent. In contrast with that, in •the paralogism's inference from• the proposition *I think* we don't take any experience as our basis; rather, we infer a conclusion merely from the concept of the relation that all thought has to the *I* as the common subject that *has* the thought. . . . The *I* is indeed in all thoughts, but this representation doesn't contain the slightest trace of intuition, distinguishing the *I* from other objects of intuition. So we can indeed perceive that this representation keeps turning up in all thought, but not that it is an abiding intuition of something that continues in existence while its transitory thoughts come and go.

Conclusion: transcendental psychology's first inference of reason, in putting forward •the constant logical subject of thought as being knowledge of •the real subject in which the thought inheres, is palming off on us something that is a mere pretence of new insight. We don't and *can't* have any knowledge of any such subject. It's true that consciousness is needed if our representations are to be *thoughts*, which

A350

implies that we'll encounter our perceptions only in the transcendental subject, .i.e. in the framework provided by 'I think'; but beyond this logical meaning of the *I*, we know nothing about the subject in itself that underlies this *I* as substratum, as it underlies all thoughts. We can allow the proposition '*The soul is substance*' to stand, as long as it's recognised that this concept of the soul as substance doesn't carry us an inch further, and so can't yield us any of the usual deductions of the pseudo-rational doctrine of the soul. . . .i.e. if we recognise that this concept signifies a substance only in idea, not in reality.

A351

Second paralogism: Simplicity

- If something *x* is such that its action can never be regarded as the upshot of several things acting in concert, then *x* is *simple*.
- The soul or the thinking *I* is such a being.
- Therefore, the soul or the thinking *I* is simple.

Critique of the second paralogism of transcendental psychology

This is the Achilles [here = 'the strong man', 'the chief pusher-around'] of all the dialectical inferences in the pure doctrine of the soul. It's not a mere sophistical trick that a dogmatist [see note on page 15] has rigged up to give superficial plausibility to his claims; rather, it's an inference that seems to withstand even the keenest scrutiny and the most scrupulously exact investigation. Here it is .with the details filled in.:

Any *composite* substance *x* is an aggregate of several substances; anything it does (or any property that it has) is an aggregate of several actions (or properties), each belonging to one or other of the several substances. Now an effect *can* be the upshot of the working together of many acting substances (as

A352

the motion of a body is the combined motions of all its parts). There's no difficulty in thinking about such compositeness when it concerns things that are *external* to the mind. But it's different when we come to thoughts—*internal* episodes belonging to a thinking being. For suppose that a thinking thing is composite; then every part of it would contribute a part of its thought, and its *whole* thought would have to come from all of its parts taken together. But this is .covertly. self-contradictory. [From here to the end of this indented passage, this version expands on what Kant wrote, in ways that the .small dots. convention can't easily indicate.] The movement of a composite body is the upshot of movements of all its parts, and they are conceptually unified as a single movement through someone's perceiving the body as a unity. Similarly, the thought of a composite thinker would have to be the upshot of thoughts of its parts; but how are those sub-thoughts to be conceptually united as a single thought? (Must they be so united? Yes. Consider a parallel case: the thought of a line of poetry. I think of hounds while you are thinking of spring, your brother is thinking of winter, and your sister is thinking of traces—but this state of affairs doesn't constitute anyone's thinking 'The hounds of spring are on winter's traces'. *That* thought has to be had by *someone*.) Any thought of a composite thinker has to be the thought of *someone*; it can't be the thought of that very composite thinker, because *every* thought of such a thinker is an upshot of many sub-thoughts, which means that we can never get down to a thought that is inherently and absolutely *one*, from which we might get going on conceptually unified composites. So a thought can't possibly be had by something that

is essentially composite; it must be had by a *single* substance, one that isn't an aggregate of substances, i.e. one that is absolutely simple.

A353 The core of this argument lies in the proposition that if many representations are to form a single thought they must be contained in the absolute unity of the thinking subject. But this can't be proved from *concepts*. The proposition

P: *A thought must be an effect of the absolute unity of the thinking being*

can't be treated as analytic. There's no conceptual contradiction in the supposition that ·P is false, i.e. that· a thought consisting of many representations might come from •the collective unity of different substances acting together (like the motion of a body coming from the motions of all its parts), rather than coming from •the absolute unity of the subject. So the necessity that (P) a composite thought must come from a simple substance can't be demonstrated through the principle of identity—i.e. can't be proved by showing that its contradictory is inconsistent·. Might P be known synthetically and completely *a priori* from mere concepts? You won't want to suggest *that* if you have understood my account of what makes it possible for synthetic propositions to be known *a priori*!

Nor will *experience* show us (P) that every thought must involve an absolutely single subject. Experience can't tell us about the necessity of *anything*, and anyway the concept of *absolute unity* is completely out of reach of experience. Well, then, what about this proposition P on which the whole psychological inference of reason depends—where can we get it from?

A354 It's obvious •that if anyone x wants to represent a thinking being y to himself he has to put himself in y's place, as it were *substituting* his own subject for y's,. . . .and •that the reason why we insist that anyone who has a thought must be

absolutely unitary ·= partless = simple· is just that otherwise we couldn't have the 'I think'. . . . For although the whole of the thought could be split up and distributed among many subjects, the subjective 'I' can't be split up and distributed, and it's this *I* that we presuppose in all thinking.

As in the first paralogism, so here too the formal proposition of self-awareness, *I think*, remains the only basis that rational psychology can rely on when it sets out to enlarge its knowledge. But this proposition is not itself an experience—it is the *form* of the self-awareness that belongs to and precedes every experience. Given that that's its status, its bearing on any possible item of knowledge is only that of a *merely subjective condition* of that knowledge; and we go wrong when we transform it into a condition—an *objective condition*—of the possibility of a knowledge of objects, i.e. into a *concept* of thinking-being-as-such. ·We don't and can't have any such concept·: the only way we can represent to ourselves *thinking-being-as-such* is by putting ourselves, along with ·the *I think* which is· the formula of our consciousness, in the place of every other thinking being.A355

So the famous psychological proof is based merely on the indivisible unity of a representation *I*, and all that *that* does is to govern the verb *think* in its relation to a person. It's obvious that in attaching *I* to our thoughts we refer to the thought-haver only transcendently; we aren't saying anything about any quality that it has; indeed we aren't acquainted with, and don't know anything about, any qualities that it may have. All the *I* refers to is a transcendental subject—a *something in general*. There is nothing determinate [here = 'detailed'] in it, which is one reason why it has to be simple. . . . But this simplicity of *the representation of a ·thinking· subject* is not knowledge of the simplicity of *the subject itself*. . . .

A356

So this much is certain: through the *I*, I always have the thought of myself as ‘simple’ in the sense of having an absolute but merely *logical* unity; but this doesn’t involve me in knowing anything about the actual simplicity of myself as a haver of thoughts. Just as the proposition ‘I am substance’ involves only the *pure* category of substance, which I can’t make any use of empirically, so here I can legitimately say: ‘I am a simple substance’, i.e. a substance the representation of which never involves a pulling together of several different elements, but. . . this proposition tells me *nothing* about myself as an object of experience, because the concept of substance is used here in a way that •doesn’t involve any underlying intuition and therefore •doesn’t have an object. . . . Now let us *test* the supposed usefulness of this proposition ‘I am a simple substance’.

The only reason why anyone has cared about the assertion of the simple nature of the soul is as a way of distinguishing this •thinking subject from all •matter, thus enabling the soul to escape from the dissolution to which matter is always liable. [That was one of Descartes’s two arguments for the immateriality of the soul: all matter is divisible, no soul is divisible, therefore etc.] That’s why the proposition in question is usually

A357

expressed as ‘The soul is not corporeal’. Well, now, suppose we

- take this top proposition of rational psychology, in the meaning that is appropriate to a judgment of pure reason derived solely from pure categories, and

- allow it full objective validity, so that it becomes the •fact–stating proposition that *everything that thinks is a simple substance*;

•even with this grotesque self-indulgence we *still* can’t get the top proposition to throw any light on the question of whether or how the soul differs from matter. That is what

I am about to show; and that will be tantamount to sidelining this supposed psychological insight, relegating it to the domain of mere ideas without the grip on actuality that would give it an objective use.

In the Transcendental Aesthetic I conclusively proved that bodies are mere •appearances of our outer sense, not •things in themselves. So we’re entitled to say that our thinking subject isn’t corporeal: it is represented by us as an object of *inner* sense, so it can’t be an object of *outer* sense, i.e. an appearance in space, •as bodies are. This amounts to saying that we can’t find thinking beings—*as thinking beings*—among outer appearances; i.e. that their thoughts, consciousness, desires and so on can’t be outwardly intuited because they all belong to inner sense. This argument seems to be so natural and so popular that even people with only average intellectual abilities have relied on it as a reason for the age-old view that souls are quite different from bodies.

A358

•Here, as so often, a genuine truth has to be watched so that it doesn’t purport to say more than it does. It is true that extension, impenetrability, cohesion, and motion—in short, everything that outer senses can give us—are different from and don’t contain thoughts, feelings, desires, or decisions, because these are never objects of outer intuition. But •let’s not let that run away with us. There is

- (1) the *Something* that underlies outer appearances, affecting our sense in ways that give it representations of space, matter, shape etc.;

and there is

- (2) the *Something* that is the subject of our thoughts.

And the above argument for saying that the soul is not a body doesn’t conflict with the view that (1) is identical with (2)—i.e. that what *underlies* outer appearances is the same noumenon (or, better, the same transcendental object) as what *underlies* or *has* our thoughts. It’s true that the way our

outer sense is affected by the Something doesn't give us any intuition of representations, of will, or the like, but only of space and space-related properties; but the Something itself isn't extended or impenetrable or composite, because those predicates have to do only with sensible intuitions that we have through being affected by certain objects that we know nothing about in any other way. In saying that the Something is 'not extended' etc., we aren't expressing any knowledge about what kind of an object it is, but only acknowledging that considered in itself—apart from any relation to the outer senses—it's not something to which those predicates of outer appearances can be applied. But there's nothing about it that is inconsistent with the predicates of inner sense, representations and thought. Thus, even if we allow that the human soul is simple in nature, that doesn't distinguish it from the substratum of matter—if matter is considered (as it should be) as mere appearance.

A359

If matter were a thing in itself (·and if the soul were also a thing in itself·), then matter as composite would have to be different from the soul, which is simple. But when we take matter to be mere outer appearance of Something that can't be known through any predicate that we can assign to it, we have to admit that this Something might be simple, even though it affects our senses in such a way as to give us the intuition of something extended and therefore composite. Nor is there any obstacle to supposing that •the substance that appears to our outer sense as extended *has thoughts*, and that •it can represent these thoughts by means of its own inner sense. If that were how things stood, a single thing would be (taken one way) corporeal while also being (taken another way) a thinking thing whose thoughts we can't intuit though we can intuit their signs in the domain appearance. And then we'd have to give up the thesis that only **souls** think, taking souls to be substances of a particular kind; we

would have to replace that by the commonplace statement that **men** think, i.e. that the very same thing that as outer appearance is extended is also (in itself) internally a simple subject of thoughts.

A360

[Kant now re-states the view he has been expressing, in several ways that aren't sufficiently different to throw much new light. Then:]

.A361

Thus the collapse of rational psychology's main support brings the whole thing crashing down. It's as true here as it is elsewhere that we can't hope to extend our knowledge through mere concepts—let alone through the consciousness that is the merely subjective form of all our concepts—in the absence of any relation to possible experience. And in our present case there is an extra reason for that general result. The basic concept of a *simple nature* can't be fitted to anything we encounter in experience, so that there's no way it can function as an objective concept.

Third paralogism: Personhood

- Anything that is conscious of •the numerical identity of itself at different times—i.e. of being the very same individual thing at different times·—is to that extent a *person*.
- The soul is conscious of the numerical identity of itself at different times, .
- Therefore the soul is a person.

Critique of the third paralogism of transcendental psychology

If I want to know through experience the numerical identity of an external object, I shall focus on the permanent element in the appearance—the element that is the subject x such that everything else in the appearance is a state of x—and I shall note *its* identity throughout the time in which the states come and go. Now, I am an object of inner sense, and all

A362

time is merely the form of inner sense. Consequently, I relate each of my successive states to the numerically identical self in all •time. . . . This being so, the proposition that the soul is a person has to be regarded not as something I *infer* but rather as an identical [here = 'trivially analytic'] proposition about consciousness of oneself in time—which is what makes it valid *a priori*! For all it says, really, is that in the whole time in which I am conscious of myself I am conscious of this time as belonging to the unity of myself. I can say

- this whole time is in me, as individual unity, or that
- I am to be found as numerically identical in all this time,

and it makes not the slightest difference which I say.

In my own consciousness, therefore, identity of person is unfailingly met with. But if I view myself from the standpoint of someone else (as an object of *his* outer intuition), it is this external observer who first represents

- me as in time;

because really all I get from my self-awareness is a representation of

- time in me.

A363 Although this observer admits the *I* that accompanies. . . .all representations at all times in *my* consciousness, he won't infer from this that I am something objectively permanent. For just as the time in which he places me is the time not of •my sensibility but of •his, so the identity that is necessarily bound up with •my consciousness is not therefore bound up with •his identity. . . .

The identity of the consciousness of myself at different times is therefore only a formal condition of my thoughts and their coherence, and in no way proves the numerical identity of myself as a thinking subject. Despite the logical identity of the *I*, there may have been a change that rules out a continuing identity. It could be that one thinking

subject is replaced by another, that by a third, and so on, while the same-sounding *I* is used all through, because each outgoing thinking subject hands over its state to its immediate successor.⁴

A364

Consider the dictum of certain ancient schools, that everything in the world is *in a flux* and nothing is *permanent*, nothing lasts. This can't be reconciled with the thesis that there are substances, because they are by definition permanent things; but it isn't refuted by the unity of self-consciousness, because our own consciousness doesn't tell us whether as souls we are permanent or not. Since we count as belonging to our identical self everything we are conscious of, we have to judge that we are one and the same throughout the whole time of which we are conscious. [Kant wrote '*only* what we are conscious of', but that was presumably a slip, because 'everything that we are conscious of is what's needed for his line of thought.] But we still can't claim that this judgment would be valid from the standpoint of an outside observer. Here is why: What we encounter in the soul is not any permanent appearance, but only the representation *I* that accompanies and connects all the inner appearances; so we can't prove that this *I*, a mere

⁴ An elastic ball that collides with another similar one in a straight line passes on to the other its whole motion, and therefore its whole state (that is, if we take account only of the positions in space). If, then, in analogy with such bodies, we postulate substances of which one passes on to another its representations along with the consciousness of them, we can conceive a whole series of substances of which the first transmits its state together with its consciousness to the second, the second to the third, and so on down the chain, with each substance handing over all its own states and those of its predecessor. The last substance would then be conscious of all the states of all the substances that had been switched into and out of the series, and would be conscious of them *as its own*, because they would have been transferred to it along with the consciousness of them. Yet it wouldn't have been one and the same person in all these states.

A365 thought, isn't in the same state of flux as the other thoughts that are strung together by means of it.

[Kant now offers a horribly difficult paragraph, about the order in which we *do* argue for

- 'the soul is permanent',
- 'the soul is a substance', and
- 'the soul is a person',

and the order in which we *could* argue for them if things were different in certain ways. The details are cloudy, and the paragraph seems not to be needed for a grasp of the main lines of Kant's thoughts about the paralogisms. He then continues:]

A366 Just as we have kept the concept of *substance* and of the *simple*, it's also all right for us to keep the concept of *person*; but we must give it its merely transcendental status as something that concerns the unity of the subject—the thinking subject about which we don't know anything else, but whose states are thoroughly inter-linked by self-awareness. Taken in this way, the concept is good enough for practical use; but we mustn't parade the proposition 'The soul is a person' as •adding something to our self-knowledge through pure reason, and as •exhibiting to us, from the mere concept of the identical self, an unbroken continuance of the subject. Why? Because if we look to this concept for leverage on any question that aims at synthetic knowledge, it will just keep spinning on its axis, giving no help. We don't know what *matter* may be as a thing in itself, but because it is represented ·to us· as external ·to us·, we can observe its permanence as appearance. But if I want to observe the mere *I* in the change of all representations, I have no other correlate to use in my comparisons except again myself, with the universal conditions of my consciousness. [Kant means: In empirically identifying matter as substantial, I compare some of my intuitions with others, comparing •the *subset* of them that do pertain

to matter with •the subset that don't. But when I come to the question of whether *I* am a substance, all I can appeal to is the omnipresent *I* that accompanies absolutely all my mental states, so that I can't show my substantial status by comparing some of my intuitions with others.] So if someone else raises the question of whether I am a continuously existing person, the only answers I can give are tautological ones in which I . . . take for granted that which the questioner wants to know. ·That is, I answer his question about what I am in the only way I *can* tackle such a question, namely by reporting on my own inner states and events; but I have to report these as *mine*, with the *I* running all through my account; that makes my subjective *I* deputise for the questioner's objective concept of substance, and so has the effect of presupposing an answer to his question without throwing any light on it·.

Fourth paralogism: ideality (in regard to outer relation)

- If the only basis for believing in x's existence is an inference to x as a cause of given perceptions, then it is open to question whether x does exist.
- The existence of outer appearances is never immediately perceived; our only basis for believing in their existence is an inference to them as causes of given perceptions.
- Therefore it is open to question whether any objects of the outer senses really exist.

367

My label for this uncertainty—this open-to-question-ness—is 'the ideality of outer appearances'; and the doctrine of this ideality, ·expressed in the conclusion of the fourth paralogism·, is called **idealism**. The opposing doctrine, which says that we *can* have certainty about ·the real existence of· objects of outer sense, is called **dualism**.

Critique of the fourth paralogism of transcendental psychology

Let's start with the premises. This paragraph and the next will give a sympathetic statement of the lines of thought that lie behind the premises of the fourth paralogism. We're justified in contending that we can't immediately perceive anything that isn't in ourselves, and that for me the only object of a mere perception—i.e. the only thing that I immediately perceive—is my own existence. So the existence of an actual object outside me. . . .is never given directly or immediately in perception. Perceiving something is having one's inner sense in a certain state; and the only way to bring an outer object *x* into the story is by thinking of *x* as the outer cause of the inner state, and thus *inferring* the existence of *x*. . . . Obviously what is external to me isn't in me; so I can't encounter it in my self-awareness or, therefore, in any perception, because the right way to see perceptions is as mere states of our self-awareness.

..A368

So I'm not in a position to *perceive* external things, but can only infer their existence from •my inner perception, taking •this as an effect of some external immediate cause. Now, the inference from a given effect to a definite cause is always uncertain, because the effect may be due to more than one cause. Thus, when we are thinking about the causes of perceptions, it always remains doubtful—open to question—whether the cause is internal or external; i.e. whether all the so-called •outer perceptions aren't a mere play of our •inner sense, or whether they are related to actual external objects that cause them. Anyway, the existence of outer objects is only inferred, and is vulnerable to all the troubles that an inference can run into, whereas the object of inner sense (I myself with all my representations) is immediately perceived, and there can't be any doubt that *it* exists.

So it's wrong to think of an 'idealist' as someone who

denies that there are any external objects of the senses. An idealist, properly so-called, is someone who won't admit that the existence of such objects is known through immediate perception, from which he infers that there couldn't be any experience that made us completely certain of the reality of external objects of the senses.

A369

Before exhibiting our paralogism in all its deceptive illusoriness, I should first remark that we must distinguish •transcendental idealism from •empirical idealism. [Kant will stay with this and related distinctions for about four pages. He won't again refer explicitly to •the fourth paralogism, but his discussion of types of idealism constitutes a critique of •it.] By *transcendental idealism* I mean this doctrine:

Appearances are *all* to be regarded as mere representations, not as things in themselves, so that time and space are only •sensible forms of our intuition, not •states given as existing by themselves and •not conditions of objects viewed as things in themselves.

To this idealism there is opposed a *transcendental realism* that regards time and space as given in themselves, independently of our sensibility. The transcendental realist thus interprets outer appearances (taking for granted that they are real) as things-in-themselves, which exist independently of us and of our sensibility, and are therefore outside us—taking the phrase 'outside us' in its most radical sense. It's this transcendental realist who afterwards plays the part of empirical idealist: after wrongly supposing that if objects of the senses are external they must have an existence by themselves, independently of the senses, he finds that from this point of view all our sensuous representations are inadequate to establish the reality of those objects.

A370

The transcendental idealist, on the other hand, can be an empirical realist—or a *dualist*, as he is called. That is, he can grant the existence of matter without •going outside his mere self-consciousness or •assuming anything more than the certainty of his representations. . . . For he regards the facts about what matter there *is*, and even about what there *could be*, as facts merely about appearance; and when appearance is separated from our sensibility it is nothing. For him, therefore, matter is only a species of representations (intuition); and these representations are called ‘external’ not because they relate to objects that are *in themselves external* (they don’t), but because they relate perceptions to the space in which all things are external to one another, although the space itself is in us.

[It may be useful to have a brief restatement of the main theses of the preceding two paragraphs: Kant has distinguished

(1) two **transcendental** theses about matter, i.e. two views about the meanings or metaphysical status of propositions about matter:

(a) **idealism**: such statements are really complex statements about our states of mind;

(b) **realism**: such statements are entirely independent of facts about our minds—they don’t imply such statements and aren’t implied by them.

And he has distinguished

(2) two **empirical** theses about the status, for us, of the proposition that there is matter in the world:

(a) **idealism**: we can’t have certainty that the proposition is true;

(b) **realism**: we can be perfectly certain that the proposition is true.

One natural pairing, Kant is saying, is

(1b) transcendental realism and (2a) empirical idealism.

Because the proposition that there is matter has a status that puts it out of our reach, we can’t be sure that there is any matter. The other natural pairing is

(1a) transcendental idealism and (2b) empirical realism.

The proposition that there is matter is a special kind of proposition about our own mental states; that puts it within our reach, enabling us to be quite sure that it is true.]

Right from the outset I have declared my acceptance of transcendental idealism; and that clears the way for me to accept the existence of matter on the unaided testimony of my mere self-consciousness, taking it to be proved in the same way that I prove ·to myself· the existence of myself as a thinking being. ·Here’s how it goes·: I am conscious of my representations; so these representations exist, and so do I, the subject that *has* them. External objects (bodies) are mere appearances, so they are only one kind of representation that I have, and representations of that kind aren’t *of* anything beyond the representations themselves. Thus •the existence of external things is as secure as •my own existence, because I know *both* from the immediate testimony of my self-consciousness. The only difference is that the •representation of myself as the thinking subject belongs to **inner** sense only, whereas the representations that signify extended things belong also to **outer** sense. [Note ‘belong *also*’: outer sense is just a *part* of inner sense.] I don’t need *inference* to establish the reality of outer objects, any more than I need inference to establish. . . .the reality of my thoughts. In both cases, the objects are nothing but representations, the immediate perception (consciousness) of which is automatically a sufficient proof of their reality.

. . . .Transcendental realism, on the other hand, inevitably runs into trouble, and finds that it has to allow empirical idealism. Here is why: It regards the objects of outer sense as distinct from the senses themselves, taking mere appearances as self-subsistent beings that exist outside us. On that view, however clearly we are conscious of our representations of these things, it’s still far from certain that if the representations exist then the corresponding objects also exist. In my

A371

A372 system, on the other hand, these external material things are. . . nothing but mere appearances, i.e. representations of whose reality we are immediately conscious.

So far as I know, all psychologists [here = 'philosophers of mind'] who adopt empirical idealism are transcendental realists; and they have certainly been consistent in parading empirical idealism as 'setting' an important problem from which human reason can't easily extricate itself. For if we regard outer appearances as representations produced in us by their objects, and if these objects are things existing in themselves outside us, it's impossible to see how we could come to know the existence of the objects other than by inferring causes from effects; and the conclusions of such inferences are *always* doubtful, even when the cause in question is *in us*. Perhaps our outer intuitions are indeed caused by something that is (in the transcendental sense) 'outside' us; but, if so, this cause isn't the kind of object we have in mind when we talk about 'matter' and 'bodies'. . . . What we are talking about is not this •transcendental object that we don't know about either through inner or through
A373 outer intuition. Rather, we are speaking, of the •empirical object, which is called an *external* object if it is represented *in space*, and an *inner* object if it is represented only *in its time-relations*. And space and time are to be found only *in us*.

The phrase 'outside us' is thus unavoidably ambiguous: sometimes it refers to

(1) something which *as a thing in itself* exists apart from us,

and at other times it refers to

(2) something belonging solely to outer *appearance*.

The psychological question about the reality of our outer intuition involves (2), and we need an unambiguous way of saying this. So I shall distinguish (2) *empirically external*

objects from (1) ones that may be said to be transcendently external, by labelling (2) as 'things that are to be found in space'.

Space and time are indeed *a priori* representations that reside in us, as forms of our sensible intuition, *before* any real object has acted on our senses through sensation and enabled us to represent the object in terms of its spatial and/or temporal relations. But the material or real element, the Something that is to be intuited *in space*, necessarily presupposes perception; in the absence of perception, no power of imagination can invent and produce that Something. So it is sensation that indicates a reality in space or in
A374 time. . . . (When a sensation is taken to be *of* something, though without giving any details about it, we call it 'perception'.) Once a sensation has been given, its internal variety enables us to picture in imagination many objects that have no empirical place in real space or time. There's no room for doubt about this: it's perception that provides the raw materials we need if we are to have thoughts about objects of sensible intuition. This holds equally for 'inner perceptions of' pleasure and pain and for the sensations of the outer senses, such as colours, heat, etc., but just now my topic is the 'outer' part of the story. *This perception represents something real in space*, and here is my three-part reason for saying so. (1) Just as space is the representation of a mere *possibility* of coexistence, perception is the representation of a *reality*. (2) This reality is represented •to outer sense, i.e. •in space. (3) Space is itself nothing but mere representation. And so we get the 'double' result:

•Only what is represented in space can count as real in space.⁵

⁵ We must take careful note of the paradoxical but correct proposition that there's nothing in space but what is represented in it. 'Why is it true?' Because space itself is nothing but a representation,

A375 •*Everything* that is represented through perception as given in space is real in it. . . .

So all outer perception provides immediate proof of something real in space—or, rather, it is itself what is real. This puts empirical realism beyond question—there *does* correspond to our outer intuitions something real in space. Of course space and all its appearances are representations, which means that they are only *in me*, but ·that doesn't abolish the distinction between inner and outer·: what is real, i.e. the material of all objects of outer intuition, is given in this space as actual and independent of all imaginative invention. And it's impossible for anything that is (in the transcendental sense) *outside* us to be given in *this* space, which is nothing apart from our sensibility. Thus, even the strictest idealist can't require a proof that our perception has a corresponding object that is 'outside' us in the strict ·transcendental· sense. If there were any such object, it couldn't be represented and intuited as outside us; because this would involve *space*, a mere representation, containing no reality that isn't in perception. . . .

A376

Knowledge of objects can be generated from perceptions, either by mere play of imagination or by means of experience. [See note on page 155 regarding 'knowledge']. And in the course of this there can indeed arise illusory representations, ones with no corresponding objects, the deception being attributable sometimes to the imagination's playing tricks (in dreams) and sometimes to the judgment's going astray (in so-called 'sense-deception'). To avoid such deceptive illusion, we have to steer by the rule:

and nothing can be in it except what is contained in that representation. . . . It must indeed seem strange to say that a thing can exist only in the representation of it, but the sense of strangeness evaporates in our present context, where the things in question are not things in themselves but only appearances, i.e. representations.

Anything connected with a perception according to empirical laws is actual.

But such deception, as well as this shield against it, has as much to say to idealism as to dualism. I'm talking about ·transcendental idealism, i.e.· our present concern with the *form* of experience. I needn't re-introduce empirical idealism because I have already refuted that and its mistaken challenge to the objective reality of our outer perceptions [and Kant briefly repeats his arguments to that effect. Then:]

A.377

·A new distinction between kinds of idealism needs to be drawn now·. On the one hand we have

(1) the dogmatic idealist, who *denies* the existence of matter.

He must base this denial on supposed contradictions in the thought of there being such a thing as matter at all. I haven't needed to discuss this so far, but I shall do so: the difficulty will be removed in the next chapter [which starts on page 206] on dialectical inferences, where I'll display reason as being at odds with itself regarding the concepts it makes for itself. . . . On the other hand we have

(2) the sceptical idealist, who *doubts* the existence of matter, thinking that it can't be proved to exist.

·While it's appropriate to brush the dogmatic idealist aside as being wholly wrong·, the sceptical idealist, ·though also in error·, is a benefactor of human reason! All he does is to challenge our basis for asserting that matter exists; we thought we could base it on immediate perception, but he criticises that as inadequate. This challenge compels us to be constantly on the watch—even in the smallest advances of ordinary experience—to ensure that we don't treat as a well-earned possession something that we may have obtained only illegitimately. Now we can see clearly the value to us of these ·sceptical·-idealist objections. [Kant goes on to say *what* this 'value' is: it turns out to consist in our being forced

A378

by sceptical idealism to keep in mind and stay true to the tenets of transcendental idealism. If we treat outer objects as things in themselves, our situation is as bad as the sceptical idealist says, and indeed even worse. So we *have to* adopt the only alternative, namely the thesis that outer things are mere representations. He continues:]

A379

The question then arises: ‘In the philosophy of mind, is dualism the only tenable position?’, and our answer has to be: ‘Yes indeed, but only when “dualism” is understood in the empirical sense.’ That amounts to taking dualism as saying that in the interconnected web of experience

- matter, as substance in the [domain of] appearance really is given to outer sense,

just as

- the thinking *I*, also as substance in the [domain of] appearance, is given to inner sense.

Further, inner and outer appearances must be connected with each other according to the rules that this category of substance brings to our perceptions—inner as well as outer—enabling them to constitute one experience. But if we try (and people often do) to extend the concept of dualism and take it in the transcendental sense, we’ll arrive at something for which there isn’t the slightest basis. Why? Because we’ll have misapplied our concepts, taking •the difference between two ways of representing objects (which, as regards what they are in themselves, still remain unknown to us) as •a difference in these things themselves. (And this fault is present not only in transcendental dualism but also in the two opponents to it—*pneumatism* on one side and *materialism* on the other.) [Pneumatism is the thesis that the soul is immaterial.] The *I* represented through inner sense in time is a specifically quite distinct •appearance from objects in space outside me, but these two shouldn’t be construed as different •things. The *transcendental object* that underlies

outer appearances is not matter; the *transcendental object* that underlies inner intuition is not a thinking being. Rather, each of these is a ground (to us unknown) of the appearances that supply us with empirical concepts of matter and of mind. . . .

A380

Consideration of pure psychology as a whole, in view of these paralogisms

If we compare **(1)** the *doctrine of souls* as the physiology [= ‘empirical’] study of inner sense, with **(2)** the *doctrine of body* as a physiology of the object of the outer senses, we find that while there’s a lot of empirical knowledge to be gained in both of them, they are notably unlike in what can be learned *non-empirically* through them. In **(2)** there is much *a priori* synthetic knowledge to be had from the mere concept of an extended impenetrable being, whereas in **(1)** there’s no comparable knowledge from the concept of a thinking being. Here is why. Although both kinds of being are appearances, the **(2)** appearance to outer sense contains something fixed or lasting, which supplies the underlying *thing* which all the transitory states are states *of*. That enables it to present a synthetic concept, namely the concept of space and of an appearance in space. In contrast with that, time—which is the sole form of our inner intuition—doesn’t contain anything lasting, so it provides knowledge only of the change of states, not of any object that they are states *of*. In the ‘soul’ (as we call it) everything flows and nothing stays still except the *I*, which is simple solely because the representation of it has no content and thus no qualitative complexity. . . .

A381

A382

What would it take for us to have, through pure reason, knowledge of the nature of a thinking being as such? We would need the *I* to be an **intuition** which, being presupposed in all thought as such (prior to all experience), could

yield *a priori* synthetic propositions. But the *I* that we have is no more an •intuition than it is a •concept of an object! Rather, it is the mere **form** of consciousness, which can accompany the two kinds of representation (•inner and outer•) but can't elevate them to the level of items of knowledge unless **something else** is given in intuition—something that provides material for a representation of an object. Thus the whole of rational psychology, as a science surpassing all the powers of human reason, collapses. The most we can do is to •study our soul under the guidance of experience, and •confine ourselves to questions that stay within the limits of what might possibly be answered by inner experience.

But although rational psychology is useless as a way of extending our knowledge, and when so used is entirely made up of paralogisms, it undeniably has considerable negative value as a critical treatment of our dialectical inferences, those that arise from common and natural reason.

A383

What leads us to resort to a doctrine of the soul based on nothing but pure principles of reason? No doubt we are primarily aiming to secure our thinking self against the danger of materialism. This sense of danger takes the form of the fear that

(1) if all matter went out of existence, all thought—and even the very existence of thinking beings—would be destroyed.

But *that* fear is dealt with by the pure concept of our thinking self that I have been presenting. What we get from it, far from the fear (1), is a clear proof that

(2) if the thinking subject went out of existence, the whole corporeal world would necessarily also vanish, because that world is nothing but an appearance in the sensibility of our •thinking• subject, a way in which its representations occur.

Admittedly this doesn't tell me anything more about the properties of this thinking self, e.g. giving me insight into whether it is permanent or not. It doesn't even throw light on whether the thinking self exists independently of the transcendental substratum of outer appearances (supposing there is one); because that substratum is as unknown to me as is the thinking self. [Kant will now speak of •other-than-speculative reasons for hoping for something. He is speaking of •practical reasons—ones connected with morality rather than metaphysical theory—for *hoping* that one's soul can survive through into an after-life.] Still, I may come to have a non-speculative reason to hope for an independent and continuing existence of my thinking nature, throughout all possible changes of my state. In that case it will be a great help if, while freely admitting my own ignorance, I can repel the dogmatic assaults of a speculative opponent, showing him that just as I can't support clinging to my hope by appeal to any knowledge of the nature of the self, he can't bring such knowledge to support his denial that the hope can be realized.

A384

The real goal of rational psychology lies in three other dialectical questions that •are also based on this transcendental illusion in our psychological concepts, and •can't be settled except through the inquiry I have been conducting. They concern

- (1) the possibility of interaction between a soul and an organic body,
i.e. the question of what it is to have an animal nature and of how the soul fits into human life;
- (2) the beginning of this interaction,
i.e. the question of the soul in and before birth; and
- (3) the end of this interaction,
i.e. the question of the soul in and after death (the question of immortality). [In the foregoing, 'interaction' translates Kant's *Gemeinschaft*, often translated as 'community'.]

A385 Some people think that these questions involve difficulties that they can use as dogmatic [see note on page 15] objections to certain beliefs or hopes concerning the soul. They want to be admired for having a deeper insight into the nature of things than the general run of us can claim to have! Well, I maintain that what they have is merely a *delusion* in which they *hypostatise* something that exists merely in thought—that is, they treat it as a real object existing. . . .outside the thinking subject. [Kant is going to use ‘hypostatise’ quite a lot, often in its basic sense of ‘treat as a thing or substance’, but occasionally in the different though related sense of ‘treat as being real independently of the mind’.] In other words, they regard extension, which is nothing but appearance, as a property of outer things that exist quite apart from our sensibility, and claim that motion is due to these things and really occurs in and by itself, apart from our senses. This is a delusion because matter, whose interaction with the soul causes so much fuss, is a mere *form*, a particular way of representing an unknown object through the kind of intuition that is called ‘outer sense’. Perhaps there really is outside us something corresponding to this appearance that we call ‘matter’; but even if there is, in its role as appearance it isn’t outside us; it is only a thought *in* us, although this thought represents it as existing outside us because it comes to us through outer sense. So ‘matter’ doesn’t refer to a kind of substance that is utterly unlike the object of inner sense (the soul), but only to the distinctive nature of certain appearances of objects. The objects are in themselves unknown to us, but we call our representations of them ‘outer’ as compared with those that we count as belonging to ‘inner’ sense, although these outer representations belong only to the thinking subject, as do all thoughts. There is indeed something deceptive about them: representing objects in space, they seem to detach themselves from the soul, so to speak, and to hover outside it.

And yet the very space in which they are intuited is nothing but a representation, and what it’s a representation of can’t be found outside the soul. So we drop the question about the interaction between the soul and other known substances of a different kind outside us; and we’re left with a question about how representations of inner sense are connected with states of our outer sensibility—the question of how these can be so inter-linked according to settled laws that they hang together in a single experience. A386

As long as we hold inner and outer appearances together in our minds as mere representations in experience, we won’t see anything absurd or strange in the thought of their interaction—the interaction between these the two kinds of senses. We run in trouble over interaction only when

we hypostatise outer appearances, come to regard them not as representations but as things existing by themselves outside us, with the same qualities that they have in us, and think of them as acting on our thinking subject in the way they (as appearances) act on one another.

We get into difficulties then, because the efficient causes outside us—material things colliding with one another—have a character that can’t be squared with their effects in us. That’s because the cause relates only to outer sense, the effect to inner sense—and although these senses are combined in one thinking subject they are extremely unlike each other: the only outer effects are changes of place, and the only forces are drives that result in changes of place; whereas within us the effects are thoughts, which don’t have any spatial features—no locations, motions, shapes etc. That’s why, when we try to trace outer causes through to their effects in inner sense, we get lost. But we should bear in mind that bodies are not objects-in-themselves that are present to us, but a mere appearance of some unknown A387

object; •that motion is not an effect of this unknown cause, but only the appearance of its effect on our senses; •that bodies and motions are not something outside us, but mere representations in us; and •that, therefore, the motion of matter doesn't *produce* representations in us, because the motion is a representation only. . . . The bottom line is that this whole self-inflicted problem boils down to this:

How and through what cause are the •representations of our sensibility so interconnected that the •ones we call 'outer' intuitions can be represented according to empirical laws as objects outside us?

This question doesn't in any way involve the supposed difficulty of explaining how our representations could be effects of utterly different efficient causes outside us. That difficulty arose from our taking the appearances of an unknown cause to be the cause itself outside us—a mistake that is bound to lead to confusion.

A388 When a judgment involves a misapprehension that has taken deep root through long custom, one can't, straight off, correct it as clearly as one can correct mistakes that aren't conceptually confused by inevitable illusion. So my freeing of reason from sophisticated theories can hardly have yet the clarity that is needed for its complete success. But I think the following comments will be a move towards complete clarity.

Objections are of three kinds: **(1)** dogmatic, **(2)** critical, and **(3)** sceptical. A **(1)** dogmatic objection is directed against a *proposition*, a **(2)** critical objection is directed against the *proof* of a proposition. To make a **(1)** dogmatic objection to proposition P about some object x, one needs an insight into the nature of x that will entitle one to maintain the opposite of what P says about x. The objection counts as 'dogmatic' because it claims to know more about how x is constituted than does the proposition it is opposing. [Re 'dogmatic', see

note on page 15.] A **(2)** critical objection doesn't say anything about whether the proposition P is any good, so it doesn't presuppose. . . . fuller knowledge concerning the nature of the object x; all it does is to attack the proof •that has been offered for P. If a critical objection succeeds, it shows only that P is unsupported, not that it is wrong. A **(3)** sceptical objection sets up P and not-P as equally matched opponents, treating each—turn about—as asserted dogmatically and objected to •dogmatically• by its opponent. This conflict, seemingly dogmatic on both sides, implies that all judgment on the topic in question is completely null and void. So dogmatic and sceptical objections both lay claim to as much insight into their object as they need for their assertion or denial. But a critical objection confines itself to •pointing out that an assertion presupposes something that's empty and merely imaginary, thereby •overthrowing the asserted theory by pulling its supposed foundation out from under it, without trying to establish any rival view about the nature of the object.

A389

When we bring the ordinary concepts of our reason to bear on •the question of• the interaction between our thinking subject and the things outside us, we are dogmatic, regarding outer things as real objects existing independently of us (in line with a certain *transcendental dualism*, which doesn't assign these outer appearances to the subject as representations, but completely separates them from the thinking subject, placing them outside us while still giving them the properties they are given in our sensible intuition. This switch is the basis of all the theories about the interaction between soul and body; they all accept without question the objective reality of outer appearances. . . . The three standard theories about this are in fact the only *possible* theories: that of **(1)** physical influence, that of **(2)** predetermined harmony, and **(3)** that of supernatural intervention. [At Kant's time

..A390

'physical' and its cognates in other European languages didn't imply any restriction to items that *we* would include in 'physics'. It comes from a time-honoured trilogy—logical (what *must be*), physical (what *is*), and ethical (what *ought to be*). So item **(1)** is simply the view that bodies and minds genuinely causally affect one another.]

The accounts **(2)** and **(3)** of the relations between the soul and matter are based on an objection to **(1)** the view of common sense. The objection is this: what appears as matter can't by its immediate influence be the cause of •representations, because •these are too different in kind from matter. [Kant goes on to say that this objection would be meaningless if the objectors regarded matter, in the way Kant does, as a mere representation produced by unknown outer objects. Then:] If their objection were to square with my principles, it would have to say that the true (transcendental) object of our outer senses can't be the cause of the representations (appearances) that we label as 'matter'. But no-one is entitled to say *anything* about the transcendental cause of our representations of the outer senses; so the objection in this form of it is entirely groundless. So we'll have to take these objectors against **(1)** the doctrine of physical influence to be •sticking to the ordinary outlook of transcendental dualism, and •supposing that matter is a thing-in-itself rather than the mere appearance of some unknown thing. So the aim of their objection will be to show that outer objects of this kind, which don't exhibit among themselves any causality except the causing of movements, can't possibly be efficient causes of representations; so that a third entity must intervene to establish if not •interaction then at least •correspondence and harmony between the two. But this objection of theirs *starts* with a basic untruth, namely •their view about• physical influence, which is built into their dualism. Thus, what their objection really refutes is not •the thesis of• natural influence between soul and

body but rather •their own dualistic presupposition. . . . A392

So the commonly accepted doctrine of physical influence can't be effectively opposed by a **dogmatic** objection. •There are two bases from which someone might try to launch a dogmatic objection•. •He could flagrantly hypostatise representations, setting them outside himself as real things. I have shown how untenable this is, and that there's no alternative to transcendental idealism. Well, then, •he could •accept that alternative•, agree that matter and its motion are mere appearances and therefore mere representations, and object to **(1)** on the grounds that the unknown object of our •outer• sensibility couldn't possibly be the cause of representations in us. But he can't justify this, because no-one is in a position to work out what an *unknown* object can or can't do!

However, a sound **critical** objection can be made against the ordinary version of **(1)** the doctrine of physical influence •between soul and body•. The supposed interaction between two kinds of substances, the thinking and the extended, is based on a crude dualism; it turns extended substances (which are really nothing but mere representations of the thinking subject) into things that exist by themselves. . . .

Let's take the notorious question of the interaction between the thinking and the extended, filter out from it any fictitious ingredients, and see what we are left with. It is simply this: A393

How is it *possible* for a thinking subject—*any* thinking subject—to have outer intuition, i.e. an intuition of space and of the filling of space by shape and motion? Which is a question that none of us can possibly answer. This gap in our knowledge *can't* be filled. The most we can do is to mark its place by referring to 'the transcendental object that causes representations of this •outer• type, though we can never know anything about it or even have a concept of it'.

We don't *need* such a concept when dealing with problems arising in the •domain of experience, for then we treat these appearances as objects in themselves, without worrying about the ultimate basis for their possibility as appearances. But we *would* need the concept of a transcendental object if we were to pass the limits •of this •domain•.

These reminders of what the inter-relation is between thinking beings and extended beings suffice to settle all the arguments about the state of the thinking nature before this inter-relation begins (i.e. prior to life) or after it ends (in death). Take the opinion that the thinking subject was able to think before becoming connected a body. This becomes the thesis that

•Before the start of the kind of sensibility through which something appears to us in space, the transcendental objects that do in fact appear •to us• as bodies could have been intuited in an entirely different manner.

And the opinion that after the end of the soul's connection with the corporeal world it could still go on thinking becomes the thesis that

•A stoppage of the species of sensibility through which transcendental objects. . . appear to us as a material world wouldn't automatically create a stoppage of all intuition of transcendental objects. It's quite possible for those same unknown objects to go on being known by the thinking subject, though not of course now intuited as bodies.

Now, no-one can give this the faintest support from any speculative principles. Even the *possibility* of what is asserted can't be established, but only assumed. But it's equally impossible to bring any valid dogmatic objection against it. None of us knows anything about the absolute, inner cause of outer corporeal appearances; so none of us can justify

claiming to know what the outer appearances in our present state (that of life) really rest on; or to know that when this state ends (in death), that will bring the end of all outer intuition or even of the thinking subject itself.

A395

Thus all strife about the nature of the thinking being and its connection with the corporeal world is sheerly a result of plugging a gap in our knowledge with paralogisms of reason, treating our thoughts as things and hypostatizing them. This gives rise to a 'science' that is entirely imaginary, on both sides of each debate, because both sides •suppose they have knowledge of objects of which no human being has any concept, or •treat their own representations as objects, and so whirl around in a perpetual circle of ambiguities and contradictions. This dogmatic delusion keeps many people in bondage to theories and systems, by tempting them with thoughts of an imagined happiness. [It's not clear whether Kant is referring to the happiness of believing in such a theory or system, or to the happiness of life after death.] And the only way of getting free from this bondage is through a sober critique that is both strict and fair, and that confines all our speculative claims to the domain of possible experience. It doesn't do this by stale scoffing at ever-repeated failures, or pious sighs over the limits of our reason, but by effectively *fixing* these limits in accordance with established principles, inscribing its 'go no further' on the Pillars of Hercules. [These marked the two sides of the straits of Gibraltar, regarded by the ancients as the furthest limit of sea voyaging.] Nature herself has erected these, so that the voyage of our reason shan't be extended further than the continuous coastline of experience lets us go—a coast we can't leave without venturing on a shoreless ocean which, after alluring us with deceptive promises, eventually compels us to abandon as hopeless all this vexatious and tedious endeavour.

A396

A394

* * *

I still owe you a clear general account of the transcendental and yet natural illusion in the paralogisms of pure reason, and also a justification of my classifying them in a way that runs parallel to the table of the categories. If I had tried to provide these at the start of this chapter, I would have risked writing obscurely, or clumsily getting ahead of myself. I'll now try to provide what I owe.

Here's an account of *illusion* in general: it consists in treating the *subjective* condition of thinking as being knowledge of the *object*. That covers the •illusions of the senses that sometimes occur in special cases—e.g. being led by your blurred vision of something to think that it has a furry surface—but that isn't relevant to our present topic of dialectical •illusion of pure reason. *That* has to involve subjective conditions of *all* thinking, not just of some special cases. What *are* these universal conditions? Well, in the Introduction to the Transcendental Dialectic I showed that pure reason concerns itself solely with the totality of the synthesis of the conditions for a given conditioned [page 167], and there will therefore be only three cases of the dialectical use of pure reason:

A397

- (1) The synthesis of the conditions of thought as such.
- (2) The synthesis of the conditions of empirical thinking.
- (3) The synthesis of the conditions of pure thinking.

In each of these, pure reason is concerned only with the absolute totality of this synthesis, i.e. with the condition that is itself unconditioned. This •trio of kinds of synthesis gives rise to the •trio of transcendental illusions (corresponding to the •three chapters of the Dialectic) and to the •trio of 'sciences' of pure reason—(1) transcendental psychology, (2) transcendental cosmology, and (3) transcendental theology. Here we are concerned only with the first.

[Back at page 171 Kant arrived at those three pretended sciences through three logical forms of inference, based on three logical types of proposition: (1) subject-predicate ('categorical'), (2) if-then ('hypothetical'), (3) either-or ('disjunctive'). He doesn't work hard at showing that or why the two trios coincide, though we'll see right away what he thinks ties the two versions of (1) together.]

[This next paragraph expands Kant's very compressed one in ways that the 'small dots' convention can't easily handle. The expanded version is offered with fair confidence that it correctly presents Kant's thought.] Thought *as such* doesn't involve any object (whether of the senses or of the pure understanding). That is, it isn't essential to thought that it be *about* something. So the synthesis of conditions of thought—i.e. the stitching together of items that must be involved in *any* thinking—can't involve any relations with objects; it's just a stitching together of •the thought with •the thinker, the person whose thought it is. And objects enter the picture only through the thinker's *mistaken* view that he has a synthetic representation of himself as an object.

It follows from this that (1) the dialectical inference to the conclusion that there is an unconditioned item that is a condition of all thought doesn't commit a •material error (for it abstracts from all content, all objects); rather, it is defective in •form alone, which is why it must be called a 'paralogism'.

A398

Furthermore, since the *only* condition that accompanies all thought is the *I* in the universal proposition 'I think. . .', *that's* what reason has to identify as the unconditioned condition. It *is* indeed a condition of all thought, but only a •formal condition, securing the •logical unity of every thought, with no *object* having any role in it. . . .

If you ask me 'What is the constitution of a thing that thinks?', you'll want a synthetic reply. (An analytic answer might explain what is meant by 'thought', but that's

all; it can't tell you anything about what makes thought possible.) Now, I have no *a priori* knowledge on which to base a synthetic reply, because that would require an appeal to intuition, but your highly general question shuts intuition out because it concerns thought *as such*, including thought that doesn't involve intuition. . . . Yet it *seems* as if I could reply to you on the basis of the proposition that expresses self-consciousness—*I think*. For this *I* is the primary subject—so it is substance, it is simple, and so on. [Kant then says something impenetrably obscure about what should lead me to have suspicions about this plausible answer to your question, though it doesn't diagnose *what* is wrong with it. Then:]

But I can learn what has gone wrong if I dig deeper into the origin of the attributes that I ascribe to myself as a thinking being as such. [Kant then provides the diagnosis that we have already met: concepts such as 'is a substance', 'is simple' and so on aren't fit for expressing any item of knowledge except in a context where there can be intuitions supporting the distinction between what is and what isn't a substance, what is and what isn't simple. He continues:] If I call a thing in the ·domain of· appearance *simple*, I mean that the intuition of it, though it is a part of the appearance, can't in its turn be divided into smaller parts. ·And in our present context, where our concern is with thinking beings *as such*, thinking beings *in general*, no role can be given to intuitions.· If I know something as *simple in concept* but not as *simple in the ·domain of·* appearance, then this isn't an item of knowledge about the object but only about the concept that I form of a 'something' that can't be intuited. My only ground for saying in this case that I think something as completely simple is that I really don't have anything to say about it except merely that 'it is something'.

Now the bare self-awareness, *I*, is in concept substance, in

concept simple, etc.; and in this sense all those psychological doctrines are unquestionably true. But this doesn't give us the knowledge of the soul that we are looking for. Why not? Because none of these predicates can be applied to anything given in intuition, so they can't have any consequences that hold for objects of experience, so they are entirely empty. The concept of substance doesn't teach me that the soul endures by itself, or that it is a part of outer intuitions that cannot itself be divided into parts, and therefore can't arise or perish by any natural alterations. These are properties that would make the soul known to me in the context of experience and might tell me something about its origin and future state; they're the kind of thing that brings the *schematised* concept of substance into play [see page 92]. But if I say, in terms of the mere ·unschematised· category, 'The soul is a simple substance', it is obvious that since the bare concept of substance (supplied by the understanding) contains nothing beyond the requirement that a thing be represented as being subject in itself, and not in turn predicate of anything else, nothing follows from this as regards the permanence of the *I*, and the attribute 'simple' certainly doesn't aid in adding this permanence. Thus, from this source, we learn nothing whatsoever as to what may happen to the soul in the changes of the natural world. If we could be assured that the soul is a *simple part of matter*—a *physical atom*—we could use this knowledge, with the further assistance of what experience teaches about such things, to deduce the permanence, and (with its simple nature thrown into the mix) the indestructibility of the soul. But of all this, the concept of the *I*, in the psychological principle 'I think', tells us nothing.

[The remaining couple of pages of the first-edition treatment of the Paralogisms are brutally difficult, and are probably not worth the trouble, given that Kant is going to re-do the

whole thing in the second edition.]

[The treatment of the Paralogisms as re-stated in the second edition of the *Critique* starts here, picking up from page 178. It is about half the size of the A-material that it replaces; it will run to page 204.]

The paralogisms of pure reason (second edition)

[Back on page 178 Kant wrote of the need to 'keep a critical eye open as we follow this procedure through all the 'four' basic concepts of pure psychology'. He now says that he'll present the material in a 'continuous' way, meaning that he won't (as he did in the first edition) deal at length with each of the four paralogisms separately, though he does start by giving them a paragraph each. That's after a preliminary paragraph which he says may help us to sharpen our wits in the remainder of the chapter. This extremely difficult paragraph, which isn't needed for what follows, is omitted from this version.]

(1) Here's a proposition that is absolutely necessary—indeed, it's analytic:

•In all judgments I am the *determining* subject of the relation that constitutes the judgment.

It has to be granted that the *I*, the *I* that thinks, can always be regarded as *subject*, and as something that doesn't occur in the thought in a merely *predicate*-role. But this doesn't mean that *I*, as *object*, am for myself a *self-subsistent* being or *substance*. The latter statement goes very far beyond the former, and demands for its proof data that aren't to be met with in thought. . . . [In the background of this is Kant's thesis, expounded in the metaphysical deduction of the categories, that our concept of *substance* is the concept of something that figures in our thought as a subject and never as a predicate.]

(2) Here is another proposition that is analytic because it merely states something that is already contained in the very concept of thought:

•The *I* of self-awareness, and therefore the *I* in every act of thought, is *one*, and can't be resolved into a plurality of subjects, and consequently signifies a logically simple subject.

But this doesn't mean that the thinking *I* is a simple *sub-* 408
stance. That proposition would be synthetic. The concept of substance always relates to intuitions that in my case •have to be sensible, and therefore •lie entirely outside the domain of the understanding and its thought. [Recall that •for Kant 'sensible intuition' means 'intuition in respect of which the person is **pas-**
sive'; he holds that all human intuition is like that (hence 'in my case'); and that •he ties 'understanding' tightly to '**active**'.] But it is of this thought—and not of anything intuitive or sensible—that we are speaking when we say that the *I* in thought is simple. •A comment on these first two paragraphs: It's very hard work to find out which of the things that intuition presents us with are substances, and which of them are simple; so it would be astonishing if results about substantiality and simplicity were just handed to me, as though by revelation, in the poorest of all representations—the mere bare empty *I think*.

(3) A third proposition that is implied by the concepts that it uses, and is therefore analytic:

•Through all the variety of which I am conscious •through time• I am identical with myself.

But this identity of the subject, of which I can be conscious in all my representations, doesn't involve any intuition of the subject that would present it as an object; so it can't signify the identity of the *person*, i.e. the . . . identity of one's own substance, as a thinking being, in all change of its states. To establish *that*, we would need various synthetic judgments, 409
based on intuition, that come to us, not a mere analysis of the proposition *I think*.

(4) A fourth analytic proposition:

- I distinguish my own existence as that of a thinking being from other things outside me—among them my body.

This is analytic because *other* things are ones that I think of as *distinct* from myself. But this proposition doesn't tell me whether this consciousness of myself would be possible if there were no things outside me giving me representations, or therefore whether I could exist merely as thinking being (i.e. without existing in human form, ·equipped with a body·).

So we see that the analysis of my consciousness of myself in thought in general, ·thought as such·, contributes *nothing* to my knowledge of myself as object. ·Those who have fallen for the paralogisms have· mistaken •the logical exposition of thought in general for a •metaphysical account of the nature of the object.

Suppose that we *could* prove *a priori* that all thinking beings are in themselves simple substances, so that personhood is inseparable from them and that they are conscious of their existence as separate and distinct from all matter. That would be a great stumbling-block—indeed *the* great stumbling-block—in the way of my whole critique. Why? Because in conducting such a proof we would have stepped outside the world of sense and entered the domain of noumena; and no-one could then deny our right to advance yet further into this domain, indeed to settle there and—with luck—stake a claim to permanent possession. The proposition

Every thinking being is, just because it is a thinking being, a simple substance

is a synthetic *a priori* proposition. •It's synthetic because it goes beyond the concept from which it starts, adding to the concept of a thinking being its way of existing. •And it's *a priori* because the predicate (namely simplicity) that it adds

to the concept of the subject can't be given in any experience. It would then follow ·from the supposition at the start of this paragraph· that *a priori* synthetic propositions are possible and admissible, not only (as I have said) in relation to objects of possible experience and indeed as principles of the possibility of this experience, but that they are applicable to things in general and to things in themselves—a result that would make an end of my whole critique and force me to go along with the status quo. But if we look closer we'll find that there is no such serious danger.

The whole procedure of rational psychology is dictated by a paralogism that is exhibited in this inference of reason:

- Anything that can't be *thought* otherwise than as subject doesn't *exist* otherwise than as subject, and is therefore substance. 411
- A thinking being, considered merely as such, can't be thought otherwise than as subject.
- Therefore a thinking being exists also only as subject, i.e. as substance.

Why is this a paralogism? What's wrong with it? The answer is that there is an ambiguity in the middle term, the one that occurs in both premises, namely: 'can't be thought otherwise than as subject'. The major premise uses this term unrestrictedly: it speaks of things that 'can't be thought other than as subject' *however* they are being mentally engaged with, including their being presented in intuition. But the minor premise concerns something that 'can't be thought otherwise than as subject' when thinking about itself as a subject of thought and the unity of consciousness and *not* when it is confronting itself through inner sense as something given in intuition. So the conclusion is reached invalidly, through a fallacy of ambiguity. 412

[In a long and very difficult paragraph Kant reminds us of claims he has defended in two parts of the *Analytic*; these,

he says, confirm that he is right in ‘resolving this famous argument into a paralogism’. Then:]

Refutation of Mendelssohn’s proof of the permanence of the soul

..413 The usual argument for the soul’s permanence ·or immortality· takes it that the soul is a simple being and argues that it therefore can’t go out of existence by *dissolution*; ·i.e. it doesn’t have *parts*, so it can’t be destroyed by being taken *apart*·. The acute Mendelssohn soon noticed that this argument doesn’t prove that the soul can’t go out of existence, because it might be supposed to go out of existence ·not by falling apart but· by *vanishing*. In his *Phaedo* he tried to plug that gap by arguing that the soul can’t undergo such a process of vanishing, which would be a true annihilation. [Kant’s point in that last clause is just that when a thing is merely dismantled that isn’t a true annihilation because its parts stay in existence, whereas what’s at issue now is a complete annihilation with nothing left behind.] His tactic was to argue that a thing that is *simple* can’t cease to exist. His argument goes like this (not a quotation):

414 The soul has no parts, so there is no plurality involved in it. So it can’t be diminished or lessened in any way, which means that it can’t *gradually* lose something of its existence, *gradually* going out of existence. If it could go out of existence, therefore, this would have to happen absolutely *suddenly*, with no time between a moment when it exists and a moment when it doesn’t—which is impossible.

But what Mendelssohn overlooked was this: even if the soul is simple, meaning that it doesn’t have parts that are external to one another, and thus doesn’t have **extensive** magnitude, we can’t deny that it like every other existing thing has **intensive** magnitude, i.e. a *degree* of reality in respect of its

faculties and indeed of all that constitutes its existence; and this degree of reality can diminish through all the infinitely many smaller degrees. In this way the substance whose permanence is at issue here might be changed into nothing not by dissolution but by gradual loss of its powers—by fading away to nothing, so to speak, ·rather than shrinking down to nothing·. For consciousness itself always has a degree, which can always be diminished; and the same must also be true of the power of being conscious of the self and likewise of all the other faculties. [In a footnote here, Kant seeks to allay an obscure worry about how the notion of •degrees of consciousness relates to the notion of •clarity of thought. The cure of the trouble is also obscure, and the footnote is omitted here.] Thus the permanence of the soul, regarded merely as an object of inner sense, remains unproved and indeed unprovable. Its permanence during life is, of course, evident in itself, because a human being is not only something that *thinks* but also something that is an object of the outer senses. But this won’t satisfy the rational psychologist, who sets out to prove from mere concepts the soul’s absolute permanence beyond this life.

· A LONG FOOTNOTE ·

[At this point Kant has a very long footnote, which it is more convenient to lift up into the main text. Here it is:]
Some philosophers think they have done enough to show that some new scenario of theirs is *possible* by defying everyone to prove that it contains a contradiction. (For example, those who think they can see the possibility of thought even after this life has stopped—although all they know about thought comes from empirical intuitions of our human life!) But those who argue in that way can be brought to a puzzled halt by the presentation of other ‘possibilities’ that are no less bold. [Kant wrote ‘no more bold’, but we’ll see below that this must have been a slip.] An example would be **(1)** the ‘possibility’ that

a simple substance might divide into several substances, and conversely **(2)** the ‘possibility’ that several substances might fuse together to form one simple substance. ·I shall comment on these in turn·.

(1) Something is divisible only if it is composite, but it might be a composite not of substances but only of degrees of the various powers of a single substance. We can certainly make sense of this thought:

All the powers and abilities of the soul, even that of consciousness, are reduced to one half in such a way that the substance still remains.

Well, there’s no *contradiction* in the thought:

All the powers and abilities of the soul are reduced to one half, with the half that the soul loses staying in existence outside it.

In that scenario, everything in the soul that is real and therefore has a degree—in other words, its entire existence, nothing omitted—has been halved; so another separate substance would come into existence outside it, ·to possess the half of everything lost by the original soul·. The many items that have been divided all existed before the division. They didn’t exist as many substances, but as many items that contributed to the reality of the substance, i.e. to how much existence it had. So its being *one* substance was therefore only a mode of existence, which in virtue of this division has been transformed into a plurality of subsistence. [That last sentence conservatively translates what Kant wrote. He may mean something like this:

What we are fundamentally talking about here are the many thoughts and powers etc. that have been split into two groups. The fact that they were all possessed by what we call ‘one substance’ is just a fact about how *they* existed—one might say that ‘they existed one-sub-ly’; whereas now, after the split-up, they exist two-sub-ly.

That seems to be the most plausible reading of the passage, though it doesn’t explain Kant’s using first ‘existence’ and then ‘subsistence’.]

(2) Several simple substances might be fused into one, with nothing being lost except the plurality of *things*, because the one substance would contain the degree of reality of all the former substances together. And perhaps the simple substances that appear to us as *matter* might produce the souls of children, i.e. producing them through a division of the parent souls considered as *intensive* magnitudes, with the parent souls making good their loss by fusing with new material of the same kind. (This division of the parent souls wouldn’t be mechanical or chemical, but rather would involve a causal influence unknown to us, of which mechanical and chemical influences were only appearances.)

I’m *not* saying that these fantasies are useful or valid; and the principles of my Analytic have warned us against using the categories (including that of substance) in any way except empirically. But if

the **rationalist** is bold enough to construct a self-subsistent being out of the mere faculty of thought, with no help from any permanent intuition through which an object might be given, doing this merely on the ground that the unity of self-awareness in thought can’t be explained in terms of something composite; instead of admitting, as he ought to do, that he can’t explain the possibility of a thinking nature *·at all·*,

why shouldn’t the **materialist**, though he can’t appeal to experience in support of his ‘possibilities’ either, be justified in being equally bold and using his principle to establish the opposite conclusion, while still preserving the formal unity *·of self-awareness·* upon which his opponent has relied?

·END OF LONG FOOTNOTE·

416 [Now we have three impossibly difficult paragraphs, which won't be paraphrased here. Here are three points that may help you to wrestle with this material yourself.

[(1) Kant is here using 'analytic' and 'synthetic' as labels not for •propositions but for •procedures, giving them senses that are *totally* different from what they have in the rest of this work. These new-to-us meanings were in fact quite standard before Kant's time (and perhaps afterwards, though his main use of 'analytic' and 'synthetic' as labels for propositions seems to have grabbed the limelight and pushed the other out of sight). The two procedures differ in *direction*: a •synthetic study of a body of doctrine starts with its final output, the theories that constitute it, and works backwards to the reasons for those theories, the reasons for those reasons, and so on back to the ultimate basis for the whole thing; whereas an •analytic study starts with what is epistemically basic—what one knows at the start—and proceeds from there to consequences, then consequences of those consequences, and so on to the final theories. [In a footnote on page 173 we saw Kant using 'analytic' and 'synthetic' in different though related senses, namely as labels for methods of exposition.]

[(2) In this present passage, Kant describes a •synthetic' approach to rational psychology as expressed in the displayed quartet of propositions on page 176 above: the approach starts from 'All thinking beings are substances' and works its way 'backwards' through the other three members of the quartet, and eventually back to the all-purpose *I think* that assures of our existence. He doesn't *make* much of this, except to remark that it commits rational psychology to 'problematic idealism', i.e. to the thesis that 'Is there a material world?' is a permanently open question.

[(3) He also describes an •analytic' approach that starts with *I think* and works its way forward to the simplicity and substantiality of the soul. He uses this to generate

another displayed quartet (which is elegant rather than helpful), and to make the claim that we now see rational psychology straining to give information about the nature of thinking beings, and being thwarted because it can't adopt the *materialist* view that thinking things are material things, or the *spiritualist* view that there are only immaterial thinking things. Then:]

..421

So there's no informative *doctrine* of rational psychology, but only a *discipline*. This sets impassable limits to speculative reason, limits that keep us from •throwing ourselves into the arms of a soulless materialism or, on the other side, •losing ourselves in a spiritualism that must be quite unfounded so long as we remain in this present life. Without providing any positive doctrine, rational psychology reminds us that we should regard

reason's refusal to give a satisfying answers to our inquisitive questions about things that are beyond the limits of this present life

as being

reason's hint that we should divert our self-knowledge from fruitless and extravagant speculation to fruitful practical use.

Though in such practical use reason is never directed to anything but objects of experience, it gets its principles from a higher source, and sets us to behave as though our destiny reached infinitely far beyond experience, and therefore far beyond this present life.

All this makes it clear that rational psychology owes its origin simply to a misunderstanding, in which •the unity of consciousness that underlies the categories is mistaken for •an intuition of the ·thinking· subject as an object, and is then brought under the category of substance. This unity is really only unity in *thought*, and on its own it doesn't present any object; so the category of substance can't be applied

422

to it, because that category always presupposes a given *intuition*. Therefore, this ·thinking· subject can't be known. The subject of the categories cannot, just by thinking them, acquire a concept of itself as an object of the categories. [The unity of consciousness 'underlies' the categories, and the thinking self is the 'subject' of them, in the sense that the only way to use any category is in a thought that one has, an *I think*, and in this context the 'subject' is the being that *has* the thought.]. . . .⁶

There's a desire to get knowledge that will extend beyond the limits of possible experience while also furthering the highest interests of humanity; speculative philosophy has claimed to satisfy it; and we can now see that the claim is based on deception. Still, the severity of my critique has

⁶ As I have already said, the *I think* is an empirical proposition, and contains within itself the proposition 'I exist'. But I can't say 'Everything that thinks exists', because that would imply that the property of thought makes everything that has it a necessary being. So I can't regard my existence as **inferred from** the proposition 'I think', as Descartes maintained; because if it were inferred from that premise there would also have to be the premise 'Everything that thinks exists'. Rather, the proposition that I exist is **identical with** *I think*. The *I think* expresses an indeterminate empirical intuition, i.e. perception. . . . But the *I think* precedes the experience that would be needed to apply the category to the object of perception. ·That doesn't make trouble for the *I exist* that is identical with *I think*, because· the existence involved in this isn't yet a category. A genuine category can't be applied to an indeterminately given object, but only to one of which we have a concept and are asking whether it exists outside the concept. In our present context an indeterminate perception simply points to something real that is given—given to thought as such and therefore not given •as appearance or •as a thing in itself (*noumenon*) but simply •as something that actually exists and is tagged as such in the proposition, *I think*. When I called the proposition *I think* an 'empirical proposition', I didn't mean that the *I* in this proposition is an empirical representation. On the contrary, it is purely intellectual, because it belongs to thought in general, ·i.e. to *all* thought, including thought that has no empirical content·. . . .

rendered reason good service by proving that it's impossible to arrive dogmatically at any results—concerning any object of experience—that lie beyond the limits of experience. ·Why is that a 'service'?· Because it secures reason against any possible assertion of an opposing view. The defence against the opposing view can be seen as having four stages: **(1)** try to prove that the proposition one is defending is necessarily true; **(2)** find that this can't be done; **(3)** explain why it can't be done, namely because of the unavoidable limits of our reason; then **(4)** make the opponent back down because he too has been trying to infringe the limits.

But this doesn't take anything away from the right, indeed the necessity, of believing in a future life in accordance with the principles of the •practical use of reason, which is closely bound up with its •speculative use. The merely speculative proof has never had the slightest influence on ordinary common-sense. It stands on the tip of a hair, so precariously that even the schools can stop it from falling only by keeping it spinning around like a top; so even they can't see it as providing an enduring foundation on which something might be built. My critique doesn't at all lessen the value of the proofs that work for the world at large; indeed it increases their clarity and natural force by stripping away those dogmatic pretensions. Here is why:

Those arguments place reason in its own special domain, namely, the order of ends ·or purposes·, which is also an order of nature. Now, because reason is in itself not only a theoretical but also a practical faculty, it isn't tied down to natural conditions and can legitimately expand the order of ends—and with it our own existence—beyond the limits of experience and of life.

·And here is *how* it does that·. When dealing with the *analogy with the nature* of living beings in this world, reason has

to accept the principle that no organ, no faculty, no impulse, no *anything* is either superfluous or disproportioned to its use, so that everything is exactly conformed to the end or purpose that is destined for it. Now, if we were to judge things on the basis of this kind of analogy, we would have to regard man. . . . as the only creature who is excluded from this order of ends. [Although he hasn't said so, Kant must here be thinking of the natural 'order of ends' as the way every feature of any organism is fitted for its 'end' of its own survival and flourishing.] Think about man's natural endowments, not merely his talents and the impulses to enjoy them, but above all the moral law within him. These go far beyond any benefit or advantage he could get from them in this present life—so far beyond that they teach him to prize the mere consciousness of a righteous attitude as being supreme over all other values, quite apart from any advantage it might bring him and apart even from the shadowy reward of posthumous fame. They make him feel an inner call to fit himself, by his conduct in this world and by renouncing many of its advantages, for citizenship in a better world that he has in his idea. [Here, as always in the Dialectic, 'idea' is a Kantian technical term, meaning 'concept of reason'.] This powerful and incontrovertible proof is reinforced by our ever-increasing knowledge of •purposiveness in everything we see around us, and by contemplation of the •immensity of creation, and therefore also of a certain •limitlessness in how far our knowledge might be extended and in our •drive to extend it accordingly. All this still remains to us after the critique has done its work; but we must give up all hope of grasping the necessary continuance of our existence merely from our theoretical knowledge of ourselves.

Concluding the solution of the psychological paralogism

The dialectical illusion in rational psychology arises from the confusion of an idea of reason—the idea of a pure intelligence—with the completely featureless concept of a thinking being as such. I think *myself*—in the all-purpose *I think*—for the sake of a •possible experience, at the same time abstracting from all •actual experience; and from my ability to do this I infer that I can be conscious of my existence even apart from experience and its empirical conditions. In doing this I am confusing •the possible *abstraction* from all the empirical details of my existence with •a supposed consciousness of a possible *separate existence* of my thinking self, and that leads me to think I have *knowledge* that what is substantial in me is the transcendental subject. But really all that I have in thought is the unity of consciousness. . . .

The task of explaining how the soul relates to the body doesn't properly belong to •the psychology I'm discussing here, because •it aims to prove the personhood of the soul even when it is not related to the body (i.e. after death), so that •it is *transcendent* in the proper sense of that term. It does indeed occupy itself with an object of experience, but only in the aspect of it in which it ceases to be an object of experience. My doctrine, on the other hand, does supply a sufficient answer to this question about how the soul relates to the body, including the question of whether and how they could act on one another causally. It's generally recognised that what makes that problem especially difficult is the belief that

(1) the object of inner sense (the soul)—the formal condition of its intuition = **time** only, is basically unlike

(2) the objects of the outer senses (·bodies·)—the formal condition of their intuition = **space as well as time**.

428 But the two kinds of objects differ from one another not intrinsically but only in so far as (2) *appears* externally to (1); whatever thing-in-itself underlies (2) the appearance of matter may after all not be so radically unlike (1) the thinking subject. When you bear that in mind you'll find that this difficulty vanishes. The only question that remains is this:

•How is it possible for *any* two substances to interact causally?

But that question lies outside the domain of psychology; and you won't hesitate to agree, in the light of what I have said in the Analytic regarding basic powers and faculties, that the question lies outside the field of all human knowledge.

Moving across from rational psychology to cosmology

The proposition *I think* or *I exist thinking* is an empirical one. So it is based on empirical intuition, and thus on how the object of the intuition—which in this case is the *I*, the thinking subject—presents itself as an appearance. It seems to follow on my theory •that the soul, even in its thinking, is completely transformed into appearance, and •that in this way our consciousness itself, as being a mere illusion, must amount to nothing.

[Kant next discusses a different *I think*—the all-purpose one that is involved in *any* thought, in thought *as such*, the *I think* that has been the focus of the Paralogisms. *This* is a logical puller-together [*logische Funktion*, 'logical function'] of whatever variety of elements intuition may present me with; it's something that I actively *do*, not something that I sensibly

= passively *encounter*. So it doesn't exhibit ·this *I*·, the subject of consciousness, as an appearance; it doesn't exhibit me *as* anything at all; it is involved in *all* the intuitions that I have, both sensible (passive) and intellectual (active), so it can't itself have any features that would tie it to one or other of those two kinds of intuition. In this *I think* I don't, of course, represent myself to myself *as I am in myself*, but nor do I represent myself to myself *as I appear to myself*. And if represent myself as (1) a subject of thoughts or as (2) a ground of thought, I am not here using the categories of (1) substance or of (2) cause. That's because the categories are operators on materials supplied by our sensible intuition, but the *I think* that we're discussing here isn't among those materials—it's the *doer* that pulls the materials together under a single consciousness. Kant ends this amazingly difficult paragraph thus:] In the consciousness of myself in mere thought I am the *being itself*, without providing any facts about myself for me to think about.

The proposition *I think*, understood as amounting to *I exist thinking*, is no mere logical puller-together; it says something *about* the subject (which in this case is also the object) regarding its existence; it requires an inner-sense intuition that presents the object not as a thing-in-itself but merely as an appearance. So here we have not simply 430 (1) activity of thought but also (2) ·passive· receptiveness of intuition—i.e. we have (1) the thought of myself applied to (2) the empirical intuition of myself. Let's pretend that I want information about how my thinking subject goes about its pulling-together work when it is applying the categories of substance, cause and so on. How can I go about enquiring into this? •I need more than merely the all-purpose *I* that accompanies all thought, because I'm looking for *information about* myself as an actively thinking subject, and the all-purpose *I* contains no information. •And my inner-sense

intuition of myself as an appearance won't do the job either. Why not? Because there would be a kind of circularity in trying to learn about the pulling-together work of the *I* from an intuition that is itself an appearance pulled together by the *I*. So what I need to carry out this pretended inquiry is an intuition of myself that enables me to know myself as a *noumenon*. And that's impossible, because the only intuition that I have of myself is a sensible one providing only data of *appearance*. . . .

Suppose that the following were true—as indeed it is:

We eventually discover—not in experience but in certain *a priori* laws of the pure use of reason (laws that are not merely logical rules but concern our existence)—grounds for regarding ourselves as *legislating* completely *a priori* in regard to our own *existence*, settling what sort of things we are. This reveals in us a *spontaneity* through which we determine our reality with no need for the conditions of empirical intuition. And we also become aware of something else. Although our existence can't be *thoroughly* determined other than through sensibility, we become aware that the consciousness of our existence contains *a priori* something that can—by virtue of a certain inner *power*—serve to determine our existence in its relations to a non-sensible intelligible world.

This wouldn't contribute *anything* to the project of rational psychology. In this marvellous *power* that my consciousness of the moral law first revealed to me, I would have for the determination of my existence a principle that is purely intellectual. But what predicates would I use in doing this? They would have to be just the ones that are given to me in sensible intuition; which means that as regards rational psychology I would be exactly where I was

before practical reason was brought into the story. I can't have knowledge of myself unless I can make use of my concepts-of-understanding such as substance, cause, and so on; and I can't give meaning to *them* except with help from sensible intuitions; and sensible intuitions can never help me to move beyond the domain of experience. Still, in my practical thinking (which is always directed to objects of experience), it is all right for me to apply these concepts to freedom and the subject that has the freedom, giving them meanings that are analogous to the meanings they have when used theoretically. In doing this, however, I would be using these concepts merely to capture the *logical* functions of

•subject and predicate, ground and consequence,
•and not the full-fledged schematised concepts of

•substance and property, cause and effect.

That would enable me to think of the acts that I perform in conformity to moral laws as always capable of being explained in terms of the laws of nature and the categories of substance and cause, although they come from an entirely different source. I needed to make these points so as to head off any misunderstanding of my doctrine about our *appearing* to ourselves in self-intuition. I'll revert to these matters later.

Chapter 2: The antinomy of pure reason

433 In the introduction to the Transcendental Dialectic I showed that all the transcendental illusion of pure reason rests on dialectical inferences that can be classified on the basis of the three forms of inference-of-reason, just as the categories can be classified on the basis of the four forms of judgment. [Kant now repeats his earlier claim—see page 176—that (1) the paralogisms are arguments starting from a premise using the subject-predicate form, (2) the antinominal fallacious arguments start from a premise that is hypothetical in form, and (3) the third (theological) fallacious arguments have a special relationship to the disjunctive form. His formulations remind us that (1) concerns *subjective* conditions while (2) concerns *objective* ones. I get into (1) by thinking about *myself*, into (2) by thinking about *the world out there*. Then:]

But there's something we should especially notice; ·it's another enormous difference between (1) and (2)·. Transcendental paralogism produced a purely one-sided illusion concerning the idea of the subject of our thought. The concepts of reason don't cause any illusion that gives the slightest support to the opposing assertion. ·i.e. to the denial of the conclusion of the paralogism, thinking especially of 'The soul is *not* simple', which would open the door to the thesis that the soul is a material thing·. So the only position that the paralogism claims to support is pneumatism [= 'the thesis that the soul is immaterial'], though of course the fiery ordeal of critical investigation makes that 'support' go up in smoke.

A completely different situation arises when reason is applied to (2) the *objective* synthesis of appearances. For in this domain, however hard reason may try to establish its principle of unconditioned unity (indeed making the principle

seem quite plausible), it also produces lines of thought that *go against* that principle, falling into such contradictions that it has to back off from its demand for such unity in the cosmological domain.

We are confronted here by a new phenomenon of human reason—an entirely *natural* antithetic into which reason stumbles •unavoidably, •quite of its own accord, •without being led on by sophisticated arguments or enticed into traps set for the unwary. It does guard reason from the slumber of a false belief such as is generated by a purely one-sided illusion ·like that of the paralogisms·; but it subjects it to the temptation either •to abandon itself to a sceptical despair or •to defend one of the two sides dogmatically and stubbornly, refusing to give the other side its day in court. Either attitude is the death of sound philosophy. . . .

434 Before ushering in the various forms of opposition and dissension to which this conflict or antinomy of the laws of pure reason gives rise, I offer a few remarks to explain and justify the method I'm going to adopt in dealing with this subject. I label as a 'world-concept' any transcendental idea that concerns absolute totality in the synthesis of **appearances**. I have two reasons for this: •the concept of the world-whole, though itself only an idea, rests on this unconditioned totality; and •such concepts concern only the synthesis of appearances, and thus only empirical synthesis. Accordingly, just as the paralogisms of pure reason formed the basis of a dialectical psychology, so the antinomy of pure reason will reveal the transcendental principles of a supposed pure rational cosmology [= 'theory of the whole world']. But it won't be trying to show this 'science' to be valid and to adopt it. As the title '*conflict* of reason' indicates well enough, the object of the exercise will be to display it in all its flashy but false illusoriness, as an idea that can never be reconciled with appearances. (·It's obvious that the label 'world-concept'

doesn't apply to the idea of the transcendental *I* or have any role in the paralogisms; but it also doesn't belong in the third of the three basic kinds of dialectical illusion either. When we are dealing with absolute totality in the synthesis of **the conditions of all possible things in general**, —as we are in the third kind of illusion.—there arises an *ideal* of pure reason which, though it may indeed stand in a certain relation to the world-concept, is quite distinct from it.)

1. System of cosmological ideas

To clear the way for enumerating these ideas with systematic precision according to a principle, I need to make two points. **(1)** The only source for pure and transcendental concepts is the understanding. Reason really doesn't generate any concept. The most it can do is to free a concept of *understanding* from the unavoidable limitations of possible experience, thus trying to extend it beyond the limits of the empirical, though still in a certain relation to the empirical. Here's how it does this: [In reading what follows, bear in mind that Kant is concerned with such condition/conditioned relations as cause/effect, part/whole, earlier-time/later-time. So one example of the absolute **totality** of the conditions of a given conditioned would be: the set of **all** the past events that are causally related, by however long a chain, to a given present event.] For a given conditioned item, reason demands *absolute totality* on the side of the conditions that. . . the understanding finds for all appearances, and through this demand it converts the category into a transcendental idea. How is that so? Well, the only way to make the tracking of empirical conditions extend as far as the unconditioned is by making it absolutely complete; and there can't be experience of any such absolute totality, which is why the unconditioned is never to be met with in experience, but only in the idea. Reason makes this

demand on the basis of the principle that

•If some conditioned item *x* is given, then the entire sum of *x*'s conditions, and consequently the absolutely unconditioned, is given (because that unconditioned totality is what has made it possible for *x* to exist).

Two important things follow from this. **(i)** Because each condition/conditioned relation is an instance of one of the categories (e.g. cause/effect), it follows that the transcendental ideas of reason are simply categories extended to the unconditioned. That enables us to set them out in a table arranged according to the four headings of the table of the categories [see page 52]. **(ii)** Only some of the categories enter into this match-up with the ideas of reason, namely the ones that pulls things together into a *series* of conditions, each member of which is subordinated to its immediate neighbour, not co-ordinated with it. To understand why, you have to grasp the basic thought that if *x* is a conditioned item subordinated to condition *y*, then *y* in some way generates or creates *x*. The absolute totality that reason demands is the totality of the

(a) ascending series of conditions related to a given conditioned *x*,

·i.e. the series consisting of the condition *y* to which *x* is subordinated, the condition *z* to which *y* is subordinated, and so on, back *up* through the series of conditions. Reason doesn't demand totality of the

(b) descending series of consequences of an item *x*,

·i.e. the series consisting of something *y* that is subordinated to *x*, *z* that is subordinated to *y*, and so on *downwards*. Nor does reason demand totality in reference to the

(c) aggregate of co-ordinated conditions of at item *x*,

·this being a set of conditions that don't fall into either an ascending or a descending series. Why? Because in case **(a)** when *x* is given, all its conditions are presupposed as

having given rise to *x*, and are considered as given together with it. In case **(b)** the downward series of consequences of *x* don't give rise to *x* or make *x* possible, so our intellectual engagement with *x* doesn't require us to give any thought to that series, e.g. worrying about whether it has a last member or not; reason simply isn't interested in that. I'll return to **(c)** co-ordinated conditions a little later.

[Kant illustrates **(a)** and **(b)** with the example of *time*. Here we are in *today*; this had to be reached through yesterday, which had to be reached through the day before, and so on backwards. So the entire series of ever-earlier times is 'presupposed' by our confrontation with today, and reason tells us to accompany our thoughts about today with a thought of the totality of that series of ever-earlier times. On the other hand, the existence of today doesn't presuppose tomorrow, nor does tomorrow presuppose the day after; so the series of ever-later times is not something reason challenges us to think about in its totality. Reason has an interest in the question 'Was there a first time?' but not in the question 'Will there be a last time?' Then:]

..438

I shall use the label 'the *regressive* synthesis' for the synthesis of the ascending series from the given appearance *x* to its nearest condition *y*, then to *z* the nearest condition of *y*, and so on; and I'll label as 'the *progressive* synthesis' the series that runs in the opposite direction. . . . So there we have it: the cosmological ideas deal with the totality of the regressive synthesis, the series of antecedents, not of consequents. You might set up a 'problem of pure reason' concerning the progressive form of totality—involving such questions as 'Will there be a last time?', 'will there ever be an effect that doesn't cause anything?'—but that would be something you *chose* to think about, not something you *had* to think about.

(1) AN IDEA SUPPOSEDLY RELATED TO THE CATEGORIES OF QUANTITY.

[The categories of quantity as announced on page 52—unity, plurality, totality—are irrelevant to what we are about to encounter, which is all about time and space. Kant papers over the gap by referring to time and space as 'quanta', i.e. items that permit of the notions of more and less.] In arranging the table of ideas in accordance with the table of categories, we first take the two original *quanta* of all our intuition, *time* and *space*. **Time** is in itself a series, and it is also the formal condition of all series—i.e. the right way to think about any series *x*, *y*, *z*, . . . is in the form 'x and then y and then z. . .'. With regard to any given time, e.g. the present, we can distinguish *a priori* the antecedents (the past) from the consequents (the future). So the transcendental idea of the absolute totality of the series of conditions of any given time refers only to all *earlier* times; and the idea of reason requires that the whole of previous time, which is a condition of the given moment, has to be thought of as being given in its entirety along with that given moment. Now in **space**, taken in and by itself, there is no distinction between progress and regress. For as its parts are co-existent, it is an *aggregate*, not a *series*. The present moment can be regarded only as conditioned by past time, never as conditioning it, because this moment comes into existence only *through* past time, or rather through the passing of the preceding time. But as the parts of space are co-ordinated with, not subordinated to, one another, one part is not the condition of the possibility of another; so space doesn't in itself constitute a series, as time does. However, when we apprehend space we mentally •pull together the different parts of space, and •that procedure is successive: it occurs in time and contains a series. [Kant now offers two obscure sentences whose gist seems to be this: any region of space *x* can be regarded as conditioned by its limits

439

(without the limits it wouldn't exist), and those limits are its shared boundary with some larger region y within which x is nested; so we can think of the sequence of regions x, y, z, \dots , of which each item contains the one before it, as a *regressive* series analogous to the series of causal ancestors of a given event. He then continues:] In respect of boundary-setting, therefore, the advance in space is also a regress; so we do have here a regressive or ascending series of conditions, so that space too falls under the transcendental idea of the absolute totality of the synthesis in the series of conditions. I can as legitimately ask about the absolute totality of appearance in space as about the absolute totality of appearance in past time. Whether we can ever *answer* such questions is something we'll look into later.

(2) ·AN IDEA RELATED TO A CATEGORY OF QUALITY·

[The categories of quality are •reality, negation, and limitation; Kant fastens on a special case of the first of these, ignoring the other two.] •Reality in space, i.e. *matter*, is conditioned. Its internal conditions are its parts. ·Consider for example a brick that can be divided into 100 cubic-inch parts; the brick as a whole is an upshot of those parts, they make it possible and indeed actual·, it is *conditioned by* them. If we think of the brick first in terms of 1-inch³ parts, then we can also think of in terms of 0.1-inch³ parts, then 0.01-inch³ parts, and so on down into ever smaller and more remote conditions of the brick. So there is here a regressive synthesis, a series of ever smaller and ever more remote conditions of the brick—the kind of series whose absolute totality is demanded by reason. The only way to satisfy reason's demand would be to produce a *completed* division, and that would have to be either •one that went on for ever, with no smallest member, or •one that ended in something simple, i.e. a thing having size but not having parts. (In the former case, matter would vanish into •nothing; in the latter it would vanish into •something

that isn't matter any more ·because all matter *must* have parts·.) Here also, then, we have a series of conditions, and an advance to the unconditioned.

(3) ·AN IDEA RELATED TO A CATEGORY OF RELATION·

[The categories of relation are substance-property, cause-effect, and interaction ('community'). Kant here fastens on cause-effect, but first explains why the other two are not relevant to ideas of reason.] As regards the categories of real relation between appearances, the •relation of a substance to its properties doesn't have the right shape for a transcendental idea to be based on it, because it doesn't offer any regressive series of conditions which reason could demand be carried to its completion. Several properties that are possessed by a single substance are co-ordinated with each other, ·are on the same level·, so they don't constitute a series. You may think 'Aren't they subordinate to the substance that has them?' The answer is No. A substance's properties or 'accidents' are *the way the substance exists*; ·it's just not the case that the substance is a condition of the properties·. [Kant goes on to say that the substance/property category might seem suitable for an idea of transcendental reason, and this would be the idea or concept of *the substantial*. That's an idea of reason all right, Kant says. It is indeed the idea or concept of *object as such*, which is involved in our thinking the transcendental subject apart from all predicates, i.e. involved in the thinking with the transcendental contentless *I* that is at work in the paralogisms. But it has no place *here*, because it doesn't involve any series of conditions which reason could demand to have completed. Then:] That holds also for substances in •interaction with one another ('community'). Among such substances there are none that are subordinate to others; so they don't form a series; so reason's demand for completeness of series of conditions gets no bite on them. There thus remains only the category

..440

441

442 of ·one-way· •causality. That does present us with a series 443
of causes of a given effect, a series that moves upwards from
the effect to its conditions, to their conditions, and so on,
enabling us to answer the question of reason. [Kant really
does say ‘answer’ (*antworten*), though one would have expected him to
say only that such a series enables us to *ask* reasons’s question.]

(4) ·AN IDEA RELATED TO A CATEGORY OF MODALITY·

The only way to get a series out of the categories of
modality—the concepts of the possible, the actual, and
the •necessary—is by fastening on •necessity, and having
the following thought: Anything that exists *contingently*
must always be regarded as conditioned by a condition
relative to which it is •necessary; if this condition also
exists contingently, then it must in turn be conditioned by
(and necessary relative to) a further condition. . . and so on
upwards, backwards, with reason demanding unconditioned
necessity—something whose existence is necessary in itself,
not necessary relative to something else.—and that can be
supplied only in the totality of the series. This requirement
of a condition for everything that exists contingently is laid
down by a rule of the understanding. [Kant doesn’t say what
rule this is. It ought to come from the so-called Postulates of Empiri-
cal Thought’ (pages 123–6); but they don’t yield any such result, being
‘nothing but explanations of how the concepts of possibility, actuality
and necessity work in their empirical employment’. Mightn’t the rele-
vant ‘rule’ be the second analogy, which says that all appearances are
caused? No! That has already been used in the preceding paragraph;
and anyway we’ll see that what Kant does with this present notion of
condition-that-makes-x-necessary is quite different from the regressive
series of causes.]

Thus, when we pick out the categories that necessarily
lead to a series in the synthesis of the manifold, we find that
there are exactly four cosmological ideas, corresponding to
the four trios of categories:

- (1) Absolute completeness of the *Composition* of the given whole of all appearances.
- (2) Absolute completeness in the *Division* of a given whole in the ·domain of· appearance.
- (3) Absolute completeness in the *Origination* of an appearance.
- (4) Absolute completeness as regards *Dependence of Existence* of the changeable [here = ‘contingently existing’] in the ·domain of· appearance.

It’s important to bear in mind that the idea of *absolute totality* concerns only. . . .*appearances*, not the understanding’s pure concept of a totality of *things as such*. . . .

And another point: What reason is really looking for in this synthesis of conditions—a synthesis that forms a series, a backwards series—is solely the *unconditioned*. The aim 444
is to have the series of premises in such a complete form that there won’t be any need for any other premises to be presupposed. This *unconditioned* is always contained in the *absolute totality of the series* as represented in imagination. But this utterly complete synthesis is only an idea, because we can’t know in advance whether such a synthesis of appearances is possible. If we represent everything only through pure concepts of understanding, leaving sensible intuition out of it, we can indeed say straight off that for a given conditioned item the whole series of conditions. . . .is likewise given. The conditioned item is given only *through* the series of its conditions. But when we are dealing with appearances, we find that a special constraint enters the picture because of the fact that conditions of appearances are given through the *successive* synthesis of the manifold of intuition—a synthesis that has to be made complete by working backward along the series. Whether *this* completeness is possible in sensibility is a further problem. Reason

has the idea of this completeness, independently of whether
 .445 we can connect it with any adequate empirical concepts. . . .
 And it pursues this completeness as a way of pursuing *the*
unconditioned.

We can think about this unconditioned item in either
 of two ways. **(a)** We can think of it as consisting in the
 entire series, in which *each* member is conditioned and
 only the totality of them is absolutely unconditioned. This
 is the *infinite* regress; it has no limits, no first member;
 and is given in its entirety. But the regress in it is never
 completed—i.e. we never complete it—and can only be
 called potentially infinite. **(b)** We can think of the absolutely
 unconditioned item as being *in* the series—a part of it to
 which the other members are subordinated but which isn't
 itself subordinated to or conditioned by any other condition.
 446 On this view, there is a first member of the series. We have
 labels for each of these first members:

- (1t)** ·ever earlier· past times—**the beginning of the world;**
- (1s)** ·ever larger regions of· space—**the limit of the world;**
- (2)** ·ever smaller· parts of a given limited whole—**the simple;**
- (3)** ·ever earlier· causes—absolute **self-activity** (freedom);
- (4)** explanations of the existence of contingent things—
 absolute **natural necessity**.

We have two expressions, 'world' and 'nature', which some-
 times coincide. Here are their meanings:

'**the world**' signifies the mathematical sum-total of all ap-
 pearances and the totality of their synthesis, both **(1)** moving
 to items that are ever larger and **(2)** moving to parts that are
 ever smaller; and

'**nature**' signifies that same world when viewed as a dynamical
 whole,⁷ ·a whole in which things *happen*—**(3)** and **(4)**·.

⁷ [Kant has a footnote here, explaining that 'nature' can be used 'adjecti-
 vally' to refer to the whole system of happenings and dependences
 and the laws of nature governing it, or 'substantivally' to the great

When we are interested in nature, we aren't concerned
 with the spatio-temporal size of the world or of its parts; our
 interest is in the unity in the *existence* of appearances, ·i.e.
 in the connecting-up and hanging-together of all the facts
 about what happens and about what contingently exists. . . .

Some pages back, I labelled the ideas we are now dealing
 with as 'cosmological' ideas—i.e. *world*-ideas—and this is
 a good label, for two reasons. One is that we use the word
 'world' to stand for the sum of all *appearances*, and that's
 what these present ideas aim at—the unconditioned in the
appearances. The other reason is that when we use the term
 'world' in its transcendental sense, it refers to the absolute
totality of all existing things, and again that's what these
 present ideas aim at—the *completeness* of the synthesis
 (even though that is reachable only in the regress through
 the conditions). These ideas are all transcendent, but in
 a special way: they don't surpass appearances by talking
 about noumena, but only by going too far for any possible
 experience to keep up with them. The mis-match between
 them and possible experience is a matter not of *kind* but of
degree. So it really is all right to call them cosmical concepts,
world concepts. . . .

2. Antithetic of pure reason

[This numbered item runs to page 225 where we'll encounter **3**.] .448
 I use the term 'antithetic' to mean 'conflict between dogmatic
 doctrines. . . .where neither side can establish superiority
 over the other'. So the antithetic ·I'm going to discuss here·
 doesn't concern one-sided assertions, but rather the conflict
 of the doctrines of reason with one another and the causes
 of this conflict. The transcendental antithetic is an inquiry

big thing to which or in which these happenings occur and these
 contingent things exist.]

449 into the antinomy of pure reason, its causes and its upshot. If in using our reason we don't—as the principles of understanding would have us do—confine ourselves to objects of experience, but venture to extend these principles beyond the limits of experience, there arise *sophistical* doctrines that can't hope for confirmation in experience and needn't fear refutation by it. Each of these doctrines •is internally free from contradiction, and also •finds in the very nature of reason conditions of its necessity; the only trouble being that the opposite doctrine is also free from self-contradiction and •seemingly• well supported.

The questions that naturally arise in connection with such a dialectic [see explanation on pages 45–6] of pure reason are the following: **(i)** In what propositions is pure reason unavoidably subject to an antinomy? **(ii)** What is this antinomy? **(iii)** Is there, despite this conflict, a way for reason to reach certainty? and, if so, what is it?

450 So a dialectical doctrine of pure reason has two features that no •other• sophistical proposition has. •It arises out of a question that human reason *has to* encounter as it goes about its work, not one that is merely *chosen* for some special purpose. •The illusion involved in such a doctrine (and in its opposite) is not the kind of *constructed* illusion that vanishes as soon as it has been detected, but a *natural* and unavoidable illusion which, even after it has stopped leading us into error, still continues to delude though not to deceive us—the illusion can be rendered harmless but it can't be eradicated.

What such a dialectical doctrine will be about is not •the unity of understanding in empirical concepts, but rather •the unity of reason in mere ideas. Since this unity of reason involves a synthesis according to rules, it must conform to the understanding; and yet as •demanding• absolute unity of synthesis it must at the same time harmonise with reason.

But the conditions of •this unity are such that when •it is adequate to reason it is too big for the understanding; and when •it's suited to the understanding it is too small for reason! So we have here a conflict that we can't avoid, try as we may.

So these sophistical assertions reveal a dialectical battlefield in which the side permitted to open the attack always wins, and the side forced onto the defensive is always defeated. It's like the situation with knights at arms who, however bad or good their cause is, can be sure of carrying off the laurels provided they arrange to be allowed to make the last attack, and don't have to withstand a new onslaught from their opponents. . . . As impartial umpires, we must set aside the question of whether the cause for which this or that contestant is fighting is good or bad; they'll have to decide *that* for themselves. . . . 451

This is an approach [*Method*] in which we watch—or rather *provoke*—a conflict of assertions, not so as to decide in favour of one of the sides but •so as to understand the conflict. Specifically•, we want to investigate whether *this* is the case:

What they are quarrelling about is a deceptive appearance that neither side could grasp even if there were no opposition to be overcome, so that their conflict *can't* lead to any result.

We could call this the 'sceptical approach'. It is nothing like *scepticism*, which is a principle of technical and scientific ignorance that undermines the foundations of all •knowledge, and tries in every way it can to destroy •its reliability and steadfastness. The sceptical approach aims at *certainty*. It tries to discover the point of misunderstanding in disputes that are sincerely and competently conducted by both sides. 452 It's like the way in which

•wise legislators study •the perplexities that judges run into when trying cases, in order to •learn about the defects and ambiguities of their laws.

Compare that with what we can do with

•our limited wisdom: study •the antinomy that occurs in the application of laws, this being the best way to •evaluate the legislation that has given rise to them.

When reason is going about its abstract business it doesn't easily become aware of its errors; our sceptical approach enables us to alert reason to what is at issue when it decides on its principles

But it's only for transcendental philosophy that this sceptical approach is essential; although it can't be dispensed with here, it can be in every other field of enquiry. It would be absurd to adopt it in •mathematics, because *there* it's impossible for false assertions to be concealed, made invisible, because mathematical proofs must always proceed under the guidance of pure intuition, with every step along the way self-evident. In •natural science a doubt may cause the scientist to pause, and that can be useful; but in that domain there can't be any misunderstanding that isn't easily removed; and the final resolution of any dispute, whether found early or late, must come from experience.

453 •Morality can also present all its principles along with their practical upshots in concrete examples drawn from the real world or at least from possible experiences; and that enables moral studies to steer clear of the misunderstandings that can come from abstraction. But it's quite otherwise with transcendental assertions that claim to report on what is beyond the domain of all possible experiences. *Their* line of abstract thought can't be given in any *a priori* intuition (•like mathematics•), and any errors they contain can't be detected through any experience (•like natural science•). So transcendental reason can't be tested in any way except

through the attempt to harmonise its various assertions, and for this we must allow a free and unhindered development of the conflicts into which they fall. Now I'll set the stage for that.

[Kant presented each 'conflict' with the Thesis material on the odd-numbered pages and the Antithesis material on the facing even numbered pages; some editions have them in facing columns on the same pages. But nothing is gained by having thesis and antithesis glaring at one another; so the present version will give the material in the order: •statement and proof of thesis, •statement and proof of antithesis, •remarks on thesis, •remarks on antithesis. The marginal numbers will be corresponding disordered.]

First antinomy

Thesis: **The world has a beginning in time, and is also limited as regards space.** 454

Proof: Suppose that the world *doesn't* have a beginning in time. From this it follows that

•up to any given moment an eternity has elapsed; an infinite series of states of affairs has happened in the world, one after another.

But what it is for a series to be infinite is that it can never be completed through any one-after-another process. So it's impossible for an infinite world-series to have occurred, •because to say that it *has* occurred is to say that it is now completed•. Therefore, the world can't exist now unless it began at some time in the past. This was the first point to be proved.

As regards the second point, once again assume the opposite:

•The world is an *infinite* given whole of coexisting things.

456 Now, when something isn't given in intuition as within certain limits, the only way we can think about how big it is is through the synthesis of its parts, and the thought of *its size* has to come from the thought of *completing* the process of going through it part by part.⁸ Thus, if we are to have the thought of the world that fills all spaces—thinking of this as a whole—we must think of the successive run-through of the parts as *completed*, and that's the thought of an infinite time's having passed in the enumeration of all coexisting things. This, however, is impossible. Therefore, an infinite aggregate of actual things can't be regarded as a given *whole*; so there can't be a thought of *all of it out there, right now*. So the world's spatial extent is not infinite, but is enclosed within limits. This was the second point in dispute. [Just to make sure this is clear: The thesis-arguer argues first that there can't be a coherent thought of a now-complete temporally past series of items, and then *infers* from this that there can't be a coherent thought of an actually now-existent infinitely large thing.]

455

Antithesis: **The world has no beginning, and no limits in space; it is infinite as regards both time and space.**

Proof: Suppose the opposite: the world has a beginning. Now, the beginning of x is a real event preceded by a time in which x doesn't exist. So if the world began, there must have been an earlier time in which the world didn't exist, i.e. an empty time. But it isn't possible for there to be an empty time

⁸ [Kant attaches two footnotes to this one sentence. In one he equates going through something part by part with measuring it. Something whose size is 'indeterminate' can't be measured, he implies; but if it is enclosed within limits, we can still have the notion of the completeness of the part-by-part run-through, because that is supplied by the limits. The second footnote says that if something has an infinite size, there can't be an intuition that would give us the concept of *all of it*; and in this case our thought of all-of-it is simply our thought of the completed synthesis or run-through of its parts—an infinite sequence that is complete 'at least in our idea'.]

at the end of which something comes into existence. Why? Because in an empty time there's no difference at all between any moment and any other; and that means that nothing could mark off one moment as the moment for something to come into existence. . . . In the world many series of things can begin, but the world itself can't have a beginning, and is therefore infinite in respect of past time.

As regards the second point, again assume the opposite: the world is finite in spatial extent. This implies that a limited world exists surrounded by an unlimited empty space, which in turn implies that as well as things' being related to one another in space, they will be related to space because the entire aggregate will be sitting there *in—surrounded by*—the empty part of space. Now, the world is an absolute whole, and there is no object of intuition outside it; so there's no correlate to which the world is related; so the supposed-relation of the world to empty space would be a relation of it to no *object*. But such a relation is nothing; so the limitation of the world by empty space is nothing; so the world can't be limited in space; i.e. it is infinite in respect of extension.⁹

457

⁹ Space is merely the form of outer intuition. It isn't a real object that can be outwardly intuited. What about *absolute* space—i.e. space thought of independently of all the things that occupy it and thus give it a detailed character? That's not a *thing*; it's nothing but the mere *possibility* of outer appearances. . . . So empirical intuition is not a composite of appearances and space, . . . with these being two things that are correlated in a synthesis. The connection between them is really just that space is the *form of* the intuitions that underlie appearances. If we try to set the two side by side—space side by side with all appearances—we'll create sorts of empty 'facts' that *couldn't* be registered in any perception. For example, the 'fact' about whether the world as a whole is moving through empty space, and, if it is, how fast.

Comment on the first antinomy

458 **On the thesis:** In stating these conflicting arguments I haven't tried to play tricks, constructing a 'lawyer's proof', as they call it. That's what you have when an advocate tries to take advantage of his opponent's carelessness—letting him appeal to a misunderstood law so as later to score points by pointing out the misunderstanding. Each of the above proofs arises naturally out of the subject-matter, and neither side has taken advantage of any openings provided by errors of the dogmatists on the other side.

I could have made a pretence of establishing the thesis in the usual manner of the dogmatists, by starting from a defective concept of what it is for a magnitude to be infinite:

A magnitude *x* is infinite if it contains so many units that there can't possibly be one that is greater, i.e. contains more units than *x* does. But however many there are of something, it's always possible to *add one*. So there can't be an infinite given magnitude; and it's therefore impossible for there to be a world that has lasted infinitely long or is infinitely large; so the world must be limited in both respects.

I could have argued like that; but that argument uses a concept of *infinitude* that doesn't fit what we actually mean by 'an infinite whole'. It doesn't represent *how great x* is, so it isn't the concept of a *maximum*. When we use that concept—
460 ·the one used in the indented argument above·—our thought about *x* is merely that •how many units *x* contains is •greater than any number. This involves *choosing* the kind of unit one wants to use—·the smaller the unit, the more of them *x* contains·—with the result that, according to this defective concept of infinity, the infinity that *x* involves is larger or smaller, depending on whether the chosen units are small or big. ·That is absurd, of course, because *x*'s size isn't really

altered by our choosing different units. . . .

The true transcendental concept of infinitude is this: the magnitude of *x* is infinite if the process of going through the units *x* contains, one by one, can never be •completed.¹⁰

So it follows with complete certainty that an eternity of actual successive states leading up to a given moment can't have elapsed, because if it *had* elapsed that would be a •completed infinity. So the world must have a beginning.

In the second, ·spatial·, part of the thesis, we don't have the problem of a completed infinite series, because the parts of an infinitely large world wouldn't form a series—they would exist *together*. But consider how we have the thought of an infinitely large world. It can't be a thought about something that is or could be given in intuition, ·e.g. about how it would look if seen from such-and-such a distance·. The only way to think about it is in terms of the process of going through its parts, one by one. But in the case of something infinite we can't do that—we can't *complete* doing it. So it's impossible that the world should be infinite in size. . . .

On the antithesis: The proof of the infinitude of the given •world-series and of the •world-whole—i.e. the world's infinite •age and infinite •size·—rests on the fact that the only alternative is for the world to be bounded by empty time and empty space. I'm aware that attempts have been made to dodge this conclusion by arguing the world could have a limit in time and in space without there being absolute ·empty· time before the beginning of the world, or absolute ·empty· space extending beyond the real world—both of which are impossible. I entirely agree with the philosophers of the Leibnizian school that empty time and empty space outside the world are impossible. Space is merely the *form*

¹⁰ So the answer to 'How many units does this quantum contain?' is 'More than any number'—which is the mathematical concept of the infinite.

of outer intuition; it's not a real object that can be outwardly intuited; it's not a thing that is related in a certain way to appearances, but the *form of* the appearances. Everything we can say about space is an upshot of things we can say about appearances in space. No facts about the size or shape of appearances are facts about how appearances relate to a self-subsistent space. . . . Thus, appearances can't be limited *by an empty space* outside them, though space, whether
 461 full or empty,¹¹ can be limited by appearances. All this applies equally to time. But it can't be denied that these two nothings, empty space outside the world and empty time before the world, have to be assumed if we are to assume a limit to the world in space and in time.

There's a line of thought that professes to show that the world could have limits in time and space without its duration and size being fixed by an infinite void ·by which it is preceded or surrounded·. But that line of thought consists in quietly switching

- from the •sensible world ·that we have been talking about· to who-knows-what •intelligible world,
- from •the first beginning (an existence preceded by a time of non-existence) to •an existence in general that doesn't presuppose any other condition in the world,
- from •limits of extension to •boundaries of the world-whole

thus getting time and space out of the way. But our topic has been the phenomenal or sensible world and its magnitude; if we set aside those conditions of sensibility, ·i.e. time and space·, we'll be destroying the very being of that world. The intelligible world is merely the general concept of *world*,

¹¹ What about *empty space* that is *limited by appearances*? That is, what about empty space *within the world*? That doesn't contradict transcendental principles; so far as *they* are concerned, we can allow it; though I am not asserting that it is outright actually *possible*.

abstracted from all conditions of its intuition; and just because of that abstraction we can't possibly say anything synthetic, whether affirmative or negative, about it.

Second antinomy

Thesis: **Every composite substance in the world is made up of simple parts, and nothing exists anywhere except the simple or what is composed of the simple.** 462

Proof: Let's assume the opposite: Composite substances are *not* made up of simple parts. Now, take some substance *x* and set aside in your thought all the composition that is involved in *x*—·i.e. think about it as raw material, filtering out all the facts about how bits of it are put together·. What will be left for you to think about? No composite parts, of course; but *x* is supposed not to have simple parts, so you aren't left with them either; so you are left with *nothing*—no substance at all. So either **(i)** it's impossible to remove in thought all composition, or **(ii)** after its removal something remains that exists without composition, i.e. ·something that has no parts·, something simple. Well now, when small substances are assembled so as to be parts of a big substance *x*, it's just a contingent fact that they are inter-related in this way; they could have been arranged differently or just scattered; and this means that the composition that *x* involves *can* be set aside in thought. It follows that if **(i)** is true, *x* isn't composed of substances; that ·implies that *x* is itself not a substance, which· contradicts our stipulation ·that it is a composite substance·. All that remains is **(ii)** the original supposition, namely that a composite of substances in the world is made up of simple parts. 464

From this it follows immediately that •*all* the things in the world are simple beings; that •composition is merely a fact about how they are related to one another; and that

•although we can't ever isolate these elementary substances so as to take them out of this state of composition, reason must *think* them as the primary subjects of all composition, and therefore as simple beings that exist prior to all composition.

463

Antithesis: No composite thing in the world is made up of simple parts, and nothing exists anywhere that is intrinsically simple.

Proof: Assume the opposite: a composite thing (as substance) is made up of simple parts. Now, all external relations amongst things, and therefore all putting together of substances to make composite substances, are possible only in space; so any composite substance *x* must occupy a region of space that has as many parts as *x* has. . . . Every part of a composite substance must therefore occupy a space. But we're supposing that the absolutely basic parts of every composite substance are simple, which implies that a simple thing can occupy a space. Now, any real thing that occupies a space is made up of a manifold of constituents side by side, which means that it is composite. And any real composite is made up of constituent substances (it couldn't be made up of properties, because *they* can't exist side by side without being in substances); so the line of thought we are exploring here implies that the world is made up of simple things, each of which is a composite of substances—which is self-contradictory.

The second proposition of the antithesis, that nowhere in the world does there exist anything simple, is intended to mean only this:

465

The existence of something utterly simple can't be established by any experience or perception, either outer or inner; so that the utterly simple is therefore a mere idea. No experience *could* show that anything in the objective world matches this idea; and because the

idea has no object, it can't be used in any explanation of appearances.

Why can't it have an object? Well, to have an object for this transcendental idea we would need to have an empirical intuition of the object that we *know* doesn't contain any complex of elements external to one another and combining to make a single composite object. Of course we can have an intuition of something in which we aren't aware of any complexity, but that doesn't prove that no intuition of this object *could* reveal it to be complex—and that's what would be the case if the object were simple. So absolute simplicity of an object can't be inferred from any perception whatsoever; an utterly simple object can never be given in any possible experience. And since we have to regard the world of sense as the sum of all possible experiences, it follows that nothing simple is to be found anywhere in it.

This second part of the antithesis goes much further than the first part. [Kant's account of why this is so is obscure and puzzling; it seems not to matter for the rest of the work.]

Comment on the second antinomy

On the thesis: When I speak of a whole as necessarily made up of simple parts, I'm referring only to a *substantial* whole, which is only item that can be 'composite' in the strict sense of the word; that is, I'm talking about items that can exist (or at least be thought of) *separately*, and that happen to be brought together and inter-connected in such a way as to constitute a single thing. Space is not 'composite' in that sense, because its parts can't exist or be thought of separately from the whole; it's the whole that makes the parts possible, not vice versa. . . . Since space isn't a composite made up of substances, . . . if I remove all compositeness from it there's nothing left (not even *points*, because a point

466

468 is possible only as the limit of a space, and so of a composite). So space and time don't consist of simple parts. And the *states* of a substance aren't composed of simple parts; and this is true even of a state that has a magnitude. An example is *alteration*. It has a magnitude, because there are big alterations and small ones; but a big alteration doesn't come about through the piling up of many *simple* alterations! The inference that is drawn here from the composite to the simple applies only to things that can exist independently of any other things, and that rules out states and properties of things and also events and regions of space. If you apply the inference to everything that could be in any way called 'composite'—and people have often done just that—it's easy to make the thesis of this antinomy look silly by coming up with things that are irrelevantly 'composite' yet not composed of simple parts.

In the thesis I am trying to prove the existence of simple substances only as elements in things that are composite, so I could call the thesis 'transcendental atomism'. But for many years the word 'atomism' has been tied to a particular way of explaining bodily appearances, a process that avails itself of empirical concepts; so perhaps a better label for the thesis would be—borrowing from Leibniz—the dialectical principle of *monadology*. But that's not very accurate either, 470 because the word 'monad', used in Leibniz's way, refers only to something that is *immediately* given as a simple substance (e.g. the *I* in self-consciousness), not to an element of composite things; and the thesis of the second 467 antinomy is concerned only with the latter.

On the antithesis: The proof of the doctrine of the infinite divisibility of matter is purely mathematical; and the monadists have objected to it on grounds that lay *them* open to suspicion. Given a really evident mathematical proof, they won't acknowledge that

•the proof is based on insight into the constitution of space, i.e. the constitution of something that is the formal condition of the possibility of all matter.

They maintain instead that

•the proof merely draws out the consequences of abstract concepts that we have chosen to construct, and so it doesn't apply to real things.

How could we possibly invent a kind of intuition other than the one that is given in the basic intuition of space? As for the properties we can attribute *a priori* to space: how *could* they fail to be properties also of things that are possible only because they occupy this space? If we listened to the monadists, we would have to suppose that there are •real-world points which are simple = partless and yet have the special privilege of being able to fill space just by being lumped together. (That's because they would be *parts* of space. Don't confuse them with •mathematical points; *they* are simple too, but they don't fill space, because they aren't *parts of* space but merely *limits in* it—a point is merely the *end* of a line, not a *part* of it.). . . . In this work of the monadists, philosophy is playing tricks with mathematics, and it does this because it forgets that the topic here is 469 *appearances* and the condition that makes them possible. Of course given the understanding's pure **concept of the composite**, we can form the **concept of the simple**, but that isn't what's needed here. For the monadists to be right, we'd have to find an **intuition of the simple** to go with the **intuition of the composite** (i.e. of matter). But the laws of sensibility rule out any intuition of the simple, so it's impossible to find anything simple in objects of the senses. The abstract thesis that

•anything composite made up of substances presupposes simples that make it up
is true when we are talking about *concepts* of composite

and simple; but it is not true when applied to *phenomenal* composites—ones given through empirical intuition in space. That's because anything given through empirical intuition in space *must* have the characteristic that no part of it is simple, because no part of space is simple. The monadists were smart enough to look for an escape from this difficulty: instead of •taking space to be a condition of the possibility of the objects of outer intuition (bodies), they •took bodies and the causal relations among substances to be a condition of the possibility of space. But •that's putting things backwards•. The only concept we have of bodies is *as appearances*, so they must presuppose space, which is a condition of the possibility of all outer appearance. So this escape-hatch is blocked, as I showed well enough in the Transcendental Aesthetic. The monadists' argument would
471 of course be valid if bodies were things in themselves.

[Kant now has a longish paragraph that is really a comment on the thesis. It says that the thesis of the second antinomy is unique among the 'sophistical assertions' in claiming to have empirical evidence for its truth. He is referring to the mistaken view that for me to be aware of the *I* of self-consciousness is for me to be empirically confronted with a certain object, myself, as simple. This has already been amply refuted in the discussion of the paralogisms, Kant says, but he whips quickly through the refutation again here.]

Third antinomy

472 Thesis: **It's not the case that absolutely all the appearances of the world •can be derived from causality according to laws of nature and •can't be derived from anything else. To explain these appearances we have to assume that there is another causality, that of freedom.**

Proof: Assume the opposite: There is no causality except the causality governed by laws of nature. This implies that everything that *happens* presupposes a preceding state of affairs from which it inevitably follows, according to a rule. But the state of affairs *x* from which *y* arose must itself be something that has happened (i.e. has come to exist having previously not done so), because if *x* had always existed then *y* would always have existed also, rather than having just *happened*. That's how it goes with causality according to the law of nature—events are caused by earlier events which are caused by still earlier events. . . and so on. Therefore, if everything that happens does so in accordance with laws of nature, there will. . . never be a first beginning, so there'll be no completion of the backward-running sequence of causes •of any given event•. But the law of nature is just this, that nothing happens without a cause sufficiently determined *a priori*. The proposition that no causality is possible except in accordance with natural laws is therefore self-contradictory; so this can't be regarded as the only kind of causality. [Two points about this paragraph: •A certain phrase of Kant's has been translated as 'a cause sufficiently determined *a priori*' in every previous translation. This version follows suit, with no firm sense of what the phrase means. •In its switches between 'laws' and 'the law', this paragraph exactly tracks Kant.] 474

So we have to assume a causality through which something *y* happens without its cause *x* having arisen from a still earlier event *z* through necessary laws. In other words, we have to assume that an event can be a self-starter, occurring *absolutely spontaneously*, thereby *starting* a series of appearances that carries on from there in accordance with laws of nature. This •second kind of causality• is *transcendental freedom*. Without it, the series of appearances on the side of the causes is never complete, however thoroughly we explore the source of nature. 473

Antithesis: **There is no freedom; everything in the world happens solely in accordance with laws of nature.**

Proof: Assume the opposite: There is 'freedom' in the transcendental sense, as a special kind of causality in accordance with which the events in the world can have happened. This causality would be a power of *beginning* •a state of affairs and, therefore, also •the whole series of all its consequences. ·Don't underestimate the strength of what is being said here·: the series of events will have had its real beginning in this spontaneity, and the spontaneous event x that kicked off the series won't itself have arisen from *any* previous event or state of affairs. There will have *been* an immediately preceding state of affairs, but x won't have been caused by it. So transcendental freedom stands opposed to the law
475 of causality; and what it it assumes about how successive states of affairs are (dis)connected makes a *unified* experience impossible. So this freedom can't be met with in any experience, and it is therefore an empty thought-entity.

So when we are trying to find the pattern and order in the world's events, we have nowhere to look but *nature*—not freedom. 'Freedom from—independence of—the laws of nature is a liberation from •constraint!' Well, yes, but it's also a 'liberation' from the •guidance of all rules. 'But *the laws of freedom* enter into the causality exhibited in the course of nature, and so take the place of natural laws.' No! If freedom were governed by laws, it wouldn't be *freedom* but simply *nature* under another name. •Nature differs from •transcendental freedom as •law-governedness differs from •lawlessness. Nature **(i)** imposes on the understanding the demanding task of always looking for the sources of events further and further back in the series of causes, with every item in the series being causally conditioned ·by something still earlier·. But it compensates for that by **(ii)** promising us a thoroughgoing law-governed unity of

experience. The illusion of freedom, on the other hand, offers to remove the **(i)** burdensome task imposed by nature, by giving to the understanding a point of rest in its climb up the chain of causes, taking it to an unconditioned causality that is a self-starter; but it undercuts the **(ii)** promise of intellectual unity by offering us a *blind* causality that breaks the guiding thread of rules that we need if our experience is to be thoroughly coherent.

Comment on the third antinomy

On the thesis: The transcendental idea of freedom is just 476 one part of the psychological concept of freedom, which is mainly empirical. It's the part that concerns absolute spontaneity considered as something that an action must have if it's to be properly imputed to the agent—i.e. if the person who acted is rightly to be held responsible for the action, perhaps blamed or praised for it. When speculative [see note on page 168] reason has tackled the question of the freedom of the will, what has always so greatly embarrassed it is the merely •transcendental question: Must we admit a power of *spontaneously* beginning a series of successive things or states of affairs? We needn't concern ourselves with the question:

(i) *How* is such a power ·of spontaneous action· possible?

Just as we don't trouble ourselves the question

(ii) *How* is causality in accordance with the laws of nature possible?

We have to settle for the *a priori* knowledge that this latter type of causality must be presupposed; we haven't the least notion of **(ii)** what could make it possible for the existence of one item to bring about the existence of a different one; in this territory reason must be guided by experience alone.

·And the claim made by the thesis is limited in another way as well·. We have established the thesis that

•a series of appearances must have a first beginning involving freedom

only as something that's required for it to be conceivable that the world *began*; for all the •states of affairs and events after that beginning, all we need are purely natural laws. Still, now that the power of spontaneously beginning a series in time has been proved (though not understood!), it's now permissible for us •to allow causal chains *within* the history of the world to be capable of beginning spontaneously, and so •to attribute to their substances—i.e. the substances involved in the initiating event—a power of acting from freedom. Don't be scared off from this conclusion by the thought:

A chain of events occurring within the world's history can have only a *relatively* first beginning, because every such chain is preceded in the world by some other states of affairs, which implies that no *absolute* first beginning of a series is possible during the course of the world.

This is a misunderstanding: what I'm talking about here is a beginning not in •time but in •causality. Suppose for example that I get up from my chair right now, doing this completely freely, without being *made* to do it by the influence of natural causes; this event will be the utterly first beginning of a new *causal* chain of events—the infinite series of all its natural consequences—although as regards *time* my getting up from my chair is only the continuation of a preceding series. . . .

Reason's insistence that we assign to the series of natural causes a *beginning* due to freedom is clearly on display when we observe that all the ancient philosophers except the Epicureans saw themselves as obliged, when explaining the movements of things in the world, to assume a *prime*

mover, i.e. a freely acting cause that spontaneously *began* this series of events. They didn't try to make the world's beginning conceivable through mere nature.

On the antithesis: Someone defending the omnipotence of nature against the sophistical arguments offered in support of the opposing doctrine of freedom would argue as follows. [The argument runs to the end of this so-called 'Comment on the antithesis'. The thesis presented an •argument for a •conclusion; the 'Proof of the antithesis' criticised the •conclusion, and the 'Comment on the antithesis' is now going to criticise the •argument.] *If you don't admit anything as being. . . .temporally first in the world, there's no need for you to look for something that is causally first.* What you have done is to think up an utterly first state of the world, and therefore an absolute start of the ever flowing series of appearances, thus providing a resting-place for your imagination by setting bounds to limitless nature—*who told you that **that** was all right?* The substances in the world have always existed (or anyway the unity of experience requires us to suppose that they have); so there's nothing problematic about assuming that the causal chains into which they enter have also always existed, so we should call off the search for a first beginning, whether temporal or causal. It's true that we have no grasp of what could make it possible for there to be such an infinite ancestry for a given event—a causal chain with no initiating member. But if you treat that as a reason for refusing to recognise this enigma in nature (·the real causal chain that never began·), you're going to be obliged reject many fundamental properties and forces that are equally impossible to grasp intellectually. You'll even have to deny the possibility of anything's *happening!* If your experience didn't assure you that things undergo alterations, you wouldn't be able to think up *a priori* the possibility of such a ceaseless sequence of being and not-being.

And even if we *did* allow a transcendental power of freedom, so as to have a beginning of events in the world, this power would have to be outside the world. . . . But it could never be right to ascribe to substances in the world itself a power that is outside the world; because that would virtually abolish the patterns among appearances, patterns created by the way appearances are causally inter-related according to universal laws; and our name for them is 'nature'. In losing nature we would also lose the criterion of empirical truth, through which experience is distinguished from dreaming. We could hardly make any sense of a 'nature' that existed side by side with such a lawless faculty of freedom acting on the world from outside the world. Freedom would keep interfering with the laws of nature, reducing the world of appearances to disorder and incoherence.

Fourth antinomy

480 Thesis: **There belongs to the world, either as a part of it or as its cause, a being that is utterly necessary.**

Proof: The sensible world, as the sum-total of all appearances, contains a series of alterations. (Why? Because without such a series we wouldn't be presented with a time-line, and there has to be such a time-line if the sensible world is to be possible.¹²)

And every alteration is subject to its condition—a condition of its existence—which precedes it in time and makes it necessary. Now, every given conditioned item x presupposes a complete series of conditions running up to something unconditioned, and *that's* the only thing that is utterly

¹² Objectively speaking, time comes before alterations; but subjectively—in actual consciousness—it's the other way around, because the representation of time, like every other representation, is given only *through* the prompting of perceptions.

necessary. So we have to accept the existence of something absolutely necessary, because the consequences of such a thing—namely alterations—certainly exist. And *this necessarily existent item belongs to the sensible world*. If it didn't, that would mean that the series of alterations in the world would derive its beginning from a necessary cause that that didn't itself belong to the sensible world; and that is impossible. Here is why: A series in time can only be made to begin by something that precedes it *in time*, so the item we are talking about—the top condition of the beginning of a series of changes—must have existed at a time when series didn't yet exist (because a •beginning of x is an •existence preceded by a time in which x didn't yet exist). So the . . . necessary cause of •all• alterations must belong to time, and—because time is possible only as the form of appearances—the necessary cause can only be conceived as belonging to the world of sense. Therefore: something utterly necessary is contained in the world—either as •the •initial• part of the series of alterations in the world or as •the series.

Antithesis: **There is no unqualifiedly necessary being anywhere (i) in the world, or (ii) outside the world as the world's cause.**

Proof: Suppose the opposite of (i): Either (1a) the world itself is necessary, or (1b) a necessary being exists in the world. Then there are two alternatives. Either

(1b) the series of alterations started with something that is unqualifiedly necessary, and therefore without a cause; or

(1a) the series itself has no first member; every item in it is conditioned •by earlier members• and is contingent; but the series as a whole is unqualifiedly necessary and unconditioned.

[The point of 'unqualifiedly necessary' is to exclude from the discussion items that are merely necessary-relative-to-cause-x.] But (1a) con-

flicts with the dynamical law of the determination of all appearances in time; and **(1b)** contradicts itself, because no set of things can exist necessarily if no single member of it is necessary.

483 Then suppose the opposite of **(2)**: An utterly necessary cause of the world exists outside the world. In that case this cause, as the highest member in the series of the causes of alterations in the world, must begin₂ that series of causes. But this cause must itself begin₁ to act,¹³ so its causality must be in time, and must therefore belong to the totality of appearances, i.e. to the world—which contradicts the hypothesis that it is ‘outside the world’. Therefore neither **(i)** in the world nor **(ii)** outside the world (though in causal connection with it) does there exist any unqualifiedly necessary being.

Comment on the fourth antinomy

484 **On the thesis:** In proving the existence of a necessary being I ought not, in this connection, to use any but the *cosmological* argument, i.e. the one that ascends from conditioned items in the domain of appearance to something unconditioned. . . .that is regarded as the necessary condition of the absolute totality of the series. [Kant adds a sentence here explaining that he is setting aside the so-called ‘ontological proof’ of the existence of a supreme being, which he will discuss separately [starting on page 272].]

The pure cosmological proof’s way of demonstrating the existence of a necessary being has to leave unanswered the question of whether this being is the world itself or something distinct from the world. To show that it is distinct from the

¹³ There are two ways of using ‘begin’—the transitive or dyadic use (‘x begins₂ y’) and the non-transitive or monadic use (‘x begins₁’). The inference in this paragraph draws a conclusion involving begin₁ from a premise using begin₂.

world we need premises that aren’t cosmological and don’t simply move up the series of appearances. That’s because we would have to use •the general concept of *contingent beings* (viewed as objects of the understanding alone) and •a principle enabling us to connect this conceptually with •the concept of •a necessary being. But all that •takes us outside cosmology; it •belongs to a *transcendent* philosophy, which I’m not yet in a position to discuss.

If we start our proof cosmologically, basing it on appearances that form a series according to empirical laws of causality, we mustn’t then suddenly **switch** from this mode of argument by bringing in something that isn’t a member of the series. If we are working our way back from some item to its condition, from that to *its* condition, and so on backwards up the series, this involves a condition-to-conditioned relation that we have to stick with the whole way up; if there is a *highest condition* it must have that status by entering into the very same condition-to-conditioned relation that we have been using all the way up to this highest one. If the relation in question is a sensible one that falls within the domain of the possible empirical use of understanding, then *every* member of the series must •belong in that domain, and so •be temporal, and that includes the highest condition or cause, the one that brings the regress to a close. That’s why the necessary being must be regarded as the highest member of the cosmical series, •i.e. as belonging to the world. 486

Yet certain thinkers have allowed themselves to make that **switch**. They have started out quite correctly: from the alterations in the world they have inferred that the alterations are empirically contingent, i.e. depend on empirically determining causes, and so have obtained an ascending series of empirical conditions. [Kant’s own sentence •contains that odd bit about inferring the contingency of the alterations *from the alterations*, and •uses *empirisch* three times.] But because they couldn’t find

in such a series any first beginning, any highest member, they abruptly dropped the empirical concept of contingency and grabbed the pure category of contingency instead; and this involved them in a strictly *intelligible* series—one that is to be handled purely in terms of concepts and intellect rather than one involving the senses—a series that could be completed only by the existence of an unqualifiedly necessary cause. And since this cause wasn't tied down to any sensible conditions, it didn't have to be in time, and so its causality didn't have to be thought of as *beginning*. But such a procedure is entirely illegitimate, and I shall now show why.

488 Taking contingency as a category—a pure concept of the intellect—something is called 'contingent' because its contradictory is possible. Now, something's being 'contingent' in this intellect-linked sense doesn't entail that it is empirically contingent. When something is altered, it comes to be at time T_2 in a state S_2 that is the opposite of a state S_1 that was actual and therefore possible at an earlier time T_1 . But S_2 is not the contradictory opposite of S_1 . The thought of a contradictory opposite of S_2 is the thought of S_1 's existing at T_2 instead of S_2 's existing then. And the possibility of *that* doesn't follow from the fact that an alteration has occurred (i.e. the fact that something that *is* in state S_2 *was* in state S_1). Let's take a simple example. We have the premise that

(i) A body x that was moving at T_1 comes to rest at T_2 ; and we are interested in reaching the conclusion that

(iii) x 's being at rest at T_2 is contingent,

which is equivalent to the proposition that

(iv) the contradictory opposite of x 's being at rest at T_2 is possible.

But we can't infer (iv) from (i). To get to (iv) we need the premise that

(ii) x could have been moving at T_2 rather than being at rest,

·which doesn't follow from the fact that x first moved and then stopped. . . . The upshot is that the fact of *alterations* doesn't imply that any of the states that things are in at given times are contingent or possible in the categorial sense; so we don't have anything here that can carry us to the existence of a being that is 'necessary', with this similarly conceived in purely intelligible terms. Alteration proves only empirical contingency; i.e. that the new state couldn't have existed at that time if its preceding cause hadn't occurred, that being what the law of causality implies. A cause that is reached by moving up the series from conditioned item to their conditioning causes of it—call it unqualifiedly necessary if you like—will be met with *in time* and will belong to the series of appearances.

On the antithesis: We run into trouble when we try to 485 assert the existence of an unqualifiedly necessary highest cause that we could encounter when ascending the series of appearances. The trouble doesn't arise out of the mere concept of *thing that exists necessarily*; it involves the causal connectedness of a series of appearances for which a condition has to be assumed that is itself unconditioned; so the trouble has to be •cosmological, relating to empirical laws, and not •ontological [here = 'relating to abstract logic']. We are bound to discover that the ascending series of causes in the sensible world can never come to an end with an *empirically unconditioned* condition—a real-world cause that has no cause—and that will show us that there's no valid cosmological argument from •the contingency of states of the world, as shown by the alterations they undergo, to •the existence of a first cause that is the utterly basic cause of the series. 487

This antinomy presents a peculiar face-off: The thesis infers the existence of a primordial being [see note on page 173] from a certain premise, and from *that same premise* the antithesis infers the non-existence of a primordial being, this

derivation being as valid as the other! The thesis told us that *a necessary being exists* because

- the whole of past time includes the series of all conditions, and therefore also includes the unconditioned (i.e. the necessary);

while now the antithesis assures us that *there is no necessary being* because

- the whole of past time includes the series of all conditions (and just because they are conditions they must *all* be conditioned).

This happens because the two arguments focus on different aspects of •the series of conditions of which each is determined by another in time. The thesis argument focuses exclusively on *the absolute totality* of •that series, and this leads it to something that is unconditioned and necessary •because without that the series goes back for ever, so that there's no absolute totality of all its members•.

489 The antithesis argument, on the other hand, focuses on the *contingency* of everything in •the series. . . .and from this point of view everything unconditioned and all absolute necessity completely vanish. [The next sentence contains the phrase *gemeinen Menschenvernunft*, which could be mechanically translated as 'common human reason', but actually means 'ordinary common-sense'. In using it, Kant may have been influenced by the fact that the phrase does contain *Vernunft* = 'reason'. In the next few pages 'common-sense' will often translate *gemeine Verstand*, which mechanically translates into 'common understanding'.] Yet each argument is entirely in line with ordinary common-sense, which often conflicts with itself through considering its object from two different points of view. Two famous astronomers got into a fight that arose from their choosing different points of view •from which to see the same set of empirical facts•. One argued that the *moon revolves on its own axis*, because it always turns the same side towards the earth. The other drew the opposite

conclusion that *the moon does not revolve on its own axis*, because it always turns the same side towards the earth! Both inferences were correct, according to the point of view that each chose in observing the moon's motion. M. de Mairan regarded this situation as so remarkable that he wrote a book about it. [We are now at the end of the switching back and forth between odd- and even-numbered pages.]

3. What's at stake for reason in these conflicts

[The above number picks up from **2.** on page 211.] So there it is—the whole dialectical play of cosmological ideas! No possible 490 experience could present an object that was congruent with those ideas; and indeed reason can't even *think* them in a way that harmonises them with the universal laws of nature. But they aren't ideas that we have simply *chosen* to think up. Our reason is *necessarily* led to them when, in the continuous advance of empirical synthesis, it tries to grasp in its unconditioned totality something that (according to the rules of experience) has to come out as *conditioned*. These sophisticated assertions—i.e. the theses of the four antinomies—are just attempts to solve four natural and unavoidable problems of reason. [Kant's indigestible statement of *why* there are precisely four of them is omitted here. It's presumably meant to be equivalent to the explanation he has already given us.]

In presenting reason's pretensions when it tries to extend its domain beyond all limits of experience, I have hidden their *glitter*. Their full splendour is on show only when they're connected with empirical matters, and I have kept that connection out of sight: I have presented only the *basis* 491 for their legal claims, doing this in dry formulas that have (as befits a transcendental philosophy) been divested of all empirical content. But when the progressive extension of the

use of reason is connected with empirical matters—starting with the domain of our experiences and steadily soaring to these lofty ideas—philosophy displays a dignity which, if only it could *keep it up*, would make it much more valuable than any other branch of human knowledge. Why? Because in that role philosophy promises a foundation for our highest expectations and prospects concerning those ultimate ends onto which all reason's efforts must ultimately converge. Look at the questions it promises to answer!

- Does the world have a beginning in time and a limit to its extension in space?
- Is there anywhere, perhaps in my thinking self, something indivisible and indestructibly *one*? or are there only things that are divisible and transitory?
- Am I free in my actions? or am I, like other beings, led by the hand of nature and of fate?
- Is there a supreme cause of the world? or must our thoughts be limited to the things of nature and their order?

492 For answers to *these* questions any mathematician would gladly trade in the whole of his science! That's because mathematics can't yield satisfaction concerning those highest ends that humanity cares most about. (Actually, mathematics—yes, even mathematics, that pride of human reason—gets its great value from allowing and encouraging a use of reason that extends beyond all experience! But its way of doing this doesn't create troubles. What mathematics does is to guide reason to knowledge of nature in its order and regularity. . . .and in the astonishing unity of its active forces, bringing reason to a level of insight far beyond anything that could be expected from a science based on ordinary experience. And in doing this it also provides natural science with excellent materials for supporting its investigations—so far as their character permits—by suitable intuitions.)

Unfortunately for theory-building, though perhaps fortunately for humanity's practical concerns, reason in the midst of its highest expectations finds itself in trouble. It is so compromised by the conflict of opposing arguments that it's not safe—and isn't honourable—for it to withdraw from the quarrel, seeing it as a mere mock fight in which it doesn't have to get involved; and it's even less in a position to cry *Peace!*, because it has a stake in the matters in dispute. All reason can do, then, is to look into the *origin* of this conflict in which it is divided against itself, to see whether this has arisen from a mere misunderstanding. If that turned out to be the case, both sides in the dispute might have to give up their grandiose claims, but a lasting and peaceful reign of reason over understanding and the senses would be inaugurated.

493

Before getting into that thorough investigation, let's consider this: if we had to choose one side or the other, which side would we prefer to take? Because we'll be approaching this question in terms of our own interests rather than the logical criterion of truth, we won't reach a decision about which side is *right*; but our enquiry will do some good: it will give us a grasp of what has led the participants in this quarrel to choose the side they have chosen, given that it wasn't any superior insight into the matter under dispute. It will also explain such facts as that one side in each conflict is upheld with passionate zeal, the other with calm assurance; and the fact that people in general warmly welcome one side and are dead set against the other.

To carry out this preliminary enquiry as thoroughly as it deserves, we need first to compare the principles from which the two sides start out. [The comparison turns out to be quite straightforward, but Kant's compact presentation of it is worth spreading out a bit, as it is here. (i) Each antithesis is wholly governed by *empiricism*, which gives uniformity to

the manner of thinking, creating a **simple and unmixed** approach to the topic. The antithesis has empiricism at work **(1a)** in explaining appearances within the world, and it stays with empiricism **(1b)** when wrestling with the transcendental ideas of the world-itself-as-a-totality. **(2)** In the assertions and arguments on the thesis side, two elements are at work: •empiricism is accepted as suitable for **(2a)** explanations of items in the series of appearances, but when the defender of the thesis comes to **(2b)** the problems created by reason's demand for *totality*, it switches off empiricism and becomes tolerant of the notion of 'intelligible' limits, i.e. ones that are to grasped purely through abstract thought. So the driving force of the thesis side is a **complex mixture**, unlike the simple unmixed empiricism on the antithesis side. But Kant says he will label the thesis side as involving 'the *dogmatism* of pure reason', thus picking on the non-empiricist element in it, the tolerance of intelligible beginnings, because that is the essential and distinguishing characteristic of the thesis side, the part that it doesn't share with the antithesis side.

..494 Kant now proceeds:]

In dealing with the cosmological ideas, we find these three things on the side of *dogmatism*, i.e. of the thesis:

•PAY-OFFS FOR DOGMATISM•

(i) First, a certain *practical interest* that every right-thinking man endorses if he knows what is truly good for him. That •the world has a beginning, that •my thinking self is simple and therefore indestructible, that •in its voluntary actions my thinking self is free and raised above the compulsion of nature, and finally that •all the order among the things that make up the world is due to a primordial being from which everything derives its unity and purposive connection—these are foundation stones of morals and religion. The antithesis knocks all supports out from under us, or at least appears to do so.

(ii) Secondly, reason has a *speculative interest* on the side of the thesis. When the transcendental ideas are postulated and used in the manner prescribed by the thesis, we can take in *a priori* the *whole* sequence of conditions and conditioned items—because we'll be starting from something that isn't conditioned. The antithesis doesn't do this, and that's a very serious disadvantage for it. When you put to the antithesis a question about the conditions of any conditioned item, and then repeat the question for any conditioning item that is also in its turn conditioned, all the antithesis can do is to go on endlessly giving answers of the same general kind. According to the antithesis, •each beginning was preceded by an earlier beginning, •each part has still smaller parts, •each event is preceded by an event that caused it, and •the conditions of existence in general are also always conditioned, so that we can never steady ourselves by coming to rest in an unconditioned and self-subsistent primordial being.

(iii) Thirdly, the thesis has also advantage of *popularity*, which is a large part of its claim to favour. Common-sense has no trouble with the idea of the unconditioned start of any series. Being more accustomed to descend to consequences than to ascend to grounds, it doesn't puzzle over whether there could be something absolutely first; on the contrary, what it gets from such concepts are •comfort and •a fixed point to which to attach the thread by which it guides its movements. The alternative is a restless ascent from conditioned items to their conditions, always with one foot in the air; and there's no satisfaction in that! [The mixed metaphor 'ascend to grounds' is Kant's. A comparable mixture occurs when—e.g. on page 209—he speaks of our 'advance' along a 'regress'.]

•PAY-OFFS FOR EMPIRICISM•

In dealing with the cosmological ideas on the side of *empiricism*, i.e. of the *antithesis*, we find the following. **(i)** There is no practical gain, from pure principles of reason,

495

496

for morals and religion. Pure empiricism seems rather to deprive morals and religion of all power and influence. If these are true—

- there is no primordial being distinct from the world,
- the world never began and therefore had no author,
- our will isn't free,
- the soul is divisible and perishable like matter,

—then *moral* ideas and principles lose all validity, and share in the fate of the *transcendental* ideas that served as their theoretical support.

(ii) But there is a speculative pay-off; reason in its speculative activities gets advantages from empiricism that are very attractive and far surpass those that dogmatism can offer. According to empiricism, the understanding is always on its own proper ground, namely the domain of truly possible experiences, investigating their laws which it then uses for the indefinite extension of its sure and comprehensible knowledge. In this domain every object—and every relation between objects—can and should be represented in intuition, or at least in concepts for which the corresponding images can be clearly and distinctly provided in intuitions. There's
497 no need to leave the chain of the natural order and resort to *ideas*, the objects of which aren't known because they are mere thought-entities and so can't be *given*. [Kant goes on to say that the understanding not only doesn't *need* to leave its domain but isn't permitted to do so. He gives some details of what such wandering would involve—they amount to an unsympathetic sketch of the things he has said about what is going on on the thesis side.]

So the empiricist will never allow (1) any epoch of nature to be regarded as the utterly first, or any extent of nature that he has discovered to be the whole of it. He won't permit (2) any shift from the objects of nature. . . .to supposedly absolutely simple objects of which neither sense nor imagi-

nation can ever present an example. He won't admit (3) the legitimacy of assuming in nature itself any power of freedom that operates independently of the laws of nature. . . . And finally he won't grant (4) that a cause ought ever to be sought
498 outside nature, in a primordial being, because all we know is nature. . . .

Suppose that this were the situation:

The empiricist philosopher's only purpose in offering his antithesis is to subdue the impertinent curiosity of those who misunderstand the right use of reason so thoroughly that they •proclaim their insight and knowledge at just the point where true insight and knowledge stop, and •represent as furthering our speculative interests something that is valid only in relation to practical interests. . . .

If *that* were the whole story, the empiricist would merely be presenting a principle that urges us to moderate our claims, to be modest in our assertions, while also extending the range of our understanding as far as possible through our assigned teacher—experience. Behaving like that wouldn't cut us off from bringing intellectual *presuppositions* and *faith* to bear on our **practical** concerns, but it wouldn't allow them to be labelled and celebrated as 'science' and 'rational insight'. All knowledge is **speculative**, and can't be
499 about anything that isn't supplied by experience. . . .

But most of the time empiricism itself becomes dogmatic about ideas, confidently *denying* whatever lies out of reach of the knowledge it can have through intuition. When that happens, empiricism shows the same lack of modesty ·that it has criticised in its dogmatic opponents·; and this fault is especially blameworthy because it does irreparable harm to reason's practical interests.

500 The contrast between Epicurus's teaching and Plato's is like that.¹⁴ Each of those two types of philosophy says more than it knows. Epicureanism encourages and furthers knowledge, though to the detriment of practical concerns; Platonism supplies fine practical principles, but to the detriment of natural science, because it allows reason to indulge in •ideal explanations of •natural appearances. . . .

(iii) Now for the third factor that might enter into one's decision about which side to take in these conflicts: It's extremely surprising that empiricism should be so universally unpopular. You'd have thought that common-sense would eagerly adopt a programme that promises to satisfy it through entirely empirical knowledge and the rational connections it reveals, in preference to the transcendental dogmatism that compels it to rise to concepts that far outstrip the insight and
501 rational faculties of the most practised thinkers. But this is precisely what makes dogmatism attractive to common-

¹⁴ Actually, it's not clear that Epicurus ever did propound his principles as objective assertions. Perhaps he meant them as merely guiding rules for the speculative employment of reason; and if that is right, he showed in this regard a more genuinely philosophical spirit than any other ancient philosopher. The rules would be:

- In explaining appearances, proceed as if the domain of your enquiry is not circumscribed by any limit or beginning of the world.
- Assume that the stuff the world is made of is such as it must be if you're to learn about it from experience.
- Explain events only in ways that will bring them under unalterable laws of nature.
- Don't bring in any cause from outside the world.

These are very sound (though much neglected) principles for •extending the range of speculative philosophy while also •discovering the principles of morality without bringing in extraneous stuff. Those who require us in our speculative activities to *ignore* the dogmatic propositions •that there is a limit and beginning to the world etc. • shouldn't be accused of meaning to *deny* them.

sense, which it puts in a position where the most learned can claim no advantage over it! If common-sense understands little or nothing about these matters, the same is true of everyone else. It can't express itself in such a scholastically correct way as others—the experts—can, but it can go on indefinitely spinning out sophistical arguments while it wanders around among mere ideas. In that territory, no-one knows anything, so everyone is free to be as eloquent as he pleases; whereas in matters that involve the investigation of nature, common-sense has to stand silent and to admit its ignorance. So convenience and vanity combine in support of these dogmatic principles. A philosopher •or scientist• shrinks from •accepting a principle that he can't justify, and even more from •using concepts without knowing whether they apply to anything; but common-sense does this all the time! It wants to set out confidently from some starting-point, and it chooses for that purpose something that frequent use has made familiar to it. It isn't troubled by its inability to *conceive* this starting-point, because it is unaware of that (it doesn't really know what 'conceiving' means). For the ordinary plain person all speculative concerns shrink to invisibility in the presence of practical concerns; and when his fears or hopes incite him to assume or believe something, he fancies that it's something that he understands and knows. [Kant then remarks that empiricism can't be popular
502 in the way that idealising dogmatism is, and that however much harm empiricism may do to the highest practical principles, it won't ever come to influence the general run of people as much as the opposing dogmatism does.]

Human reason is by nature *architectonic*, meaning that it regards all our knowledge as capable of being fitted into a *system*; so the only principles it will accept are ones that don't make it outright impossible for all our items of knowledge to be combined into a system. But the propositions

of the antithesis do make the completion of the edifice of knowledge quite impossible. According to them, •behind every state of the world there is a still earlier one ·that still isn't the first·, •in every part there are still smaller parts that have parts in their turn, •before any event there is an earlier event ·that caused it and· that was itself also caused, and •in existence in general everything is conditioned, so that any discovery of conditions is at the same time a discovery of more things that are conditioned ·and are therefore subjects of further enquiry·. So there we have it: the antithesis won't admit any beginning or a starting-point—won't admit anything that could serve as a *foundation* for a complete edifice of knowledge—so it makes such an edifice altogether impossible. Thus reason's architectonic interest. . . carries with it a natural recommendation for the assertions of the thesis.

If someone •could disown all such interests, and consider reason's assertions solely in the light of how good the grounds for them are and irrespective of their consequences, and if •his only escape from the throng ·of competing doctrines· was to subscribe to one or other of the ·two· opposing parties, his state would be one of continuous vacillation. Today he would be convinced that the human will is *free*; tomorrow, reflecting on the indissoluble chain of nature, he would hold that freedom is mere self-deception and that everything is simply *nature*. But if he were called upon to *act* in some way, this play of merely speculative reason would vanish like a dream, and he would choose his principles purely on the basis of his practical interests.

·That's enough about the *prima facie* attractions of the two sides of the antinomial conflict·. For a reflective and enquiring being—such as you and I are·—it's only honest to devote a certain amount of time to examining his own reason, divesting himself of all partiality and openly submitting his

results to the judgment of others. So no-one should be blamed for, let alone prohibited from, presenting for trial the two opposing parties, leaving them. . . to defend themselves as best they can before a jury of. . . fallible men. 504

4. The transcendental problems of pure reason, considered as downright *having to be soluble*

To claim to solve all problems and answer all questions would be impudent boasting, and would show such extravagant self-conceit that one would instantly forfeit all confidence. But there are sciences whose very nature requires that every question arising within their domain should be completely answerable on the basis of what is known. Why? Because in these sciences it isn't permissible to plead unavoidable ignorance, because in each case the materials that generate the question also supply the answer. •Morals provide an **example**·: We must be able in every possible case to know in a rule-guided way what is *right* and what is *wrong*, because this is a question about what we're obliged to do, and we have no obligation to do something if we can't know ·that we're obliged to do· it. •Natural science provides a **counter-example**·: When we are explaining natural appearances, much must remain uncertain and many questions must remain unanswerable, because what we know of nature sometimes falls a long way short of explaining everything that there is to be explained. •Then what about this one·? In transcendental philosophy is there any question concerning an object presented to pure reason that we can be excused for not decisively answering because the answer can't be extracted from this same reason? In giving this excuse, we would have to show •that any knowledge we can get will still leave us completely unsure about what to say on the topic in question, and •that while we're conceptually equipped to 505

raise the question we don't have the ·conceptual· means to answer it.

Now I maintain that transcendental philosophy is unique among all domains of speculative knowledge, in that every question about an object given to pure reason can be answered by this same human reason. We can never shrug off the obligation to give a thorough and complete answer to such a question on the grounds that we are unavoidably ignorant or the problem is unfathomably deep. The very concept that puts us in a position to ask the question must also equip us to answer it, because (as in the case of right and wrong) the object—the subject-matter of the question—isn't to be met with outside the concept.

What I have just been saying applies not across the whole of transcendental philosophy but only in **(2)** its cosmological part, ·i.e. the topics of the antinomies·. [Kant goes on to explain why the story doesn't apply to the parts of transcendental philosophy that involve **(1)** the paralogisms or **(3)** the theological 'ideals'. The explanation is that in **(1)** and **(3)** there isn't an *object* to ask about in the first place. The nearest Kant gets to being clear about why that is so is in the following paragraph about **(1)**, which he presents in a footnote:]

Faced with a question about the constitution of a transcendental object, we can't give an answer saying *what it is*; but there's something we can say, namely that the *question* itself is *nothing*, because no object of it—no item that it can be *about*—has been given. Thus, all questions dealt with in **(1)** the transcendental doctrine of the soul are answerable—and indeed answered—in this second way, ·namely by saying that no real question has been asked·. The topic here is the transcendental subject of all inner appearances, ·the omnipresent *I*·, which isn't an appearance and

consequently isn't *given* as an object. That means that it doesn't satisfy the conditions needed for any categories to be applied to it, and that's what the initial question was really asking—which *categories* apply to this transcendental item? This is a case where the old saying holds true, that *no answer is itself an answer*. A question about the constitution of something that can't be thought through any definite predicate—because it's completely outside the sphere of objects that can be given to us—is null and void.

[In contrast with that, Kant says, the cosmological ideas raise questions that really *are* about something, really *do* have an object, because each of those ideas involves a taking-to-the-limit of a concept that can be used empirically. He continues:] The cosmological ideas are the only ones that can presuppose their object as being *given*, along with the empirical procedure that it conceptually involves—the procedure or 'empirical synthesis' of exploring earlier and earlier times or larger and larger regions of space, finding smaller and smaller parts of things, probing further and further back into the causal ancestry of an event, digging deeper and deeper into explanations for states of affairs·. The question arising from these ideas concerns this ordinary empirical procedure, asking merely whether it is to be carried so far as to contain *absolute totality*. That's what takes us from the empirical to the transcendental: the point is that this totality can't be given in any experience and therefore isn't empirical. [Kant's central point up to here is that in the case of **(2)** the cosmological ideas we are shifted from something comfortably empirical to something disturbingly transcendental by a shift from *some* to *all*; whereas with **(1)** the psychological and **(3)** the theological ideas it's not a matter of *shifting* in an intelligible way from something empirical to something transcendental; with **(1)** and **(3)** what we're dealing with is something that is transcendental in a more radical way.] Since we are

here dealing solely with a thing as an object of a possible experience, not as a thing in itself, the answer to the transcendent cosmological question can't lie anywhere except in the idea. We aren't asking about the constitution of any object in itself; and possible experience comes into our question *not* because we are asking

- What actual fully detailed experiences could we have ·in pursuing the empirical synthesis·?

but only because we are asking

- What is the content of the transcendental idea to which the empirical synthesis is a mere approximation?

508 And since the idea is a mere creature of reason, reason can't duck its responsibility and pass it on to the unknown object.

So here we have a science that is in a position to demand and expect clear and assured answers to *all* the questions that arise within its domain, even if they haven't yet been found. This isn't as extraordinary as it seems at first, and transcendental philosophy isn't the only example of it. Consider the other two pure rational sciences, . . . pure mathematics and pure ethics. •Has it ever been suggested that, because of our necessary ignorance of the conditions, it must remain uncertain what exact relation, in rational or irrational numbers, a diameter bears to a circle? . . . •There can't be anything uncertain in the universal principles of morals, because the principles, if they aren't altogether void and meaningless, must flow from the concepts of our reason. •In natural science, on the other hand, there are countless conjectures that can't be expected ever to become certain. Why not? Because natural appearances are objects that are given to us independently of our concepts, so the key to them lies not in us and our pure thinking but outside us, and in many cases the key is not to be found and so an assured solution is not to be expected. . . .

So we are faced with questions that reason propounds to itself, questions for which we are obliged to provide at least a critical solution:

- (1) Has the world existed from eternity rather than having a beginning? Does the world stretch out infinitely far in space rather than being enclosed within certain limits?
- (2) Is anything in the world simple, rather than everything's being infinitely divisible?
- (3) Does anything come about through the exercise of freedom, rather than everything's depending on the chain of events in the natural order?
- (4) Is anything completely unconditioned and intrinsically necessary, rather than everything's being conditioned in its existence and therefore dependent on external things and intrinsically contingent?

We can't evade these questions by pleading the narrow limits of our reason and confessing, under the pretext of a humility based on self-knowledge, that it's beyond the power of our reason to answer them. These are all questions about an object that can be found only in our thoughts—the object being *the utterly unconditioned totality of the synthesis of appearances*. If our own concepts don't enable us to say 510 anything for sure about such an object, we mustn't blame the object—'It's hiding from us!' Such a thing isn't to be met with anywhere except in our idea; it can't be *given*; ·so it can't in any reasonable sense be *hidden* either. We must look for the cause of failure in our idea itself. The idea is *a problem*, and it can't be solved if we go on obstinately assuming that there is an actual object corresponding to the idea. A clear account of the dialectic that lies within our concept itself would soon give us complete certainty about what we should think regarding the above questions.

If you maintain the pretext that certainty regarding these problems can't be had, I put to you a question that you must answer clearly: These ideas that are giving us so much trouble here—*where do you get them from?* [Kant follows this with something obscure that he may mean as a *prima facie* possible answer to this question. It leads on to something easier to grasp, namely this:] Suppose that the whole of nature were spread out in front of you, with nothing. . . . concealed from your senses or your consciousness, this still wouldn't provide you with a concrete empirical instance of any of the ideas. To have *that* you would need not merely •this intuition-of-everything but also something that empirical knowledge couldn't give you, namely, •a completed synthesis *and the consciousness that it is absolutely complete*. So your question doesn't have to be raised in the explanation of any given appearance, which means that it's not a question imposed on us by the object itself. You can never encounter the object, because it can't be given through any possible experience. In all possible perceptions we are always caught up among *conditioned* items, whether •conditioned• in space or in time; we don't encounter anything *unconditioned*, which would raise the question of whether the unconditioned item consists in •an absolute beginning of synthesis or rather •an absolute totality of a series that has no beginning. In its empirical meaning, the term 'whole' is always only comparative—as in 'I saw over the whole house, not just the ground floor'. As for the absolute wholes. . . . involved in the four cosmological questions: they have nothing to do with any possible experience. Suppose we're explaining the appearances of some body, and it occurs to us to wonder whether it is made up of simple parts or rather is infinitely divisible. Answering that question wouldn't enable us to explain the body any •better—it wouldn't even enable us to explain it •differently—because

neither answer to it could ever come before us empirically. . .512

Thus the solution of these problems can never be found in experience, and that's why you shouldn't say that it's 'uncertain' what should be said about the object •of our idea•. The object is only in your brain, and can't be given outside it; so all you need is to be consistent in your thoughts and avoid the trouble-making ambiguity that would transform your idea into a supposed representation of an object that is empirically given and thus knowable according to the laws of experience. Thus, the dogmatic answer •to a transcendental cosmological question• isn't 'uncertain'—it's *impossible!* What can be completely certain is the critical treatment •of the questions•. It doesn't tackle the questions objectively, but •subjectively, i.e.• in relation to the foundation of knowledge on which the question is based.

5. A sceptical look at the cosmological questions raised by the four transcendental ideas

We would give up demanding that our questions be answered dogmatically if we realized from the outset that a dogmatic answer, whatever it turned out to be, would serve only •to make us even more ignorant, and •to plunge us from one inconceivability into another, from one darkness into an even blacker one, and perhaps even into contradictions. If our question asks for a simple Yes or No, it would be smart of us to postpone the search for grounds for an answer, and first ask ourselves: what we would gain from the answer Yes? what we would gain from the answer No? If we find that in each case the answer is 'We would get nothing but nonsense', that will give us a good reason •to stop thinking about the •answers Yes and No to our original question, and• to starting thinking critically about the •question, looking into whether it assumes something that is groundless and

fools around with a defective idea (one whose falsity can more easily be seen by putting it to work and seeing where it leads than by looking at it in the abstract). That's what is so
514 useful about the sceptical way of dealing with the questions that pure reason puts to pure reason. It enables us, at a small cost, to keep clear of an enormous dogmatic tangle and engage instead in a sober critique, which as a true cathartic [= 'laxative'] will happily purge us of delusion and of the know-it-all punditry that it leads to.

A cosmological idea has to do only with an object of experience, which ·of course· has to fit a possible concept of the understanding. Suppose, then, that in preparing to tackle some cosmological idea I could see in advance that

•the ·relevant· kind of conditioned item in the synthesis of appearances must be, depending on how you look at it, either *too large* or *too small* for any *concept of the understanding*.

That would teach me that the idea in question must be entirely empty and senseless, because it can't be made to fit its object, however hard I work to get them to agree. The reason why holding onto the world-concepts [= 'cosmological ideas'] is bound to get us caught in an antinomy is that they all have this 'too-large-or-too-small' feature. ·Let's see this, case by case·.

(1t) If *the world has no beginning*, then it is *too large* for your concept, which consists in a successive regress that can never reach the whole eternity that has passed. If *the world has a beginning*, it will cut off the necessary empirical regress, making *too small* for the concept of the
515 understanding. That's because a beginning is still something that is conditioned, because it presupposes an earlier *time*; and the law of the empirical use of the understanding requires you to look for a higher temporal condition. So the ·temporally limited· world is clearly *too small* for this law.

(1s) This also holds for the two answers to the question about the world's magnitude in space. If it is *infinite* and unlimited, it is *too large* for any possible empirical concept. If the world is ·spatially· *finite* and limited, you are entitled to ask what sets these limits. ·The answer can't be that it is set by empty space, i.e. that the limit of the world is the surface that has only the world on one side of it and only empty space on the other·. Empty space isn't an independently existing entity that can stand in some relation to *things*, so it can't be a condition at which you could stop ·in your thinking about the world's size·. Still less can it be an empirical condition, something that you could encounter in experience (how can there be any experience of something that is utterly empty?); yet absolute totality in an *empirical* synthesis always requires an empirical concept of the unconditioned item. Consequently, a *limited* world is *too small* for your concept.

(2) If every appearance in space (every specimen of *matter*) consists of infinitely many parts, the process of dividing and redividing and. . . etc. will always be *too great* for our concept; while if the division of space is to stop at some member of the division (the simple), the division process will be *too small* for the idea of the unconditioned. For this ·supposedly
516 end-of-division member will always still allow of a regress to further parts contained in it·.

(3) If we suppose that every event in the world happens in accordance with the laws of nature, every event will have a cause that is also an event, so that you'll have to keep working back to earlier and earlier causes, with no end ·to your process·. Thus, nature considered as working always through efficient causes is *too large* for any concept that you can use in the synthesis of events in the world.

If you sometimes accept the occurrence of self-caused events, i.e. production through *freedom*, then the question

Why? will still pursue you. The law of causality that governs experience will compel you to look behind the supposedly free event, trying to discover what caused it; so you'll find that the totality of connection that you are allowing is too *small* for your necessary empirical concept.

(4) If you admit *an utterly necessary being* (whether it be the world itself, or something in the world, or the cause of the world), you'll be setting it in a time *infinitely* remote from any *given* point of time; because if you don't, the supposedly necessary being would be dependent on some other being that preceded it, and what's absolutely necessary doesn't depend on anything. So this 'absolutely necessary being' is too large for your empirical concept: you can't reach it through any process, however long you may keep it up.

If your view is that everything belonging to the world. . . is *contingent*—meaning 'contingent on something else', which means 'dependent on something else.'—then any existence that is given to you is too *small* for your concept. For that existence will force you to look around for some other existence on which it depends.

I have said that in each case the cosmological idea is either too large or too small for. . . any possible concept of the understanding. That found fault with the idea, saying that **it is too big or too small** for its job, namely fitting possible experience. Why didn't I make my points the other way around, finding fault with the empirical concept by saying that **it is too small or too large** for the idea? Here is why: It's only through possible experience that our concepts can have any reality; without it, a concept is a mere idea, without truth and without applying to any object. So the possible empirical concept is the standard by which we must answer the question:

•Is this idea *merely* an idea, a thought-entity, or does it apply to something in the world?

If it's right to say that x is too large or too small for y, it must be the case that x is required for the sake of y and has to be adapted to y. Among the questions that the ancient dialectical Schools played around with was this:

•If a ball can't pass through a hole, should we say that the ball is too large or that the hole too small?

In a case like this, it *doesn't matter* which we choose to say, because we don't know which exists for the sake of the other. In other cases there's a right answer: we don't say that a man is too tall for his coat, but that the coat is too short for the man.

This has led us to what is at least *a well-grounded suspicion* that the cosmological ideas, and with them all the mutually conflicting sophistical assertions, are based on an empty and tricked-up concept of how the object of these ideas is to be given to us. This suspicion may put us on track for exposing the illusion that has for so long led us astray.

6. Transcendental idealism as the key to sorting out the cosmological dialectic

I have sufficiently proved in the Transcendental Aesthetic that everything intuited in space or time, and therefore all objects of any experience we could possibly have, are nothing but appearances. That means that they are mere representations, having no independent existence outside our thoughts; and this applies when they are material things as well as when they are sequences of events. My label for this doctrine is 'transcendental idealism'. The 'transcendental realist' is someone who turns these states of our sensibility into independently existing things, i.e. turns *mere representations* into *things in themselves*.

518

519

It wouldn't be fair to credit me with accepting *empirical* idealism, that doctrine that has been so unpopular for so long. It admits the genuine reality of space, while denying—or at least finding doubtful—the existence of extended things in space; so that it doesn't make room for a well-grounded distinction between truth and dreams. As for the appearances of inner sense in time, empirical idealism has no problem regarding *them* as real things. It says indeed that this inner experience is a sufficient proof, and indeed the only proof, of the actual existence of its object (meaning the existence of its object as a thing in itself, complete with all its temporal features!).

520

As against that, my transcendental idealism accepts the reality of the objects of outer intuition just as they are intuited in space, and the reality of all changes in time as they are represented by inner sense. Space is a *form* of the intuition that we call 'outer', and without objects in space there would be no empirical representation whatsoever; so we can and must regard the extended beings in space as real; and the same holds for inner events in time. But this space and this time, and along with them all appearances, are not in themselves *things*; they are nothing but representations, and can't exist outside our mind. Even the inner and sensible intuition of

our mind, as the object that we are conscious of when we are conscious, and that is represented to us as having a sequence of different states through time, is not the real self as it exists in itself—i.e. is not the transcendental subject—but is only an appearance of that unknown being, an appearance that has been given to our sensibility. We *can't* admit this inner appearance as something that exists in itself, because it is temporal, and no thing in itself can be in time. But the empirical reality of appearances in space and time is secured well enough, and is thoroughly

separated from dreams, if both dreams and genuine appearances cohere truly and completely in one experience, in accordance with empirical laws.

521

The objects of experience, then, are never given in themselves but only given in experience, and have no existence outside it. Of course the moon may have inhabitants that no human being has ever perceived; but that means only that in the possible advance of our experience we could encounter them. . . . They are real if they are empirically connected with my real consciousness, though that doesn't mean that they are real in themselves, i.e. real apart from this advance of experience.

Nothing is really given to us except perception and the empirical advance from this to other possible perceptions. . . . Calling an appearance 'a real thing' when we haven't yet perceived it is either •saying that in the advance of experience we must meet with such a perception or •not saying anything. [Kant goes on to say that all this applies only to appearances—things in space and time—and not to things in themselves. Then:]

The faculty of sensible intuition is strictly only a receptivity, a capacity of being affected in a certain way with representations. [Kant means not that the representations affect us, but that the representations are effects upon us, What does the affecting? Read on.]. . . . The non-sensible cause of these representations is completely unknown to us, and we can't intuit it as an object. Why not? Well, such an object would have to be represented as not being in space or time, because these are merely conditions of sensible representation; and we can't conceive of any intuition that doesn't involve space or time. Still, we can use the label 'the transcendental object' for the purely intelligible cause of appearances as such, merely so as to have something corresponding to sensibility viewed as a receptivity. [In calling it 'purely intel-

ligible' Kant means that we can have the utterly abstract *thought* of whatever-it-is-that-causes-all-our-appearances. The notion of *causing* enters the story because in respect of all our intuitions we are *passive* = acted-on = acted-on-*by-something*. It will have occurred to you that this use of 'cause' can't involve the regular *category* of cause, which Kant has insisted is usable only in connecting appearances with other appearances. He will deal with that point in the 'concluding note' on page 261, saying that our thought of the intelligible causes of experience is 'analogous to' our thoughts about cause-effect amongst appearances.] We can regard this transcendental object as what produces

523 all our possible perceptions—it's responsible for how far they stretch, and how they hang together—and we can say that it is *given in itself prior to all experience*. But appearances, that are as this transcendental object makes them, aren't given •in themselves but only •in this experience; they are mere representations, and the only thing that enables them to mark out a real object is •not

- their relation to the transcendental object, their intelligible cause,

but•

- their hanging together *with one another* according to the rules of the unity of experience.

So we can say that the real things of past time are given in the transcendental object of experience; but they aren't •objects for me, aren't •real in past time, unless the light of history or the tracks left by causes and effects lead me to think that

- a regressive series of possible perceptions in accordance with empirical laws leads—i.e. •the course of the world leads—to a past time-series as a condition of the present time;

though this series can be represented as real only •in the connection of a possible experience, not as real •in itself. Thus, all the events that have occurred in the immense

periods that have preceded my own existence really mean only the possibility of extending the chain of experience from the present perception back to the conditions that determine this perception in respect of time.

So if I give myself the thought of all existing objects of the senses in the whole of time and the whole of space, I don't set them in space and time prior to experience. All 524 I am having is the thought of a possible experience in its absolute completeness: the objects are nothing but mere representations, so they are given only in such a possible experience. To say of something that it exists before I have had any experience of it is only to say that it is to be met with if, starting from perception, I advance to the part of experience it belongs to. The *cause* of all the details of what happens in this advance—settling how far I can go and what episodes I'll encounter along the way—is transcendental, so that I can't possibly know it. But that's not my concern. What I care about is the rule of the experiential journey in which objects are given to me—meaning that appearances are given to me. In the upshot it simply doesn't matter whether I say that

- (i) in the empirical advance in space I could meet with stars a hundred times further away than the most distant stars that I now see,

or instead say that

- (ii) such stars are perhaps to be met with in cosmic space even though no human being ever did or ever will perceive them.

For even supposing those stars were given as things in themselves, without reference to possible experience, they still wouldn't be anything to me, and therefore wouldn't be objects. To be objects for me they would have to be contained in the series of the empirical regress. [The rest of this paragraph is expanded from what Kant wrote, in ways that the •small dots• convention

can't easily indicate.] Thinking in terms of **(ii)** rather than **(i)**, and thinking of these 'stars' as things in themselves, won't do any harm because there's no content to such thoughts; they can't do harm because they have nothing to bite on. But isn't it sometimes harmful to think in transcendental terms about something that is really a matter of appearances? Yes indeed: harm comes about when we think in the wrong way about the cosmological idea of an *absolute whole* of appearances of some kind, and get ourselves pulled into raising a question that oversteps the limits of possible experience. *That's* where we must be on our guard against misinterpreting our own empirical concepts.

7. Critical solution of reason's cosmological conflict with itself

The whole antinomy of pure reason rests on the dialectical argument:

- If a conditioned item is given, the entire series of all its conditions is also given;
- Objects of the senses are given as conditioned;

therefore, etc. Through this inference of reason, the first premise of which seems so natural and evident, as many cosmological ideas are introduced as there are differences in the conditions (in the synthesis of appearances) that constitute a series. [Kant always calls the first premise 'the major premise'. That technical term in the logic of syllogisms contributes nothing here, and indeed isn't here being used correctly.] The cosmological ideas postulate absolute totality of these series, and *that's* how they put reason into unavoidable conflict with itself. We'll be better placed to detect what is deceptive in this sophistical argument if we first correct and tighten up some of the concepts used in it.

[In the next sentence, Kant will connect something's being 'given' (*gegeben*) with something's being 'set' (*aufgegeben*), meaning set as a task or a challenge. As you can see, it's neater in German than in English.] In the first place, it's obvious beyond all possibility of doubt that if the conditioned item is given, then a regress in the series of all its conditions is *set as a task*. If something is conditioned, then it has a condition (that's what being 'conditioned' *means*), and if that condition is conditioned in its turn, then. . . and so on through all the members of the series. So the above proposition—the 'set as a task' one—is analytic, and has nothing to fear from a transcendental criticism. It is reason's logical demand that we track as far as we can a concept's connection with its conditions—I mean the connection that directly results from the concept itself.

And if the conditioned item as well as its condition are **things in themselves**, then when the conditioned item is given the regress to its condition is not merely *set as a task* but already really *given*. And since this holds of all members of the series, the complete series of the conditions. . . is given (or rather presupposed) along with the initial conditioned item. Why? Because the conditioned item is given, and it is possible only through the complete series. The synthesis of the conditioned item with its condition is here a synthesis of the mere understanding, which represents things *as they are* and doesn't consider whether and how we can get in touch with them. But if what we're dealing with are **appearances**, the story changes. Because they are mere representations, appearances can't be given 'to me' except through my arriving at knowledge of them (or rather my arriving at *them*, for they *are* just empirical items of knowledge). I can't say that if the conditioned item is given then all the appearances that are its conditions are—in the same sense of the word—*given*. So I'm utterly unable to infer from the fact that a conditioned item is given the absolute totality of

the series of its conditions. That's because the appearances are. . . nothing but an empirical synthesis in space and time, and are given only in *this synthesis*. So we can't infer that if a conditioned item (in the domain of appearance) is given, the synthesis that constitutes its empirical condition is therefore given along with it. . . . This synthesis comes into being *in the regress*, and never exists without it. But there is something we *can* say about a •regress to the conditions, i.e. about a continuing empirical synthesis running up through the conditions, namely that •it is *set as a task*, and that it can't ever be brought to a halt by a lack of conditions.

528 This makes it clear that the first premise of the cosmological inference means 'conditioned' in the transcendental sense of a pure category, while the second premise takes it in the empirical sense of a concept of the understanding applied to mere appearances. So the argument commits the dialectical 'fallacy of equivocation', as they call it. There's nothing •artificial about it; it's a quite •natural illusion of common-sense. When something is given as conditioned, this illusion leads us to assume (in the first premise) the series of its conditions, assuming them *uninspected*, so to speak. This assumption is just the logical demand for adequate premises for any given conclusion. Also, the conditioned item's connection with its condition doesn't involve any time-order; they are presupposed as being in themselves given *together*. And in the second premise it's just as natural as it is in the first to regard appearances as things in themselves and as objects given to the pure understanding, abstracting from all the conditions of intuition under which alone objects can be given. Yet in this—i.e. in treating the second premise in that way—we have overlooked an important difference between the concepts. **(i)** In the first premise, the synthesis of the conditioned item with its conditions (and the whole series of conditions) doesn't carry with it any temporal con-

straint or any concept of succession. **(ii)** But the empirical synthesis—i.e. the series of the conditions in appearance that the second premise is talking about—is necessarily successive, the members of the series being given one after another in time; so I can't assume the absolute *totality* of the synthesis and of the series represented through it. In 529 the first premise all the members of the series are given in themselves, without any temporal condition of time, but in this second premise they are possible only through the successive regress—an actual *procedure* whose episodes are given only by being carried out.

Once we have pointed out this error in the argument on which both parties base their cosmological assertions, we can fairly dismiss them both on the grounds that they can't justify their claims. But that won't end the quarrel—as it would do if one or both of the parties were proved to be wrong in their actual doctrines—not just in their •arguments but in their •conclusions. Granted, neither of them has argued soundly for his conclusion, but it seems utterly clear that, since one asserts that •the world has a beginning and the other says that •the world has no beginning and has existed from eternity, one of them must be right! But even if that's the case, it's impossible to decide which one that is, because the arguments on the two sides have equal *Klarheit* [usually = 'clarity'; perhaps here = 'persuasiveness']. The parties can be told to keep the peace before the tribunal of reason; but the dispute still drags on. The only way for it to be settled once and for all, to the satisfaction of both sides, is for the very fact that they can so splendidly refute one another to win them over to the view that they are really quarrelling about nothing, and that a certain transcendental illusion has mocked them with a reality where none is to be found. That's the path I shall 530 now follow in putting an end to this undecidable dispute.

* * * * *

Zeno of Elea, a subtle dialectician, was severely rebuked by Plato as a mischievous sophist who showed off his skill by setting out to prove a proposition through plausible arguments and then immediately overthrowing it by other arguments that were equally strong. For example: Zeno maintained that God (probably conceived by him as simply the world) is

(i) neither in motion nor at rest,

(ii) neither similar nor dissimilar to any other thing,

(iii) neither finite nor infinite.

His critics saw him as intending, absurdly, to deny both of two mutually contradictory propositions; but I don't think this was justified. As for **(i)**: if by the word 'God' he meant the universe, he would certainly have to say that it doesn't stay in one place (rest) and doesn't change its location (motion) either, because all places are in the universe, so the universe isn't itself in any place. As for **(ii)**: if the universe includes in itself everything that exists, it can't be either similar or dissimilar to any *other* thing, because there aren't any *other* things—things outside it—with which the universe could be compared. If two opposed judgments presuppose an inadmissible condition, then the failure of that condition brings them both down. . . .

[Kant gives a homely example involving 'smells good', 'smells bad', and the 'condition' of each of these, namely 'has a smell'. He tries to harness this to a conflicting pair of judgments, and doesn't provide enough detail to see that he has failed. Still, you can get the general idea. Then:]

As for **(iii)**: The propositions

(a) The world is infinite in extent, and

(b) The world is not infinite in extent

are contradictory opposites, so that if I assert the falsity of **(a)** I am committed to the truth of **(b)**. But notice that in denying that the world is infinite I am not affirming that

(c) The world is finite in extent.

The propositions **(a)** and **(c)** could *both* be false. In merely denying **(a)** we are merely removing the infinitude, which we might do by denying the whole separate existence of the world. What the assertion of **(c)** does is to remove the infinitude while asserting the existence of the world in itself as something with a determinate size. And that assertion could be false along with **(a)**, because it could be that the world is not given as a thing in itself, and therefore not given as being either infinite or finite in size. Let me call this kind of opposition *dialectical*, and the opposition of contradictories *analytical*. Then I can say: two dialectically opposed judgments can both be false, because one is not a mere contradictory of the other, but says something more than is required for a simple contradiction.

If I regard **(a)** and **(c)** as contradictory opposites rather than dialectical opposites, I am assuming that the world (the complete series of appearances) is a thing in itself; that the world is still *there*, even if I suspend my infinite or finite regress in the series of its appearances. But if I reject this assumption—or rather this transcendental illusion—and deny that the world is a thing in itself, the contradictory opposition of the two assertions is converted into a merely dialectical opposition. Since the world doesn't exist in itself, independently of the regressive series of my representations, it doesn't exist *in itself* as an *infinite* whole or exist *in itself* as a *finite* whole. It exists only in the empirical regress of the series of appearances, and isn't to be met with as something in itself. So if this series is always conditioned, it can't ever be given as complete; and the world thus isn't an unconditioned whole, and doesn't exist as such a whole,

532

533

either of infinite or of finite size.

What has been said here about the spatial half of **(1)** the first cosmological idea, i.e. about the absolute totality of spatial magnitude in the realm of appearance, applies also to all the other cosmological ideas. [If you need a reminder about 'regressive synthesis', see the passage starting on page 208.] The series of conditions is to be met with only •in the actual process of the regressive synthesis itself, not •in the domain of appearance viewed as a thing given in and by itself, independently of any regress. Thus, faced with **(2)** the question 'How many parts does it have?', asked of a given appearance, we have to say 'Neither finitely nor infinitely many'. For an appearance isn't something existing in itself. Its parts are first given in and through the actual process of going from a thing to its parts, then to their parts, then to *their* parts, and so on; and this process is never completely finished—so it never provides a finite total or an infinite total. This also holds for **(3)** the series of subordinated causes, and for **(4)** the series that goes from something conditioned to unconditioned necessary existence. These series can never be regarded as being, *in themselves in their totality*, either finite or infinite. Because they are series of suitably inter-related *representations*, each exists only in the process associated with it; it can't exist independently of this process, i.e. exist in itself as a self-subsistent series of things.

534

Thus reason's conflictedness in its cosmological ideas vanishes when it is shown •that it is merely dialectical, and •that it is a conflict due to an illusion that arises from our applying •an idea of absolute totality (that holds only as a condition of things in themselves) to •appearances (that exist only in our representations). . . . Still, we can turn this antinomy—this conflictedness—to our advantage, not a dogmatic advantage but a critical and doctrinal one: namely, providing an indirect proof of the transcendental ideality of

appearances—a proof that ought to convince anyone who isn't satisfied by the direct proof I gave in the Transcendental Aesthetic. This present proof consists in the following dilemma:

—If the world is a whole existing in itself, it is either finite or infinite.

—It isn't finite (shown in the proof of the antithesis).

—It isn't infinite (shown in the proof of the thesis).

Therefore

—The world (the sum of all appearances) is not a whole existing in itself. 535

From this it follows that appearances in general are nothing independently of our representations—which is just what it means to call them 'transcendentally ideal'.

This is important. It lets us see that the proofs given in the fourfold antinomy aren't mere glittering tinsel; they are *grounded* on •the supposition that the appearances of which the sensible world is composed are things in themselves. On that basis we can derive each of the two conflicting propositions; this conflict shows that there is an error in •this assumption, which in turn leads us to the discovery of the true constitution of things as objects of the senses. The transcendental dialectic doesn't *at all* favour scepticism, but it certainly favours the sceptical method, which can point to such dialectic as an example of how useful the method can be: when reason's arguments roam free and tangle with one another, the sceptical method can always extract from the situation something useful and likely to help us correct our judgments—even if that's not what we set out to do!

8. Applying the regulative principle of pure reason to the cosmological ideas

536 The cosmological principle of totality doesn't **give** [*geben*] a maximum of the series of conditions in a sensible world, regarded as a thing in itself, but only **sets it as a task** [*aufgeben*]*—*the task of going through the process of working one's way back through the series of conditions. So the principle of pure reason has to be amended along these lines; and then it is still valid, not as the *axiom* that we think the totality as actually in the object, but as a *problem* for the understanding, and therefore for the person whose understanding it is. Given any conditioned item x, the task or problem is set by this command:

Look into the conditions of x, then the conditions of those conditions, and so on backwards through the series of conditions; and in doing this, think of yourself as pursuing the completeness prescribed by the idea.

[Kant repeats that because x is an appearance, not a thing in itself, this completeness won't ever actually be achieved. Then:] The principle of reason is thus properly only a *rule*, ordering us to work back through the series of the conditions of given appearances, and forbidding us to bring this process to an end by treating some item in it as utterly unconditioned. It isn't a principle of the possibility of experience and of empirical knowledge of objects of the senses, so it's not a principle of the understanding. Why not? Because the understanding's business is with experiences in space and time, and those are always enclosed within limits. It isn't

•a **constitutive** principle of reason*—*i.e. one that tells us what is the case*—*enabling us to extend our concept of the sensible world beyond all possible experience.

Rather, it is

•a **regulative** principle of reason, which serves as a *rule* or *regulation*, telling us how to behave when working back through a series of conditions. Specifically, it tells us to continue and extend our experience *as far as we can*, never accepting that we have reached an absolute empirical limit. [The link between 'rule' and 'regulative', via the Latin *regula* = 'rule', is even clearer in German where the words are *Regel* and *regulativ*.]

It doesn't say *in advance of any empirical exploration**—*i.e. prior to the regress*—*what is present in the object as it is in itself. I call it a 'regulative principle' to distinguish it from a 'constitutive' cosmological principle, which would be one that speaks of the absolute totality of the series of conditions, viewed as actually present in the empirical object. My point in making this distinction is •to bring out the fact that there isn't any such constitutive principle, and so •to prevent us from ascribing objective reality to an idea that serves merely as a rule. [Without his intervention, Kant says, that mistake would be inevitable. He calls it a 'transcendental subreption', meaning roughly 'a transcendental bait-and-switch act'.]

This rule of pure reason can't tell us *what the object is*, 538 but only *how the empirical regress is to be carried out* so as to arrive at the complete concept of the object. [Kant repeats his reasons for all this, associating the would-be constitutive principle with believing that the subject-matter exists 'in itself' and therefore has properties independently of our experiencing them. The crucial point is that nothing unqualifiedly unconditioned is to be met with in experience'. Then:]

So the first thing we have to do *in obeying the command of the regulative principle* is to settle what we are going to say about a synthesis of a series*—*a process of empirical exploration*—*that won't ever be complete. [Kant now

introduces the terms ‘infinite’ and ‘indefinite’, of which the former has been favoured by mathematicians and the latter by philosophers. He declines to explore the concerns *they* were dealing in using this terminology.] I want only to define these concepts precisely enough for my purposes.

It is all right to say of a straight line that it can be extended ‘to infinity’. To distinguish between an •infinite advance and an •indefinitely great advance, in a case like this, would be mere nit-picking. When we say, ‘Draw a line’, it does indeed sound more correct to add **(i)** ‘... making it indefinitely long’ than to add **(ii)** ‘... making it infinitely long’. Whereas **(ii)** means that you *mustn’t* stop extending it—which is not what is intended—**(i)** means only that you may extend it as far as you like. And if we are talking only about what one *can* do, then **(ii)** is quite correct, for we can always make the line longer, without end. So is it in all cases in which we speak only of *the progress*, i.e. of the *advance* from the condition to the conditioned: this possible advance proceeds, without end, in the series of appearances. From a given pair of parents the **descending** line of generation may proceed without end, and we can quite well regard the line as actually continuing without end in the world. For in this case reason never requires an absolute totality of the series, because it doesn’t presuppose that totality as a condition that is *given*, but only something giveable that can be endlessly added to.

Now consider the question of how far the **regress** goes in an **ascending** series, running from something given as conditioned back up to its conditions, then to *their* conditions, and so on. Can we say that **(ii)** the regress runs to infinity, or only that **(i)** it extends indeterminately or indefinitely far? For example, can we **(ii)** ascend infinitely from the men now living through the series of their ancestors? Or can we only say that **(i)** we have never had empirical evidence that such-and-such

is the first or top item in the series, and that we therefore may and indeed should *search for* the parents of each ancestor we come across, though we shouldn’t presuppose them?

[Kant gives different answers for two different kinds of case. **(a)** One concerns the series that goes from an empirically given material thing to its parts, then to *their* parts, and so on; parts are ‘inner conditions’ of the thing they are parts of, and the series running from the thing through all its parts is infinite. His thought is this: suppose the first item in the series is a brick, which I hold in my hand; then in a good sense I hold the entire series in my hand; so the series is complete, rounded off, contained, in a way that makes ‘indefinite’ inappropriate and therefore makes ‘infinite’ appropriate. **(b)** When the series involves a condition-to-conditioned relation where the condition is a totally distinct thing from the item that it conditions, then the series has an indeterminate or indefinite character, because nothing rounds it off in the way the infinite series of brick-parts is rounded off by the whole brick’s being in a limited space. Then:]

In neither case, whether the regress is infinite or indefinite, is the series of conditions seen as being given as infinite in the object. The series are not things in themselves, but only appearances linked by the ‘x is a condition of y’ relation, so they are given only in the regress itself, i.e. in the actual process of discovering them. So we aren’t facing the question

•How long is this series in itself? Is it finite or infinite? That question doesn’t arise, because the series ‘in itself’ is *nothing!* The question we do face is this:

•How are we to go about conducting the empirical regress? And how far we should continue it?
 ...When **(a)** the whole series is empirically given, it is *possible* to proceed back to infinity in the series of its inner conditions. When **(b)** the whole is not given from the outset,

being given only through the empirical regress, we can only say that the *search* for still higher conditions of the series is *possible* to infinity. In case **(b)** we could say: 'There are always more members, empirically given, than I can reach through the regress of decomposition', ·i.e. the process of investigating smaller and smaller parts·. In case **(a)** we *can* always proceed further in the regress, because no member is empirically given as utterly unconditioned; so a higher member of the series is always possible; so the enquiry regarding it is necessary. In **(b)** we necessarily *find* further
 534 members of the series; in **(a)**. . .we necessarily *enquire* for them. . . .

The next section will show these observations in their proper light by putting them to work.

9. Putting the regulative principle of reason to work empirically, in connection with the cosmological ideas

I have already shown, more than once, that no transcendental use can be made of the pure concepts either of the understanding or of reason. The ·thought of the· complete totality of the series of conditions in the sensible world rests entirely on a transcendental use of reason, in which reason demands this unconditioned completeness from something
 544 it assumes to be a thing in itself. Since the sensible world doesn't contain any such completeness, we should never ask, concerning the over-all size of a series in the sensible world, whether it is limited or *in itself* unlimited. The only question concerns the empirical regress in which we trace experience back to its conditions, and it is this: If we do this in conformity with the rule of reason, not stopping except with an answer to reason's questions that fit the object, *how far will that take us?*

What still has to be shown is **(i)** that the principle *is* valid as a rule for *continuing*. . . a possible experience. I have shown well enough **(ii)** that the principle of reason *is not* valid as a constitutive principle of appearances ·viewed as things· in themselves. If we can keep **(ii)** steadily in view, reason's conflict with itself will be entirely at an end. [Translators have given different accounts of what should be kept 'steadily in view'. Müller **(2)**. Kemp Smith **(1)** and **(2)**. Pluhar **(1)**. Guyer and Wood **(1)**. The pronoun Kant uses favours **(1)** rather than **(2)**, but this fits so badly with the rest of the paragraph (as you'll see in a moment) that Müller has to be right—Kant's pronoun was a slip.] That's because this critical solution will both •destroy the illusion that put reason at odds with itself and •reveal the sense in which reason is in harmony with itself—the conflict having arisen solely through misunderstanding of this. In this way a principle that would otherwise have been *dialectical* is turned into something *doctrinal*—·i.e. a threatening source of error and confusion is converted into a solid bit of true theory·. In fact, if this principle holds good in its subjective role as leading to the greatest possible empirical use of understanding *in conformity with the objects of experience*, the upshot will be much the same *as if* it were an axiom that determined *a priori* the objects in themselves (though of course such an axiom couldn't possibly come from reason). Why? Because the only way such an axiom could have any influence in extending and correcting our knowledge of the objects of experience is by busying itself in producing the widest possible empirical use of the understanding, ·which is just what the regulative principle does·.

1. Solution of the cosmological idea of the totality of the composition of the appearances of a cosmic whole

..545 Here, as in the other cosmological questions, the regulative principle of reason is based on this proposition:

•In the empirical regress we can't experience *an absolute limit*; we can't experience any condition as being *empirically* absolutely unconditioned.

That's because such an experience would have to involve perceiving a limitation of appearances by nothing, i.e. by the void, and it's impossible to perceive a void.

546 This proposition, which says in effect that the only conditions I can reach in the empirical regress must be regarded as empirically conditioned in their turn, contains the rule that however far along the ascending series I may have gone, I must always enquire about a still higher member of the series, whether or not I find it.

For the solution of the first cosmological problem, therefore, all that's needed is to decide whether, in the regress to the unconditioned spatial and temporal magnitude of the universe, to call this never-limited ascent a regress to •infinity or only an •indefinitely continued regress.

The only general thought I have of •the series of all past states of the world, or of •the series of ever larger spheres of things that coexist in space, is merely an indeterminate •or indefinite• thought of a possible empirical regress. . . .¹⁵

547 Now, I have a •concept of the world-as-a-whole but I could never have an •intuition of it. So I can't argue from the size of the world-whole to the size of the regress; that would

¹⁵ So this world-series can't be bigger or smaller than the possible empirical regress on which its concept rests. Since this regress can't provide us with a determinate infinite or a determinate finite. . . it's clear that the spatial and temporal size of the world can't be taken to be finite or to be infinite. The regress through which the world's size is represented rules out both.

be back to front; it's only by reference to the size of the empirical regress that I can even have a *concept* of the size of the world. . . . Since the world is not given to me in its totality through any intuition, its size isn't given to me independently of the regress. So we can't say anything at all about the world's size, not even that it contains a regress that proceeds to infinity. Saying the latter would be anticipating members that the regress hasn't yet reached, implying that *there are so many of them that no empirical synthesis could reach them all*; and this would be determining the size of the world (although only negatively) independently of the regress—which is impossible. . . .

..548

So I can't say that the world is *infinitely* old or *infinitely* large. That concept of magnitude involves the thought of *a given infinitude*; that is empirically impossible, and so in reference to the world as an object of the senses it is unqualifiedly impossible, •i.e. impossible *period*•. Nor will I say that the regress from a given perception to everything in its series backwards in time or outwards in space proceeds to *infinity*, because that would imply that the world has infinite magnitude. And I won't say that the regress is *finite* either, because an absolute limit is likewise empirically impossible. So I can't say *anything* about •the spatial or temporal size of• the whole object of experience, the world of sense; all I can talk about is the rule concerning how experience is to be obtained and further extended. . . .

Thus the first and **negative** answer to the cosmological question about the size of the world is that the world has **no** first beginning in time and **no** outermost limit in space.

To see why, suppose the opposite: the world is limited in one way by empty time and in another by empty space. 549 It can't be limited in either way *in itself*, because it's an appearance and not a thing in itself; so these •supposed• limits of the world would have to be given in a possible expe-

rience, i.e. we would have to have a perception of limitation by utterly empty time or utterly empty space. But such an experience, being completely empty of content, is impossible. Thus, an absolute limit of the world is impossible empirically, and therefore impossible *period*.¹⁶

Out of all this we also get an **affirmative** answer: the regress in the series of appearances, which is what gives us our grip on the notion of the world's size, **does** go on *indefinitely*. This is tantamount to saying that although the sensible world has no absolute spatial or temporal size, the relevant empirical regress. . . has its own rule, namely:

•From each conditioned item x in the series, •move back along the series to one that is more remote, namely a condition of x (doing this by means of your own experience or the guiding-thread of history or the chain of effects and causes), and •never slack off from widening the range of the possible empirical use of your understanding.

•The second half of that is justified by the fact that such extension of the scope of one's understanding is the main thing—the *only* thing—that reason's principles are *for*.

[Kant goes onto say that this rule doesn't require an endless regress, ruling out in advance (for example) finding ancestors that had no ancestor, or a star that is further away than any other star. But the rule does require that in carrying out the regress we must always go from appearances to

¹⁶ Notice how different this proof is from the dogmatic proof of the antithesis of the first antinomy [page 214]. In that argument the sensible world was taken to be what the common and dogmatic view says it is, namely a thing given in itself in its totality, independently of any regress; and the argument said that unless the world occupies all time and all places, it cannot have any determinate position in either of them. So the conclusion of that argument was different from the conclusion I have just reached here, because the dogmatic proof concluded that the world is actually infinite.

appearances; and this means that the regress won't ever take us to something that *we recognize as* a limit or boundary. Kant repeats his reasons for this, through a couple more paragraphs.]

2. Solution of the cosmological idea of the totality of the division of a whole given in intuition

If I take a whole thing that is given in intuition and divide it, I am going from a conditioned item to conditions of its possibility; and if I go on dividing and subdividing, I am pursuing a regress through the series of these conditions. The absolute totality of this series would be given only if the regress could reach *simple* parts, i.e. parts that didn't in their turn *have* parts. If there aren't any simple parts, so that all the parts I encounter as I work through the regress of divisions are themselves also divisible, then the regress of divisions runs to *infinity*. [Kant repeats here the explanation reported in items (a) and (b) on page 243. Then:] But we aren't entitled to describe a whole that is divisible to infinity as *made up of infinitely many parts*. For although the intuition of the whole contains **all the parts**, it doesn't contain **the whole division**. All there is to the division is the continuous pulling-apart, i.e. the regress through which the series first becomes actual. Since there is no end to this regress, all the members or parts at which it arrives are contained in the given whole, viewed as an *aggregate*. But the whole *series of the division* is not so contained, because it is an infinitely long procedure, so it never constitutes a *whole*, so it isn't something of which we can say 'How many are there?'—'Infinitely many'.

This general point about items with parts, considered in the abstract, is easy to apply to •space. Every space intuited as within limits—i.e. every limited region of space—is a whole whose parts, as obtained by decomposition, are

553 always themselves spaces. So every limited space is infinitely divisible.

And from that we can naturally infer a second application of the position taken in the first paragraph of this section, namely the application of it to outer appearances enclosed within limits, i.e. to •bodies. The divisibility of every body is based on the divisibility of every region of space, for space is just the possibility of a body as an extended whole. So body is also infinitely divisible, though it doesn't consist of infinitely many parts.

One might think that the notion of divisibility applies to •bodies in a quite different way from how it applies to •space, because body has to be represented in space as *substance*. The thought goes like this:

We certainly can agree that decomposition can never remove all compositeness from space; i.e. we can't make sense of the idea of a region of space that isn't made up of small regions. That's because there's nothing self-subsistent about space; a region of space doesn't, metaphysically speaking, stand on its own feet; so that **(1)** if you think away the region's being made of smaller regions, you have thought away *everything*. But it isn't similarly true that **(2)** if you think away all the compositeness from a portion of matter, you are left with nothing. What makes **(2)** false is the concept of a substance, because a substance is meant to be the subject of all compositeness—when something is composite, that means that some smaller things or substances have been composed or put together to make it up—and these substances must survive even if they are taken apart so as to dismantle the body that they make up.

But while this account of compositeness in relation to substance is true of a thing in itself, as thought through a

pure concept of the understanding, it doesn't hold for what we call *substance* in the domain of *appearance*. For this latter isn't an absolute subject, a metaphysically basic *thing* that *has* various properties and relations; rather, it is a permanent sensible image; the only way it is anything at all is in intuition, and in intuition nothing unconditioned—such as a thing that has properties and isn't itself a property of some more basic thing—is to be found.

[Kant now says firmly that the notion of subdividing matter to infinity is all right when applied to matter regarded merely as **(i)** *stuff that fills space*, but is *not* all right when applied to **(ii)** an organised body—at any rate it's not all right if it means that however far we go in pulling apart the organised body we will always find *organised* parts of it. Leibniz thought that every animal is made up of smaller animals which are made up of still smaller animals . . . and so on to infinity; and Kant, without mentioning Leibniz, declares that 'this is not a thinkable hypothesis'. His point is this:

In the case of **(i)**, the infinitely many parts come into existence *as parts* only through the process of division; since they are merely *portions of stuff*, there's nothing to mark them off from one another until *we* mark them off. But in the case of **(ii)**—an organism as conceived by Leibniz—the infinity of parts are all there already, marked off from one another by the facts of how they are organised. If there were such an infinity of already-demarkated items, the answer to the question 'How many of them are there?' is 'Infinitely many', and yet their how-many-ness, their cardinality, is perfectly determinate or definite. Kant says that this is self-contradictory.

He goes on to say that a 'how many' that is determinate or definite is 'equal to some *number*', and he clearly thinks that 'infinite number' is a contradiction in terms.]

554

[This version's awkward wrestlings with 'how many' are attempts to handle Kant's use of *Menge*. Its main dictionary meaning is as a concrete noun meaning 'multitude' or 'mob' (there was a *Menge* gathered in the plaza); but Kant often uses it as an abstract noun that is fairly like *Zahl* = 'number' in its meaning and exactly like it in its grammar. Thus, a phrase like

the *Menge* of parts in a given appearance

means

how many parts there are in a given appearance.

•Two translators have used 'number' to translate both words, but that's wrong because Kant explicitly distinguishes them. A *Zahl* is a definite, determinate, sharp-edged *Menge*; but there can be a *Menge* that isn't a *Zahl* because it is indefinite, indeterminate, fuzzy, or because it is infinite. •A third translator uses 'multitude', which is quite wrong because it's a concrete noun. •Two others translate *Menge* by 'multiplicity', which is better, but still not right because 'multiplicity' means *many-ness*, not *how-many-ness*. No one English word does the job; hence the awkwardness.]

Transition from the mathematical to the dynamical transcendental Ideas

556 When I presented the antinomy of pure reason in a list based on the transcendental ideas [see page 210], I showed what the source was of this conflict and showed that the only way to remove it is by declaring both of the opposed assertions to be false. Throughout all this I was making the common-sense assumption that all the conditions are *spatio-temporally* related to the conditioned items; and the conflict comes solely from that. It implies that all the members of the series of conditions for a given conditioned item—the series whose *totality* made all the trouble—are of the same sort throughout: a condition is always a member of the series along with the item that it conditions, and so is *homogeneous* with it. In such a series the regress was never thought of as completed; that would require thinking of some

member of it as a first member, i.e. as unconditioned, and this would always be false because all the series' members are conditioned. That's how it came about that even when there was no special interest in the *size* of the conditioned items, the size of the series of its conditions was crucial. It was the series' size that created the difficulty: reason made the series either •too long or •too short for the understanding. And there was no room for compromise there; the difficulty had to be resolved by cutting the knot.

557

But in all this I was setting aside an essential distinction that divides into two pairs the four concepts of understanding that reason promote to being ideas. According my list, two of these concepts imply a *mathematical* synthesis of appearances, and the other two imply a *dynamical* synthesis of appearances. Until now it has been all right to ignore this distinction, but now we must attend to it because reason's troubles with the dynamical transcendental ideas are open to moves that couldn't be made with the mathematical ones. [Kant explains this somewhat elaborately, using a law-court metaphor; but his basic point can be put more simply and directly: In each of the mathematical kinds of series, the thought of a termination of the series had to be the thought of something that is *in* the series but isn't conditioned as everything else in the series is—(1) a first event or outermost shell of stars, (2) simple portions of matter. But in the dynamical series—the series of (3) ever-earlier causes, and of (4) ever-more-general-and-basic-states-of-affairs—there is at least a possibility that a series is terminated (or started) by something that is *not* itself a member of it, i.e. is not homogeneous with the members of the series. In Kant's words:] The heterogeneous can be admitted as at least possible in the case of dynamical syntheses, both (3) that of causal connection and (4) that of the connection of the necessary with the contingent.

..558

Thus, in the **(1,2)** mathematical series of appearances the only conditions that we can come to are *sensible* ones, i.e. ones that are themselves parts of the series; but in the **(3,4)** dynamical series there can be conditions that are *merely intelligible* and are therefore not themselves parts of the series. In this way reason obtains satisfaction, and the unconditioned item is posited independently of the appearances, without obscuring the always-conditioned nature of the appearances or cutting the series of them short in a way that violates the principles of the understanding.¹⁷

Because the dynamical ideas allow that an appearance may have a condition that is not itself an appearance, something happens here that is altogether different from the upshot of the mathematical antinomy. In **(1,2)** the mathematical cases we were forced to denounce the opposed dialectical assertions as **both false**. In **(3,4)** the dynamical series, on the other hand, it may be that the opposed dialectical assertions are **both true**. Here is why: If we trace a series back to some **(a)** unconditioned item that **(b)** isn't sensible and so doesn't belong *in* the series, **(a)** enables us to satisfy reason's demand for something unconditioned, and **(b)** enables us to satisfy the understanding's insistence that everything sensible is conditioned. . . .

3. Solution of the cosmological idea of the totality of the derivation of cosmical events from their causes

[The above **3.** will be followed by **4.** on page 259.] When we are dealing with events, there are only two kinds of causality that

¹⁷ Understanding doesn't admit *among appearances* any condition that is itself empirically unconditioned. But if for some conditioned item that is an appearance we can conceive an *intelligible* condition x (one that isn't a member of the series of appearances), doing this without in the least interrupting the series of empirical conditions, x may be accepted as *empirically unconditioned*, without interfering with the continuity of the empirical regress.

we can conceive: causality •according to *nature* and causality •arising from *freedom*. The former is the connection in the sensible world of one state with a preceding state on which it follows according to a rule. If that preceding state had always existed, it couldn't have produced an effect right now; so it must also be •an event•, something that has happened, which implies that it must have been caused. And so we get the general result, required by a principle of the understanding, that every cause must also be an effect.

561

By 'freedom' in its cosmological sense I understand a thing's power to begin a state *on its own*, •without help or stimulus from anything else•. So •an exercise of• the causality of freedom won't result from a temporally prior cause such as is required by the law of nature. •The concept of• freedom in this sense is a pure transcendental idea, •its transcendental nature being secured by two facts about it•: First, there's nothing *in* the concept that is borrowed from experience. Second, the freedom that the concept *refers to* can't be given in any experience; because the very possibility of experience depends on its being a universal law that *every event has a cause* that is itself an event and therefore also has a cause. . . , and so on, so that the whole domain of experience, however big it is, is transformed into a sum-total of the merely natural. In this way, however, it isn't possible to get an absolutely complete causal chain, so reason creates for itself the idea of a spontaneity that can begin to act on its own, without having to be kicked into action by an antecedent cause in accordance with the law of causality.

It's especially important that this •*transcendental* idea of freedom is the basis for •the practical concept of freedom, and is the source of the difficulties that people have always had over whether practical freedom is possible. To be 'free' in the practical sense is to have a will that isn't compelled by sensuous impulses. A will

562

- is **sensuous** to the extent that it is affected by sensuous motives; and
- is it is **animal** if it can be *necessitated* by sensuous motives.

The human will is certainly *sensuous*, but rather than being *animal* it is *free*; because its actions aren't *necessitated* by sensibility—a man has a power of self-determination, independently of any coercion by sensuous impulses. [Kant gives Latin labels to the three kinds of will.]

It's easy to see that if all causality in the sensible world were mere nature, then every event would be determined by a preceding event in accordance with necessary laws. The actions of the will would be the natural effects of appearances that were their causes, which means that the will's actions would be necessary. So the abolition of transcendental freedom would carry with it the elimination of all practical freedom. Why? The short answer is that if there were no transcendental freedom, all causality would be 'mere nature'. Here is a longer answer: Practical freedom presupposes that something x that *hasn't* happened *ought* to have happened; and *this* implies that x's natural-world cause didn't determine x in such a way as to exclude a causality of our will—a causality that can act independently of and even contrary to the influence of those natural causes, •producing something that is determined in the time-order in accordance with empirical laws, •thus beginning a series of events *entirely of itself*.

563

The question of whether freedom is possible poses a challenge to psychology, but the problem about it isn't a •physiological one [see note on page 1], .i.e. it's not a problem that could be solved by an empirical study of how human minds work. Why not? Because it rests on dialectical arguments of pure reason, so that its treatment and solution belong exclusively to •transcendental philosophy. (This is an

example of a general fact: *whenever* reason gets into conflict with itself through venturing beyond the limits of possible experience, the problem that arises is *transcendental* and not *physiological*.) Transcendental philosophy can't decline the task of solving this problem, but before I get into that I must specify in more detail how it's going to go about the job.

[Kant begins this by saying that if appearances were things in themselves, all the series of conditions—the dynamical as well as the mathematical ones—would be homogeneous, so that in all four cases the trouble would concern series that were either too large or too small for the understanding. But in fact the dynamical ideas—our topic in this subsection and the next—differ from the mathematical ones in that they don't involve any issue about the *size* of the regress. They do raise an issue about whether in each case there is something unconditioned, but if there is it's something right outside the realm of appearances—neither a series that is cut off somewhere along its length nor a series that continues for ever. Kant continues:] So we can abstract from the size of the series of conditions, and consider only the dynamical relation of the condition to the conditioned.

564

In this dynamical area, we won't have any difficulty about a series's being too big or too small; our concern will be purely with the question of whether anything in the series is conditioned by something that isn't in it. So our present question about nature and freedom is this:

- Is freedom possible at all? If it is, can it co-exist with the universality of the natural law of causality? Is it right to say that every effect in the world must arise *either* from nature *or* from freedom, meaning that it can't arise from both? Shouldn't we rather say that a single event can arise in different ways from both?

All events in the sensible world are thoroughly inter-connected

in accordance with unchangeable laws of nature—that's an established principle of the Transcendental Analytic, and no exceptions are allowed. Our present question concerns whether freedom is completely excluded by this unbreakable rule, or whether an effect that is thus determined in accordance with nature might not *also* be grounded in freedom. This is a case where the common but deceptive assumption of the *absolute reality* of appearances—i.e. the assumption that they are things in themselves—exerts its harmful influence, throwing reason into confusion. If appearances are things in themselves, freedom can't be saved, for in that case nature will be the complete and sufficient determining cause of every event. . . . If, on the other hand, appearances are taken for what they actually are—not things in themselves, but merely representations connected according to empirical laws—they must themselves have grounds that aren't appearances. The effects of such an intelligible cause are appearances, so they can be determined through other appearances, but the causality of the intelligible cause is not determined in that way. [The apparent equation of non-empirical grounds of appearances with intelligible causes of appearances is Kant's, not a by-product of any liberties taken in this version. For Kant those are two ways of talking about the thing-in-itself that a given appearance is an appearance of.] While the effects are to be found in the series of empirical conditions, the intelligible cause along with its causality is outside the series. Thus the effect can be regarded as

- being free in regard to its intelligible cause

while also

- resulting from appearances according to the necessity of nature.

Expressed in this general and abstract manner, this distinction is bound to seem extremely subtle and obscure, but it will become clear when I put it to work. All I have wanted to

do here is to point out that because it's an unbreakable law that in a context of nature all appearances are thoroughly causally interconnected, the inevitable upshot of obstinately insisting on the transcendental reality of appearances is to destroy all freedom. . . .

Possibility of causality through freedom, in harmony with the universal law of natural necessity

If an appearance x in the sensible world has in itself a faculty or power that isn't an object of sensible intuition but through which x can be the cause of appearances, x's causality can be regarded from two points of view: regarded as the causality of a thing in itself, it is **intelligible in its action**; regarded as the causality of an appearance in the world of sense, it is **sensible in its effects**. (I label as 'intelligible' anything having to do with an object of the senses that isn't itself appearance.) So we would have to form both an empirical and an intellectual concept of the causality of x's faculty, with a single effect, y, falling under both concepts. That is: we can say that x has a power or faculty to produce_{emp} y, and a power to produce_{int} y. This two-sided way of conceiving a faculty possessed by an object of the senses doesn't conflict with any of our indispensable concepts of appearances or of possible experience. Here is why. Any appearance x, not being a thing in itself, is an appearance of some transcendental object that gives x the features that it has as an appearance; so the way is clear for us to ascribe to this transcendental object, besides the features it has as an appearance, a causality—a way of producing—that is not itself an appearance though its effect is to be met with in appearance. Every cause must have a character, i.e. a law of its causality, without which it wouldn't be a cause. [Kant means that if A causes B it must do so because of some facts about A's nature, facts that hook into a law

dictating that anything whose nature is like A's in the relevant respect will have something like B as an effect.] In the case we are now envisaging, there would be a subject x belonging to the sensible world which had

- (i) an *empirical character* through which its actions are thoroughly connected up with other appearances in accordance with unvarying laws of nature. . . .; this is x's character as an appearance; and
- (ii) an *intelligible character*, through which x is indeed the cause of those same actions, but which is not itself an appearance; this is x's character as a thing in itself.

(These two kinds of causality or production, one an appearance and the other not, both have effects that *are* appearances: ·we are talking about empirical and intelligible causes of, for example, someone's uttering certain words or pulling a certain trigger·.)

568 Now this acting subject x would not, in its intelligible character, have any temporal features, because time is a condition only of appearances and not of things in themselves. In x ·in its intelligible character· no *action* would *begin* or *cease*; so it wouldn't have to conform to the law governing everything that does happen in time, namely that every *event* must have its cause in the *appearances* that precede it. In short: x's intelligible causality wouldn't have a place in the series of empirical conditions through which the event is made to be necessary in the world of sense. Of course this intelligible character can never be immediately known, for nothing can be perceived except in so far as it appears. It would have to be *thought* in accordance with the empirical character—just as we can't help *thinking* a transcendental object as underlying appearances, though we know nothing of what it is in itself.

Thus, the subject x in its empirical character—i.e. in its role as an appearance—would have to conform to all the laws

of causal determination. All it would be is a part of the world of sense, and its effects must, like all other appearances, be the inevitable outcome of nature. They can in principle be completely determined by and explained through outer appearances in accordance with the laws of nature. . . .

569

In its intelligible character (though all we have of *that* is a general concept), this same subject x must be considered to be free from all influence of sensibility and from all determination through appearances. Because it is a noumenon, nothing *happens* in it; so it can't involve any change that would have to come from a prior cause, and therefore it doesn't causally depend on appearances. Therefore, because natural necessity is to be met with only in the sensible world, this active being must in its actions be free from all such necessity. No action begins *in* this active being itself; but we can quite correctly say that the active being *of itself* begins its effects in the sensible world. That isn't to say that the effects in the sensible world can begin of themselves; *they* are always predetermined—though solely through their empirical character (which is merely the appearance of the intelligible character)—by antecedent empirical conditions, so that their occurrence is just another link in the natural causal chain. That is how •freedom and •nature, in the full sense of these terms, can exist together in the same actions, according as the actions are related to their intelligible or to their sensible cause.

How the cosmological idea of freedom connects with universal natural necessity

I thought I should sketch this outline of the solution of our transcendental problem so as to give a better view of the course that reason takes in solving it. I'll now present the various factors involved in this solution, considering each in detail.

570

Here's a law of nature:

•Every event has a cause; •what a cause *C* does to cause an effect *E* must occur earlier than *E*, and must be something that has *happened* rather than a state of affairs that has always obtained, so that *C* in turn must have been brought about by a still earlier cause in the realm of appearances; and therefore •all events are empirically determined in an order of nature.

It's only because of this law that appearances constitute a *nature* and become objects of experience. It's a law of the understanding, and *every* appearance falls under it with *no* exceptions. (If we allowed that some appearance wasn't bound by this law, we would be putting that appearance
571 beyond the reach of any possible experience, turning it into a mere thought-entity, a figment of the brain.)

From this it looks as though there can't be an *absolute totality* of any back-tracking causal chain; but we don't have a problem with that, because the point has already been dealt with in the general discussion of reason's conflictedness when in the series of appearances it proceeds to the unconditioned. . . . The only question here is this: Admitting that in the whole series of events there is only natural necessity, is it possible to regard a single event as being on one hand merely an effect of •nature and on the other hand an effect due to •freedom? Or are these two kinds of causality inconsistent with one another?

[Kant now has a paragraph insistently reiterating that events, appearances, can't contribute to a causal chain without having first been produced through a causal chain.
..572 It's no use looking to *them* for instances of freedom. Then:]

Given that effects are appearances and that their causes are appearances too, is it necessary that *the causality* of their cause is exclusively empirical? Mightn't the following alternative state of affairs be the real one?

Although every effect in the ·domain of· appearance must be connected with its cause in accordance with the laws of empirical causality, this empirical causality is—without the least violation of its connection with natural causes—an effect of a causality that is not empirical but intelligible.

[On the preceding page Kant has said that the empirical character is 'merely the appearance of the intelligible character; now he is saying that empirical causality is (not the appearance of, but) an effect of intelligible causality.] [Kant continues with some stunningly obscure remarks whose general tenor is that this causality of freedom is a self-starter that doesn't have a preceding cause (and indeed doesn't occur in time), though its effect in the realm of appearance is itself an appearance that fits into an entirely natural causal chain. Then:]

We *need* the principle of the causal connection of appearances if we are to be able to explore and learn about the •natural conditions of natural events, i.e. •events' causes in the ·domain of· appearance. If we accept this principle and don't allow any exceptions to it, •physical [see note on page 193] explanations can proceed on their own lines without interference, and •the understanding gets everything it can demand—I'm talking about how the understanding in its empirical use rightly insists on seeing nothing but nature. Nothing gets in the way of any of this if we assume the following (even if we adopt it only as a fiction):

Some natural causes also have a faculty ·= power· that is ·not· •sensible but· only •intelligible, because it is activated solely by •grounds in the understanding and never by •empirical conditions, though the action of these causes in the ·domain of· appearance conforms with all the laws of empirical causality. In this way the acting subject as a phenomenal cause is tied in with nature through the unbroken dependence

of all its actions ·on their natural causes·, and it's only by ascending from the empirical object to the transcendental one that we find that this phenomenal subject contains, along with all its causality in the ·domain of· appearance, certain conditions that must be regarded as purely intelligible.

574 This won't interfere with our understanding's going about its legitimate business of determining what causes what in the domain of appearances, following the rules of nature, because in doing that we needn't *raise* the question 'What kind of ground for these appearances and their connections must exist in the transcendental subject that is empirically unknown to us?' This intelligible ground doesn't threaten our •empirical enquiries, and is solely the business of the •pure understanding. The *effects* of what the pure understanding thinks and does are to be found among the appearances, but they ·won't interfere with disciplined empirical investigations because they· must be capable of complete causal explanation through other appearances in accordance with natural laws. Our explanations of them must be utterly based on their strictly empirical character; their intelligible character (i.e. the transcendental cause of their empirical character) won't come into it because it is completely unknown to us except in so far as the empirical is a sensible sign of it.

Let us apply this to experience. Man is one of the appearances in the sensible world, and therefore one of the natural causes whose causality is subject to empirical laws. Like everything else in nature, man must have an empirical character. We come to know this character through the powers and faculties that it reveals in its effects. In inorganic or sub-human animal nature we don't find any reason to think that there's a faculty at work that is conditioned in any but a non-sensible manner. But man ·is different: he· knows all the rest of nature solely through his senses, but

knows *himself* also through pure self-awareness; and this knowledge concerns acts and inner states that he can't regard as impressions of the senses. He is thus to himself (on the one hand) •phenomenon, and (on the other hand) •a purely intelligible object because of certain faculties ·= powers· whose action can't be ascribed to the receptive- 575
ness of sensibility—faculties that we call 'understanding' and 'reason'. In particular we distinguish *reason* in a quite special and prominent way from all empirically conditioned powers. That's because reason views its objects exclusively in the light of *ideas*, and in accordance with them it shapes up the understanding, which then proceeds to make an empirical use of its own similarly pure concepts.

In all matters of conduct we use *imperatives*, which we impose as rules on our active powers; and that makes clear that our reason is causally active, or at least that we represent it to ourselves as being so. The word 'ought' expresses a kind of necessity, and a kind of connection with grounds ·or reasons·, that isn't found anywhere else in the whole of nature. All the understanding can know in nature is what

•is, has been, will be.

It's impossible that anything in nature

•ought to be

different from how it actually is at its given moment in history. Indeed, when it's only nature that we are dealing with, 'ought' has no meaning whatsoever. It's as absurd to ask what *ought* to happen in the natural world as to ask what properties a circle *ought* to have. The only legitimate questions are 'What *did* happen?' and 'What properties *does* the circle have?'

This 'ought' expresses a possible action, the reason for which is nothing but a mere •concept; whereas the reason for a merely natural action must always be an •appearance. 576
[In that sentence 'reason' translates *Grund* = 'ground'. Kant is using

it to contrast reasons or grounds like this: **(i)** 'Why ought I to help him?' 'Because that's the honourable thing to do'; **(ii)** 'Why did it burn?' 'Because lightning struck it'.] The action to which the 'ought' is being applied must indeed be possible under natural conditions, but these play no part in determining the will itself, only in determining what effect the act of the will has in the ·domain of· appearance. No matter how many natural grounds or how many sensuous impulses impel me to *will*, they can't give rise to the 'ought' but only to a willing that is far from being necessary and is always conditioned. The 'ought' *confronts* that kind of willing, limiting it, steering it, indeed outright forbidding or authorizing it. Whether what is willed belongs to mere sensibility (the pleasant) or to pure reason (the good), reason won't back down in face of any ground that is empirically given. In these situations reason doesn't follow the order of things as they show up in appearance. What it does instead, with absolutely no outside prodding, is to make for itself •an order of its own according to ideas; it adapts the empirical conditions to •this order, and on the basis of •it declares actions to be necessary even if they haven't happened and perhaps never will. Through all this, reason assumes that it can have causality in regard to all these actions, because otherwise no empirical effects could be expected from its ideas.

577 Now let's stick with this, and regard it as at least *possible* for reason to have real causality with respect to appearances. If it does, then reason. . . must exhibit an empirical character. Why? [The gist of Kant's explanation seems to be this: If the will is a *cause*, then it must operate according to rules of the 'if same-cause then same-effect' sort; so the will has to have features that enable it to fit under such a rule, and to the extent that these features have to show up empirically we can call them the 'empirical character' of the will. Kant continues:] This empirical character doesn't change, but the

effects of it do, because of changes in the environment.

Thus every man's faculty of will has an empirical character, which is nothing but the facts about the causality of his reason that show up in a regular way in his reason's effects in the ·domain of· appearance. Because of this regularity or rule, people other than the man himself can draw conclusions about. . . •what his reason does and •why, thereby making an estimate about the subjective principles of his will. This empirical character has to be discovered from the appearances that it gives rise to and from the rule to which experience shows them to conform; and from this it follows that:

All men's actions in the ·domain of· appearance are causally determined by their empirical character and by other cooperating causes. If we could get right to the bottom of all the appearances of men's wills, there wouldn't be a single human action that we couldn't predict with certainty, and recognise as necessarily flowing from its antecedent conditions.

578

[That's the first time in this work that Kant has brought *predictability* into his statement of determinism—and the last.] As regards this empirical character, then, there is no freedom; yet it's only in the light of this character that a man can be studied—if we are simply *observing* him, like anthropologists, conducting a physiological [see note on page 1] investigation into the effective causes of his actions.

But when we consider these very same actions in the light of the man's practical or moral *reasons for* them, rather than the natural *causes of* them, we find a rule and order altogether different from the order of nature. ·That this practical order is different from the natural one is shown by something I said earlier.: It could be that everything that *has happened* in the course of nature (happening inevitably because of empirical causes) *ought not to have happened*.

·But the difference between them is real even when they don't diverge·: we sometimes find—or at least think we find—that •the ideas of reason have actually proved their causality in respect of men's action considered as appearances, and that •these actions have occurred not because they were determined by empirical causes, no, but because they were determined by grounds of reason.

Granted, then, that reason can be said to have causality in respect of appearance, can its action still be said to be free, given that its •sense-related empirical character is completely and necessarily determined in all its detail? This empirical character is itself determined in the •thought-related intelligible character. But we don't know the intelligible character; our only indication of it is given by appearances; and the only immediate knowledge that these give us is of the sense-related empirical character.¹⁸ ·We nevertheless *do* bring intelligible characters into our ways of thinking about people's behaviour, but let's understand what we're doing when we do this·. In attributing an action to a thought-related cause, ·i.e. to the person's intelligible character·, we aren't saying that the action *follows* from the intelligible character in accordance with empirical laws. The action isn't *preceded* by •the conditions of pure reason, but only by •their effects in the ·domain of· appearance of inner sense. Pure reason is a purely intelligible faculty, so there's nothing temporal about it—it doesn't enter into sequences of events. The causality of reason in its intelligible character does not,

¹⁸ So the real morality of actions (merit and guilt), even that of our own conduct, remains entirely hidden from us. When we pass moral judgment on someone we can only be concerned with his *empirical* character. We can never know to what extent this character is to be attributed to the pure effect of freedom, and to what extent it's a matter of mere nature—innocent faults of temperament or sheer good luck in having a good temperament. So we can't make any perfectly just judgments about this.

in producing an effect, *arise* or come into play at a certain time. If it did, it would be subject to the natural law of appearances, according to which causal series are stretched out through time, in which case its causality would be nature, not freedom. So we can say this: If reason can have causality in respect of appearances, it is a faculty *through* which a sensible condition ·= cause· comes into play. The condition that lies within reason isn't sensible, so it doesn't come into play itself. In this way something comes into view that we couldn't find in any empirical series, namely that the condition of a successive series of events may itself be empirically unconditioned. For here the condition is *outside* the series of appearances—it's in the intelligible ·domain, not the sensible one·—so it isn't subject to any sensible condition or to having a temporally prior cause.

Yet this same cause does, in another relation, belong to the series of appearances. A man is himself an appearance. His will has an empirical character, which is the empirical cause of all his actions; and all those causings are contained in the series of natural effects and are subject to the law according to which everything that happens in time has an empirical cause. This implies that no given action. . . . can begin entirely *of itself*, ·without any temporally prior cause·. But we ·can't talk in this way about pure reason. We· can't say that the state in which *it* determines the will is preceded and caused by some other state. That's because reason isn't an appearance, so it isn't subject to any conditions of sensibility, so even as regards its causality it isn't temporal, and the dynamical law of nature—embodying the rules about what temporally and causally follows from what—doesn't apply to it.

Reason is the permanent condition of all the voluntary actions by which a man takes his place in the domain of appearance. Each of these actions is, before it actually

happens, settled in the empirical character—i.e. settled as something that is bound to happen. In respect of the intelligible character (of which the empirical character is only the sensible schema) there's no role for *before* and *after*; and every action—no matter how it relates temporally to other appearances—is the immediate effect of the intelligible character of pure reason. So reason acts freely; it isn't acted on by temporally preceding natural causes, outer or inner. Don't think of this freedom only •negatively—'pure reason is *not* subject to empirical conditions'. Looked at only in that negative way, the faculty of reason would lose its role as a cause of appearances. It should also be described in •positive terms, as the power to originate—to start up, without being prodded to do so—a series of events. So nothing begins *in* reason itself; as an *unconditioned* condition of every voluntary act, it can't have conditions that predate it. An effect of reason does have a beginning in the series of appearances, but it—the effect—never constitutes an utterly first beginning in the series.

582

Let's look at an example. I don't mean to *confirm* this regulative principle of reason by showing it at work empirically, because you can't prove transcendental propositions by examples. I want the example simply as an illustration. Let it be a malicious lie through which social harm has been done. We try first (i) to discover the motivating causes of the lie, and then in the light of these (ii) to determine how far the man who told the lie can be held accountable for the action and its consequences. In connection with (i) we trace the liar's empirical character to its sources, finding these in a bad upbringing, evil company, partly also in shameless viciousness of his natural disposition, and in frivolity and rashness; and we don't forget to look also into the on-the-spot causes that helped cause the lie. In all this we proceed just as we would in *any* inquiry into the

causal chain leading to a natural effect. But although we believe that the action was determined by all these causes, we still (ii) blame the man. We *don't* blame him for his unfortunate natural make-up, or for the circumstances that have influenced him, or even for his previous way of life. We adopt the supposition that we can •entirely set aside any facts about how his life has unrolled, can •regard the past series of conditions as not having occurred, and can •see his act as completely unconditioned by any preceding state, as if by this action the man had started up an entirely new series of consequences, doing this all by himself •and without being caused to do so by any preceding cause. Our blame is based on a law of reason according to which reason is to be regarded as a cause which—irrespective of all the above-mentioned empirical conditions—*could* have and *ought to* have made the man act otherwise. This causality of reason is to be regarded not as merely a co-operating agency but as complete in itself, even when the sensible (•empirical•) drives don't favour it but are directly opposed to it. The action is ascribed to the agent's intelligible character; at the moment when he utters the lie, the fault is entirely his. Whatever the empirical conditions of the act, •his• reason is completely free, and its failure is to be given the whole blame for the lie.

583

This judgment of accountability clearly shows us as being in a frame of mind where we think that reason

- isn't affected by those sensible influences;
- isn't liable to alteration, although its appearances—
i.e. the ways it shows up through its effects—do alter;
- doesn't have earlier states and later ones;

584

and therefore

- doesn't belong to any causal chain in which appearances necessitate other appearances in accordance with laws of nature.

Reason is present in all the actions of men at all times and under all circumstances, and is always the same; but it isn't itself in time, and doesn't fall into any new state that it wasn't in before. In respect to new states, it is *determining*, not *determinable*. So the question 'Why hasn't reason determined *itself* differently?' is illegitimate because reason hasn't been determined by anything, itself or anything else; whereas the question 'Why hasn't reason through its causality determined the *appearances* differently?' is legitimate, except that no answer to it is possible! For a different intelligible character would have given a different empirical character, because the empirical character is just the appearance of the intelligible character. When we say that in spite of his whole previous course of life the agent 'could have' refrained from lying, this only means that the act is under the immediate power of reason, and that reason in its causality is not subject to any conditions of appearance or of time. Although difference of time makes a basic difference to appearances in their relations to one another—for appearances are not things in themselves and therefore not causes in themselves—it can't affect how the action relates to reason.

585

Thus in our judgments concerning the causality of free actions, we can get as far as the intelligible cause, but not beyond it. We can know that it is free, i.e. that it is determined independently of sensibility, and that in this way it may be the sensibly unconditioned condition of appearances. But as for the question

Why in these circumstances does the intelligible character give just *these* appearances and *this* empirical character?

—that's something that our reason has no power to answer, and indeed no right to ask. (It's like asking 'Why does the transcendental object of our outer sensible intuition

give us intuition in *space* only and not some other mode of intuition?') But the problem that we had to solve doesn't require us to raise any such questions. Our question was just this: 'Can freedom and natural necessity exist together without conflict in one and the same action?' and I have sufficiently answered this. I have shown that the conditions of •freedom are quite different from the conditions of •natural necessity, so that the law of •natural necessity has no affect on •freedom, which implies that •both can exist together, without interfering with each other.

• • •

Please understand that in these remarks I haven't been trying to establish that the transcendental causes of the appearances of our sensible world *really do* include a faculty of freedom. Investigating whether *that* is so involves more than just concepts, so it couldn't be a transcendental inquiry. And anyway it couldn't have succeeded, because we can never •infer from experience anything that we can't •think in accordance with the laws of experience. I haven't even been trying to prove the *possibility* of freedom; because I couldn't have succeeded in that either: we can't from mere concepts show *a priori* the possibility of any causality, any real basis for anything. Freedom is here being treated only as a transcendental idea through which reason •plans to use something that isn't sensibly conditioned to start up a series of conditions in the domain of appearance, and so •becomes tangled in a conflict—an antinomy—with the very laws that reason itself prescribes for the empirical use of the understanding. All I have shown, all I have *wanted* to show, is that this antinomy rests on a sheer illusion, and that causality through freedom is at least *not incompatible with* nature.

558

**4. Solution of the cosmological idea of the totality of
the dependences of appearances as regards their
existence in general**

587 [The above **4.** follows on from **3.** on page 249.] In subsection 3 we were looking at changes in the sensible world in their role as a dynamical series, with each member subordinate to another as effect to cause. Now we're going to use this series of states only to guide us in our search for a being that can serve as the highest condition of everything that is changeable, i.e. in our search for the *necessary being*. What this is about is not •unconditioned causality but the •unconditioned existence of substance itself. That is, it's not about events' being caused by something that wasn't itself caused, but about states of affairs' depending on the existence of something that doesn't itself depend on anything. So our topic is not a series of •intuitions in which one intuition is the condition of the next, but rather a series of •concepts. [That contrast between intuitions and concepts presumably echoes the unannounced switch from a series of 'changes' in the first sentence of this paragraph to 'this series of states' in the second.]

[Kant now offers an obscure paragraph, which owes its difficulty partly to a misplaced and distracting argumentative flourish about what would be the case 'if appearances were things in themselves'. The relevantly working content of this paragraph is just the point that if we're looking for something whose existence is unqualifiedly necessary (i.e. not merely necessary *for* such-and-such), we won't find it •in the series of all appearances, because *they* are all 'conditioned in their existence', meaning that they all exist only contingently. The following paragraph says that perhaps we can find it •outside
..588 the series of appearances. Thus:]

But the dynamical regresses differ in an important way from the mathematical ones. Any mathematical regress is

concerned only with **(1)** combining parts to form a whole, or **(2)** dividing a whole into parts; so the conditions of such a series must always be regarded as parts of the series—i.e. as appearances and as homogeneous with the rest. In a dynamical regress, on the other hand, we are concerned not with wholes and parts but with **(3)** the derivation of a state from its cause or **(4)** derivation of the contingent existence of substance itself from necessary existence. In **(3)** and **(4)**, therefore, the condition doesn't have to be a part of an empirical series along with the conditioned. [Just to make sure that it's clear: **(3)** concerns facts about *alterations in* substances that stay in existence throughout, while **(4)** concerns facts about the existence of those substances.—Not long ago we saw Kant writing in terms of **(3)** 'changes' versus **(4)** 'states'; now he is writing in terms of **(3)** 'states' versus **(4)** 'substance'. This is bad behaviour, but it probably has no doctrinal significance.] So there remains a way of escape from this apparent antinomy: perhaps the two conflicting propositions are both true when placed in different contexts. The situation may be this:

All things in the world of sense are contingent, and thus have only an empirically conditioned existence; but there is a non-empirical condition of the whole series; i.e. an unconditionally necessary being.

This necessary being, as the *intelligible* condition of the series, wouldn't be a member of the series, whose members are all empirical; so it wouldn't affect the *empirically conditioned* status of each member of the empirical series. . . . This way of **(4)** laying an unconditioned being at the basis of appearance differs from the approach in **(3)** involving the empirically unconditioned causality of freedom. In **(3)** it was the causality of a certain thing that was intelligible and unconditioned; the thing itself, the substance, that *had* the freedom was thought of as a member of the series of conditions—it was in fact just a plain empirical-world person such as you or me.

Whereas here in (4) the necessary being must be thought of as itself lying outside the series of the sensible world, and as purely intelligible. That's the only way to save it from falling under the law that declares all appearances to be contingent and dependent.

The *regulative principle of reason*, in its bearing on (4) our present problem, says this:

•Everything in the sensible world has an empirically conditioned existence, and isn't unconditionally necessary in respect of any of its qualities. For every member of the series of conditions we must •expect, and as far as possible •seek, an empirical condition in some possible experience. We never have any right to derive an existence from a condition outside the empirical series, or to regard anything *in* the series as utterly independent and self-sufficient.

590 Yet this principle doesn't *at all* debar us from recognising that the whole series may rest on some intelligible being that is free from all empirical conditions and itself contains the ground of the possibility of all appearances.

I'm not trying to prove the unconditionally necessary existence of such a being; I'm not even trying to prove the *possibility* of a purely intelligible condition of the existence of appearances in the sensible world. •My point is merely that if such a being *is* impossible, its impossibility can't come from any considerations concerning the sensible world. We lay down two limitations. • On the one hand:

•We set limits to *reason*, preventing it from leaving the guiding-thread of empirical conditions and straying into the *transcendent* and explaining things in terms that can't be cashed out empirically.

And also, on the other hand (•this being my present point):

•We set limits to *the law of the purely empirical use of the understanding*, preventing it from •making deci-

sions about the possibility of things in general—i.e. of things of *any* sort—•and •ruling that intelligible things are not merely incapable of explaining appearances but are *impossible*.

What I have been arguing is that •the thoroughgoing contingency of all natural things and of all their empirical conditions is quite consistent with •the. . . .assumption of a necessary though purely intelligible condition; and that as there is no real contradiction between the two assertions, *both may be true*. •This is a claim of the form 'Q is true and P is consistent with it', and *not* one of the form 'it is possible that P'. Perhaps an unqualifiedly necessary being. . . .*is* in itself impossible, but its impossibility can't be inferred from •the universal contingency and dependence 591 of everything belonging to the sensible world, or from •the principle that forbids us to stop at any member x of the contingent members and appeal to a cause outside the world as explaining x. Reason goes along one path in its empirical use, and along its own special path in its transcendental use.

The sensible world contains nothing but appearances; these are mere representations that are always sensibly conditioned; the objects we encounter in this domain are never things in themselves. So it's not surprising that in dealing with a member—*any* member—of the empirical series we're never justified in making a leap out beyond the sensible network. Making such a leap would be treating appearances as if they were things in themselves that exist apart from their transcendental ground and can be left standing while we look for an outside cause of their existence. That •procedure of leaping outside and looking around• is what we would eventually have to do if we wanted to explain the existence of contingent *things*; but with mere *representations* of things the procedure isn't legitimate. The point is that

592 their contingency is *itself* merely one of the phenomena, so it can be dealt with only in terms of the regress that governs the phenomena, i.e. solely in terms of the *empirical* regress. But the thought of an *intelligible* ground of the appearances. . . .as being free from the contingency of appearances doesn't conflict with •the unlimited empirical regress in the series of appearances or with •their thoroughgoing contingency. That's all I had to do in order to dispose of the apparent antinomy; and this is the *only* way to do it. [Kant goes through the argument again, and again insists at length that allowing for a purely intelligible condition of appearances doesn't interfere in the slightest with the regulative principle that orders us always to expect and seek empirical conditions for empirical conditioned items.]

Concluding note on the whole antinomy of pure reason

..593 So long as our business with our concepts of reason has to do only with •the totality of conditions in the sensible world and •the question of what they can do to satisfy reason, our ideas are at once transcendental and *cosmological*—i.e. transcendental and *about-the-world*. But as soon as we posit something unconditioned (and that's what all this is really about) in something that is entirely outside the sensible world and thus outside all possible experience, the ideas become *transcendent*. Until that happens, they serve only for completing the empirical use of reason—an idea •of completeness• that can't ever be fully achieved though it must always be pursued. But *now* the ideas cut loose entirely from experience, and make for themselves objects for which experience supplies no material, and whose objective reality is based not on completion of the empirical series but on pure *a priori* concepts. [To say that they 'make for themselves objects which. . .' is to say that they purport to be ideas *of* something which. . .] Such transcendent ideas have a purely intelligible object. It's

all right for us to admit this object as a transcendental object about which we know nothing else; but we can't have a determinate thought about it, picking it out in our thought as 'the item that is F and G and H', where those letters stand for predicates expressing what the object is intrinsically like. That's because we don't have *either* of the things that would be needed for such a thought: because this object is independent of all empirical concepts, •we are cut off from any reasons that could establish that the object is even possible, and •we haven't the least justification for assuming that there is such an object. So it's a mere thought-entity. But we're pushed into risking this step by just one of all the cosmological ideas, namely the one that gives rise to the fourth antinomy. That's because the existence of appearances is never self-explanatory; it is always conditioned •by something else•, so we have to look around for something different from all appearances, i.e. for an intelligible object in which this contingency may terminate. But once we have allowed ourselves to assume a self-subsistent [= 'self-explanatory'] reality entirely outside the domain of sensibility, appearances can only be viewed as contingent ways in which beings that are themselves intelligences represent intelligible objects. So •in our attempt to get some sort of hold on things that are only intelligible•, all we are left with is *analogy*, through which we can use the concepts of experience to form *some* sort of concept of intelligible things—without knowing anything about these things as they are in themselves. Since anything contingent can be known only through experience, and we're concerned here with things that are not to be in any way objects of experience, we'll have to derive our knowledge of them from that which is in itself necessary, i.e. from pure concepts of things in general. Thus the very first step that we take outside the world of sense requires us to begin our search 594 595

for new knowledge of intelligible things by investigating the unqualifiedly necessary being, and to derive from the concepts of *it* the concepts of purely intelligible things in general. That's what I aim to do in the next chapter.

Chapter 3

The ideal of pure reason

1. The ideal in general

We have seen above that no objects can be represented through pure *concepts of understanding*—i.e. through the categories—apart from the conditions of sensibility, because without sensibility there's nothing to give the concepts objective reality and all they have to offer is the mere *form* of thought without any *content*. But when the categories are brought to bear on appearances, we can encounter concrete instances of them—e.g. having not merely

- abstract thoughts about *if-then-relatedness*

but also

- contentful thoughts about *this event's causing that one*

and so on. But *concepts of reason*—i.e. *ideas*—are even further removed from objective reality than the categories are, because there are no appearances that could be concrete instances of them. They involve a certain *completeness* that outruns anything that empirical knowledge could possibly achieve. All reason is doing with its ideas is aiming at systematic unity—a unity that it won't ever completely achieve, but will try to get as close to it as it's empirically possible to get.

What I call 'ideals' of reason seem to be even further removed from objective reality than other ideas. An 'ideal' in my sense is an idea of **(2)** some individual thing that could be (or even *is*) **(1)** fully specified just by that idea. [In this context, *bestimmen* and its cognates, usually translated by 'determine' etc., are translated by 'specify' etc. The meaning is the same, but we needed a rest from 'determine' etc., which Kant uses 900 times in this work.] The **(1)** 'full specification' feature is not enough on its

own to make an idea an ideal; there has also be the feature that the idea **(2)** picks on an individual. The difference is an intellectual analogue of the difference between **(1)** a complete adjectival description of something and **(2)** a proper name of something. [Kant wrote this in terms of 'the ideal', as though there were only one, but that isn't his view; before long we'll see him writing of something's being 'an ideal'. His considered view is that **(a)** 'ideal' is a general term that could apply to several items, and that **(b)** each ideal is a concept that purports to apply to just one item. His ways of using the singular phrase 'the ideal' may reflect a tendency to let **(b)** suppress **(a)**.]

The thought of *humanity* in its complete perfection contains not only

- (1)** all the essential qualities of human nature, the ones that constitute our concept of it—with these extended to the point where they completely conform with humanity's ends and thus constitute our **idea of perfect humanity**,

but also

- (2)** everything else, additional to **(1)**, that is required to make the thought in question completely specific, with every detail filled in in such a way as to make this our **idea of the perfect man**—this being not merely an idea but an **ideal**.

(The filling in of details is logically straightforward: from each pair of contradictory predicates, select one.) What is an 'ideal for us' was in Plato's view an *idea in the divine understanding*, an individual object of the divine mind's pure intuition, the most perfect F for every possible value of F, and the archetype of which all the F things in the domain of appearance are copies.

Without flying *that high*, we have to concede that human reason contains not only ideas, but also ideals; they don't have creative power, as Plato's do according to him; but they have *practical* power (as regulative principles), and form

the basis of the possible perfection of certain *actions*. [In this context, 'practical power' = 'moral power'.] Moral concepts involve something empirical (pleasure or unpleasure), which stops them from being completely pure concepts of reason. And yet they can serve as examples of pure concepts of reason, doing that through their formal features, in connection with the principle through which reason sets bounds to an intrinsically lawless freedom. *Virtue* is an **idea**, and so also is *human wisdom in its complete purity*. But the Stoics' *wise man* is an **ideal**, i.e. a man existing only in thought but completely fitting the idea of wisdom. Just as

- the idea gives the *rule*,

so also in this sort of case

- the ideal serves as the *archetype* that completely specifies the copy.

Our only standard for our actions is the conduct of this divine man within us: we compare ourselves with him, judge ourselves in terms of him, and so reform ourselves—though we can't match up with him completely. Such ideals don't have objective reality, but that doesn't mean that they're figments of the brain. They supply reason with a standard that is indispensable to it. Reason needs a concept of *that which is entirely complete in its kind*, as a basis for judging things that are incomplete—measuring how far and in what ways they fall short. How about having an example of the ideal in the ·domain of· appearance? for example a wise man in a novel? It can't be done; and even to try is rather absurd and not very edifying, because any attempted portrayal of an ideal man will naturally fall short, thereby constantly eroding the completeness of the idea and making it useless as an illusion at which one might morally aim. This can cast suspicion on *the good* itself—the good that has its source in the idea—by creating the impression that it's just a fiction.

598

[Then a paragraph in which Kant distinguishes an ideal of reason, which is essentially precise and definite, from products of the imagination, which are fuzzy assemblages of left-overs from past experience. He is impolite about painters who carry these in their heads and claim to use them in producing and judging paintings. Then:]

..599

In contrast with that, what reason aims at with its ideal is complete specificity [= 'detailedness'] in accordance with *a priori* rules. So reason thinks for itself an object that it regards as being completely specifiable in accordance with principles. But experience won't supply the conditions that are required for such specificity; so this concept is a transcendent one.

2. The transcendental ideal

Every *concept* is indeterminate because of what it doesn't contain, and is subject to this principle of determinability:

- Of every pair of contradictory predicates, only one can belong to a concept.

This principle is based on the law of contradiction. So it's a purely logical principle—it abstracts from the entire content of knowledge and is concerned solely with its logical form.

Every *thing* x is possible only because it conforms also to this principle of complete determination:

If all the possible predicates of things are set alongside their contradictory opposites, then one of each pair of contradictory opposites must belong to x. 600

This principle doesn't rest merely on the law of contradiction; for, besides considering each thing in its relation to the two contradictory predicates, it also considers it in its relation to *the sum of all possibilities*, i.e. to the sum-total of all predicates of things. Presupposing this sum-total as being an *a priori* condition, the principle represents everything as deriving its own possibility from the share that it has of

this sum of all possibilities. So this principle of complete determination concerns content, not merely logical form. It is the principle of

- the *synthesis* of all predicates that are to constitute a thing's complete concept,

and not merely the principle of

- analytic* representation of a thing through one of two contradictory predicates.

601 It contains a transcendental presupposition, namely the material for all possibility, with that being regarded *a priori* as containing the data for the particular possibility of each thing.

The proposition 'Everything that exists is completely determinate' doesn't mean only that each existing thing has

- one out of every given pair of contradictory predicates,

but that each existing thing has

- one out of every contradictory pair of possible predicates.

What this proposition does is not merely •to set predicates off against one another logically, but rather •to set the thing itself off, in transcendental fashion, against the sum of all possible predicates. So what it says is this: knowing a thing x completely would involve knowing every possible predicate P and characterizing x as either having or lacking P. The concept of the **complete** nature of a thing is thus one of which there can't be a concrete instance; so it's based on an idea that resides only in our reason. . . .

This idea of the *sum of all possibility*, in its role as what's needed for the complete specification of every individual thing, is itself unspecific regarding the predicates that may make it up; our only way of thinking of it is through the utterly unspecific thought 'the sum of all possible predicates, whatever that may be'. But if we look closer and

harder, we find that many predicates can be excluded from it for either of two reasons: (1) they are *derivative* from other predicates (and so don't belong in this idea which is a *basic* concept); (2) they are incompatible with one another. With these exclusions, this idea does indeed turn itself—refine itself—into a concept that is *a priori* completely specific, thus becoming the concept of an individual object that is completely specified by the mere idea; so the idea must be labelled an 'ideal' of pure reason. 602

When we consider all possible predicates, not merely logically but transcendently (i.e. in terms of the content that can be thought *a priori* as belonging to them), we find that through some of them a being is represented, through others a mere not-being. Logical negation, indicated through the little word *not*, doesn't properly refer to a concept but only to the relation between two concepts in a judgment; so it's nowhere near to being able to specify a concept in terms of its content. . . . A transcendental negation, on the other hand, signifies not-being in itself, and is opposed to transcendental affirmation, which is a Something the very concept of which in itself expresses a being. Transcendental affirmation is therefore called 'reality' [German *realität*, from Latin *res* = 'thing'], because through it alone, and so far only as it reaches, are objects something (things); whereas its opposite, transcendental negation, signifies a mere lack—all it yields is the cancellation of every thing. 603

The only way to have a specific thought of a negation is to base it on the opposed affirmation. Someone born blind can't have the least notion of darkness because he has none of light. The savage knows nothing of poverty, because he has never encountered wealth. An ignorant person has no concept of his ignorance because he has none of knowledge,

and so on.¹⁹ All concepts of •negations are derivative in this way; it's the •realities that contain the data and the material—or the transcendental content—for the possibility and complete specification of all things.

604 So we get this result: If the complete specification ·of any individual thing· is based in our reason on a transcendental substratum that contains the whole store of material from which all possible predicates of things must be taken, this substratum can't be anything but the idea of an *all-of-reality*. All true negations are nothing but *limits*—and we couldn't call them that if they we'd based on the unlimited (the *all*).

[Kant's next paragraph is horribly difficult. In it he introduces the term *ens realissimum*, which is Latin for 'most real being' This phrase occurred widely in mediaeval and early modern philosophy; it was often understood, as it is here by Kant, as the concept of *being that has all positive properties*, i.e. a being with nothing even slightly negative in its nature. What Kant has been saying is that any individual thing must have one property out of each basic pair of properties of the form *F/not-F*, and in this paragraph he identifies the *ens realissimum* as the individual that has, out of each such pair, the *positive* one, the one that 'belongs to being absolutely'. Only one thing can answer to that description, so

the concept of an *ens realissimum*
is in fact

the concept of **the** *ens realissimum*;
which means that this concept counts as not just an 'idea'
but an 'ideal' of reason [see page 263 for the explanation of 'ideal'

¹⁹ Of the many wonderful things we can learn from the observations and calculations of astronomers, the most important is the depth of our *ignorance*. Without the help of the astronomers, common sense would never have given us an adequate sense of how much we don't know. Reflecting on this has to make a big difference to our decisions about how to employ our reason.

in terms of individual things]. Furthermore, Kant thinks of this concept as the basis for every other individual's completely determinate nature: the complete story about the properties of any individual thing *x* is the story of *which selection* of the properties of the *ens realissimum* *x* has. Thus, this 'ideal' is the basic condition of the possibility of every individual thing that exists; which means that it is a *transcendental* ideal [see pages 25–26 for the explanation of 'transcendental' in terms of 'making knowledge possible']. Furthermore, the concept of the *ens realissimum* is the only genuine ideal that human reason is capable of, because this is the only case in which we can have a universal concept *C*—a concept of

being that is thus-and so

—and know *a priori* that only one thing falls under the concept, so that although it is in form a universal concept it is in fact 'the representation of an individual'. Kant continues:]

The logical specification of a concept by reason is based on a disjunctive (·either-or·) inference of reason, in which the first premise contains a logical division (the division of the sphere of a universal concept), the second premise limiting this sphere to a certain part, and the conclusion specifying 605 the concept by means of this part. [In the remainder of this paragraph, Kant makes some remarks about disjunctive inferences, i.e. ones of the form 'P or Q, Not P, therefore Q' as a basis for his claim that when reason uses the transcendental ideal as the basis for its 'specification of all possible things', it is proceeding in a manner that is 'analogous' to what it does in disjunctive inference. And he reminds us that he has already made this connection. The details of his obscure account of the logical-inference side of this analogy are not needed for what follows, namely:]

It goes without saying that reason's purpose of representing the necessary complete specification of things doesn't involve it in presupposing the *existence* of a being that 606

corresponds to this ideal; all it needs is the *idea* of such a being, as a basis for its thought of the absolutely complete nature of this or that limited thing. So the ideal is **the archetype of all things**, which are all imperfect copies of it. Each individual thing is infinitely far from being a perfect or complete copy of this ideal; but each of them approximates to it to some degree, and the source of any thing's possibility is such overlap as it has with the idea ·I'm talking about, the idea of the *ens realissimum*·.

So all possibility of things. . . must be regarded as •derivative, with the sole exception of the possibility of the thing that includes in itself all reality, and *that* must be regarded as •original ·= non-derivative·. That's because the only way anything else can be distinguished from the *ens realissimum* is through negations; and a negation is merely a limitation, a blockage to a thing's having greater reality than it does have; so every negation presupposes the reality that it is a negation *of*; so that the whole story about the intrinsic nature of any individual thing is derived from the *en realissimum*. ·For example,

•some predicates that are also predicates that fit the *ens realissimum*,

together with

•negations of all the other predicates that fit the *ens realissimum*

express the entire intrinsic nature of *you*·. All variety among things consists in the many different ways of limiting the concept of the highest reality—the *ens realissimum*—that is their common substratum, just as all shapes are the many different ways of limiting infinite space. This object of reason's ideal can therefore be labelled 'the primordial Being', or 'the supreme Being', or 'the Being of all beings';
607 but these labels don't signify

the objective relation of •an actual object to •other things,

but just

the relation of •an idea to •concepts.

They don't tell us *anything* regarding the existence of a being of such outstanding pre-eminence.

We can't say that a primordial being is made up of a number of derivative beings, because the derivative beings presuppose the primordial one and therefore can't themselves constitute it. So the idea of the primordial being must be conceived as simple.

In my first rough outline I said something that isn't strictly correct. It is in fact never really right to speak as I did of the derivation of some limited possibility from the primordial being as a *limitation* of its supreme reality, as though it were dividing it up (·e.g. speaking of *your* intrinsic nature as what we get by slicing out from the *ens realissimum* a certain subset of its properties·). If that *were* correct, then the primordial being would be a mere aggregate of derivative beings, and I have just shown that that's impossible. The real truth of the matter is that the supreme reality must underlie the possibility of all things not as their *sum* but as their *basis*; and the source of the variety among things is not ·different ways of· limiting •the primordial being itself, but ·different ways of· limiting •everything that follows from it. ·That really is a different story, because· what follows from it includes. . . everything that is real in the domain of appearance, and there's no way *that* could be an ingredient in the idea of the supreme being.

608

If we follow through on this idea of ours by hypostatizing it [here = 'thinking of it as standing for something objectively real'], we'll be able to specify the primordial being through the mere concept of the highest reality—picking it out as being that is •one, •simple, •all-sufficient, •eternal and so on. . . . The

concept of such a being is the concept of *God*, taken in the transcendental sense; and therefore (as I said before) the ideal of pure reason is the object of a transcendental *theology*.

609 However, to use the transcendental idea in that way would be going beyond the limits of its purpose and validity. When reason used the idea as a basis for the complete specification of •things, it was using it only as the *concept* of all reality, without requiring that this reality to be objectively given and itself to be a •thing. We have no right to think that this ideal—a thing-like upshot of our bringing together the manifold of our idea—is itself an individual being; we have indeed no right to assume that it is even possible. And none of the ·theological· consequences that flow from ·treating· such an ideal ·as a real thing· have any bearing on *the complete specification of things*; yet *that* is just what the idea has been shown to be necessary for.

It's not enough just to *describe* the procedure of our reason and its dialectic; we must also try to discover the sources of this dialectic, so as to be able to *explain* the illusion it involves as a phenomenon of the understanding. [Of the *understanding*? But hasn't Kant been saying over and over again that the illusion is a pathology of *reason*? Good question! But wait!] ·And it certainly *can* be explained·, because the ideal that we're talking about is based on a natural idea, not an artificial one that we have simply *chosen* to construct. So this is my question: How does it come about that reason •regards all possibility of things as being derived from one single basic possibility, namely that of the highest reality, and then •supposes that this one possibility is contained in one special primordial being?

The discussions in the Transcendental Analytic provide the answer. For an object of the senses to be possible is for it to relate to our thought in a certain way. And *how* it

relates to our thought is a two-part story: •its empirical *form* can be thought *a priori*, and •the remainder has to be given through sensation. That 'remainder' constitutes the *matter* of an experience, it corresponds to *reality* in the domain of appearance; and it has to be •given, because otherwise we couldn't even •think about it as a possibility. A complete specification of an object of the senses involves checking it against all the empirical predicates there are, specifying with each predicate whether the object in question is a yes or a no. [Kant then gives a very obscure reason for saying that for this procedure to work, the sum of all predicates that it appeals to must be thought of as possessed by 'experience, considered as a single all-embracing item'; the characters of empirical objects, and their differences from one another, must be based on their different selections from the set of all the predicates of this single item, *experience*. Then:] The fact is that •the only items that can be given to us are objects of the senses, and •they can be given only in the context of a possible experience; so ·we get the principle that·

(a) nothing is an object *for us* unless it presupposes *the sum of all empirical reality* as the condition of its possibility.

This principle applies only to things given as objects of our senses, but a natural illusion kicks in, making us regard the principle as holding for things in general, things as such. ·That amounts to replacing it by this:

(b) nothing is an object *of any kind* unless it presupposes *the sum of all reality* as the condition of its possibility.

(Notice the disappearance of 'for us'.) ·And so by omitting this limitation ·to sensible things· we mistake •the empirical principle of our concepts of the possibility of things viewed as appearances for •a transcendental principle of the possibility of things in general.

We go on from there to hypostatise [see note on page 191] this idea of the sum of all reality. Here's how we go about that. **(1)** We replace the thought of the *distributive* unity of the empirical use of the understanding by the *collective* unity of experience as a whole; **(2)** then we think of this experience-as-a-whole as being one single *thing* that contains all empirical reality in itself; and then finally **(3)** by means of the switch from **(a)** to **(b)** we switch from the concept of *that* 'single thing' to the concept of a thing that stands at the pinnacle of the possibility of *all* things, and supplies the real conditions for their complete specification.

3. Speculative reason's arguments for the existence of a supreme being

Although reason has this pressing need to presuppose something that can provide the understanding with a basis for completely specifying its concepts, it doesn't infer from this need that the 'something' in question is a real being—it's much too aware of the presupposition's ideal and merely fictitious nature for *that*. But there's another direction from which reason is pressured to think of the *ens realissimum* as a real being, namely: reason is impelled to seek a resting-place in the regress from given conditioned items to the unconditioned. This unconditioned item still isn't given as being in itself real, or as having a reality that follows from its mere concept; but it's the only thing that can complete the series of conditions when we track these back to their bases. That's the natural route that our reason leads us all to follow—even the least reflective of us—though not everyone sticks with it. It doesn't start from concepts but from common experience, so it is based on something actually existing. But if this basis—this ground floor—doesn't rest on the immovable bedrock of the absolutely necessary, it

subsides. And the 'rock' won't provide stability either if there's empty space beyond and under it, in the form of unanswerable 'Why?'-questions that are raised by it. Its way of avoiding that is to *fill everything up* so that there's no room for any further 'Why?'—which it does by being infinite in its reality.

If something exists—no matter what—then a place must be found for something that exists *necessarily*. Why? Because a contingent item exists only under the condition of another contingent item as its cause, and from this we must infer yet another cause, and so on until we are brought to a cause that is not contingent, its existence being unconditionally necessary. That's the argument reason relies on in its advance to the primordial being.

Now, reason looks around for a concept that would fit a being that exists in this noble way—existing with unconditioned necessity. It isn't aiming to infer *a priori* from the concept that the thing it stands for really exists (if *that's* what it was up to, it wouldn't have to look any further than mere concepts, with no need to start from a fact about something's existing). All it wants is to find, among all the concepts of possible things, the concept that is perfectly compatible with absolute necessity. In reason's view, the first step in the argument has already established that there must be *something* whose existence is unqualifiedly necessary. If after setting aside everything that isn't compatible with absolute necessity it is left with just one thing, *that* thing must be the unqualifiedly necessary being. It makes no difference whether its necessity can be comprehended, i.e. whether its existence can be inferred from its concept alone.

Something that contains in its concept the 'Because. . . ' for every 'Why. . .?', that is not defective in any part or any respect, that is in every way sufficient as a condition, seems to be just the thing to count as existing with absolute

necessity. For one thing, because it *contains* the conditions of everything that is possible, it can't in its turn be conditioned by anything else; so it satisfies at least that much of the concept of unconditioned necessity. No other
 614 concept can match up to this, because each of the others lacks something that it needs for completion, so that it can't have this characteristic of independence from all further conditions. Given that something x doesn't contain the highest and in all respects complete condition, we can't infer
 •that x is itself conditioned in its existence; but we *can* infer
 •that x doesn't have the unique feature through which reason can know *a priori* that some thing is unconditioned.

Thus, of all the stock of concepts of possible things it's **(a)** the concept of a *most real being* that is the best candidate for the role of **(b)** concept of *an unconditionally necessary being*; and though **(b)** may not be completely adequate to **(a)**, we have no choice in the matter: we see that we have to stick with **(b)**. We can't just drop **(a)** the existence of a necessary being; and if we are to retain it, we need a candidate for the role, and in the whole field of possibility we can't find a better one than **(b)** the most real being = *ens realissimum*.

That's the natural way in which human reason goes about this. It starts by convincing itself of the existence of *some* necessary being. It recognizes this as having an unconditioned existence. It then looks around for the concept of

•that which is independent of all conditions,
 615 and finds it in the concept of

•that which is the sufficient condition of everything else,

which is the concept of

•that which contains all reality.

Now, this total-without-limits is absolute unity, and carries with it the concept of an *individual* being—namely the supreme Being. In this way reason concludes that the

supreme Being, as the primordial ground of all things, exists by absolute necessity. [The point of the repeated 'that which' was to keep 'thing' or 'individual' out of sight until Kant was ready to argue his way to it. German has a way of doing this that is less clumsy than our 'that which'.]

How we evaluate that procedure depends on what we're trying to do. **(1)** If the existence of some sort of necessary being is taken for granted, and it's also agreed further that we must reach a decision about *what* being this is, then the procedure described in the preceding paragraph obviously has a certain cogency. That's because the best choice (really there is no *choice* because the other candidates are non-starters) is the absolute unity of complete reality as the ultimate source of possibility. [The phrase 'the absolute unity of complete reality' conservatively translates what Kant wrote. He is referring to the *ens realissimum* = the most real being, perhaps intending his phrase to mean something like 'an individual thing that in some way encompasses the whole of reality'.] **(2)** But if we aren't under pressure to come to any decision, and prefer to leave the issue open until the full weight of reasons compels assent—i.e. if our present task is merely to judge how much we really know about this problem, and what we merely flatter ourselves that we know—then the procedure I have described appears, when looked at with an impartial eye, in a much less favourable light.

It is in fact defective even if the two claims that it makes are granted. •First, the claim that from any given existence
 616 (e.g. my own existence) we can correctly infer the existence of an unconditionally necessary being. •Secondly, the claim that the what is needed for a concept of a thing to which we can ascribe absolute necessity is provided by the concept of a being that

- contains all reality and therefore
- contains every condition and therefore
- is absolutely unconditioned.

Granting both those claims, it still doesn't follow that the concept of a *limited* being that *doesn't* have the highest reality is logically debarred from absolute necessity. As between these two concepts:

- (a) a *limited* being that *doesn't* have the highest reality,
- (b) being that contains all reality,

although (a) doesn't contain the element *unconditioned* that is involved in (b), we shouldn't infer that anything falling under (a) must be conditioned. . . . On the contrary, we are entirely at liberty to hold that *all* limited beings are unconditionally necessary, despite the fact that we can't infer their necessity from the universal concept we have of them, i.e. from the concept *limited being*. So this argument hasn't given us the least concept of the properties of a necessary being; it's a complete failure. [The 'argument' in question is the 'natural' procedure of human reason that Kant expounded on the preceding page.]

617 And yet the argument still has a certain importance, and it carries some authority that can't be summarily stripped from it just because of its logical short-fall. Suppose that the following is the case:

Certain moral obligations are laid upon us by the idea of reason, but they don't have any *reality* when applied to us, i.e. they aren't accompanied by any *incentives*, unless the law expressing them is made effective and given weight by a supreme being.

If that's how things stand, we are obliged to follow the best and most convincing concept of the supreme being that we can find, even if it does fall short logically. The stand-off in the speculative sphere, with neither side able to secure

its position logically, is broken by a practical consideration, namely our duty to *decide*. Granted that reason can't make a conclusive case for either answer to the question of whether there is a supreme being, it does here have a pressing *incentive* to go one way rather than the other; and the case for doing so is at least better than any other that we know; if reason didn't go along with this and judge accordingly, it would be open to criticism from itself.

This argument rests on the intrinsic insufficiency of the contingent, which means that it is transcendental; but it's so simple and natural that it is found convincing by the plainest common-sense when that comes into contact with it. We see things alter, come into existence, and go out of existence; so there must be a cause for their existence or at least for their changes of state. But any cause that can be given in experience raises the same causal question. If we are to think there's an end to the series of causal questions we must postulate some highest cause—a cause that isn't an effect. Where can we more neatly locate this *highest* causality than where there also exists the *supreme* causality?

[The two adjectives translate *oberste* and *höchste* respectively. They don't have clearly different meanings; but in this context they seem to express the notions of a cause that is the 'highest' member of some causal chain and causality that is 'supreme' in the sense of being at the top of every causal chain.] That is to locate it in the being that contains primordially in itself the sufficient ground of every possible effect, a being that we can easily manage conceptually by thinking of it as *the being that has all-embracing perfection*. We then go on to regard this supreme cause as unqualifiedly necessary, because we find it utterly necessary to ascend to it, and find no reason to pass beyond it. And so it is that in all peoples there shine amidst the most benighted polytheism some gleams of monotheism, not by reflection and deep theorizing but simply by the natural course of the

common understanding as it gradually comes to grasp its own requirements.

There are only three possible ways of proving the existence of God by means of speculative reason.

619 All the paths leading to this goal either **(1)** begin from determinate experience in which we learn about the specific constitution of the world of sense, and ascend from that through the laws of causality to the supreme cause outside the world; or **(2)** have experience as their empirical basis but without any details about it, starting from the bare fact that something exists; or **(3)** set all experience aside and argue completely *a priori*, from mere concepts, to the existence of a supreme cause. These are **(1)** the *physico-theological* argument for God's existence, **(2)** the *cosmological* argument, and **(3)** the *ontological* argument. There are no others. There *can't be* any others.

I'm going to show that reason can't get any further along the empirical path than it can along the transcendental path, and that its no use it's stretching its wings so as to soar above the world of sense by the sheer power of speculation. In the preceding paragraph, I took the three theological arguments in the order in which gradually expanding reason takes them; but now I'll take them in the reverse of that order. The reason for that is something that I shall show in due course, namely: although experience is what first prompts this enquiry, it is the *transcendental concept*—the one highlighted in the ontological argument—that reason is aiming at in the other two arguments as well. So I shall start by examining the transcendental ('ontological') argument, and will then look into the question of what if anything can be done to strengthen it by adding an empirical factor.

4. There can't be a successful ontological argument for the existence of God

From things I have already said it's obvious that the concept of *absolutely necessary being* is a concept of pure reason, i.e. a mere idea whose objective reality is emphatically *not* proved by the fact that reason requires it. This latter claim goes for all ideas of reason, of course, not just this one. An idea of reason only directs us towards some kind of completeness that we can't actually achieve, so it serves to •set boundaries for the understanding rather than •extending it to new objects. But now we're faced with a strange and bewildering fact, namely, that while the inference from •'Something exists' to •'An utterly necessary being exists' seems to be compelling and correct, when we try to form a concept of such a necessity—i.e. a concept of something's necessarily existing—we find that we can't overcome the obstacles that the understanding puts in our way through its requirements for what such a concept would have to be like.

All down the centuries men have spoken of an *absolutely necessary* being; and they've tried to prove that such a thing •exists without bothering to consider whether and how such a thing is even •conceivable! Of course it's easy to provide a verbal definition of this concept, namely as 'something whose non-existence is impossible'. But this tells us nothing about what would require us to regard something's non-existence as unqualifiedly unthinkable. If we don't know about *that*, we can't know whether in using this concept we are thinking anything at all. . . .

It gets worse. This concept—at first ventured on blindly, and then become familiar—is now supposed to have its meaning exhibited in a lot of examples, so that there's no need for any further enquiry into its intelligibility. Every

geometrical proposition, e.g. *a triangle has three angles*, is unqualifiedly necessary, and this led people to apply 'unqualifiedly necessary' to an object that lies entirely outside the sphere of our understanding, as though they understood perfectly what they were saying.

All the supposed examples—all of them—are taken from •judgments and not from •things and their existence. But the unconditioned necessity of a judgment is not the absolute necessity of the thing. The absolute necessity of the judgment is only a *conditioned* necessity of the thing, or of the
622 predicate in the judgment. The proposition about triangles doesn't say that three angles are utterly necessary; all it says is that under the condition that there is a triangle. . . . three angles will necessarily be found in it. This logical necessity has had so much power to delude that this has happened:

People have thought that by forming an *a priori* concept of a thing and building *existence* into the concept, they were entitled to infer that the object of the concept necessarily exists.

[Kant comments on this in a compressed, very difficult sentence, the gist of which is this: The familiar and legitimate use of the concept of necessity is of the form 'Given that there is an F, there must be a G'—given that there's a triangle there must be a trio of angles. So the procedure described in the above indented passage ought to lead only to: Given that there is a being which blah-blah-blah and exists, it must exist. But this is trivial and uninteresting, and doesn't give people what they want, namely the conclusion that the item they purport to be talking about necessarily exists—exists unconditionally—exists absolutely—doesn't merely exist given such-and-such.]

If in an analytic proposition I cancel the predicate while retaining the subject, contradiction results; which is why I say that that predicate belongs necessarily to that subject. [In

this context, 'cancel' translates a word that could mean 'reject', 'annul', or the like.] But if I cancel both the subject and the predicate, there's no contradiction because there's nothing left that could be contradicted. Consider the analytic proposition *Every triangle has three angles*. If I say of something that 'it is a triangle and doesn't have three angles' I contradict myself; but there's nothing contradictory about cancelling both the subject and the predicate, saying 'This thing isn't a triangle and doesn't have three angles'. This holds true of the concept of an absolutely necessary being x. If you cancel
623 x's existence you cancel x itself with all its predicates—and how could *that* involve a contradiction? [Notice Kant's sudden switch to 'you'. As you'll see, he really is here imagining himself as addressing a defender of the ontological argument for God's existence.] There's nothing •outside x that would be contradicted, because x is not supposed to have derived its necessity from anything else; and there's nothing •intrinsic to x that would be contradicted, because in cancelling x you have at the same time cancelled all its intrinsic properties. *God is omnipotent* is a necessary judgment—indeed, it's analytic. The omnipotence can't be cancelled if you posit a deity, i.e. an infinite being, because the concept of omnipotence is part of the concept of deity; which means that 'There is a God who is not omnipotent' is a contradiction. But if you say 'God doesn't exist' there's nothing even slightly contradictory, because the statement has cancelled God's omnipotence (and all his other properties) in the act of cancelling God.

So you see that if I cancel the predicate of a judgment along with its subject, no internal contradiction can result, whatever the predicate may be. Your only escape from this conclusion is to say that some subjects can't be cancelled, and must always be left standing. But that's just another way of saying that there are unqualifiedly necessary subjects—which is the very thing I have been questioning and you have

624 been trying to defend! I can't form the least concept of a thing such that if it is cancelled along with all its predicates the result is a contradiction; and my only way of judging impossibility through pure *a priori* concepts is in terms of contradiction.

No-one can deny the general points I have been making, but you challenge them by claiming that there is a counter-example to them. There's just one concept, you say, where the non-existence or cancelling of the thing it applies to is self-contradictory, namely the concept of the most real being, 'the *ens realissimum*'. The most real being possesses **all reality**, you say, which you claim justifies you in assuming that such a being is possible. (I'll let you have that assumption in the meantime, though you really aren't entitled to it, because a concept's not being self-contradictory doesn't prove that it's possible for it to apply to something.)²⁰ Your argument proceeds from there: **all reality** includes existence; so existence is contained in the concept of a certain possible thing x. Thus, if x is cancelled then the intrinsic possibility of x is cancelled—which is self-contradictory.

I reply: You have •taken the concept of a thing that you purported to be using only in thinking about the thing's possibility and have •introduced into it the concept of existence; and *that* is a contradiction. It's contradictory when existence

²⁰ A concept is always possible if it isn't self-contradictory. That's the logical criterion of possibility. . . . But a concept might be 'possible' by that standard and yet be empty, 'i.e. a concept that doesn't apply to anything'. That may be the case if the objective reality of the synthesis through which the concept is generated has not been properly worked out; and *that*, as I have shown above, rests on •principles of possible experience and not on the •principle of analysis (the law of contradiction). This is a warning against arguing directly from the logical possibility of concepts to the real possibility of things.

is brought in openly, and it's equally contradictory when it is smuggled in ('as you have done') under a label such as 'all reality'. And apart from the point about contradiction, there's another way of showing that what you are doing doesn't achieve anything. If we allowed 'existence' to occur in a concept in the way you want, it may look as though you have won the game but actually you'll have 'achieved nothing because you'll have *said* nothing, producing a mere tautology. Here is a challenge for you. Consider any true proposition of the form *x exists* (let x be anything you like; I shan't quarrel over that), and answer this question: Is this proposition **(1)** analytic or **(2)** synthetic?

(1) If you say 'analytic', then there are two options. **(1a)** Because the mere thought of x guarantees x's existence, x itself must *be* a thought—something inside you—in which case it couldn't be the most real being! Or **(1b)** you have built x's •really existing into your notion of x's •possibility. [The passage between *asterisks* expands Kant's words in ways that the small dots convention can't readily indicate; but it expresses his thought.] *Now, *anything* we say of the form 'x is F' (where F is some predicate) tacitly assumes that x is *possible*; so it could always be expanded to 'If x is possible then x is F'. It follows that you, by equating

'x is possible'

with something of the form

'blah-blah and x exists',

are in your statements about x always implicitly saying something of the form 'If blah-blah and x exists, then x is F'. So any assertion of something's existence will, for you, always be equivalent to the corresponding statement

If blah-blah and x exists, then x exists,*

which is nothing but a miserable tautology. Apply this now to the x that concerns us here, namely x = the most real being. The word 'real' in the concept of the subject

sounds different from the word 'exists' in the concept of the predicate, but that doesn't affect the crucial fact that, on this account of what it is for something to be 'possible', any existential statement involves assuming in the subject
 626 concept something that is merely repeated in the predicate.

(2) And if you say that *x exists* is synthetic—and every reasonable person must agree that *all* existential propositions are synthetic—then you'll have to give up your contention that 'in the special case of *the most real being exists*' it is a contradiction to deny that predicate of that subject. The feature can't-be-denied-without-contradiction is a privilege that only analytic propositions have—indeed it's just what constitutes their analytic character.

I would have hoped to obliterate this deep-thinking nonsense in a direct manner, through a precise account of the concept of existence, if I hadn't found that the illusion created by confusing a •logical predicate with a •real predicate (i.e. a predicate that characterizes a thing) is almost beyond correction. Anything we please can be made to serve as a logical predicate; the subject can even be predicated of itself; for logic abstracts from all content. But a *characterizing* predicate is one that is added to the concept of the subject and fills it out. So it mustn't be already contained in that concept.

Obviously, 'being' isn't a real predicate; i.e. it's not a concept of something that could be added to the concept of a thing. It is merely the positing of a thing, or of certain state or property. Logically, it is merely the copula of a judgment. The proposition 'God is omnipotent' contains two concepts, each with its object—God and omnipotence. The
 627 little word 'is' doesn't add a new predicate but only serves to posit the predicate *in its relation* to the subject. If I now take the subject (God) with all its predicates (omnipotence among them), and say 'God is', or 'There is a God', I'm not attaching

any new predicate to the concept of God, but only positing the subject with all its predicates, positing the *object* in relation to my *concept*. The content of both 'object and concept' must be exactly the same: the concept expresses a possibility, and when I have the thought that its object exists I don't *add* anything to it; the real contains no more than the merely possible. A hundred •real dollars don't contain a cent more than a hundred •possible dollars. If there *were* something in the real dollars that isn't present in the possible ones, that would mean that the concept *hundred dollars* wasn't adequate because it didn't capture everything that is the case regarding the hundred dollars. A hundred real dollars have a different effect on my financial position from the effect of the mere concept of them (i.e. of their possibility). For the existing object isn't analytically contained in my concept; it is added to my concept. . . .; and yet the conceived hundred dollars are not themselves *increased* through thus acquiring existence outside my concept.

628

When I think of a thing through some *or all* its predicates, I don't make the slightest addition to the thing when I declare that this thing *is*, 'i.e. that it *exists*'. If this were wrong—i.e. if saying that the thing exists were characterizing it more fully than my concept did—then what I was saying *exists* wouldn't be exactly what in my concept I had been thinking of as possible. If I have the thought of *something that has every reality except one*, the missing reality isn't added by my saying that this defective thing exists. On the contrary, it exists with something missing, just as I have thought of it as having something missing; otherwise the existing thing would be different from the one thought of 'through my concept'. So when I think a being as the supreme reality (nothing missing), that still leaves open the question of whether it exists or not. Although my concept contains the whole possible real content of a thing as such,

there's something that it doesn't contain. . . ., namely the possibility of knowing this object *a posteriori*. And here we find the source of our present difficulty. If we were dealing with an object of the senses, we couldn't muddle the thing's •existence with the mere •concept of it. That's because through the •concept the object is thought only as conforming to the *universal conditions* of possible empirical knowledge as such, whereas through its •existence it is
629 thought as belonging to the context of experience as a whole. In being thus connected with the *content* of experience as a whole, the concept of the object is not added to in any way, but a possible perception has been added to our mental life. . . .

[Kant goes on to say that •with *any* kind of object x, the existence of x is different from the concept of x; that •when x is a sensible object the difference can be stated—as he has just stated it—in terms of what is implied for our perceptions; but that •if x is not a sensible object—e.g. if x is the *ens realissimum*—perception doesn't come into it, and indeed x's existence can't be cashed out in terms of any facts about our knowledge. This amplifies Kant's recent suggestion that the concept/object muddle is easier to make for non-sensible objects than for sensible ones. He continues:]

The concept of a supreme being is in many ways a very
630 useful idea; but just because it's a mere *idea* it can't, unaided, enlarge our knowledge about what exists. It can't even teach us anything about what is *possible*. We have to grant that it satisfies the analytic criterion of possibility, meaning that it isn't self-contradictory, because there can't be any contradiction in an accumulation of realities, i.e. of positives. [Kant wrote, more literally, 'i.e. of *positings*'; but for him 'positing' something is always *affirming* it, and in the present context the core notion is that of affirmation-without-denial or positive-untouched-by-anything-negative. In the background is Leibniz's argument: (1) the

concept of the *ens realissimum* is the concept of something that is as real as it's possible to be; so (2) it's the concept of something having all positive attributes, with nothing negative in its make-up; but (3) a contradiction involves something's being combined with its own *negation*; so—putting together (2) and (3)—there can't be anything contradictory in the concept of the *ens realissimum*. Thus, the notion of positiveness has to be uppermost in Kant's use here of 'positings', because of the 'positings' role as guarantors of consistency, their being equated with 'realities', and their connection with Leibniz.] But we can't specify *a priori* whether a supreme being •or *ens realissimum*• is possible. For one thing, we aren't told anything about what these 'realities' *are*; and even if we were, we still couldn't judge whether such a being is possible, because the criterion of possibility in synthetic knowledge is found only in experience, and there can't be experience of the object of an idea. So the celebrated Leibniz is far from having succeeded in what he prided himself on achieving—an *a priori* grasp of the possibility of this sublime ideal being.

So much for Descartes's famous ontological argument for the existence of a supreme being—it's all just wasted effort! We can no more extend our stock of knowledge by mere ideas than a merchant can better his position by adding a few zeros to his cash account.

5. There can't be a successful cosmological argument for the existence of God

There's something quite unnatural about taking •an idea 631 that we have *chosen* to form and trying to extract from it •the existence of an object corresponding to it. It's just a new-fangled product of scholastic cleverness. The attempt would never have been made if reason hadn't previously created an apparent *need* for it, as follows.

Reason has a need to assume, as a basis for the existence of *anything*, something whose existence is necessary, so as to have a terminus for the backward search for reasons, reasons for reasons, and so on. This necessity of existence has to be unconditioned or absolute, and we have to be *a priori* certain about it. So reason was forced to look for a concept that would satisfy this demand (if it could be satisfied), a concept enabling us to know in a completely *a priori* manner that something exists. That's the concept that was supposed to have been found in the idea of a supremely real being, an *ens realissimum*; so that idea was used only to give us a more definite knowledge of the necessary being—a being of whose existence we were already convinced or persuaded on other grounds, i.e. grounds other than the ontological argument.

632 But this *natural* procedure of reason was concealed from view; and instead of ending with this concept, philosophers tried to start with it. Instead of offering a different argument for the necessary existence of something, and then using the concept of the *ens realissimum* to flesh it out, they tried to make that concept the whole basis for a different argument for the same conclusion. That is the pedigree of the misbegotten ontological argument, which doesn't satisfy the natural and healthy understanding or academic standards of strict proof.

The *cosmological argument*, which I'm now about to examine, still connects absolute necessity with supreme reality, but whereas the ontological argument reasoned

from the supreme reality to necessity of existence,
the cosmological argument reasons in the reverse order,
from the (previously given) unconditioned necessity

of some being to its unlimited reality.

Whether the argument is rational (*vernünftigen*) or sophistical (*vernünftelnden*), it is at least following a natural path, the one that is most convincing not only to the man in the street but also to the philosophical theorist. It sketches the outline of *all* the arguments in natural theology, an outline that has always been and always will be followed, however much the arguments are decorated and disguised by frills and curlicues. This argument, which Leibniz called 'the argument from the world's contingency', I shall now proceed to expound and examine.

It goes like this:

(1) If anything exists, a totally necessary being must also exist.

(2) I (at least) exist.

Therefore

(3) An absolutely necessary being exists.

Premise (2) contains an experience, while premise (1) presents the inference from there being any experience at all to the existence of something necessary. So the argument really begins with experience, and isn't wholly *a priori* or 'ontological'; so we need another label for it. For this purpose 'cosmological' has been selected, because the object of all possible experience is called *the world*. . . .

The argument proceeds from there as follows:

(4) The necessary being can be specified in only one way, i.e. by one out of each possible pair of opposed predicates. So

(5) The necessary being must be *completely* specified through its own concept.

(6) The only possible concept that completely specifies its object *a priori* is the concept of the *ens realissimum*.

Therefore, putting (3) together with (5) and (6), the only concept through which a necessary being can be thought

634 is the concept of the *ens realissimum*. In other words, a supreme being necessarily exists.

This cosmological argument brings together so many sophisticated principles that speculative reason seems in this case to have mustered all the resources of its dialectical skill to produce the greatest possible transcendental illusion! I'll set aside for a while the testing of the argument, because I want first to expose the trick through which an old argument is dressed up here as a new one—the trick of appealing to the agreement of 'two witnesses', one from pure reason and the other with empirical credentials. What makes this a *trick* is the fact that there's really only one witness, the one from pure reason, which then changes its clothes and alters its voice in order to pass itself off as a second witness. In order to put firm ground under its feet, this argument takes its stand on experience, giving itself a different look from the ontological argument, which puts its entire trust in pure *a priori* concepts. But the cosmological argument uses this experience only for a single step, the one that infers the existence of *some* necessary being. The empirical premise can't tell us what properties this being has; so reason leaves experience and tries to discover from mere concepts what properties an absolutely necessary being must have. . . . It thinks that the requirements for *existing with absolute necessity* are to be found in the concept of an *ens realissimum* and nowhere else, and thus concludes that the *ens realissimum* is the absolutely necessary being. But this involves presupposing that

- the concept of the *ens realissimum* is completely adequate to the concept of absolute necessity of existence; which is to say that

- the concept of absolute necessity of existence can be inferred from the concept of the *ens realissimum*;

which is just what the ontological argument said! The cosmological argument was to have •managed without the ontological argument, but now we find that it's •based on it! ·In case that's not clear enough, I'll go through the crucial part of it more slowly·. For a thing to be absolutely necessary is for its existence to be secured by mere concepts—that's what absolutely necessary existence is·. If I say that •the concept of the *ens realissimum* is one (indeed the only one) that is appropriate and adequate to •necessary existence, I must admit •that necessary existence can be inferred from •that concept. Thus the so-called cosmological argument really owes any force it may have to the ontological argument from mere concepts. The appeal to experience is idle. Perhaps experience leads us to the concept of absolute necessity, but it can't show us what it is that has such necessity. The moment we try to specify *that*, we have to abandon all experience and search among pure concepts for one containing the conditions of the possibility of an absolutely necessary being. And if we find it, we thereby establish the being's existence. . . . 636

[Kant now uses a technicality from the theory of syllogisms to justify his claim that the cosmological argument needs a step that involves the ontological argument, which means that the cosmological argument really has nothing to offer. And he says that the cosmological argument is as deceptive as the ontological argument, and has a further fault all of its own, namely deceptiveness about the path it is following. Then:]

I remarked a little way back that hidden in this cosmological argument is a whole nest of dialectical [= 'illusion-producing'] assumptions; the transcendental critique can easily reveal and destroy them. All I'll do now is to list these deceptive principles; by now you know enough to explore and extirpate them for yourself. ·There are four of them·.

638 **(1)** There's the transcendental principle of inferring a cause from anything contingent. This has work to do in the sensible world; outside that world it doesn't mean anything. That's because the merely intellectual concept of *contingency* can't generate any synthetic proposition such as the principle of causality. . . . Yet in the cosmological argument that principle is used just precisely as a way of getting us outside the sensible world. **(2)** Then there's the inference to a first cause, from the impossibility of an infinite causal chain in the sensible world. The principles of the use of reason don't entitle us to make this move even within the world of experience; still less to make it beyond this world in a realm that the causal chain can never reach. **(3)** Reason's unjustified complacency about having completed this series. What it has really done is to •remove all the conditions, •find that it can't conceive anything further, and •construe *this* as 'completing the series'. Whereas the removed conditions are required for there to be any concept of necessity! **(4)** Muddling two questions about the 'possibility' of the *ens realissimum*—•is it logically possible? i.e. is its concept free from contradiction? and •is it transcendently possible? To answer the second question we would need a principle that in fact is applicable only to the domain of possible experiences. And so on.

The trick the cosmological argument plays is to let us off from having to prove the existence of a necessary being *a priori*, through mere concepts. If we were to prove this we'd have to do it in the manner of the ontological argument, and we don't feel up to doing that. So we take as the starting-point of our inference an actual existence (an experience as such), and advance as best we can to some absolutely necessary condition of this existence. [Starting from 'an experience as such' is starting from the bare fact that some experience occurs, without caring about *what* experience it is.] We don't have to show that this

condition •is possible, because we have just proved that it 639
 •exists. If we now want to learn more about the nature of this necessary being, we don't try to do this in the manner that would in fact be effective, namely by discovering *from its concept* the necessity of its existence. 'If we could do *that*, we wouldn't have needed an empirical starting-point!' No, all we look for is the *necessary* condition—the *sine qua non*—for something to be absolutely necessary. This move would be legitimate in any inference from a given consequence to its ground, but in this one case it doesn't serve the purpose. That's because the condition that is needed for absolute necessity is to be found in only one individual thing; so this thing must contain in its concept everything that is required for absolute necessity, and consequently it enables me to infer this absolute necessity *a priori*. That means that I can run the inference in the opposite direction, contending that anything to which this concept of supreme reality applies is absolutely necessary.

[Here's a more abstract statement of Kant's line of thought here: In the cosmological argument we have

an inference from an empirical premise to the conclusion that something x exists absolutely necessarily.

Wanting to discover what sort of thing x is, we ask 'What would a thing have to be like to exist necessarily?' This is a perfectly normal procedure. Compare: the data convince us that there was an earthquake in Bam at time T; we want to know more about it, so we consider 'What would an earthquake have to be like to do what this one did to the city of Bam?' But in our present case, we discover that what a thing would have to be like to exist necessarily is to have a concept that guaranteed necessary existence; **(1)** falling under such a concept is not only required for necessary existence, it is also *sufficient* for necessary existence. Add to this the further discoveries **(2)** that the only concept giving such a guarantee is the concept of *supremely real thing*, and **(3)** that one and only one thing can fall under this concept. Putting **(1)**, **(2)** and **(3)** together, we get all we need for

an inference from the concept of *supremely real thing* to the conclusion that something, namely the supremely real thing, exists necessarily.

But that is the ontological argument! We wanted to argue from •an empirical premise to •the conclusion that something exists necessarily, and then to fill in details about what this necessarily existing thing is like. In the course of doing this we stumbled onto an inference from •a purely conceptual premise to •the conclusion that something exists necessarily; which puts our initial argument out of business.]

If I can't make this inference (and I certainly *can't* if I'm to keep the ontological argument out of the picture), I have come to grief in the new way I've been following, and am back again at my starting-point. The concept of *the supreme being* answers all the *a priori* questions that can be raised about a thing's intrinsic nature; and it has the unique feature of being a universal concept that applies to only one possible thing (·the universal concept of *supreme being*; contrast with the universal concept of *human being*, which can have any number of instances·); and all this makes it an *ideal* that is unmatched. But it doesn't answer the question of whether the supreme being exists; ·the ontological argument *says* that it does, but we've seen that the ontological argument isn't valid·. Yet that's just what we were trying to find out about, ·and now we see that in proceeding in this way we have achieved nothing·. . . .

It may be all right for us to *postulate* the existence of a supremely sufficient being as the cause of all possible effects, wanting this to ease reason's search for the unity in the grounds of explanation. But if we go so far as to say that such a being *necessarily exists*, we have moved from •modestly expressing an admissible hypothesis to •boldly claiming absolute certainty. Why? Because someone who claims to know that it is unqualifiedly necessary that P must himself be absolutely certain that P.

The whole problem of the transcendental ideal comes down to this:

- Given absolute necessity, find a concept that has it.
- Given the concept of something x, find x to be absolutely necessary.

If either task is possible, then so must the other be, because the only way reason acknowledges for something to be absolutely necessary is for it to follow necessarily from its concept. But we are utterly unable to perform either task, whether to satisfy our understanding in this matter or to reconcile it to its not being satisfied. 641

Unconditioned necessity, which we utterly need as the basic supporter of everything, is for human reason the veritable abyss. Eternity itself in all its terrible sublimity. . . .is nowhere near as dizzying; for it doesn't support things, but only measures how long they last. Consider this thought:

- A Being that we represent to ourselves as supreme amongst all possible beings might be in a position to say to itself: 'I exist from eternity to eternity, and everything other than me exists only through my will; *but then where do I come from?*'

It's an unavoidable thought, but also an unbearable one. ·When we try to cope with it·, everything sinks under us. *The greatest perfection* is seen by our speculative reason as hovering without support, and the same is true of *the least perfection*; speculative reason has nothing to lose by letting them both vanish entirely.

Many •natural forces that declare themselves through certain of their effects remain inscrutable to us because we can't track them down by observation. And •the transcendental object lying at the basis of appearances—·the reality-in-itself that appearances are appearances *of*·—is and remains inscrutable to us; we know that it exists, but we don't and can't have any insight into its nature. (That cuts us off from, among other things, the reason *why* the conditions to which our sensibility is subject are just the 642

ones they are and not others.) But an •ideal of pure reason can't be called inscrutable. The only certificate of 'reality' that it has to produce is reason's need to use it to complete all synthetic unity. It's not given to us as a thinkable *object*, so it can't be inscrutable in the way an object can. On the contrary it can be investigated (•it is 'scrutable') because it is a mere idea that is located in and explained through the nature of reason. For what makes reason *reason* is our being able to give an account of all our concepts, opinions, and assertions—the account being in subjective terms for the illusory ones, in objective terms for the others.

DISCOVERY AND EXPLANATION

of the dialectical illusion in all transcendental arguments for the existence of a necessary being

Both of the above arguments were transcendental, i.e. were attempted independently of empirical principles. The cosmological argument is based on •an experience as such—i.e. on the mere fact of there being *some* experience—but not on •any specific property of this experience. What it relies on are pure principles of reason as applied to an existence
643 given through the sheer fact of empirical consciousness; and before long it abandons this guide-line and relies on pure concepts alone. Well, then, what is it in these transcendental arguments that causes the dialectical but natural illusion that •connects the concept of necessity with that of supreme reality, and •turns what is really only an *idea* into a *real thing*? Why is it inevitable that we'll assume that some one existing thing is intrinsically necessary, while also shrinking back from the existence of such a being as from an abyss? And how are we to get reason to understand itself in this matter, bringing it to a settled insight instead of its state of wobbling between (1) timid assertions and (2) retractions of them?

(1) Once we assume that something exists, we can't get out of concluding that something exists necessarily—how very remarkable! This is a quite natural inference (which isn't to say that it is sound), and the cosmological argument is based on it. (2) And yet if I help myself to the concept of anything—*anything*—I find that I can't think of the existence of this thing as absolutely necessary. Let x be any existing thing you like—nothing prevents me from thinking of x as not existing. Thus, while (1) I'm obliged to assume something necessary as a condition of anything's existing, (2) I can't think, of any particular thing, that *it* is necessary.

644

[In this paragraph we'll meet the useful word 'heuristic' (German *heuristisch*), which means 'having to do with methods of investigation and discovery'.] From the truth of (1) and (2) together it follows—there's no escaping this conclusion—that necessity and contingency don't concern the things themselves; otherwise there would be a contradiction. Thus, neither of these two principles—the principles that are at work in (1) and (2)—can be objective; at most they are subjective principles of reason; with (1) one telling us to seek something necessary as a condition of everything given as existent, i.e. not to stop until we reach an explanation that is *a priori* complete; and (2) the other telling us

- not to hope for this completion,

- not to treat anything empirical as unconditioned, thereby letting ourselves off from further explanations of it.

When the two principles are in this way seen as merely heuristic and *regulative*, i.e. as merely guides to intellectual behaviour,they can very well stand side by side. (1) One tells us to philosophise about nature [here = 'to do natural science'] as if there were a necessary ultimate basis for everything that exists, doing this solely so as to bring systematic unity into our knowledge by always pursuing such an idea, i.e. the idea of the imagined ultimate basis. (2) The other

645 warns us not to regard any fact about any existing thing as constituting such an ultimate basis, i.e. as absolutely necessary; it tells us to keep the way always open for further explanations, thus treating every single fact as conditioned in its turn. . . .

[Kant now makes the point that when **(1)** tells us to postulate a thing whose existence is absolutely necessary, and **(2)** says that we should never regard any empirical item—anything *in the world*—as being such a thing,] it follows that we must regard the absolutely necessary as being *outside* the world.

[Kant reports that the ancient philosophers thought that the existence of *matter* is basic and necessary, while all its *forms*—its states or properties—are contingent. His comments on this are mainly based on distinguishing

(1) matter as encountered empirically from **(2)** matter considered as a thing in itself;

but some of his turns of phrase, as well as the sheer fact that he is connecting this with ancient philosophers, suggests rather the distinction between

(3) matter considered as stuff that is extended, impenetrable etc. from **(4)** matter considered as the sheer naked substratum that *has* these properties etc.

Let's set **(3)/(4)** aside and focus on the other distinction. If the ancients had focused on **(2)** matter, Kant says, they wouldn't have thought of it as existing necessarily; given any thing at all, there's nothing to block reason from annihilating it *in thought*, and that settles that, because *thought* is the home territory of absolute necessity. So the ancients must
646 have been thinking of **(1)** matter; and their belief that it exists necessarily must have arisen from their feeling the force of a certain *regulative principle* that should guide our thoughts about empirical matter. The idea of a necessarily existing primordial being can't be cashed empirically; if such

a being could be identified empirically the whole show would come tumbling down. So the item in question must be thought of as 'outside the world', as merely the topic or focus or imaginary goal of a regulative principle. Setting the necessarily existing being outside the world, Kant says, leaves us free to •explain appearances in terms of other appearances, as confidently as if there were no necessarily existing being in the picture, while also being free to •keep pushing on with our explanations, always driving towards *completing* the chain of explanations, just as if we thought that completion could actually be achieved through our arriving at a necessarily existing being in the world. Then:] 647

Thus, the ideal of the supreme being is nothing but a *regulative principle* of reason, telling us to look on the whole way the world hangs together *as if* it originated from an all-sufficient necessarily existing cause. In this procedure we use the ideal to guide us when we are explaining the hanging-together of the world in a systematic way, showing parts of it to be **law-of-nature necessary**; but we aren't asserting that the existence of anything is **necessary in itself**, •absolutely necessary•. Still, we can't avoid the transcendental *switch* through which this formal principle is represented as (a) constitutive, and by which this unity is (b) hypostatized [i.e. through which this regulative principle is seen as (a) a fact-stating proposition, and the sought-for unity is seen as embodied in an individual (b) thing]. Compare this with •the switch we perform with• space. Because space is what makes shapes *possible* (a shape is just a way in which space is limited), although it's only a formal feature of sensibility we take it as something absolutely necessary, existing in its own right, and as an object given *a priori* in itself. Similarly with our present topic. Because the systematic unity of nature can't be prescribed as a •regulative• principle for the empirical use of our reason except through our presupposing

the idea of an *ens realissimum* as the supreme cause, it's only natural that this idea should be represented as an actual object which, being the supreme condition, is also
 648 necessary. And in this way we change a *regulative* principle into a *constitutive* one. Here's a clear indication that a substitution has indeed been made: This supreme being was utterly (unconditionally) necessary ·in its role in a regulative principle· with respect to the ·empirically given· world; but when we take it to be a thing that exists in its own right, we can't form any concept of this ·supposed· necessity. So this necessity must be something we encountered in our reason, as a formal condition of thought, not as a contentful thing-related condition of existence.

6. There can't be a successful physico-theological argument ·for the existence of God·

Well, then, if we can't satisfy the demand ·for a proof of God's existence· from •the concept of things as such, or from •experience telling us that *something exists*, it remains only for us to see where we can get if we start from •experience of *detailed facts about what exists*, i.e. our experience of the things of the present world, what they are like and how they are organised. Perhaps that will help us on our way to a secure belief in a supreme being. An argument of that sort is what I label 'the physico-theological' argument. If it can't succeed either, ·we'll have to conclude that· unaided speculative reason can't come up with a satisfactory argument for the existence of a being corresponding to our transcendental
 649 ·theological· idea.

In view of what I have been saying, we don't expect it to take long for this inquiry to be conclusively settled. How *can* there be any •experience that is adequate to an •idea? The special feature of ideas that marks them off as

ideas is precisely the fact that no experience can ever be equal to them. The transcendental idea of a necessary and all-sufficient primordial being is so overwhelmingly vast, so high above everything empirical, ·that we can't fill it out with empirical material·. For one thing, experience doesn't present *enough* stuff to fill this enormous concept; for another, it doesn't present *the needed kind* of stuff, because everything empirical is conditioned, and we'll get nowhere rummaging around in *that* for something matching up to ·the concept of· the unconditioned supreme being: no law of any empirical synthesis gives us an example of, or gives any help in the search for, any such unconditioned item.

If the supreme being stood in this chain of conditions, it would be a member of the series, and like its subordinates in the series it would call for further enquiry as to the still higher ground from which it follows. One might suggest: 'Let's separate the supreme being from the chain, and conceive it as a purely intelligible being that exists outside the series of natural causes.' But then what bridge can reason use to get across to it? All •laws governing inferences of causes from effects—indeed all episodes of •synthesis and extension of our knowledge—are concerned only with possible experience, and therefore relate solely to objects of the sensible world, apart from which those laws
 650 and syntheses can't mean a thing.

This world presents to us such an immeasurable display of variety, order, purposiveness, and beauty, exhibited both on the indefinitely large scale and the indefinitely small, that even the scanty knowledge of this that our weak understanding provides us with puts us into a frame of mind where our thoughts slide all over the place, speech loses its force, and numbers lose their power to measure. We're reduced to a state of speechless wonder—*eloquently* speechless wonder! Everywhere we see a network of effects and causes, of ends

and the means to them, regularity in how things come into and go out of existence. Nothing has *put itself* into the condition in which we find it to exist; we always look for a prior cause, which in turn commits us to looking for *its* cause, and so on backwards. This whole universe would sink into the abyss of nothingness if we didn't assume, over and above this infinite chain of contingencies, something to support it—something that

- exists in its own right without being conditioned by anything else,
- caused the universe to come into existence, and
- secures the universe's continuing survival.

651 This supreme being—higher than anything else in the world—how *big* should we think of it as being? [Kant is presumably thinking of this metaphorically, but the word he uses is *gross* = 'big'.] We are not acquainted with the whole content of the world, still less do we know how to estimate its size by comparison with everything that is possible. But since in our causal thinking we can't do without an ultimate and supreme being, what's to stop us from supposing this being to have a degree of perfection that sets it *above everything else that is possible*? We can easily do this—though only with the skimpy sketch provided by an abstract concept—by representing this being to ourselves as **a single substance that combines in itself all possible perfection**. This concept ·has many virtues·:

- it respects our reason's demand for parsimony of principles;
- it isn't self-contradictory;
- it is never decisively contradicted by any experience;
- by directing ·our inquiries· towards order and purposiveness, it helps to extend the use of reason within experience.

The physico-theological argument always deserves to be mentioned with respect. ·Of all the arguments for God's

existence·, it is the oldest, the clearest, and the best fitted to common-sense . It enlivens the study of nature, just as it gets *from* the study of nature •its very existence as well as •its ever-renewed vigour. It brings ends and purposes into ·parts of natural science· where our unaided observation wouldn't have detected them, and extends our knowledge of nature by means of the guiding-thread of a special unity that is driven by something outside nature. This knowledge reflects back on its cause—i.e. on the idea that led us to it—thus strengthening the belief in a supreme author ·of nature· to the point where it has the force of an irresistible conviction. 652

Trying to lessen the authority of this argument—what a bleak prospect! and anyway there's no chance of succeeding. Reason is constantly upheld by this body of material for the premise of the argument, material that increases in reason's hands; though only empirical, it is powerful—too powerful to be eroded by the doubts that subtle and abstruse speculation suggest. ·When such doubts threaten·, reason is at once aroused from brooding indecision, as from a dream, by *one glance* at the wonders of nature and the majesty of the universe—ascending from greatness to greatness right up to the all-highest, from the conditioned to its condition, up to the supreme and unconditioned author ·of everything·.

This procedure is reasonable and useful; far from objecting to it on those scores, I applaud and encourage it. But this type of argument wants to claim that its conclusion is •absolutely certain and •based just on the ·physico-theological· argument, without outside help; and *that* is something we can't approve. ·Let's not be hesitant about our disapproval·. It can't harm the good ·that the argument can do· if the dogmatic language of the intellectually reckless sophist is toned down to the measured and moderate requirements of a belief that is strong enough to quieten our doubts

653 though not to command unconditional submission. So I say this: The physico-theological argument cannot unaided establish the existence of a supreme being; it must always fall back on the ontological argument (to which it only serves as an introduction) to fill this gap. So the ontological argument is *the only possible one* that human reason can't ignore (insofar as any *speculative* argument for God's existence is possible at all).

Here are the main steps of the physico-theological argument: **(1)** All through the world. . . we find clear signs of an order that has been imposed with great wisdom in the furtherance of a definite purpose. **(2)** This purposive order is quite alien to the things of the world, and belongs to them only contingently; i.e. the various things couldn't have worked together, through such a great combination of different means, towards the fulfillment of definite final purposes; that is, they couldn't have done it unaided, rather than having been chosen and designed for these purposes by an ordering rational principle on the basis of ideas. **(3)** So there is a sublime and wise cause (or more than one), which must be the cause of the world, not merely as a blindly working all-powerful nature but as an intelligence, not merely through *fecundity* but through *freedom*. **(4)** That this is just *one* cause can be inferred from the unity of the inter-relations between the parts of the world, making them-

654 members of one skillfully arranged structure; this being an inference we can make with certainty as far as our own observations stretch, and with probability beyond those limits, in accordance with the principles of analogy.

Reason naturally argues from the analogy between certain natural products and things like houses, ships and watches—things produced by our human skill when we push nature around, making it work towards our ends rather than its own—inferring that the natural products are caused in the

same way as the artifacts, namely by understanding and will; and that it's possible that a freely acting nature (which is what makes possible all art, and perhaps even reason itself) is derived from a superhuman art.

·EDITORIAL INTERVENTION·

In this context, 'art' relates to what is *artificial*, in contrast to *natural*. Kant is describing a frame of mind in which everything natural is seen as a product of a higher-than-human art; and (in the parenthetical bit) all human art is seen as a product of nature. The two theses are expressed in the Shakespeare's *The Winter's Tale*, so aptly that the temptation to quote is irresistible. Perdita has said that she doesn't want 'carnations and streak'd gillyvors, which some call nature's bastards', in her garden:

Polixenes:

Wherefore, gentle maiden,
Do you neglect them?

Perdita:

For I have heard it said
There is an art which in their piedness shares
With great creating nature.

Polixenes:

Say there be;
Yet nature is made better by no mean
But nature makes that mean: so, over that art
Which you say adds to nature, is an art
That nature makes. . .

. . . This is an art

Which does mend nature, change it rather, but
The art itself is nature.

·END OF EDITORIAL INTERVENTION·

Perhaps this line of reasoning couldn't stand up under the most rigorous transcendental criticism, but let's not give

reason a hard time over that just now. It must be admitted that if we are to specify any cause of the universe, the safest way to go is by analogy with the only things whose cause and mode of action are fully known to us, namely the purposive productions of human art. There would be no excuse for reason's abandoning this causality that it knows, in favour of some other basis for explanation that is obscure, unprovable, and not directly known.

655 What could be explained by this argument from the purposiveness and harmonious adaptation of so much in nature? Only the form of the world, not the matter (i.e. not the substance). That is, this line of argument might explain what things are like and how they behave, but it can't explain the fact that they exist in the first place. [That is what Kant meant, but he expressed it by saying that the argument from purposiveness could 'prove the contingency of the world's form but not of its matter. In the paragraph 'After the proponent. . .' starting on page 286 you'll see why he dragged contingency into this; but at our present stage in the argument it is a distraction.] The latter task would require an argument to show that the things in the world wouldn't, unaided, be capable of such order and harmony in accordance with universal laws unless they were in their substance the product of supreme wisdom, i.e. unless a wise supreme being had brought them into existence. But an argument for that would require very different premises from those of the argument from the analogy with human art. The most that the physico-theological argument can argue for is an architect of the world whose work is limited by the recalcitrance of the material he works with, not a creator of the world whose thoughts are in command of everything. But such an argument-to-an-architect is not nearly good enough for the purpose we have had in mind, namely the proof of an all-sufficient primordial being. For an argument explaining why matter exists, we would have to resort to a

transcendental argument, which is just what we are trying to avoid here.

So the inference goes from the order and purposiveness everywhere observable throughout the world—with its existence being left unexplained—to the existence of a cause that is proportioned to it. The concept of this cause must enable us to know something quite definite about it; so it has to be the concept of a being which, as all-sufficient, possesses all power, all wisdom, etc.—in short, all perfection. For predicates such as 'very great', 'astounding', and 'immeasurable' in power and excellence give no determinate concept at all, and don't really tell us what the thing is in itself. All they do is to express how much greater the being in question is than the speaker, and that's the language not of description but of eulogy. . . . The only way to say something definite and non-relational about this being is by saying that it has all possible perfection. 656

Now, I hope you won't think you can see how the size and ordered complexity of the world, as you observe it, relates to its author's being omnipotent, supremely wise, absolutely one, and so on! Obviously, no-one can see such a thing. So the physico-theological argument can't lead to any definite concept of the supreme author of the world; which means that it can't lead to a theology that a religion could be based on.

It's utterly impossible, therefore, to get by the empirical road to the absolute totality of reality, of perfection, etc. Yet that's what the physico-theological argument tries to do. Well, then, how does it go about getting across this wide gap? 657

After the proponent of the physico-theological argument has led us to the point of admiring the greatness, wisdom, power, etc. of the world's author, and can't get us any further, he simply drops the argument to a theological conclusion from empirical premises, and goes back to the

early stage of his argument, where he inferred *contingency* from the order and purposiveness of the world. With

- this contingency

as his only premise, he then advances, by means of transcendental concepts alone, to

- the existence of an utterly necessary being;

and then from the concept of the absolute necessity of the first cause he takes the final step to the completely determinate or determinable concept of that necessary being, namely, to

- the concept of an all-embracing reality.

[Why 'determinate or determinable'? Kant's thought is that when you say of the absolutely necessary being that it 'is an all-embracing reality' or that it 'is real in every possible way', you have either •stated the whole detailed truth about it ('determinate') or •said something from which the whole detailed truth about it can be inferred ('determinable').] So what has happened is this: the physico-theological argument got stuck in its project, and dealt with this by suddenly switching to the cosmological argument; but this, ·as we have seen·, is only a disguised ontological argument; so really the physico-theological argument has reached its goal by pure reason alone. This despite the fact that it started off by denying any kinship with pure reason and claiming to base everything on convincing empirical evidence!

658 So the physico-theologians aren't entitled to adopt such a prim attitude towards the transcendental line of argument [= the] ontological argument, complacently posing as clear-sighted students of nature who are looking down on the cobwebby output of obscure speculators. If they would just *look at themselves* they would find that after getting a fair distance on the solid ground of nature and experience, and finding themselves no closer to the object that beams in on their reason, they suddenly leave this ground and pass over into the realm of mere possibilities, where they hope

on the wings of ideas to draw near to the object that has eluded them in their empirical search. This tremendous leap takes them to a place where they think they have firm ground under their feet, a place where they now have a determinate concept ·of the object they've been pursuing· (though they don't know how they came by it); and they extend this concept over the whole sphere of creation. So they reach the ideal, which is entirely a product of pure reason, and they explain it by reference to experience! The explanation is a pretty miserable affair, and far below the dignity of its object; ·but their biggest fault is that· they won't admit that they have *arrived at* this item of knowledge or this hypothesis by a road quite other than that of experience.

Thus the physico-theological argument for the existence of a primordial supreme being rests on the cosmological argument, which rests on the ontological argument. And those three are the only speculative arguments there can be for the existence of such a being. So we get the result that if there can be a proof of a proposition that's so far exalted above all empirical use of the understanding, it must be the ontological argument.

7. Critique of all theology based on speculative principles of reason

Taking 'theology' to stand for 'knowledge of the primordial being', theology is based either on reason alone or on revelation. Theology based on reason also divides into two, depending on what concepts it applies to its object: 659

- transcendental** theology, which uses only transcendental concepts such as those of primordial being, *ens realissimum*, being of beings;

- natural** theology, which uses a concept borrowed from nature (specifically, from the nature of our soul),

thinking of the primordial being as a supreme *intelligence*.

Someone who accepts only a transcendental theology is a *deist*. He allows that unaided reason can tell us of the *existence* of a primordial being, but he holds that our only concept of this being is transcendental—it's the concept of a being that possesses all reality, to which we can't add any details. Someone who *also* makes room for a natural theology is a *theist*. He maintains that reason can add detail to its account of the primordial being through analogy with nature—*our* nature—by describing it as a being that contains in itself the ultimate ground of everything else, doing this *through understanding and freedom*. For the deist this being is only a *cause of the world* (with nothing said
660 about whether it does this through necessity of its nature or through freedom); for the theist this being is the *author of the world*.

Transcendental theology itself divides into two: As *cosmo-theology* it aims to deduce the existence of the primordial being from an experience as such—i.e. from the mere fact that some experience occurs—without bringing in any facts about what *kind* of world the experience belongs to. As *onto-theology* it thinks it can know the existence of such a being through mere concepts, without the slightest help from any experience.

Natural theology infers the existence and the properties of an author of the world from the what the world is like, the order and unity found in it, this being a world in which we have to recognise two kinds of causality with their rules, namely •nature and •freedom. From this world natural theology ascends to a supreme intelligence, as the source either of all •natural order or of all •moral order and perfection. In the former case it is called *physico-theology*, in the latter

moral theology.²¹

What we ordinarily understand by the concept of *God* is not merely •a blindly operating eternal nature that is the root-source of things, but •a supreme being who through understanding and freedom is the author of things; it's
661 only in this sense that the concept interests us. So we could, strictly speaking, deny that the *deist* believes in *God*, and credit him only with maintaining the existence of a *primordial being or supreme cause*. But. . .we could say more charitably that the deist believes in a *God* while the *theist* believes in a *living God*. Now let us investigate where these different attempts on reason's part come from.

For present purposes we can define 'theoretical knowledge' as knowledge of what *is*, and 'practical knowledge' as the representation of what *ought to be*. [The rest of this difficult paragraph identifies a certain topic only so as to *set it aside*, reserving it for a later work. It is the question: If a theology is accepted because it is needed by moral doctrines that are themselves *a priori* necessary, what is the status of that theology? With that out of the way, Kant goes on, in a further paragraph, to say that when we are dealing merely with what *is* the case, the empirical conditioned items we are dealing with are always thought of as being contingent, which implies that their conditions are also contingent. So:]
The only way we could know that something in the domain
..662 of theoretical knowledge is utterly necessary would be on the basis of *a priori* concepts; we couldn't know such a thing about something posited as a cause •or condition• of something given in experience.

²¹ Not theological morality, which contains moral laws that *presuppose* the existence of a supreme ruler of the world; whereas moral theology is a conviction of the existence of a supreme being—a conviction based on moral laws.

663 An item of theoretical knowledge is *speculative* if it concerns an object which—or concepts of which—can't be reached in any experience. [This is an abrupt switch from Kant's meaning for 'speculative' up to here, namely as the antonym of 'practical' or 'moral'. The first meaning occurs mainly in the phrase 'speculative reason', whereas here we have 'speculative items of knowledge'. Quite soon, however, we shall encounter 'speculative reason' with the adjective used in this new sense.] This stands in contrast to *knowledge of nature*, which concerns only objects or predicates of objects that could be given in experience.

An example is the principle by which from an empirically contingent event we infer some cause of it. That principle belongs to the knowledge of nature, not to speculative knowledge, because its validity depends entirely on its being a condition of all possible *experience*. Try keeping experience out of the picture, and just look at the bare principle: *Every contingent event is caused by some prior event*. This is a synthetic proposition that connects a given item with some *other* item; and there's not the slightest justification for it when divorced from conditions of possible experience. . . .

664 Our causal thinking in the knowledge of nature involves treating as contingent, and looking for causes of, the *states* of substances and the *events* into which they enter—not looking for causes of the substances themselves, i.e. causes of their existence. If we infer from the existence of *things* in the world the existence of their causes, we are using reason in speculative knowledge. It would have to be purely speculative knowledge that told us that substance (matter) is contingent in its existence. And even if we were trying to explain only events and states, explaining how the world hangs together and the changes it undergoes, if we tried to infer from all this a cause that was entirely distinct from the world, this would again be a judgment of purely speculative reason, because the object we were inferring is not an object

of a possible experience. The principle of causality is ·of course· valid only within the domain of experience; you're diverting it from its proper role if you use it outside that domain, where there's nothing to apply it to and where indeed it is meaningless.

Now I maintain that any attempt to use reason in theology in a merely speculative manner is utterly useless and intrinsically null and void. There are principles governing the ·legitimate· use of reason in the study of nature, but they don't lead to any theology. So the only *theology of reason* that there can be is one that is based on, or seeks guidance from, moral laws. That's because the synthetic principles of reason are usable only immanently [= 'within the domain of experience'; see page 156], whereas to give us knowledge of a supreme being they would have to be used in a transcendent ·= experience-transcending· manner—which is impossible. If we could reach •the primordial being through the empirically valid law of causality, •that being would have to belong to the series of things encountered in experience—so •it would be conditioned in its turn, ·meaning that it wasn't primordial after all! And in any case, even if it *were* all right for us to use dynamical effect-to-cause reasoning to jump across the boundary of experience, what sort of a concept could we obtain by this procedure? *Not* the concept of a •supreme being, because that (·the supreme cause·) would have to be inferred from •the greatest of all possible effects (·the supreme effect, so to speak·)—and experience would never confront us with *that!* Couldn't we fill the great gap in our concept—the concept we *did* arrive at by effect-to-cause reasoning—by bringing in a mere idea of highest perfection and primordial necessity? Well, that might be granted as a favour; it can't be demanded as a right on the strength of a compelling proof. Perhaps the physico-theological argument's pairing of speculation with intuition could serve to

665

add weight to other theological arguments (if there were any), but all it can do, unaided, is to prepare the understanding for theological knowledge, tilting it in that direction; it can't complete the job on its own.

The moral is clear: transcendental questions have to be given transcendental answers, i.e. ones entirely based on *a priori* concepts, with nothing empirical added to the mix. But our present question is obviously synthetic; an answer to it would have to extend our knowledge •beyond all limits of experience, i.e. •to the existence of a being corresponding to a mere *idea* of ours, an idea that can't be matched in any experience. As I have shown, synthetic *a priori* knowledge is possible only as an upshot of what is needed for *experience* to be possible; so synthetic *a priori* principles are valid only within the given world, i.e. are applicable only to objects of empirical knowledge, appearances. That's why nothing comes of any attempt to achieve a theology through the transcendental use of purely speculative reason.

Perhaps there's someone who would rather •cast doubt on all my arguments in the Analytic than •let himself be robbed of his trust in the validity of the •theological• arguments that he has relied on for so long. Well, I have a challenge that he isn't entitled to duck:

Explain how—by what kind of inner illumination—you think you are capable of soaring so far above all possible experience, on the wings of mere ideas!

New arguments? new attempts to improve the old ones?—spare me! In fact he hasn't much room for choice, because all the merely speculative •theological• arguments eventually come back to a single source, the ontological argument; so I needn't fear being burdened by the fertile ingenuity of the dogmatic champions of reason-unconstrained-by-the-senses. Anyway, bring them on: though I don't regard myself as a quarrelsome person, I shall meet the challenge to examine

any theological argument of this sort that anyone comes up with, to show *where* it fails, and thus to nullify its claims. But •that cleansing task will never be completed, because—however long I keep it up, those who are used to dogmatic modes of persuasion will keep hoping to have better luck next time! So I confine myself to one little demand, namely that the dogmatists justify their position by answering, in terms that are universal and based on the nature of the human understanding and of all our other sources of knowledge, this question:

How we can even *begin* to extend our knowledge entirely *a priori*, carrying it into a realm where we can't have any experience or, therefore, any way of establishing the objective reality of any concept that we have thought up?

However the understanding arrives at a concept, the existence of its object can't be discovered (through analysis) *in* the concept. Why not? Because **(1)** the object of a concept—the item that it is the concept *of*—has to be something that exists independently of the concept, exists *outside* the thought the concept expresses; and **(2)** a concept can't lead us to something outside it. . . .

But although reason, in its merely speculative use, is not up to the great task of demonstrating the existence of a supreme being, it's still very useful as a *corrective* for any knowledge of this being that we get from other sources, making it consistent with •itself and with •every intelligible purpose, and cleansing it of everything •incompatible with the concept of a primordial being and everything •that would bring in empirical limitations. [This is a good place to remember that a thought doesn't have to be *true* to merit being called an *Erkenntnis*, here translated as '(an item of) knowledge'.]

So transcendental theology, despite its insufficiency, has an important negative role: it can serve as a permanent

sensor of our reason, when it is dealing strictly with pure ideas and therefore can't steer by anything that isn't transcendental. Suppose that on some other basis, e.g. on practical [here = 'moral'] grounds, the *presupposition* of **a supreme and all-sufficient being, as the highest intelligence** can establish its validity beyond all question. Then it will be of the greatest importance •to make sure that this concept is correct on its transcendental side, as the concept of **a necessary and supremely real being**; •to free it from any inappropriate empirical content (any anthropomorphism, broadly construed), and •to sweep away all counter-assertions, whether (1) *atheistic*, (2) *deistic*, or (3) *anthropomorphic*. The sweeping-away exercise won't be very difficult, because the same grounds on which we are shown that human reason can't establish the existence of such a ·supreme· being must also suffice to disqualify all counter-assertions:

- (1) There is no supreme being that is the primordial ground of all things.
- (2) The supreme being has none of the properties we attribute to it on the basis of an analogy between its output and our own.
- (3) The supreme being has all the limitations that sensibility inevitably imposes on the intelligences of which we have experience.

What premises would enable us to get, through a purely speculative use of reason, to any one of *those*?

For the merely speculative use of reason, therefore, the supreme being remains a mere *ideal*, but it's a *flawless* ideal, a concept that completes and crowns the whole of human knowledge. Speculative reason can't prove its objective reality, but it can't *disprove* it either. And if there should be a moral theology that can fill this gap, transcendental theology will be promoted from *problematic* to *indispensable*. It will be needed to specify the concept of this supreme being, and

constantly to run tests on reason, which is so often deceived by sensibility and sometimes not even in harmony with its own ideas. Because

- necessity,
- infinity,
- unity,
- existence outside the world (not as the world-soul),
- eternity as free from conditions of time,
- omnipresence as free from conditions of space, 670
- omnipotence, etc.

are purely transcendental predicates, the purified concepts of them that every theology needs so much can be obtained only from transcendental theology.

·A two-part· appendix to the transcendental dialectic

1. The regulative use of the ideas of pure reason

Pure reason's dialectical endeavours •confirm what I showed in the Transcendental Analytic, namely that all the inferences that claim to lead us beyond the domain of possible experience are deceptive and ungrounded; and they also •teach us something else. This further lesson is that human reason has a natural tendency to overstep these boundaries, and that •transcendental ideas are just as natural to •reason as the •categories are to •understanding, though with this difference: whereas the categories lead to •truth, i.e. to our concepts' fitting the object, the ideas create mere •illusion—an irresistible illusion that we can hardly cure ourselves of even by means of the severest critique.

671 Everything that is grounded in the nature of our faculties must be appropriate to and consistent with the faculties' proper use—as long as we can guard against a certain misunderstanding and so discover the direction these faculties ought to take. So the transcendental ideas presumably have their own good, proper, and therefore *immanent* use, though when their meaning is misunderstood and they are taken for concepts of real things, they get used in a transcendent way which makes them delusive. What we have here are not two sorts of *ideas* but only two ways of (mis)using ideas:

- the roaming or transcendent use, in which the idea is taken beyond the range of possible experience and taken to apply directly to some object that is supposed to correspond to it;
- the homebody or immanent use, in which the idea is aimed solely toward the use of understanding as such,

and has to do only with objects that fall within the understanding's compass.

[Note the contrast between 'apply directly to' and 'have [something] to do with'.] All errors of subreption [see note on page 242] are due to a failure of judgment, never of understanding or reason.

Reason never relates directly to an object. All that it immediately relates to is the understanding; and it's through the understanding that it gets its own empirical use. So it doesn't •create concepts of objects, but only •organizes them, giving them the unity that they can have when used in their widest possible application, i.e. in connection with the *totality* of this or that series ·of conditions·. The understanding pays no attention to this totality; all it cares about is the connecting-up by which such series of conditions come into existence and are held together by concepts. So reason's only 'object' is the understanding and the right way to use it. 672
Just as

the understanding uses concepts to pull the manifold ·of sensibility· together in the object,

so also

reason uses ideas to pull the manifold of concepts together by presenting a certain collective unity as the goal of the understanding's activities, which would otherwise be concerned solely with distributive unity.

[This language of 'collective/distributive unity' occurs in only one other place in the work, namely on page 269. Neither there nor here does Kant say clearly how 'distributive unity' differs from 'disunity', but we can perhaps gather what 'collective unity' is meant to be. Making the understanding aim at collective unity, it seems, is making it aim at constructing some *single unified* intellectual item; in the earlier passage Kant focuses on reason's error in taking that item to be an *object* such as •the being that has all reality, or •the whole of past time. In our present passage he evidently holds that there's nothing wrong with the urge-towards-constructing-a-grand-single-something, as long as we

don't perform a bait-and-switch act and convince ourselves that we are talking about a grand non-empirical object.]

So my view is this: transcendental ideas are never to be used •constitutively, posing as concepts of certain objects. When they *are* so used, they're merely sophisticated (dialectical) concepts. On the other hand, they have an excellent •regulative use, and we *need* them in that role, in which they direct the understanding towards a certain goal, setting directional lines along which all its rules converge as though on their point of intersection. Of course this point •isn't anything real; it• is a mere idea, a *focus imaginarius* [= 'imaginary focus'. Kant has just spoken of reason's ideas as directing the understanding **towards** (*zu*) this focal point; that's the direction indicated by several things in this paragraph up to here. But he immediately goes on to write as though it were something the understanding might be thought of as moving **from**. Thus:] Because this •focal point• lies quite outside the bounds of possible experience, the concepts of the understanding don't really emanate from it; yet it serves to give to these concepts maximal unity combined with the maximal scope. This is the source of the illusion that the directional lines radiated out from a real object lying outside the field of empirically possible knowledge—just as objects reflected in a mirror are seen as behind it. [*That's* why Kant replaced 'towards' by 'from'! He wanted to bring in that neat comparison with the apparent position of something seen in a mirror. Kant will mention that comparison once more, but from now on reason's role will always be described in terms of what it directs the understanding towards.] We don't have to let this illusion actually deceive us, •but we can't get rid of it, because• it is indispensably necessary if we are to direct our understanding to keep extending its range as far as it possibly can. Analogously, the object-behind-the-mirror illusion doesn't have to deceive us, but it can't be got rid of as long as we are using a mirror to see things that are behind us.

If we survey the entire range of knowledge that our understanding brings to us, we find that reason's special concern with this range is •to demand that this knowledge be *systematic*, hanging together under a single principle, and •to try to bring this about. This unity demanded by reason always presupposes an idea, namely the idea of the form of a knowledge-whole that •precedes the specific items of knowledge of the parts and •contains the conditions that settle in advance the place of each part within the whole. [Kant doesn't mean that the knowledge-whole is achieved *before* we know any of the details. The 'preceding' of which he speaks is logical rather than temporal.] So this idea postulates a complete unity in the understanding's knowledge, a unity in which this knowledge isn't a mere contingent heap of items but is a system held together in accordance with necessary laws. This idea is a concept—what's it a concept *of*? Not of any object! Rather, it is the concept of the thoroughgoing unity of concepts of objects, with this unity serving as a rule for the understanding. These concepts of reason aren't derived from nature; on the contrary, we interrogate nature in accordance with them, and regard our knowledge as defective so long as it isn't adequate to them. •Here is an example•. It is agreed that *pure earth, pure water, pure air* etc., are scarcely to be found; but we need the concepts of them in order properly to determine the share that each of these natural causes has in producing appearances. •Why does the fluid in this flask behave as it does? What is the effect of its including bromine? of its containing common salt? of its component of *pure water*? So the concept of *pure water* is empirically serviceable, despite the fact that• its *pure* element comes solely from reason. . . .

Taking reason as a faculty for deducing the particular from the universal, •its work falls into two classes•. (1) In one, the universal proposition is already *certain in itself* and

given, so that only *judgment* is needed to bring the particular under it, and this inference reveals the particular as also being necessary. I call this the ‘apodeictic’ use of reason. (2) In the other, the particular is certain, but the universal proposition from which it is derived is being accepted only as *problematic*, so that the universality of the rule from which the particular is inferred is still a problem. Several particular instances, each of them independently certain, are tried out on the rule to see whether they follow from it. If it turns out that all the particular instances we can come up with do follow from the rule, we infer ·upwards· from this to the universality of the rule, and then from the rule ·downwards· again to the particular instances—all of them, even those that are not themselves given. I call this the hypothetical use of reason.

675

The hypothetical use of reason, based on ideas viewed as problematic concepts, isn’t really *constitutive*, because in following it we don’t strictly prove the truth of the universal rule that we have adopted as an hypothesis. If every possible consequence of it really did follow from it, that would indeed *prove* its universality, but how are we to know them all? The hypothetical use of reason is only regulative; its aim is to unify our items of knowledge as much as possible, thereby approaching universality for the rule.

So the hypothetical use of reason aims for the systematic unity of the understanding’s items of knowledge, and this unity is the *criterion of the truth* of reason’s rules. On the other hand, this systematic unity (as a mere idea) is only a *projected* unity, to be regarded only as a problem and not as something given. This unity—i.e. this *idea of* unity—aids us in discovering a principle governing the various special doings of the understanding, a principle that will lead the understanding to cases that are not given, thereby making it more coherent.

676

But you can see from this that the systematic or reason-demanded unity of the manifold knowledge of understanding is a *logical* principle. Its role is to •deploy ideas to help the understanding in cases where it can’t establish rules on its own, while also •giving to the many different rules of the understanding a systematic unity under a single principle, thereby doing all it can to produce coherence. Should we accept this?—

Systematic unity is *right*, given **how objects are constituted**. We can confidently postulate this unity *a priori*, irrespective of any special interest of reason; so we’re in a position to maintain with certainty that all the understanding’s items of knowledge (empirical knowledge included) *have* the unity required by reason, and fall under common principles from which, despite their variety, they can all be derived.

No! That asserts a *transcendental* principle of reason, something claiming to be an objectively valid truth, not merely a logical, subjective rule of method. And that holds not only for the position as stated above, but also a different version of it that says:

Systematic unity is *right*, given **the nature of the understanding** that knows objects as objects . . . and so on.

I’ll illustrate this with an example of the use of reason. Among the various kinds of conceptual unity that the understanding has dealings with is the unity of the different causal powers of a single substance. The many appearances of a single substance look at first sight to be so unlike that we start out with the assumption that they are effects of correspondingly many different powers of the substance; as with sensation, consciousness, imagination, memory, ingenuity, discrimination, pleasure, desire, and so on ·as supposedly different powers or faculties· of the

677

human mind. Now there's a logical [here = 'methodological'] maxim telling us what to do right from the start, namely to reduce this seeming diversity as much as possible, by comparing these ·effects or these supposed powers· and detecting their hidden identity—for example investigating whether imagination combined with consciousness may not be the same thing as memory. . . .and so on·. Though logic can't decide whether a *basic power* actually exists, the idea of such a power is the problem posed for [here = 'the challenge to be met in'] a systematic treatment of the multiplicity of powers. The logical principle of reason demands that we bring about this unity as completely as we can; and the more the appearances of power x and power y are found to be exactly alike [Kant writes *identisch*], the more probable it becomes that they are merely different expressions of a single power; and we could call this a relatively basic power, the one that is the basis of powers x and y. And similarly with the other powers.

The relatively basic powers must in turn be compared with one another, with a view to discovering their harmony and so bringing them nearer to a single absolutely basic power. But this reason-demanded unity is purely hypothetical. The claim is not •that such a power must be there, but only •that we have to *look for it* in the interests of reason, i.e. for the setting up of certain principles for the various rules that experience may supply to us, trying in this way to bring as much systematic unity as possible into our knowledge.

When we look at the transcendental use of understanding, we find that this idea of a basic power is not •treated merely as a problem ·or task· that is to be used hypothetically, but •claimed to have objective reality, as declaring that the various powers of a substance are systematically unified and yielding an absolutely necessary principle of reason. For without having tried to show the harmony of these various

powers, or even having tried and always failed, we still take it that such a unity does actually exist. And we take this line in connection not only •with the different powers of **a single substance** (as in the cited case ·of the human mind·) but •with the powers of **a kind of stuff**—such as *matter*—where we find ·in different samples of the kind· powers that are different from one another though they have a certain amount in common. [Kant doesn't have a phrase corresponding to 'a kind of stuff', but his example of *matter* shows that what he has in mind is the distinction between countable substances and undifferentiated kinds of stuff.] In all those cases reason says that the various powers ·under investigation· *are* systematically unified because special natural laws *do* fall under more general laws. Parsimony in principles is one of nature's own laws; it's not merely something that reason requires in the interests of good management.

Actually, one can't see how there *can* be •a logical [= 'methodological'] principle of unity-of-rules unless there is also •a transcendental principle whereby such a systematic unity is *a priori* assumed to be something that the objects necessarily *have*. 679 Reason in its logical use calls on us to treat the variety of powers exhibited in nature as a disguised unity and to derive this unity, as far as possible, from a basic power. How could reason be entitled to make this demand if it were free to admit that in fact all the powers are different and that nature doesn't permit them to be systematically unified? If reason made that admission, it would be opposing its own vocation, striving for an idea that was inconsistent with the constitution of nature. You might say:

·Perhaps reason doesn't have to *presuppose* this unity in nature·. Perhaps while proceeding in accordance with its own ·methodological· principles reason *learns about* this unity from the facts about how nature *happens to be* constituted.

No! The law of reason that tells us to look for this unity is a *necessary* one, because

- without it we wouldn't have reason, and
- without reason we wouldn't have any coherent use of the understanding, and
- without that we wouldn't have any adequate criterion of empirical truth.

Conclusion: if we are to have such a criterion we have to presuppose the systematic unity of nature as objectively valid and necessary, which means that we have to accept this as a transcendental principle.

Although philosophers haven't always acknowledged this transcendental principle, even to themselves, or indeed been conscious of using it, we still find it wonderfully *buried* in the principles on which they proceed:

680

The multitude of ways in which individual things differ don't rule out identity of species; the various species must be regarded as merely different special cases of a few genera, and these in turn of still higher genera, and so on; in short, we must seek for a certain systematic unity of all possible empirical concepts by deriving them from higher and more general concepts—this is a logical principle, an academic rule, without which there couldn't be any use of reason. Why not? Because we can't infer particulars from universals—which is reason's basic activity—except where we credit things with having universal properties that are the foundation of the particular properties.

Philosophers presuppose that such unity is to be found in nature when they accept the familiar academic rule that rudiments or principles mustn't be needlessly multiplied (*entia praeter necessitatem non esse multiplicanda*). [The Latin sentence means that *entities* aren't to be multiplied beyond necessity. This is famous under the title 'Occam's Razor'.] This says that the

nature of things provides reason with what it needs for its purposes, and that the seemingly infinite variety in phenomena shouldn't dissuade us from assuming that behind this variety there's a unity of basic properties from which all the variety can be reached as a multitude of special cases. Although this unity is a mere idea, it has always been so eagerly pursued that there has been a need to moderate rather than to encourage the desire for it. It was a big step when chemists succeeded in reducing all **salts** to two main genera, namely acids and alkalis; and now they're trying to show that there's just one basic material of which acids and alkalis are merely special cases. They have worked at gradually bringing the number of basic kinds of **earths** (the material of stones and even of metals) down to •three, and eventually to •two; but, not content with this, the chemists can't get rid of the thought that these two are just special cases of •one genus, a single basic kind of earth; and that even basic salt and basic earth may be special cases of something lying still deeper. You might think that this is merely reason being economical, saving itself from trouble—adopting an hypothesis that will gain probability by any success that it achieves. But that is not so; it's easy to distinguish the •idea from •a procedure in which reason is merely catering to its own interests. Anyone working with the idea presupposes that the unity demanded by reason squares with nature itself, though admittedly it can't say *how far* this unity goes. Reason isn't *asking*—it's *commanding*.

681

If among the appearances that we encounter there was so much variety. . . in content that even the acutest human understanding couldn't see the slightest similarity among them (which is perfectly conceivable), the logical law of genera would have no sort of standing; we wouldn't even have the concept of a genus, or indeed any other universal concept; and there would be no such thing as the understanding,

682

because it deals only with such concepts. Thus, if the logical principle of genera is to be applied to nature. . . ., it presupposes a transcendental principle. And according to that principle there has to be sameness of kind in the manifold of experience (though we can't tell *a priori* how much of it there is), because if there weren't any samenesses of kind there couldn't be any empirical concepts, and so there couldn't be any experience.

The logical principle of genera, which demands identity, is balanced by the principle of species. This •calls for complexity and variousness in things (despite their sharing the same genus), •tells the understanding to attend to the diversity as much as to the identity. This principle of species (of discrimination and acuteness) stops the principle of genera (of breadth of thought) from going too far. So reason turns out to have two interests that are in tension with one another. **(1)** On the one hand there's an interest in *extent* (universality) in respect of genera, leading the understanding to get more *under* its concepts. **(2)** On the other hand, there's an interest in *content* (determinateness) in respect of the multiplicity of the species, leading the understanding to get more *into* its concepts. This twofold interest shows up in scientists' different patterns of thought. Those who are most given to general theories are hostile (as it were) to qualitative differences and are always on the look-out for the unity of the genus; while those who are most empirical in their approach keep busily trying to split nature into so much variety that one might almost despair of ever being able to bring its appearances under universal principles!

This diversity-seeking mode of thought is evidently based on a logical principle that aims at the systematic completeness of all knowledge, telling us that if we start with the genus we should come down to the level of the manifold that falls under it, thus ensuring the system's •scope; just

as the other principle has us going up to the level of the genus, trying to secure the system's •unity. No amount of knowledge about the range of the concept that marks out a genus will tell us how far we could go in dividing it up into species, just as our knowledge of the space that a body occupies won't tell us how far we could go in dividing it up into parts. Consequently, every genus requires diversity of species, and these in turn require diversity of subspecies; and since each of these subspecies has a domain that is covered by a general concept, reason. . . .demands that no species be regarded as being intrinsically a lowest species, .i.e. one that can't be split up any further. That is because any species—even a sub-sub-...species with as many subs as you like—is always a concept, containing only what is common to different things, so that it isn't completely specified. So it can't be directly related to an individual, and other concepts must always be contained under it. . . .

But it is easily seen that this •logical law would be senseless and useless if it didn't rest on a •transcendental law of specification, not one demanding an actual *infinity* of differences in the things that can be objects to us. . . .but one requiring the understanding when it has found a species to look for subspecies under it. . . . For if there were no *lower* concepts, there couldn't be *higher* ones. Now, the understanding deals only in concepts; so this process of division, however far it goes, never divides anything through intuition but always only through lower and lower concepts. The knowledge of appearances in all their detail, which is possible only through the understanding, demands an *endless* process of fine-graining our concepts. . . .

This law of specification can't be derived from experience, which can't reveal to us any such discovery as that *every* species has sub-species! The empirical process of identifying smaller and smaller species soon comes to a stop. . . .if it

isn't guided by the above-mentioned transcendental law of specification which, as a principle of reason, leads us always to look for further differences and to suspect that they are there even when the senses can't find them. [Kant then gives an example of species-division from the history of chemistry, saying that it wouldn't have happened if the law of specification hadn't been at work. He repeats that the possibility of concept-use, and thus the possibility of employing the understanding, depends on the assumption of differences and likenesses in nature. Then:]

Thus, reason prepares the ground for the understanding: **(1)** through a principle of the *homogeneity* of the manifold under higher genera; **(2)** through a principle of the *variety* of the homogeneous under sub-species; and **(3)** in order to round out the systematic unity, the further law of the *kinship* of all concepts—a law that prescribes that we proceed from each species to every other through a gradual increase of the diversity. We can call these the principles of **(1)** *homogeneity* **(2)** *specification*, and **(3)** *continuity* of forms. You get **(3)** by combining **(1)** and **(2)**, because the idea of **(3)** complete systematic connection involves the thought of **(1)** ascending to higher genera and **(2)** descending to lower species [e.g. ascending from *man* to *vertebrate* and then down from there to *vertebrate*]. That gets all the manifold differences to be related to one another because they all descend from one highest genus down through all degrees of specification.

[Kant now offers a spatial model to illustrate 'the systematic unity prescribed by the three logical principles'. The notion of a *space* of concepts, a logical space, is one that he used effectively back in the *Analytic* [pages 49–50], but his present use of it is an obstacle to understanding. The model is intrinsically clumsy; and to grasp how it works (insofar as it *does* work) you have *first* to grasp firmly the 'principles' that it's supposed to model—so that the model

doesn't help. In expounding it, Kant repeats and emphasizes the three logical principles: **(1)** The 'law of homogeneity', which says that there is one concept (that of the 'highest genus') which is an ingredient in *every* other concept. **(2)** The 'law of specification', which says that *every* general concept is an ingredient in some other concepts, ones that are more specific than it is. **(3)** The 'law of the continuity of forms', which says: Given any two concepts, there is some conceptual ingredient that they both have. (In his handling of this in the context of the model, Kant does rightly say that the journey—as it were—from one concept to another may involve going up before going down.) Here is the story again, as given by Kant after the model:]

So **(1)** keeps us from extravagantly allowing many different basic genera, and points us towards homogeneity; **(2)** restrains this tendency towards unity, and commands us not to apply any universal concept to individuals until we have distinguished subspecies within it. **(3)** combines these two laws by prescribing that even amidst the utmost manifoldness there is homogeneity that allows stepwise transition from one species to another, thus recognizing the kinship of the different branches that all spring from the same stem.

[That paragraph was reached by skipping over something that should now be mentioned, namely Kant's taking **(3)** to imply that

- Between any two concepts there is at least one intervening concept,

from which of course it follows that

- Between any two concepts there are infinitely many intervening concepts.

In his words: 'There is a continuity of forms. . . . You never get from one species to another by a •jump, but only •by gliding through• all the smaller degrees of difference that

688 come between them. In short, reason doesn't allow that any two species or subspecies x and y are the nearest possible to each other; there can always be still other intermediate species that are less different from x and from y than x and y are from one another.' Kant announces this thesis abruptly and without argument; he seems to have been seduced into it by a couple of suspect features of his spatial model. But we needn't go into that, and can now let Kant continue:]

This **(3)** logical law of the •continuity of species presupposes the **(3t)** transcendental law of •continuity in nature [both phrases are given in Latin]. Without **(3t)**, the logical law **(3)** would only lead the understanding astray, sending it along a path that may be quite contrary to the path that nature itself prescribes. So the grounds for **(3)** must be purely transcendental, not empirical. If the grounds for **(3)** were empirical, this law would come later than the systems •through which the empirical materials were made available; whereas in actual fact •it's the other way around: **(3)** has given rise to all that is systematic in our knowledge of nature. Something we can do with an hypothesis that we think up is to test it experimentally; if it survives the tests, that's evidence for its truth; and our present three laws *can* be handled in that way, and do perform some service in that role. But that's not what they are *for*. It is *not* the case that we have formulated them, thinking them up out of our own heads, as hypotheses to be tentatively put forward to be experimentally tested. It's obvious from looking at these laws that they regard **(1)** the parsimony of basic causes, **(2)** the manifoldness of effects, and **(3)** the consequent kinship of the parts of nature as being in agreement with reason and with nature. So these principles carry their credentials with them; they are not to be valued merely as procedural rules. [Notice that in this paragraph our three laws concern the •causal structure of the world, not its •qualitative structure which is how Kant

689

first introduced them. He returns to qualitative structure in the next paragraph.]

But it's easy to see that **(3)** this continuity of forms is a mere idea, and can't be cashed out by anything discovered in experience. •There are two reasons for this. **(a)** The species in nature are actually separated from one another—they are discrete, not smoothly continuous. If the tracing out of the kinship between two species—•man and fish, say—were truly *continuous*, there would be a true infinity of intermediate species between any two given species, which is impossible. [Kant isn't contradicting himself here. His continuity thesis is about concepts, corresponding to *possible* species, whereas the point he makes here concerns *actual* species.] **(b)** We couldn't make any determinate empirical use of this law, because all it does is to tell us in broad terms to *seek kinship* •among species; it says nothing about how we are to recognize kinship, about how far it goes, or about how to look for it.

. . . Reason starts from the understanding's items of knowl- .690
edge, which are immediately related to experience, and engages in an idea-guided search for the unity of this knowledge—a unity that goes far beyond possible experience. The kinship of the manifold. . . has to do with •things, but it has still more to do with their •properties and •powers. Here's an example •in which, you'll notice, what is at stake is the affinity not of the •planets but of their possible •orbits:

Our imperfect experience presents the orbits of the planets to us as circular. Then we find deviations from that. We suppose that these non-circular orbits approximate more or less closely to a circle, and that there's a fixed law [here = 'mathematical formula'] that covers the circle, these non-circular figures, and all the infinity of figures that come between them. And so we come on the *ellipse*.

•And then a further application of the same procedure:

So far as we can see, *comets* follow paths that are even more divergent from circles, because they seem not to return, i.e. not to have paths that are closed loops. We handle that by looking for something that ·mathematically· unites those paths with ellipses, and so we come upon the *parabola*. This is akin to the ellipse; indeed, an ellipse with a long enough major axis can't be observationally distinguished from a parabola.

691

Thus, under the guidance of these principles we discover •a unity in the generic shapes of these paths ·of the planets and comets·, and through that •a unitary cause of all the laws of planetary motion, namely •gravitation. From there we extend our conquests still further, trying to explain by the same principle all variations and seeming deviations from these rules. Eventually we make additions that experience can never confirm: the rules of kinship lead us to conceive of comets as following hyperbolic paths, in the course of which they entirely leave our solar system and—passing from sun to sun—unite the most distant parts of the universe, a universe that is unlimited so far as we can tell but is held together by a single moving force.

The only feature of these principles that concerns us here is a remarkable one, namely: *they seem to be transcendental*. All they contain are mere ideas to guide the empirical use of reason—ideas that reason follows only asymptotically, i.e. ever more closely without ever reaching them—and yet •they are synthetic *a priori* propositions that have objective but indeterminate validity, •they serve as rules for possible experience, and •they can also be very useful as guides to procedure in the advance of science; but •they can't be ·legitimised by being· given a transcendental deduction—I showed earlier [page 172] that such a deduction can never be given for any ideas ·of reason·.

692

[The next short paragraph is tiresomely difficult. In it Kant takes us back to Analytic and then forward again through some flourishes that aren't essential to what follows. The upshot of all this is the firm assertion that the principles of pure reason can't possibly be brought to bear directly on experience; which prompts Kant to ask:] If we thus disallow such empirical use of the principles of reason as constitutive principles, how can we secure for them a regulative use and thereby some sort of objective validity? And what would such a regulative use *be*?

[This paragraph will considerably amplify what Kant wrote, in ways that the small-dots convention can't easily indicate. But the core of the paragraph is there in Kant's words.] **(a)** Just as sensibility is an object for the understanding, so also **(b)** the understanding is an object for reason. It's the understanding's job to

(a) work on the manifold of the appearances by means of concepts, and to bring it under empirical laws, and it's reason's job to

(b) work on all possible empirical acts of the understanding, bringing systematic unity to them.

It's to be expected that there will be some analogy between **(a)** how the understanding works on appearances and **(b)** how reason works on the doings of the understanding; and one part of that analogy comes to our attention now. Back in the Analytic [see pages 93–4] we saw that

(a) Thoughts by the understanding were transformed from such indeterminate (vague) items as

•if-then propositions

•subject-predicate propositions

to the determinate items

•causal propositions

•propositions about substances;

this change being produced by adding to each basic concept of the understanding *a sensible schema*,

which was tantamount to building the notion of *time* into it.

693 Now, the concepts that reason deals with are also indeterminate; this can't be cured by adding anything *sensible* to them, but something analogous to that does happen, namely:

(b) Commands by reason are transformed from such indeterminate (vague) items as

- look for causal explanations
- look for common features

to the determinate items

- look for *complete* causal explanations
- look for *the greatest possible* qualitative unity among things.

Thus, reason's analogue of the understanding's sensible schema is just the notion of a *maximum*.

The notion of *greatest* or of *absolutely complete* is perfectly determinate; when it is built into reason's commands, they tell us *exactly* what we should do. (If the commands were less stringent—'Look for as much causal explanation as meets this or that qualification'—indeterminacy will come in via the qualification.) There is also a disanalogy: **(a)** when you amplify a category by adding its sensible schema, you add to the content of what's said about the object; the statement

•The water's freezing *causes* it to harden

says *more* than does the statement

•The water's freezing *is if-then related in some way* to its hardening.

But **(b)** the notion of a *maximum* doesn't add anything to what reason implies about the world. Reason's output does have *some* implications for the experienced world: any principle that *a priori* •prescribes to the understanding that it should produce thoroughgoing unity in its use also indirectly •says something about the object of experience; so the principles of pure reason must have objective reality in respect

of that object. But it doesn't imply anything *determinate* about its object, because the element that makes reason's principles determinate—the element that is analogous to the schematism of the categories—is the notion of *maximum* or *greatest possible*, which has to do not with •what the world is like but with •how our understanding ought to behave. 694

I use the label '*maxim* of pure reason' for any subjective principle that comes not from •the constitution of an object but from •reason's interest in achieving a certain possible completeness in its knowledge of the object. Using this label, then: there are maxims of speculative reason, which rest entirely on its speculative interest though they may seem to be objective principles.

When merely regulative principles are regarded as constitutive, then as objective principles they can conflict with one another. But when they are regarded merely as *maxims*, there's no real •conflict but merely. . . •different methods of trying to satisfy reason's one and only interest; and •we get an impression of conflict because •these methods can get in one another's' way.

So it may happen that one incompetent thinker is especially interested in *manifoldness* (in accordance with the principle of specification), while another cares more about *unity* (in accordance with the principle of homogeneity); they think they are disagreeing about the nature of the object, 695 whereas really it's a difference in which of the two principles each puts uppermost. And since neither of these principles is based on objective grounds, but solely on the interest of reason, it would really be better to call them 'maxims' rather than 'principles'. When I see intelligent people disputing about the characteristic properties of human beings—

for example, with some assuming •that there are certain special hereditary characteristics in each nation, or certain well-defined inherited differences in

families, races, etc., while others insist •that nature has made the same provision for everyone, and that the differences are due to external accidental conditions

—I have only to consider what sort of object they are talking about (•namely, *human nature*•) to realise that it's hidden far too deeply for them to be in any position to base their dispute on insights into its nature. What we see here is just the twofold interest of reason, with one party •to the dispute• embracing or at least going along with one of the interests, and the other party the other; and it's easy to •bring peace to the dispute when it is understood in *that* way, because it's easy to •reconcile the maxims of manifoldness and of unity in nature. But as long as the maxims are taken to be objective propositions about nature, we'll have disputes that will be impede science because research will be held up until they can be settled. (I stated this in terms of a fight about human nature; but it could as well have been a dispute about animals or plants—or even minerals.)

Another example of the same thing is the dispute over Leibniz's widely discussed law of the *continuous gradation* of created beings. . . . This law is simply the following out of the principle of kinship that rests on the interest of reason; it couldn't possibly be based on observation and insight into the constitution of nature. The differences between things that we encounter in our experience are much too big •to suggest that there's a continuous difference-bridge across the gap•; and even when we encounter differences that seem tiny to us, they won't be tiny from nature's point of view. There's no chance at all of our reaching a decision about the law of continuity by the empirical study of nature. . . . On the other hand,

•the method of looking for order in nature in accordance with such a principle,

and the

•maxim that prescribes that we regard such order as grounded in nature as such (without specifying where and how far it goes)

is certainly a legitimate and excellent regulative principle of reason. In its regulative role it goes far beyond anything that experience or observation could verify, but not by •stating facts that are inaccessible to experience. What it does is to •mark out the path towards systematic unity.

2. The final purpose of the natural dialectic of human reason

The ideas of pure reason can't ever be dialectical [= 'illusion-creating'] in themselves; any deceptive illusion involving them must be due solely to their misuse. Why? Because we get them from the very nature of •our reason; and it's impossible that that •supreme court for the rights and claims of speculation should itself generate deceptions and illusions. It's to be expected, then, that the ideas have their own good and appropriate role in the natural conduct of our reason. But the rabble of sophists are up to their old tricks: they scream 'Absurdity!' and 'Contradiction!' against reason; they can't penetrate to its innermost designs, but that doesn't stop them from judging and condemning it. What makes it possible for them to stand on their own feet and assertively blame and condemn what reason requires of them? It is a culture that comes from the beneficent influences of reason!

We can't confidently use an *a priori* concept unless we have first given a transcendental deduction of it, •i.e. a demonstration that the concept is a legitimate one to use. The transcendental deduction of the •categories (**concepts of pure understanding**) legitimised them by showing that they *must* fit the items they are meant to fit. The •ideas (**concepts**

698 **of pure reason**) can't be legitimised in *that* way. But *some* deduction of them must be possible, even if it's very different from the transcendental deduction of the categories. If we can't provide that, then the ideas won't have any objective validity—not even a small degree of very vague objective validity—and they'll have to be written off as mere empty thought-entities. I am now going to present the needed deduction; that will complete the critical work of pure reason.

There's a big difference between something's being given to my reason •as an object *period* and something's being given to my reason merely •as an object *in the idea*. In the former case my concepts serve to determine •= specify• the object; in the latter case there's actually only a schema, and no object for it is directly given, even in a hypothetical manner. All it does is to represent other objects indirectly, through how they are unified by means of their relation to this idea. Thus I say that the concept of a highest intelligence is a mere idea; i.e. its objective reality doesn't consist in its referring point-blank to an object (if it did, we could never show that it *is* objectively valid). It's only a schema constructed in accordance with the conditions of the greatest unity of reason—the schema of the concept of *thing*, the concept of a thing as such. And its role is just to secure the greatest systematic unity in the empirical use of our reason, which it does through our regarding the object of experience as being based on, or having been caused by, the imaginary object of this idea. We say, for instance, that the things of
699 the world must be viewed *as if* they got their existence from a highest intelligence. The idea is thus really only a heuristic concept, not an ostensive one [= 'a concept that guides discovery, not one that shows anything']: it doesn't show us how an object is constituted, but how, under the guidance of this idea, we should *try to discover* how the objects of experience are constituted and inter-connected. So if it can be shown that

the three transcendental ideas (psychological, cosmological, and theological), although they *don't* directly latch onto and specify any corresponding object, nevertheless *do*—in their role as rules for the empirical use of reason—lead us towards systematic unity by presupposing such an *object in the idea*; and in this way broaden our empirical knowledge without ever being able to run counter to it,

then •we can conclude that• it's a necessary maxim of reason to proceed always in accordance with such ideas. And *that* is the transcendental deduction of all ideas of speculative reason—not as *constitutive* principles for the broadening our knowledge to more objects than experience can give, but as *regulative* principles for bringing into the manifold of empirical knowledge a systematic unity that it couldn't achieve without the aid of these principles.

700

I'll clarify that. When we follow the above three ideas as principles we'll do three things. **(1)** In psychology, under the guidance of inner experience, we'll connect up all the appearances—all the inputs and outputs of our mind— *as if* the mind were a simple substance that stays in existence with personal identity (in this life at least), while its states. . . .continually change. **(2)** In cosmology, we must track the conditions of both inner and outer natural appearances in a never-completed enquiry—*as if* the series of appearances were itself endless, having no first or top member. (This needn't involve us in denying that the series of appearances has •purely intelligible causes—i.e. ones that don't themselves belong to the series—but we mustn't bring •these into any of our explanations of nature, because we don't know a thing about •them.) **(3)** In the field of theology, we must view everything that can belong to the fabric of

possible experience

as if this experience constituted an absolute unity, but one that is dependent through and through, and conditioned within the world of sense;

and yet also at the same time

as if the sum of all appearances (the sensible world itself) had a single, highest and all-sufficient basis lying outside its own territory, namely a self-subsistent, primordial, creative •reason;

701 in the light of which we guide the empirical use of *our* •reason to give it the broadest extent, by viewing all objects

as if they drew their origin from such an archetype.

In other words, **(1)** we oughtn't to derive the inner appearances of the soul from a simple thinking substance, but should derive them from one another under the guidance of **the idea of** a simple being. **(3)** We oughtn't to derive the order and systematic unity of the world from a supreme intelligence, but to get from **the idea of** a supremely wise cause the rule that guides our reason in making the best possible job of connecting empirical causes and effects in the world. . . .

Now there's not the slightest obstacle to our *assuming* that the **(1)** psychological and **(3)** theological ideas are objective, i.e. to our hypostatizing them. (Not so with **(2)** the cosmological ideas: if reason treats *them* as objective it falls into antinomy, which the other two don't.) So how can anyone quarrel with us about their objective reality? Anyone who denies that they are possible has no more *knowledge* to back up his denial than we have to back up our affirmation! But there not being 'the slightest obstacle' to assuming something doesn't automatically make it all right for us to assume it; and it's *not* all right for us to introduce thought-entities that transcend all our concepts (without contradicting them) as being real and determinate

objects, merely on the say-so of a speculative reason that wants to complete its work. They oughtn't to be assumed as existing in themselves; the only reality they are entitled to is the reality of a *schema* for the regulative principle of the systematic unity of all knowledge of nature; their legitimate status is: •analogues of real things, not: •real things. We strip from the object of the idea the conditions that •constrain the concept-of-the-understanding of it, and also •are needed for us to have a determinate concept of anything. What that leaves us with is the thought of a Something of whose intrinsic nature we have no concept whatsoever, but which we represent to ourselves as relating to the totality of appearances in a way analogous to how appearances relate to one another. 702

When we accept ideal beings in this way, we aren't stretching our knowledge out beyond the objects of possible experience. What we're doing is to increase the empirical unity of our experience through the systematic unity for which the idea provides the schema—so that the idea's legitimate status is that of a regulative principle, not a constitutive one. In positing a thing (a Something, a real Being) corresponding to the idea, we aren't claiming to use transcendental concepts to extend our knowledge of things; because this Being is posited only •in the idea and not •in itself, so that all it does is to express the systematic unity that is to guide the empirical use of reason. It doesn't say what this unity is based on, i.e. what the intrinsic nature is of the Being that causes the unity. 703

So the transcendental concept—the only determinate concept—that purely speculative reason gives us of God is in the strictest sense *deistic* [see pages 288 etc.]; i.e. reason doesn't guarantee the objective validity of this concept, but only gives us the idea of something that is the basis for the supreme and necessary unity of all empirical reality.

The only way we can think about this ‘something’ is on the analogy with

- a real substance that causes everything, in accordance with laws of reason.

·Contrast that with this—

- a real substance that causes a change in another substance, in accordance with the laws of the understanding

—which we can think of directly, and not only by analogy. If we want to think of it as a special object, we have to think of it in this ·analogical· manner. ·*Must* we think of it as a special object? No.; the alternative is to settle for the mere idea of the regulative principle of reason, setting aside ‘the completion of all conditions of thought’ as going beyond the limits of the human understanding. This alternative, however, doesn’t square with the pursuit of complete systematic unity in our knowledge to which reason at least sets no limits.

704 This, then, is how matters stand: When I posit a divine Being, I haven’t the slightest conception of its supreme perfection as intrinsically possible, or of the necessity of its existence; but I *am* in a position to answer satisfactorily all those questions that relate to contingent matters, and to give reason the most complete satisfaction regarding •the highest unity that it pursues in its empirical use, but not regarding •the posited Being itself. This shows that what justifies reason in thus setting off from a point that lies so far above its sphere, and trying in this way to survey its objects as constituting a complete whole, is the speculative •interest of reason, not any •insight.

We now meet a difference between two ways of viewing a single assumption; it’s rather subtle, but is important in transcendental philosophy. I may have sufficient ground to assume something •in a **relative** way without being entitled to assume it •**outright**. [To mark Kant’s stress on this distinction,

from here on ‘relative’ etc., when they are translations of Kant’s *relativ* etc., will always appear in bold type.] This distinction comes into play when we’re dealing with a merely regulative principle, knowing that it is necessary but not knowing why; in assuming that it has a supreme ground—e.g. thinking of a mere idea, and a transcendental one at that, as having •an existing being corresponding to it—we’re doing this only so as to give ourselves a more *definite* notion of the principle’s universality. I can’t suppose that •this thing exists in itself, because it can’t be reached by any of the concepts through which I can have a definite thought about any object. That is because the idea itself slams the door on all the conditions that are required for any of my concepts to be objectively valid. The only way concepts of •reality, •substance, •causality, and even •necessary existence can have a meaning that lets them say something definite about an object is through their work in making empirical knowledge of an object possible. So they can be used to explain the possibility of *things in the world of sense*, but not to explain the possibility of *the world as a whole*. To explain *that* you’d need the item that did the explaining to be outside the world, meaning that it couldn’t be an object of a possible experience. Still, I can assume such an inconceivable being—the object of a mere idea, ·not of a concept of the understanding·—**relatively** to the world of sense though not in itself. ·I’ll explain why·. If

- the greatest possible empirical use of my reason rests on an idea. . . .that can’t itself be adequately exhibited in experience but is inescapably necessary if I’m to approximate to the highest possible degree of empirical unity,

then

- I’m not only entitled but compelled to realise this idea, i.e. to posit for it a real object.

[Here, as in some other contexts, ‘realise’ = ‘real-ise’ = ‘thing-ise’ = ‘treat

as standing for a thing'.] But I'm to posit this 'real object' only as a Something that I don't at all know in itself, positing it as a basis for that systematic unity, relating it to this 707 unity in a manner analogous to how things are related in the empirical domain by the concepts used by the understanding. Accordingly, by analogy with realities in the world—i.e. with substances, with causality and with necessity—I think a Being that possesses all this in the highest perfection; and because this idea depends merely on my reason—and isn't answerable to any factual constraints from experience.—I can think this Being as *self-subsistent Reason*, which through ideas of the greatest harmony and unity is the cause of the universe. [The phrase 'self-subsistent Reason' means something like 'Reason existing as a thing, not as a power or property of a thing'. The thought is of God as Reason rather than of God as having reason.] So I leave out all conditions that would limit the idea, because what I want it for is this:

To make possible, under the shelter of this ·thought of the· primordial Basis ·for the world as a whole·, the systematic unity of the manifold in the universe, and in that way providing for the greatest possible empirical use of reason.

I do this by representing all connections *as if* they were laid down by a supreme Reason of which our reason is merely a faint copy. I go on to think about this supreme Being solely through concepts that strictly apply only in the world of sense; but all I am using this transcendental assumption ·of a supreme Being· *for* is the **relating** task of providing the substratum, ·the ground, the basis·, for the greatest possible unity of experience; and that makes it all right for me to think of a Being that I put outside of the world of sense through properties that belong solely inside world. It's all right for me to do this because I'm not claiming to know this object of my idea according to what it may be in itself; and I had *better*

not be doing that, because I have no concepts for it; even 707 the concepts of reality, substance, causality—and indeed the concept of necessary existence—lose all significance and become empty concept-*labels* when I take them outside the domain of the senses. All I am doing is to give myself the thought of the **relation** of a completely unknown Being to the greatest possible systematic unity of the universe, wanting this Being solely in the role of a schema of the regulative principle of the greatest possible empirical use of my reason.

We can see at a glance that the transcendental object of our idea can't be thought of as having an intrinsic nature to which the concepts of reality, substance, causality etc. are applicable, because these concepts haven't the least bearing on anything that lies outside the world of sense. When reason supposes a supreme being as the highest cause, this is a merely **relative** supposition, devised solely for the sake of systematic unity in the world of sense—a mere ideal Something of whose intrinsic nature we have no conception. . . .

Now we can command a clear view of the upshot of the whole Transcendental Dialectic, and give a precise account of 708 what the ideas of pure reason are ultimately *for*—ideas that become dialectical only through careless misunderstandings. Actually, pure reason is busy only with itself—that's the only business it *can* do! [Kant now says it all again, through three paragraphs: Reason is concerned with bringing systematic unity to our scientific knowledge, not in carving out a branch of knowledge of its own. In doing this subjective work it has to be thought of as 'objective', but only in a vague *as-if*-ish way that doesn't transform its principles from regulative to constitutive.]

The first object of such an idea is the *I* itself, viewed simply as thinking nature or soul. If I want to know what the intrinsic properties are of a thinking being, I must put

the question to experience; the only categories I can apply to this object—i.e. to any thinking being in the world—are ones whose schema is given in sensible intuition; and I'll never get a systematic unity of all the appearances of inner sense in *that* way. What the soul actually *is* is captured by the empirical concept of the soul; but that won't take us far in our pursuit of systematic unity; so what reason does is this:

•It takes the concept of the *empirical* unity of all thought and, by thinking of this unity as unconditioned and basic, it makes out of the empirical concept a concept of reason (an *idea*) of a simple substance that is always the identically same thing through time, but is in ·changing· interactions with real things outside it—in short, the idea of a simple self-subsistent intelligence [See the 'self-subsistent Reason' note on the preceding page.]

711 What reason is *up to* in doing this is just to get principles of systematic unity in the explanation of the appearances of the soul—a way of seeing •all states as ·united in· a single subject, •all powers (so far as possible) as derived from a single basic power, •all change as alterations in the states of one and the same permanent being, and all appearances in space as completely different from the actions of thought. The simplicity etc. of the substance is intended to be only the schema of this regulative principle, and isn't being presupposed as being the actual basis for the properties of the soul. For these properties may have some altogether different source that we don't know about. Even if we allowed these predicates of simplicity etc. that we have taken on board to count as plainly valid for the soul in itself, the soul still couldn't be known through them, because they constitute a mere *idea* that can't be cashed in by concrete examples. Such a psychological idea can do nothing but good, *provided*

that we are careful to see it only as a mere idea, regarding it as valid only **relative** to the systematic use of reason in thinking about the appearances of our soul. ·By sticking to its status as an *idea*· we'll

- prevent any empirical laws of bodily appearance (which are of a totally different kind) from getting mixed into the explanation of what belongs exclusively to *inner sense*;
- keep out all windy hypotheses about the generation, extinction, and transmigration of souls;
- keep our thinking about this object of inner sense completely pure, not mixed in with properties that don't belong here;
- direct reason's investigations towards reducing the grounds of explanation on this topic to (as far as possible) a single source.

712

The best way—actually it's the *only way*—to achieve all this is by treating such a schema *as if* it were a real being. . . .

The second regulative idea of merely speculative reason is the concept of *the world as such*. [Kant's paragraph about this is notably ill-written and hard to understand. In it he says that nature—the world—is 'the only given object in regard to which reason needs regulative principles', and then he proceeds to explain what that need is. He remarks that nature is a two-fold affair, comprising **(1)** the world of thought and **(2)** the world of bodies. Kant has just finished with **(1)** under the label of 'the first regulative idea', and he doesn't mean to get back into that topic here; but he wants to distinguish *its* use of regulative principles from **(2)**'s. In the case of **(1)**, he says, the basic psychological concept (*I*) plays an *a priori* role in all our thinking; but in our everyday intellectual management of **(2)** the corporeal world—in our applying the categories to it, and so on—we don't need help from any *idea*, i.e. any 'representation that

transcends experience'. Don't need it and couldn't use it, because in dealing with **(2)** corporeal nature we're guided solely by sensible intuition. Kant continues:] So what's left
 713 for pure reason 'to work on in the territory of **(2)**' is just nature as such, and the completeness of the conditions in nature in accordance with some principle. (Obviously, this is *not* routine everyday thinking about parts and aspects of the world of bodies.) This does provide work for the idea of *the absolute totality of the series of these conditions. . . .*: we can't ever encounter such a totality in our empirical use of reason, but the idea works for us as a rule that prescribes how we ought to conduct ourselves when dealing with such series. The rule tells us that in explaining appearances by working back up the causal chain, earlier and earlier, we ought to

•treat the series *as if* it were in itself infinite, i.e. *as if* it went on indefinitely.

And it tells us that when in the context of practical principles we are regarding reason itself as the determining cause, this being an exercise of freedom, we ought to

•proceed *as if* we were dealing with an object not of the senses but of the pure under-standing, so that there are conditions *of* the series of appearances that themselves lie outside the series, which can therefore be regarded *as if* it had an absolute beginning through an intelligible cause.

All this shows that the cosmological ideas are nothing but regulative principles, and are far from positing—in the manner of constitutive principles—an actual totality of such series. This is all dealt with in more detail in the chapter on the antinomy of pure reason.

The third idea of pure reason, which contains a merely **relative** supposition of a Being that is the sole and all-sufficient cause of all cosmological series, is the idea of *God*. The

object of this idea is something that we haven't the slightest
 reason to assume outright (·as distinct from assuming it
 in a **relative way**·); for what makes it possible—let alone
 legitimate!—to believe in a Being of the highest perfection, 714
 existing necessarily by its very nature, merely on the basis
 of its concept? It's only in relation to the world that this
 supposition can be necessary; which clearly shows that
 the idea of such a being, like all speculative ideas, merely
 expresses reason's command that we look at all connection
 in the world. . . .*as if* it had its source in one single
 all-embracing Being, as the supreme and all-sufficient cause.
 So it's obvious that reason's only purpose here is to prescribe
 its own formal rule for •extending its empirical use, not for
 extending itself •beyond all limits of empirical use; so that
 this idea is not a disguised vehicle for some principle that
 tries to apply to possible experience in a constitutive way.

This highest formal unity, which rests solely on concepts
 of reason, is the *purposive* unity of things. Reason's speculative
 concerns require us to regard all order in the world
as if it had arisen from the purpose of a supreme Reason.
 When our reason is at work in in the field of experience, this 715
 principle •gives it entirely new prospects for connecting up
 the things of the world according to teleological laws, and
 through that •enables it to arrive at their greatest systematic
 unity. In this way the assumption of a supreme intelligence
 as the exclusive cause of the universe—though in the idea
 alone—can always benefit reason and can never harm it.
 ·Here's an example·. If in studying the shape of the earth
 or of the mountains or the oceans or the like, we view it
 as the outcome of the wise purposes of an Author of the
 world, this enables us to make a good many discoveries.²²

²² The advantage of the earth's ·approximately· spherical shape is well known. But not many people realize that its being ·more exactly· a slightly flattened sphere ·brings further advantages. Such as

716 Provided we restrict ourselves to a merely regulative use of this principle, even *error* that it leads us into can't do us any serious harm. The worst that can happen is to expect a teleological connection but find only a mechanical or physical one. In such a case, we merely fail to find the additional unity; we don't destroy the unity on which reason insists in its empirical use. And even a disappointment of this sort doesn't do any harm to the generally teleological approach. Suppose an anatomist assigns to some organ of an animal body an end that it can be clearly shown not to have—what of it? Well, he was wrong; it was an error; but it's perfectly impossible to prove in any given case that an arrangement of nature, whatever it is, has no end at all. So medical physiology isn't running any risks when it extends its very limited empirical knowledge of the functions of the parts of organisms by resorting to a principle handed to it by mere pure reason; and it carries this principle so far as to assume—confidently and with the approval of the experts—that everything in an animal has a function, a good purpose. If this assumption is taken as constitutive, it goes far beyond anything supported by observations that have so far been made; which shows that it's nothing but a regulative principle of reason that is meant to help us to get the greatest possible systematic unity by means of the idea of the purposive causality of the supreme cause of the world—as if this Being, as supreme intelligence with the wisest purposes, were the cause of all things.

717 what? Well, it's only because of the flattening that the earth's axis stays steady through all the movements of material caused by earthquakes. The bulge around the equator forms so vast a mountain that the impetus of all the other mountains thrown up by earthquakes can't have any observable effect in changing the position of the earth's axis. And yet, wise as this arrangement is, we unhesitatingly explain the shape of the earth non-teleologically, in terms of the equilibrium of the formerly fluid mass of the earth.

But if we deviate from this restriction of the idea to a merely regulative use, our reason will be thoroughly led astray. That's because it will be leaving the ground of experience, which is the only territory with legible route-markers, and venturing out into the realm of the incomprehensible and inscrutable, and up in those heights it is bound to become dizzy because it will have cut itself off from any experience-related way of steadying itself.

The first error that arises from misusing the idea of a supreme being by using it constitutively rather than merely regulatively is the error of **lazy reason**.²³ That's a fair label [Kant gives it also in Latin] for *any* principle that makes us regard all our investigations into nature as utterly complete, laying reason to rest as if it had entirely succeeded in its tasks. When the **psychological** idea is used as a constitutive principle to explain the appearances of our soul, and thereby to extend our knowledge of the soul beyond the limits of experience (i.e. to its state after death), it does indeed make things very comfortable for reason; but it wreaks havoc with our use of reason in dealing with nature under the guidance of our experiences. [In the next sentence a 'dogmatic spiritualist' is someone who believes that minds are non-corporeal substances or 'spirits', and regards this as straightforward doctrine that can be maintained without any 'critical' (opposite of 'dogmatic') concern with what is needed for such knowledge to be possible.] That's what happens when the dogmatic spiritualist explains the abiding and unchanging unity of a person throughout all changes

²³ This was the title that the ancient dialecticians gave to a sophistical line of thought which ran thus: 'If it is your fate to recover from this illness, you will recover, whether you employ a physician or not.' Cicero says that this was called 'lazy reason' because if we went along with it we would have no work for our reason to do in life; which is just why I give the title 'lazy reason' to the sophistical argument of pure reason.

of state in terms of the unity of the thinking substance that he thinks he directly encounters in the *I*; or •explains our concern with things that can't happen until after we are dead in terms of our •supposed• consciousness of the immaterial nature of the thinking subject; and thus •dispenses with all empirical investigation of the actual working causes of these inner appearances. . . . This kind of bad upshot is even more obvious in the dogmatic treatment of our idea of a supreme intelligence, and in the **theological** system of nature that is falsely based on it. The dogmatist in this field of enquiry

719 •fastens on all the examples of purpose that show up in nature (many of them involving 'purposes' that we invented so as to make our explanatory work easier), and, instead of looking for causes in the universal laws of the mechanism of matter,

•attributes all those purposes directly to the inscrutable decree of supreme wisdom.

and thus

•regards as *completed* the work of reason that hasn't even *begun*.

•Why do I say that this dogmatist hasn't been using reason here? Because• the use of reason has to be guided by the order of nature and the causal chains that occur in accordance with the universal laws of nature; •and this dogmatist has ignored all that in favour of sweeping theological 'explanations' of purposes in nature•. The way for us to avoid this •dogmatic• error is to bring the idea of a supreme purposeful intelligence to bear not merely on •certain *parts* of nature (the distribution and structure of dry land, the make-up and location of the mountains, the organisation of the vegetable and animal kingdoms) but on •the systematic unity of nature *as a whole*. For then •we'll be treating •all of• nature as resting on a law-governed purposiveness from which no special subsystem is exempt, though for many of

them it may be hard to us to discover what the purpose is; •we'll have a regulative principle of the systematic unity of teleological [= 'purposive'] connection. Without being able to say in advance what any of the •teleological connections are, we'll be able to wait for them to emerge from the work we'll be doing when we track down the •physico-mechanical connections in accordance with universal laws. That's the only way in which the principle of purposive unity can help us to extend the use of reason in reference to experience without ever doing any intellectual harm. 720

The second error arising from the misinterpretation of the principle of systematic teleological unity is that of **back-to-front reason** [Kant gives this also in Greek as well as Latin]. The procedure done •in the correct order goes like this:

•We use the idea of systematic unity as a regulative principle to guide us in seeking for such unity in the connection of things, according to universal laws of nature; and •how far we have come along the empirical path will be our measure of how near we are to completeness in our use of the idea (of course we'll never get the whole way there).

And this is what people do when they get it •back to front:

•They start by presupposing the reality of a principle of purposive unity, and they hypostatise it, •i.e. think of it as being some kind of *thing*•; but since they haven't the faintest conception of what a supreme intelligence (•the *thing* in question•) would be like in itself, they characterize it in an anthropomorphic manner; •crediting it with the sorts of purposes humans have•. That leads to their imposing ends on nature, forcibly and dictatorially, instead of pursuing the more reasonable course of searching for them by investigating what actually goes on in the world.

•This makes thing go wrong for teleological thinking, in two

ways. •Teleology was supposed to *widen the scope of* our unified explanations nature in accordance with universal laws; but the back-to-front approach *suppresses* such explanations. •And this approach prevents reason from doing what it set out to do, namely to prove from nature, in conformity with universal laws, the existence of a supreme intelligent Cause. [The rest of this paragraph expands what Kant wrote—not grossly, but in ways that can't easily be handled through the ·small dots· convention.] That proof is supposed to lead stepwise to something close to the supreme perfection of an Author of all things, who is supposed to be absolutely necessarily perfect, and therefore knowable *a priori* as perfect. But that conclusion can't be reached from a premise about purposiveness in nature unless the premise is an *a priori* statement of purposiveness in nature, meaning that purposiveness is part of the *essence* of nature. That need is met by the regulative principle about purposiveness, because it requires nature to have a systematic ·teleological· unity that is not merely empirically known but is presupposed *a priori*. . . .and consequently as following from the essence of things. But the back-to-front approach doesn't have such a premise. If I follow *it*, I'll think I have a constitutive thesis that nature is in fact purposive, and I'll hold that the source of this purposiveness is not nature's essence but the will of a supreme purposive Being; which means that I'll have to regard nature's teleological unity as contingent, ·as something *added* to nature from the outside·, and therefore as not knowable from its own universal laws. So I'll be reasoning in a vicious circle, assuming the very thing that is supposed to be proved.

The regulative principle of the systematic unity of nature serves, merely in idea, as the underlay of a consistent use of reason. If you take it as being constitutive, and as asserting the existence of a *thing* that causes this unity, all you do

is to confuse reason ·by running it backwards. When it is used the right way around·, the investigation of nature takes its own independent course, tracking the chain of natural causes in accordance with their universal laws. Admittedly it does this also in accordance with the idea of an Author of the universe, but not

•to see this Author as the source of the purposiveness that reason is constantly on the watch for,
but rather

•to obtain knowledge of the existence of such an Author from this purposiveness that reason looks for in the essence of the things of nature (and as far as possible in the essence of things as such), which will involve knowing the existence of this supreme being as absolutely necessary.

This right-way-around project may fail; but anyway, success or failure, it lets the idea remain always true in itself, and justified in its use, by restricting it to the conditions of a merely regulative principle.

Complete purposive unity is *perfection*. . . . If we don't find this perfection in the essence of the things that make up the entire object •of experience, i.e. •of all our objectively valid knowledge, and therefore don't find it in the universal and necessary laws of nature, how can we extract from it the idea of a primordial being who is supreme and absolutely necessary and the source of all causality? . . .

..723

In discussing the antinomy of pure reason I said that *all* the questions raised by pure reason must be answerable, and that we can't shrug them off by pleading the limits of our knowledge. With many questions arising in natural science that plea is as •unavoidable as it is •relevant; but ·not here (I said), because· our present questions aren't about the nature of things; rather, they arise from the very nature of reason, and concern solely its own inner constitution. I'm

now in a position to confirm this seemingly bold assertion in connection with the two questions that are of most concern to pure reason; and that will complete my discussion of the dialectic of pure reason. [The ‘two questions’ are **(a)** the cluster of questions about the significance of the transcendental *I* and **(b)** the cluster concerning the idea of God. As Kant explains in the next footnote, he will make his points only regarding **(b)** theology, leaving the reader to work out what the corresponding discussion of **(a)** psychology would look like. Thus:]

With regard to a transcendental theology,²⁴ three questions can arise, and they are all answerable.

724

(1) Is there anything distinct from the world that contains the ground of the world’s order and of its hanging together in accordance with universal laws?

Yes, certainly! For the world is a sum of appearances, i.e. of *all* appearances; so it must have a ground that is transcendental, i.e. thinkable only by the pure understanding.

(2) Is this being a substance, does it have the greatest reality, is it necessary (and so on)?

This question is entirely without significance. That’s because all the categories through which we might try to formulate answers can be used only empirically, and have no sense except when applied to objects of possible experience, i.e. to the world of sense. Outside this domain they are merely labels for concepts; we may allow them, but we can’t understand anything through them.

²⁴ Given what I have already said about the psychological idea and its proper status as a principle for the merely regulative employment of reason, I needn’t dwell at any length on the transcendental illusion through which •the systematic unity of all the manifoldness of inner sense •is hypostatized—i.e. through which •the way in which my inner states are unified by the fact that they can all be accompanied by ‘I think’ is •understood as showing that ‘I’ stands for a spiritual substance that possesses all those states. The procedure is very similar to the one involved in my critique of the theological ideal.

(3) Is it all right for us at least to think of this being, distinct as it is from the world, as an object on an *analogy* with the objects of experience?

Certainly, but not as an object •in reality. We may think of it as an object •in the idea, i.e. as an unknown substratum of the systematic unity, order, and purposiveness of the arrangement of the world—an idea that reason has to form as the regulative principle of its investigation of nature. And we can go further: we won’t get into trouble if we allow this idea to have certain touches of anthropomorphism that will help the principle to do its regulative work. For it will still be only an idea, which isn’t related directly to a being distinct from the world. It does relate directly to the regulative principle of the systematic unity of the world, but only by means of a schema of this unity—namely, a supreme Intelligence which acts wisely in originating the world. That tells us nothing about what this primordial ground of the unity of the world is in itself; all it does is to tell us how we should use our idea of this Being in **relation** to the systematic use of reason in respect of the things of the world.

But you may want to ask: ‘Can we, on those grounds, assume a wise and omnipotent Author of the world?’ There’s no doubt about it—we may and indeed we *must*. ‘Will that have us extending our knowledge beyond the field of possible experience?’ *No way!* All we’ll have done is to presuppose a Something, a merely transcendental object, of whose intrinsic nature we have no concept whatsoever. (We call it an *Intelligence*, but that’s an empirical concept and doesn’t strictly apply.) It’s only in relation to the systematic and purposive ordering of the world (which we have to presuppose if we are to study nature) that we have thought this unknown being, *by analogy* with an intelligence properly so-called. We have done this by noting the purpose and perfection that are to be based on it, and attributing to it just the properties

that our reason says can be regarded as containing the basis for such systematic unity, i.e. the basis for that purpose and perfection. So this idea is valid in respect of the *use* of our reason *in reference to the world*. If we credited it with being just plain absolutely and objectively valid, we would be forgetting that what we are thinking is only a *Being in our idea*. . . .

You will have another question: 'Can I make any such use of the concept and of the presupposition of a supreme Being in rationally considering the world?' Yes, that's just what reason has resorted to this idea *for*. 'Then is it all right for me to regard •seemingly purposive arrangements as •purposes, and so derive them from the divine will. . . .?' Yes, you can do that, as long as you regard it as a matter of indifference whether we say

•Divine Wisdom has arranged everything to conform to its supreme purposes

or rather

•The idea of supreme Wisdom has a regulative role in the investigation of nature; it's a principle of nature's systematic and purposive unity. . . .

That is, when you encounter such purposive unity you must see yourself as having a choice between

'That's what God in his wisdom has willed' and 'That's what nature in its wisdom has arranged',

and you must think that it doesn't matter in the slightest which you choose to say. For what entitled you to adopt the idea of a supreme Intelligence as a schema of a regulative principle in all your natural science was just precisely this greatest possible systematic and purposive unity. So the more purposiveness it guides you to find in the world, the more fully is the legitimacy of your idea confirmed. But the sole aim of that •regulative• principle was to guide our search for the necessary and greatest possible unity of nature; and

whenever we find such unity we'll owe that to our idea of a supreme being; but we mustn't

ignore the universal •non-teleological• laws of nature, and look on this purposiveness of nature as contingent and supernatural, •imposed on nature from outside the world by a divine Intelligence•.

If we do that, we'll be contradicting ourselves, because the theological idea was adopted in the first place as an aid to discovering the laws of nature. We are entitled to •assume that above nature there is a Being with those qualities •of wisdom, power, etc.• but only •to adopt the *idea* of such a being as an aid to viewing appearances as systematically connected with one another. •We don't think of the supreme Being as *causing* the orderliness of nature, but only as relating to it in a manner that is *analogous to* causation.

For the same reasons, in our thoughts about the world's cause we're entitled not only to

•represent it in our idea in terms of a certain subtle anthropomorphism (which we *have to* have if we're to think about it at all), namely as a Being that has understanding, feelings of pleasure and displeasure, and desires and volitions corresponding to these,

but also to

•ascribe to it an *infinite* perfection that goes far beyond any perfection that our empirical knowledge of the order of the world can justify us in attributing to it.

That's because the regulative law of systematic unity tells us to study nature *as if* systematic and purposive unity along with the greatest possible complex variety were to be met with everywhere, *in infinitum*. We won't succeed in actually finding much of this world-perfection, but our reason lays down the law that we must go on looking for it and expecting to find it; and it must always be beneficial and never harmful for us to direct our investigations into nature in accordance

728

727

729

with this principle. But it's obvious that in this way of representing the basic idea of a supreme Author, I'm not basing anything on the existence and knowledge of such a Being, but only on the idea of it; and that I don't really derive anything •from this Being, but only •from the idea of it—i.e. •from the nature of the things of the world in accordance with such an idea. A certain undeveloped consciousness of the true use of this concept of reason seems indeed to have inspired the modest and reasonable language of the philosophers of all times, when they have spoken of 'the wisdom and providence of nature' and of 'divine wisdom' as though these were equivalent expressions. Indeed, while they have been dealing solely with speculative reason, they have given preference to 'the wisdom etc. of nature', because it •lets us stop short of saying something stronger than we are justified in saying, and •directs our reason to its own proper domain, namely nature.

And so it is that pure reason, which at first seemed to promise nothing less than to extend our knowledge beyond all limits of experience, contains (when properly understood) nothing but regulative principles. Admittedly, the degree of unity that these principles tell us to aim for is greater than any that could be reached through the empirical use of the understanding; yet just because they have placed the goal
730 so far away they give to the understanding a high degree of internal consistency through systematic unity. But if they are misunderstood, and treated as constitutive principles of transcendent knowledge, they give rise through a dazzling and deceptive illusion to imaginary 'knowledge' leading to contradictions and never-ending disputes.

• • •

Thus all human knowledge starts with •intuitions, goes from them to •concepts, and ends with •ideas. Although

in respect of all three elements we have *a priori* sources of knowledge that seem at first blush to scorn the limits of all experience, a thoroughgoing critique convinces us that •reason in its speculative use can't ever get beyond the domain of possible experience with any of these three elements, and that •the proper role of this supreme faculty of knowledge is to use all methods, and the principles behind them, solely for the purpose of penetrating to the innermost secrets of nature by tracking every possible sort of unity—with purposive unity being the most important of them—but never to soar beyond nature's limits, out where for us there is nothing but empty space. •Strictly speaking, the Analytic sufficed to show this, without bringing in the Dialectic. The Transcendental Analytic's critique of all •propositions that
731 can extend our knowledge beyond *actual experience* shows well enough that •they can never lead to anything more than a *possible experience*. If people weren't so suspicious of even the clearest abstract and general doctrines, and if plausible and alluring prospects didn't tempt them to resist the force of those doctrines, we could have spared ourselves the laborious interrogation of all those dialectical witnesses that a transcendent reason brings forward in support of its inflated claims; because then it would have been known with complete certainty right from the start that all such claims, even if honestly meant, must be utterly empty because they relate to a kind of knowledge that men can't ever have. As things stand, however, the *talk* will go on and on unless and until people get through to the true cause of the illusion by which even the wisest are deceived. Also, analysing all our transcendent knowledge into its elements is a worthwhile contribution to the study of our inner nature, as well as being something the philosopher is obliged to do. So we *had to* track all these attempts of speculative reason, fruitless as they are, back to their sources. And because dialectical

732 illusion doesn't merely deceive us in our judgments, but also—because of how some of these judgments connect with our interests—the illusion *attracts* us and will always do so. That's why I thought it advisable, with a view to heading off such errors in the future, to draw up in full detail the court transcript of the trial, and to deposit it in the archives of human reason.

[That concludes the Transcendental Doctrine of Elements. The remaining one-sixth of the *Critique of Pure Reason*—namely the Transcendental Doctrine of Method—will not be offered on the website from which the present text came.]