

## CHAPTER III

### FREEDOM IN KANT'S REVOLUTION

#### 1 Knowledge

- 1.1 Kant's revolution.
- 1.2 Forms of sensibility and categories of thought
- 1.3 The knowing and deciding self
- 1.4 The scope of understanding and the limits of reason
- 1.5 The antinomy of pure reason
- 1.6 Religion: denying knowledge to make room for faith

#### 2 Freedom

- 2.1 The causality of freedom: autonomy
- 2.2 Spontaneity and receptivity
- 2.3 Empirical and intelligible
- 2.4 Negative and positive freedom: autonomy
- 2.5 Free thought and the epistemology of reason

#### 3 Morality and religion

- 3.1 Morality and practical reason
- 3.2 The Categorical Imperative: (i) Universal law
- 3.3 (ii) Rational nature as an end in itself
- 3.4 (iii) Autonomy and the kingdom of ends: self-legislation and impartiality
- 3.5 The honour code of humanity
- 3.6 The postulates of freedom, immortality and God
- 3.7 Christianity as a moral faith
- 3.8 Politics and freedom

#### 4 Kant: reflection and assessment

- 4.1 Idealism and Critical philosophy
- 4.2 Two heroic illusions about freedom
- 4.3 The epistemology of morality
- 4.4 Kantian ethics without the illusions
- 4.5 Beyond the enlightenment

We now turn to the single-handed revolution brought about by Kant. Over the long term its impact on the philosophical and ethical culture of the West has been even greater than that of the revolution in France has been on its politics. The grounding idea was a new philosophy, which Kant famously described as “transcendental idealism.” Only this philosophy, he held, could provide a firm basis for knowledge and freedom.

Kant's revolution bears on just about everything that will be discussed in this book. It is also relevant to many things beyond its scope, so the account in this chapter will not be a

comprehensive treatment. It will focus on those Kantian themes that are especially relevant to late modern ethics, themes we shall find ourselves returning to in later chapters.<sup>1</sup>

Seen in this perspective, Kant's analysis of freedom as autonomy becomes the central element. It is fundamental to Kantian ethics, to the famous claims about categorical and universal reason, about rational beings as absolute ends, and to the doctrine of respect for every human being as the ground of liberal social and political order. Adapting terms from John Stuart Mill (0.0), one could say that autonomy, in the late modern period, came to be the a priori German argument for liberalism, as happiness was the a posteriori English one.

Apart from the idea of autonomy there is another reason for Kant's importance in late modern ethics. The transcendental idealist setting in which he places his account of freedom is highly relevant to what I called, in the Introduction, the Problem of Being. How does self relate to Being? In what way if any is Being fundamentally spiritual? One way to pose this question, though not the only one, is in terms of our relation to God, where Kant famously denies religious knowledge to make room for faith (1.6, 4.5). His account has been highly influential, yet, to many, unsatisfying. No one has taken it in precisely its own terms. To some it has seemed to vindicate an existential ground of intimate, personal religious commitment; to others it alienates us from the divine, as indeed more generally transcendental idealism alienates us from Being. Yet others have sought to exclude these questions from their updates of Kant altogether. But in the late-modern period this story of alienation was crucial. It was the reverse side of the story of autonomy.

The historical fortunes of Kantian ethical theory have been striking. It came under attack from a very wide variety of angles in its own time and throughout the nineteenth century, for example from Schiller, Hegel, Schopenhauer, Mill and Nietzsche. Some of the nineteenth century criticisms – that it was empty formalistic, and/or psychologically naïve – overlapped with similar reactions to the theorists of the French Revolution discussed in the previous chapter. Other critics, notably Nietzsche, saw it differently, as an attempt to keep Christianity going minus God. Of course to criticise Kant is not to ignore him; rather, much ethical thought in the nineteenth century defined its own positions in reaction to him, or at least in dissenting awareness of him. In the earlier twentieth century, by contrast, Kantian ethics was widely ignored. The idea that morality could be derived from Kant's 'Categorical Imperative' seemed dead. Yet, towards the end of the late-modern period there were important revivals of Kantian ethics. There is no question that 'for,' 'against,' and 'back to' Kantian ethics is an integral part of our story. We shall assess this historical pattern at the end of the book.

Transcendental idealism arises in the first place as an account of the possibility of empirical knowledge. Kant then sees it as also indispensable to the possibility of freedom, and in turn to the possibility of morality and religious faith. We shall trace this argument in the following three sections, starting with knowledge (section 1), moving on to freedom (section 2), and then to morality and religion (section 3). The final section (4) attempts some provisional stock-taking.

---

<sup>1</sup> This also means that our focus will be on his philosophical thinking from the *Critique of Pure Reason* onwards.

## 1 Knowledge

Key works by Kant appeared from the beginning of the 1780s into the 1790s. Undoubtedly, however, the basic text of the Kantian revolution is *The Critique of Pure Reason* ('the first Critique'), first published in 1781.<sup>2</sup> It is here that transcendental idealism is fully laid out.

### 1.1 *Kant's revolution*

The first Critique sets out to establish the possibility and the limits of our knowledge. Possibility and limit were, Kant thought, two sides of the same coin: the same analysis shows how knowledge of the world is possible and what limits on such knowledge there are. Vindicating common cognition and undermining metaphysics turn out to be the same task. The limits are limits on Pure Reason's metaphysical itches; explaining and assuaging them is the task he assigns to "Critique" or "Critical philosophy."<sup>3</sup> Alas, 'critique' and 'critiquing' have become hackneyed words. They hardly hint at the extraordinary philosophical work Kant found himself embarked on in trying to carry out the task.

"Common cognition" is one of Kant's expressions for the way we think when we think methodically, with care, but in a certain sense pre-philosophically. It refers to *good ordinary thinking*: the common-sense observation and methodical science that continually gives us a shared knowledge of our world.<sup>4</sup> When we philosophise, we initially accept some things that common cognition takes for granted. The world consists of facts that are independent of our thought; we are a proper part of that world; it goes on in us and about us; it does not depend on us. Common cognition takes for granted the world's *cognition-independence*. Yet, soon the question arises of how ordinary observation and scientific hypothesis can give us such knowledge. If the world is independent of our cognition how can we know that cognition correctly represents the world, or how even make sense of 'correct representation'? Philosophy is launched on a classic set of problems.

This is where Kant draws a distinction that he takes to be basic to any Critical philosophy. On the one hand, he says, there is a straightforward sense in which the notion that reality is independent of our cognition is both indisputably correct and compatible with knowledge. For example, whether water boils at such and such temperature at such and such pressure does not causally depend on us. It does what it does irrespective of what we think or do. Those facts about water are *there*; we can only take account of them, and we had better do so. Common cognition is realistic in this straightforward way: it is, as Kant puts it, "empirically" realist. On the other hand, however, as soon as we philosophise an extremely tempting *metaphysical* reading of this notion of cognition-independence obtrudes. Kant calls it "transcendental" realism, contrasting it with empirical realism. If true, it would have drastic consequences: it would undermine the ordinary thinking from which it itself naturally arises. The task of the Critical philosopher, Kant accordingly holds, is to vindicate ordinary thinking

---

<sup>2</sup> There were important revisions in the second edition of 1787. The *Groundwork of the Metaphysics of Morals* appeared in 1785. Then there were the two Critiques respectively of *Practical Reason* (1788, 'the second Critique') and of *Judgement* (1790, 'the third Critique'); *Religion Within the Bounds of Reason Alone* (1793); and *The Metaphysics of Morals* (1797).

<sup>3</sup> See *Critique of Pure Reason*, A xii ff. (References to Kant will follow the convention of citing A/B editions for the *Critique of Pure Reason*, and the volume/page number of the *Academie* edition, together with an abbreviated title, for other works. See the Bibliography for more information.)

<sup>4</sup> Related terms are "common sense", "natural consciousness," ...

by a critique of the transcendental realism to which ordinary thinking itself tempts us when we philosophise – to expose it as an illusion about the role of reason. Hence ‘The Critique of Pure Reason.’

Critical philosophy diagnoses the illusion and vindicates our conviction that we are free knowers and free agents. How? It is here that Kant’s revolutionary framework – “transcendental idealism” as against transcendental realism – comes into play:

Up to now it has been assumed that all our cognition must conform to the objects; but all attempts to find out something about them *a priori* through concepts that would extend our cognition have, on this presupposition, come to nothing. Hence let us try whether we do not get farther with the problems of metaphysics by assuming that the objects must conform to our cognition ....

If intuition [i.e. cognition of objects] has to conform to the constitution of objects, then I do not see how we can know anything of them *a priori*; but if the object (as an object of the senses) conforms to the constitution of our faculty of intuition, then I can very well represent this possibility to myself.<sup>5</sup>

The revolutionary idea, then, is this: the very constitution of objects is structured by our ways of receiving and processing information about them. What leads us to this conclusion? A variety of things, but the main and most direct line of argument is that knowledge of objects is possible only if we *can* know “something about them *a priori*,” i.e. independently of empirical evidence. Here Kant is not concerned with a class of *a priori* propositions which he terms “analytic:” definitional truths which make no substantial claim about how things are and thus “do not extend our cognition.” His point is not about them. It is much more controversial: if knowledge is possible then some propositions that *do* make substantial claims about how things are – ‘synthetic’ propositions – must be true *a priori*.

It is, Kant holds, only the doctrine that “all our cognition must conform to the objects” that makes synthetic *a priori* truths, and hence knowledge of nature as such, impossible. This is the doctrine he calls transcendental realism. It leads us to empiricism, the view that any possible substantive knowledge of the world is *a posteriori*. When this is thought through, however, it leads us on further, to the conclusion that knowledge is impossible. So on the one hand, Kant agrees with those who argue that empiricism is just a staging post to total scepticism. But on the other, he disagrees with those who claim that there is some metaphysical route to knowledge that transcends ordinary empirical knowledge.

The point of Kant’s terminology, then, is that he seeks to defend empirical realism – not the philosophical doctrine of empiricism but rather the simple empirical realism of the ordinary, common-sense point of view – by rejecting *transcendental* realism. The latter, he says, is what makes the ordinary knowledge of the world we take for granted seem to be impossible. From this emerges Kant’s own doctrine, *transcendental* idealism. Transcendental idealism is the way Kant explains how – compatibly with empirical realism – objects “conform” to our cognition.

### 1.2 *Forms of sensibility and categories of thought*

He follows two lines of thought. The first concerns the structure of our sensibility – the way we experience the world through our senses (including memory and “inner sense,” i.e. experience of our sensations and feelings). Here ‘the world’ refers to the world of things as they really are, the world of “noumena,” as he calls them, as against “phenomena” – things as

---

<sup>5</sup> *Critique of Pure Reason*, Bxvi-xvii. It is also here that Kant famously compares this new approach to the Copernican revolution.

they are as against things as they appear to us. Sensibility receives information from the noumenal world. It transforms that input into representations – “intuitions” – by structuring it into a certain form: a spatial and temporal form in the case of outer experience, a temporal form only in the case of inner experience. Space and time are these “forms of intuition,” and hence a priori features of our experience. This, Kant argues, is why arithmetic and geometry are both synthetic and a priori: arithmetic because it expresses the form of temporal intuition, geometry because it expresses that of spatial intuition.

Doesn't this picture simply enforce the point that we cannot have knowledge of the world as it really is? All we can know about the world is how it appears to us. True. But at this point Kant makes a crucial move: it is to treat the *world-as-it-appears-to-us*, the phenomenal world, not as subjective experience but as a genuine shared *world* which is independent of any particular subjective experience, and in which we ourselves appear.<sup>6</sup> We collectively construct the conceptual structure that constitutes the phenomenal world – the empirical world of common observation and science. The phenomenal world is the world of nature; natural science is the science of the world-as-it-appears-to-us. Science reaches conclusions about nature which go beyond our direct sensible experience, nonetheless these remain conclusions about the world-as-it-appears-to-us, the empirical, or phenomenal, world. As Kant puts it in a representative passage:

everything is real which stands in connection with a perception in accordance with the laws of empirical advance.<sup>7</sup>

Here he means “empirically real:” nature is the empirically real world. It is of nature so understood – the theoretical extension and filling out of empirical experience – that we have some a priori knowledge, based on the forms of our intuition. “Empirical realism” is the view that the objects common-sense-cum-science deals with are just as real as our own sensory experience; equally that this subjective experience, given in inner sense, is just as much a part of empirical reality, just as open to scientific investigation, as the physical objects given by outer sense.

Our knowledge, then, is of reality in its phenomenal aspect, as it appears to us, rather than its noumenal aspect, as it absolutely is. Transcendental idealism holds that the phenomenal world, the world of nature as we know it, is a joint product. There is the input from things in themselves, as they absolutely are, the noumena, and there are the forms of our sensible intuition, which give spatial and temporal form to that input. Space and time, as Kant puts it, are empirically real, but transcendently ideal.

This is not an easy picture to grasp, let alone accept. It involves a double explanation of our experience, one at the empirical level and one at the transcendental level. There is the empirical, scientific account of our experience, which sees us and our sensory modalities as part of nature, as phenomena among the phenomena. And there is the transcendental account, in which ‘we’ as knowing subjects stand in a receptive relation to the noumena. At that level, it seems, the knowing subject is noumenal, interacting with noumena; space and time are grasped as ideal, not real, i.e. not a property of things as they really are. At that noumenal level there is transcendental receptivity to noumenal objects, while at the empirical level of human beings in the natural world there is empirical receptivity to empirical objects.

Surely an attitude more consonant with common cognition is simply that nature just *is* reality – with no distinction made between ‘empirical’ and ‘noumenal’ reality – that the

---

<sup>6</sup> More on this in 4.1. The reading of Kant here I give here is what is there referred to as the ‘objective’ reading.

<sup>7</sup> B521



natural facts are all the facts there are, in a unitary and unambiguous sense of ‘the facts there are’? Kant’s case is that this naturally tempting account undermines itself, because it cannot deliver the a priori framework that is necessary for knowledge. To put it his way, *transcendental* realism (the empirically real IS the Real) brings about the collapse of empirical realism. It is a sense of the very real force of this argument that has made Kant’s rejection of transcendental realism so influential.

Kant’s has a further line of thought. Here he argues from the conceptual structure of our understanding: the “categories” of our understanding as against the “forms” of our sensibility.<sup>8</sup> What he says is notoriously difficult, involving a battery of dense, hard-to-follow arguments. But it has always seemed to be of great importance.

A basic idea is that I can think of my experience as *my* experience only if I can think of it as experience of something *other* than me. Awareness of self requires awareness of something other than self. And in fact experience does not come as a mere succession of brute sensations, nor could it do so to a subject that makes judgements. It comes as a coherent set of *judgeable experiential contents*, a holistic totality of *seemings-to-me that so-and-so*, presenting an apparently unified spatio-temporal world of which I am a part. But, Kant argues, if experience is inherently judgeable content, it falls under a certain system of a priori concepts under which any judgeable content necessarily falls. These concepts are the categories of the understanding.<sup>9</sup>

We can distinguish a thinking, experiencing and acting subject on the one hand, and an objective world on the other, if and only if we think our experience under the categories. From one of these, the category of conditionality, Kant particularises to the principle of causation, that all events have a cause. He argues that the latter principle applies a priori to any possible object of experience, that is, to the empirical world as such.

Even this very truncated summary of Kant’s analysis is enough however to raise a basic question. Grant, for the sake of argument, that Kant has shown that self-awareness presupposes the ability to frame judgements whose content is about a purported empirical world. How does this show that those judgements are true: that there really is an empirical world? We need that stronger conclusion if we are to argue a priori that the categories apply to it.

This question has been raised by many critics. Kant’s reply can only be that the distinction to which the critic appeals – between what must be true of our *judgements* about the world, and what must be true of the *world* – is misguided. The reply implies a ‘constructive’ view’ of the empirical world, that is, of the world of nature. Transcendental idealism grants, indeed insists, that we cannot know things in themselves. Nature, however, is not the noumenal but the phenomenal world, and that world is constructed from noumenal inputs to our sensibility by the operations of our own mind – the forms of our sensibility and the categories of our understanding:

we ourselves bring into the appearances that order and regularity in them that we call *nature*, and moreover we would not be able to find it there if we, or the nature of our mind, had not originally put it there.<sup>10</sup>

Of course Kant does not deny that the order and regularity that our mind “puts” there must be filled in by empirical perception and scientific inquiry. Nonetheless on this constructivist

<sup>8</sup> The first line of argument is in the section of the *Critique of Pure Reason* called the ‘Transcendental Aesthetic’, the second is in the ‘Transcendental Analytic.’

<sup>9</sup> They form a distinctive conceptual structure which Kant sets out in a table at A80/B106.

<sup>10</sup> A125. The passage is omitted in B edition but I take it to express his settled view.

view of nature the question whether the conceptual structure supplied by the mind, within which we necessarily think of nature, fits the way nature actually *is* simply does not arise. Other implications of Kant's constructivism about nature will emerge below.

At this point we should distinguish the ideas of Critical philosophy in general, from those of transcendental idealism in particular. The generic ideas of Critical philosophy are that common cognition is perfectly in order; that what it needs is elucidation and vindication rather than proof; that what that in turn requires is an explanation of how synthetic a priori knowledge is possible; and that the explanation is blocked by a natural tendency towards a philosophically misconceived realism. Transcendental idealism is Kant's way of delivering these Critical ideas. One reason, however, why the first Critique is such a watershed is that many have thought the Critical idea can survive, even if Kant's way of implementing it fails. Kant answers the question of how synthetic a priori knowledge is possible in a particular way. His way is full of insights; yet the soundness of the basic Critical idea may not depend on the soundness of Kant's 'transcendental' answer to the question. Certainly this should be born in mind. But for the moment we are concerned with Kant's version, as we shall later be concerned with the way Kant's contemporaries understood it and reacted to it. We shall come back to a review and assessment of Critical philosophy in 4.1.

### 1.3 *The knowing and deciding self*

When it comes to the nature of the self, transcendental idealism faces deep questions. Since the knowing subject is receptive to the noumena, it must itself be noumenal. This interaction, according to the account, is a transcendental precondition of the possibility of empirical knowledge – in particular of *my* empirical knowledge. So it seems that *qua* knowing subject 'I' am noumenal. But at the same time 'I' am a part of the empirical world, along with other human beings. Here 'I' am a phenomenon, an empirical object open to scientific study.

It is not hard to see problems. What is the relationship between the empirical me and the noumenal me? Is the empirical me an 'appearance' of the noumenal me? Should we see the experience of each empirical self as the appearance of a distinct noumenal self?

Given Kant's account of the synthetic a priori status of arithmetic, it is unclear that concepts of numerical identity and non-identity or distinctness can even apply at the noumenal level. Certainly we cannot know whether they apply, since we cannot know the noumenal as against the phenomenal properties of reality. It would seem inadmissible then to ask whether there is a plurality of knowing subjects or just one, in the empirical sense in which there is one computer on my desk and two cups, etc. Perhaps it would be better to abstract to an ideal 'mind' or 'knowing subject.'<sup>11</sup> But if we abstract to an ideal knowing subject then it would also seem inadmissible to ask whether *subject* and *object* are numerically distinct or identical, as against analytically distinguishable. At this point a path to the absolute idealism of some of Kant's successors comes into view.

However Kant does not take that path. Rather, and in particular in his ethical and religious philosophy, he regularly assumes that each empirical self is the appearance of a distinct noumenal self. Further, he distinguishes a person's thinking and 'willing' (choosing, deciding) from their sensibility, their emotions and their inclinations. This distinction aligns, though only roughly and misleadingly, to that between what 'I do' and what 'happens to me'. *I* reason, come to conclusions, make decisions; in contrast, sensations *occur in* me, inclinations *arise in* me, emotions *are felt* by me. It also aligns, equally roughly and

---

<sup>11</sup> More on this in 4.1

misleadingly, with the distinction between states of me or acts of mine which can be assessed as reasonable or unreasonable, as against states of me which cannot be so assessed, because they are not in any way responses to reasons. Kant interprets both these distinctions in terms of transcendental idealism. He takes them to be expressions of the distinction between persons as they are and as they appear to themselves. What I truly do, what can be ‘imputed’ to me, is just my thinking and willing. The *real* me, as one is tempted to put it, is the thinking and willing me. My inclinations, emotions, and sensible experience are appearances. Persons – rational beings – have both noumenal and phenomenal aspects, and this is how Kant thinks of them in ethical and religious contexts.

To think of noumenal persons in the plural raises the problem just considered. Further, if transcendental idealism is right to claim that our knowledge must be entirely restricted to the phenomena, and if thinking and willing is noumenal, it seems to follow that we cannot know what we are thinking or willing. Here, however, Kant has a reply. While I cannot know something *other* than me except by its appearance to me, or by inference from some other appearances to me, *my own* thoughts and decisions are not something other than me. They are my own acts. I do not know them by some form of receptivity; *I think, I decide*. The self-awareness inherent in thinking and deciding cannot be thought of on the binary model of a subject knowing an object. This in itself, Kant could say, shows that the thinking and acting self cannot be merely an empirical object in the empirical world. Kant draws telling implications from this point, about the self as a centre of unified experience; however on the arguments he gives the right question to ask may be not whether the self is a *noumenal* object, but whether it is an *object* at all.<sup>12</sup> This is another case where people have been very impressed by Kantian arguments without accepting his framework of transcendental idealism.

#### 1.4 *The scope of understanding (Verstand) and the limits of reason (Vernunft)*

By ‘metaphysics’ (in that sense in which he aims to criticise metaphysics) Kant means the attempt to achieve knowledge of things as they really are through purely a priori reasoning. Transcendental idealism rules out any such knowledge. One might then ask how we know the truth of transcendental idealism itself. Kant’s reply is that the way we know it is by showing it to be presupposed by something that is *not* itself a priori: the simple fact that we have knowledge of the world. Transcendental idealism is thus not itself metaphysics in the specific sense in which it rules metaphysics out: it does not rule itself out. Rather, there is a ‘transcendental argument’ to it from the fact of empirical knowledge, which shows first, that such knowledge is possible only if there is synthetic a priori knowledge, and then second, that synthetic a priori knowledge of an empirical world is possible only in virtue of the forms of sensible experience and the categories of the understanding that go into the ‘construction’ of that world.

But how (a recalcitrant metaphysician might ask) has Kant positively shown that pure reason cannot establish absolute truths – truths about how things are absolutely, in themselves, and not merely relatively to us? He has, perhaps, shown that common cognition neither demonstrates nor assumes such absolute truths. But why should that preclude a *metaphysical* cognition of the absolute? Kant’s deepest assumptions become clear here. The first is that the mind cannot know existent entities that are distinct from itself, that are not its own products, except through some form of receptivity, which for us takes the form of sensible experience. On this Kant agrees with empiricists like Hume. The second is that pure

---

<sup>12</sup> Paralogisms, unity of apperception.



reason is not a form of receptivity.<sup>13</sup> It is a form of *spontaneity*. Reasoning is something we *do*; thought recognises its principles by self-reflection. Pure thinking is not a receptive device through which we receive information from something other than ourselves. By the same token, it cannot achieve knowledge of independent existents solely by its own powers. The contrast between receptivity and spontaneity is basic in Kant's thought. We shall see that it is crucial in much of what follows.

Kant has other striking arguments. In the section of the first Critique called the Transcendental Dialectic he tries to show that the attempt to achieve knowledge of absolute truths by pure reasoning, when it does not collapse into confusion, leads into contradiction. This is a very strong claim, stronger in fact than he needs for his critique of metaphysics. But it greatly influenced the subsequent development of views on religion, faith and human freedom.

Kant distinguishes in the Dialectic between the Understanding (*Verstand*) and Reason (*Vernunft*). They are not separate: the Understanding is just Reason in a particular, legitimate and indispensable role, the role it plays when it applies categories to experience – deploying concepts, reasoning deductively and inductively, constructing theoretical posits, all within the a priori framework of the categories. The Understanding is Reason at work in common cognition.

Understanding is not Reason's only legitimate role; as well as this theoretical role there is also a crucial practical role, which we shall come to. However Reason seeks to go beyond both of these: or rather we do, in our employment of reason. This arises from what Kant calls a "transcendental illusion." The illusion consists in a very natural and tempting misconception about the scope of two rules of reasoning that play a legitimate role in inquiry. We misconceive their scope, Kant thinks, because of our natural proneness to transcendental realism.

One of these principles tells us that whatever is or occurs does so in virtue of *conditions* – conditions on which its existence or occurrence depends and which therefore explain it. Call this the principle of Conditionality. The other principle says that whatever is conditioned (dependent on something else) must ultimately depend on the existence of something unconditioned, non-dependent, in other words, something *absolute*. Call this the Absolute Ground principle. (These are my names, not Kant's.)

*Prima facie* the two principles conflict – at least if any explanation of the existence or occurrence of something, A, must be in terms of the existence or occurrence of something *else*, B. Yet it looks as though both are inherent in our ideal of explanation. On the one hand we want a *final* explanation, on the other hand anything produced as final seems to require further explanation. Thus, for example, if the final explanation is couched in terms of fundamental laws of physics, we ask why do these equations have the form they have? Why do their parameters have the values they have?

Kant seems to have more than one diagnosis of these two principles, inviting more than one reading of his overall view.

One diagnosis proposes that both principles should be seen as "regulative" rather than "constitutive." At first sight this means that they should be seen as rules for the guidance of our construction of theories about the world *rather* than as constituents of the construction. On this diagnosis

---

<sup>13</sup> In a number of places Kant seems to envisage the possibility of a purely "intellectual intuition." (For discussion see Guyer, in Ameriks 2000 ed., pp 49 – 54.) This notion would become a focus of attention for later idealists, as we shall see in 3. 00.

both principles can very well coexist with one another, as merely heuristic and *regulative* ... For the one says that you should philosophise about nature *as if* there were a necessarily first ground for everything belonging to existence – solely in order to bring systematic unity into your cognition by inquiring after such an idea, namely an imagined first ground; but the other warns you not to regard any single determination dealing with the existence of such things as such a first ground, i.e., as absolutely necessary, but always to hold the way open to further derivation and hence always to treat it as still conditioned by something else.<sup>14</sup>

Such a reading of the two principles would be acceptable to a good empiricist; the “transcendental illusion” would then consist in mistaking what is merely a heuristic rule for an a priori truth about the world. This diagnosis requires no transcendental idealism; it also fits quite a lot of what Kant says about the “regulative use of the ideas of pure reason.”<sup>15</sup>

However a somewhat different line of thought fits much of what Kant says better. For he has earlier argued that the Conditionality principle applies within the phenomenal world as a synthetic a priori condition of objective empirical knowledge.<sup>16</sup> It can thus be thought of in two ways. If we think of it as a rule of construction, it says something like ‘Whenever you encounter an unexplained phenomenon you should always construct the best possible explanation of it.’ You should do so in accordance with your best theories; if your theories are improved by better information, you reconstruct your explanation in the light of the new theories. But now since the empirical world *is* a construction, this rule or norm can also be seen as a necessary truth. It then says: every phenomenon has a phenomenal cause. So it is, after all, in that way *constitutive* of the empirical world: it is the basic synthetic a priori principle deployed in constructing a phenomenal world indefinitely extended in space and time, and then filled out in accordance with empirical theories of physics. Kant is able to treat it both as a truth empirically speaking, and as a rule of construction normatively speaking; and he could give this same account of his conception of arithmetic and geometry. Or rather, that could have been the result if he had pursued this particular line of thought throughout.<sup>17</sup>

Both of these lines of thought could be drawn out of Kant’s thinking; neither invokes the distinction between phenomena and noumena. What then of the transcendental idealism which, according to Kant’s argument, is the precondition of empirical knowledge?

Transcendental idealism is not purely constructive, nor does it treat the Absolute Ground principle as either a heuristic rule or a norm of construction. True, it is constructive about the *phenomenal* world, but it takes that construction to work on given materials – materials provided by things as they really are. The phenomenal world does not float freely in mid-air: it is necessarily grounded in the noumenal. Noumenal reality is the absolute, unconditioned, ground of dependent, conditioned appearance. This argument to the noumena assumes that there must be some absolute ground of the phenomena outside the phenomena. “Pure Reason always has its dialectic, whether it is considered in its speculative or its practical use; for it requires the absolute totality of conditions for a given conditioned, and this can be found only in things in themselves.”<sup>18</sup>

---

<sup>14</sup> A617/B645

<sup>15</sup> See A642/B670 ff.

<sup>16</sup> See especially the Second Analogy, B233 – A211/B256

<sup>17</sup> Rather than the analysis of the mathematical a priori presented in the Aesthetic, where he bases it on the spatial and temporal form of intuition.

<sup>18</sup> 5:107. Couldn’t the noumenal world itself satisfy the Conditionality rather than the Absolute principle? If we assumed that noumena fall under Conditionality there would be no

Beyond the *necessity* of a noumenal ground, however, theoretical Reason can deduce nothing about its *constitution*. This leaves a possibility that Kant firmly grasps, namely, that the needs of *practical* reason might give us some sort of basis for believing or accepting that the noumenal absolute is God. They might give a like basis for accepting something about the noumenal self. It is important to note that such a line of thought would be precluded both by transcendental realism and by *pure* constructivism, since these both reject the distinction between phenomena and noumena. Evidence that this is Kant's overall line of thought is provided by his treatment of the antinomy of pure reason, and we shall see it fully at work in his treatment of freedom, morality and religion.

### 1.5 *The antinomy of pure reason*

It is transcendental realism, Kant thinks, that drives us into metaphysics and leads us into contradictions. His account of these contradictions occurs in a section of the Transcendental Dialectic called "The Antinomy of Pure Reason." Here he discusses four specific conflicts, each of which is cast in the form of a thesis and a contradictory antithesis, where reason seems to have a sound argument for each.

The first two conflicts involve the notion of infinity. In the first, the world is infinitely extended in space and time *and* it is not; in the second it is infinitely divisible *and* it is not (consisting instead of ultimate indivisible parts). Kant's overall idea is that we find ourselves in these contradictions because, on the one hand, we apply the Conditionality principle, according to which every empirical object is explicable in terms of its constitution and its causes, what it is *made out of* and what it is *caused by*, while, on the other hand, we apply the Absolute Ground principle and infer that there must be some first cause and some first constituents. Conditionality generates an infinite series of constituents and causes, while the Absolute Ground principle tells us that conditioned objects must have an unconditioned ground.

His response to both these apparent conflicts is that we cannot apply the notion of actual, as against potential, infinity<sup>19</sup> in the empirical world of space and time. Equally, we cannot apply the Absolute Ground principle *in* that world. The basic mistake is to think of the empirical world as an absolute or completed "whole existing in itself."

If the world is a whole existing in itself, then it is either finite or infinite. Now, both of these alternatives are false ... Hence it is also false that the world (the sum of all appearances) is a whole existing in itself. From this it follows, then, that appearances as such are nothing apart from our representations – which is precisely what is meant by their transcendental ideality.<sup>20</sup>

Kant's constructivism about nature, i.e. the empirical world of space and time, is very clear here. The world is not '*already there*,' an independent whole existing *in itself*. But the position is subtle. We evidently do have *an* idea of the world as a whole existing in itself; it is after all what leads us into contradiction. This idea, like the ideas of God and of freedom to

---

gain in postulating them. They are the absolute basis of the hypotheticals that constitute the phenomenalistic construction: absolute, it would seem, both in the sense of being what is, as against what appears to us, and in the sense of satisfying the Absolute principle.

<sup>19</sup> Actual infinity: an infinite series every one of whose elements actually exists. Potential infinity: a series that can at every point be extended by 'constructing' the next element according to some determinate procedure.

<sup>20</sup> A507/B535. Cambridge translation.

be encountered below, cannot have phenomenal application and thus outruns the concepts of the understanding. But whereas in the case of God and of freedom we can conceive, though we cannot know, that they have noumenal application, the idea of a phenomenal world existing in itself as a whole cannot by definition have *noumenal* application, since it is already structured by phenomenal forms and categories. And yet it cannot have phenomenal application either. We may, it is true, envisage the phenomenal world as a whole for *heuristic* purposes, in the process of seeking the widest possible holistic explanations of the phenomena; at no point however can we *predicate* the idea of absolute totality, finite or infinite, of experience. This does not mean that we have to deny that the phenomenal world is *potentially* infinite. We can arrive at the notion of a potentially or constructively infinite series by allowing that at each step of explanation a further step is possible “in accordance with the laws of empirical advance”: to earlier causes, or smaller parts, or other spaces and times. For, to quote again: “everything is real which stands in connection with a perception in accordance with the laws of empirical advance.”

The laws of empirical advance are the laws physics comes up with. In effect, then, Kant’s view is that a class of things is empirically real if the best empirical explanation of appearances would warrant positing it. Physics is a theoretical representation of the appearances: it may postulate infinite series as part of the representation, or it may represent the world as finite. The point is that whether physics posits that space, time or divisibility are infinite, or finite but unbounded, it cannot be interpreted, philosophically, as positing the empirical existence of an *actually* infinite or an *actually* finite world.<sup>21</sup>

The fourth conflict also arises from an apparent contradiction between the two principles. (The third conflict, which is about freedom and determinism, will be discussed in 2.1.) The thesis is “There belongs to the world, either as its part or as its cause, a being that is absolutely necessary,” while the antithesis is that “An absolutely necessary being nowhere exists in the world, nor does it exist outside the world as its cause.”<sup>22</sup> Again Conditionality argues against the existence of such a being while the Absolute Ground principle argues for it. And once again Kant argues that transcendental idealism resolves this antinomy by opening the possibility that there is a noumenal, necessary being that is the absolute ground of the empirical world.

Overall then the pattern is as follows. Contradictory theses are apparently generated by principles that are dialectically attractive to Reason:<sup>23</sup> the Absolute Ground principle and the Conditionality principle. But these arguments turn out to have an unstated premise – transcendental realism. This doctrine, Kant thinks, is committed to taking the two principles to be, straightforwardly, truths about the world. In contrast Kant argues that the way to avoid contradiction is to distinguish, as the realist does not, between noumena and phenomena. With that distinction, we can argue that Conditionality is indeed true of the empirical world, while the Absolute Ground principle may be true (and given the structure of transcendental idealism can only be true) outside that world.

---

<sup>21</sup> On this reading Kant can allow for a construction of the physical world that makes it finite but unbounded, although he would have to give up, or at least modify, the claim that Euclidean geometry synthetic a priori.

<sup>22</sup> A453/B481. (Kant means that there is nothing “outside the world” that causes the empirical world in the ‘empirical’ sense of cause, in which a cause must occur at a point in time.)

<sup>23</sup> i.e. attractive as principles to rely on in developing an argument, whether or not they ultimately survive.

Even if the “indirect” arguments of the Transcendental Dialectic failed Kant would, as he himself notes, still have the “direct” arguments to transcendental idealism from the possibility of knowledge. But as we shall see when we turn to Hegel, there is another approach that *rejects* transcendental idealism while nonetheless *endorsing* the arguments of the Transcendental Dialectic. Accept that Absolute Ground *and* Conditionality are truths ‘so far as they go’ accept that they are contradictory, conclude that the very constitution of reality, as it in itself, and not ‘just’ our thought about it, contains contradictory elements in dialectical tension. This may seem too fantastic to be worth considering – but it is one of the paths to Hegel’s dialectic.

### 1.6 Religion: denying knowledge to make room for faith

In the Preface to the second edition of the *Critique of Pure Reason* Kant famously said that he had to “deny *knowledge* in order to make room for *faith (Glaube)*.”<sup>24</sup> This does not mean that he subscribed to what we now think of as a ‘de-ontologised’ kind of Christianity, that is, a Christianity which seeks to maintain Christian ethical and spiritual commitment while freeing it from the ontological assertion that God – the supreme personal being, with the properties of perfection which Christianity attributes to Him – *exists*. This de-ontologising theological project developed in response to directions taken by late modern philosophy as widely different as absolute idealism in the 19<sup>th</sup> Century and hermeneutics and positivism in the twentieth-century. It will be important to consider it later. But it was not Kant’s view.

His point, rather, was twofold. First, transcendental idealism shows that knowledge is possible only of the phenomena. If this is so, then to claim *knowledge* of God is to place Him among the phenomena: a knowable empirical object *in* the phenomenal world. This is not our true conception of God; such a god would be an idol.<sup>25</sup> And second, if the god in question *were* merely another empirical object then the kind of considerations that Kant thinks make religious faith reasonable would no longer make it so. The reasonableness of faith, in Kant’s view, arises from *practical*, not from *theoretical*, reason and as such could give no basis for beliefs about empirical features of the world. So here too transcendental idealism, far from undermining religious belief, is supposed to provide a new way of defending it.

Nonetheless, it is quite understandable that Kant’s philosophy of religion should have acted as a major impetus on the road to de-ontologised Christianity. It is easy to make Kant seem more radical about religion than he is. It can, in the first place, be quite plausible to read Kant as claiming not merely that the noumena are unknowable but that any would-be substantive assertion about them is strictly meaningless – beyond the “bounds of sense”. And indeed the tendency to treat the bounds of experience as the bounds of sense is one of the tendencies in Kant’s thought; it fits with the view that the Absolute Ground principle is merely heuristic, and might be thought to explain such remarks as that the principle of causality “holds only within the field of possible experience and outside it is

---

<sup>24</sup> Bxxx.

<sup>25</sup> cite



without any use or indeed without any meaning [*Bedeutung*].”<sup>26</sup> This heuristic view of the Absolute Ground principle comes to the fore in Kant’s criticisms of the rational arguments for God, especially the cosmological argument.<sup>27</sup> As we have noted, it could, worked out fully, lead to a *purely* constructivist treatment of empirical reality, that is, one that dispenses entirely with the transcendental-idealist distinction between phenomena and noumena.

However this cannot be the dominant tendency in Kant’s thinking about religion, practical reason, the self and freedom. He takes the existence of God to be a postulate of practical reason, along with freedom and the immortality of the soul. We shall come to what he means by this in 3.6 and 4.5, and in particular try to assess what right he thinks we have to believe that God, freedom and the soul exist and that the soul is immortal. But certainly he takes the concepts of freedom, self and God to be *thinkable*, even though transcendental. Nor is this inconsistent with his overall epistemology. He treats the categories as essential to any thinking about reality whatsoever – he does not restrict *thinking*, as against *knowing*, to the empirical domain. When applied in empirical judgement the abstract categories are “schematised.” In particular (as we have already noted) the abstract category of dependence becomes the empirical notion of causation in space and time. Thus while we cannot think of noumena as standing in relations of empirical, spatio-temporal causation we *can* think of them in terms of the abstract notion of dependence.<sup>28</sup> It is what we do when we think of the empirical world as dependent on some noumenal entity which is not itself dependent on anything. It is what we do when we think in this way of God. But here we do more – we think of God not just as the absolute (non-dependent) ground of the dependent, but as the perfect personal being. That there is such a being is, according to Kant, something we cannot know by theoretical reason alone – yet he does seem to argue that if we know our duty then we know that God exists.

But before we come to the latter claim let us consider his powerful criticism of the attempt to prove the existence of God by *theoretical* reason.<sup>29</sup> It has been as influential as that other great critique, Hume’s *Dialogues concerning Natural Religion*.

There are, Kant thinks, three possible proofs of the existence of God: they are the ontological proof, the cosmological proof and the “physico-theological” proof (more often known as the argument from design). Of these, Kant holds the ontological proof to be fundamental. The other two, even if sound, aren’t strong enough to show the existence of a perfect personal being without it. In contrast the ontological argument, if sound, would be strong enough on its own without the other two.

---

<sup>26</sup> A636/B664 Cp A609/B637, where he is discussing the cosmological argument: “The transcendental principle of inferring from the contingent to a cause ... has significance only in the world of sense, but ... outside it does not even have a sense ... “the principle of causality has no significance at all and no mark of its use except in the world of sense; here, however, [i.e. in the cosmological argument] it is supposed to serve precisely to get beyond the world of sense.” See also Kant’s attempt to strike a balance at A696/B724– A698/B726.

<sup>27</sup> See the section entitled “Discovery and explanation of the dialectical illusion in all transcendental proofs of the existence of a necessary being,” A614/B642 – A620/B648

<sup>28</sup> The point is highlighted by Ameriks 2000, 2006. On Kant’s notion of schematism see A137/B176 – A147/B187. On the thinkability of *causa noumenon* see CPR 5:54 – 57.

<sup>29</sup> in the chapter entitled “The ideal of pure reason,” the final part of the Transcendental Dialectic, where Kant’s target is “rational theology” which he sees as the third of the metaphysical pseudo-sciences. The German translation of Hume’s dialogues came to late for Kant to see it.

The ontological argument can be put thus. We start from the claim that we have the idea of a perfect being. We note that there is no contradiction in this idea, and infer that it is possible that a perfect being exists. Now such a being must possess, among its perfections, the perfection of non-contingent existence – it must have necessary existence as part of its essence. But if it necessarily exists, then it exists.

Kant's main reply (though not the only thing he says) is that "*Being* is obviously not a real predicate, i. e. a concept of something that could add to the concept of a thing." "It was entirely unnatural, and a mere novelty of scholastic wit, to want to take an idea contrived quite arbitrarily and extract from it the existence of the corresponding object itself."<sup>30</sup>

This reply has acquired some prestige, partly through debateable philosophical readings of developments in modern logic which are said to show that 'existence is not a predicate'. Yet that is not the issue. The central question is whether a necessary being is possible. On the face of it Kant is not in a position to reject such a possibility. He does indeed deny that a necessary being could exist in the empirical world, but seems to allow that such a being is possible in the noumenal world. But if there is a possible world in which there exists a being that exists in all possible worlds, then it exists in the actual world, since the actual world is one of the possible worlds.

Various replies to this are available: obviously they require careful consideration of the notions of possibility, necessity, and existence. Still, it is certainly not true that Kant 'killed the ontological argument.' He did not do so philosophically speaking, nor did he do so historically speaking. It remained important to Hegel, and later to Heidegger; it remains integral to a certain kind of religious faith. All of this gives it importance in the story of late modern ethics; none of it, of course, means that the argument is sound. Kant was right to dismiss it, and right also in his claim that the other two arguments can only reach their target, God, if they are supplemented by the ontological argument. I have put discussion of these issues into Appendix 4.

The core of the cosmological argument is that "If something exists, then an absolutely necessary being also has to exist."<sup>31</sup> This is an appeal to the Absolute Ground principle, and thus closely related to the fourth conflict of pure reason (1.5). Kant emphasises that the Absolute Ground principle falls well short of establishing the existence of a perfect personal being, as against the abstract notion of an unconditioned noumenal condition that grounds the empirical world. But he is also tellingly disturbed by the abstract notion itself (tending to fall back on a purely regulative reading) – in a way that throws light on his difficulty in reaching a stable view of the Absolute Ground principle:

The unconditioned necessity, which we need so indispensably as the ultimate sustainer of all things, is for human reason the true abyss... one cannot resist the thought of it, but one also cannot bear it ... Here everything gives way beneath us  
<sup>32</sup>  
 ...

In contrast, metaphysical bafflement disappears when Kant turns to the argument from design. He describes it as "the oldest, clearest, and the most appropriate to common human reason."<sup>33</sup> But he says that it

---

<sup>30</sup> A598/B626; A603/B631.

<sup>31</sup> A604/B632

<sup>32</sup> A613/B614

<sup>33</sup> A623/B65.

could at most establish a highest *architect of the world*, who would always be limited by the suitability of the material on which he works, but not *a creator of the world*, to whose idea everything is subject.<sup>34</sup>

This is a remarkably similar assessment to that which John Stuart Mill later made (0.0). Neither philosopher thinks that the argument from design really does establish such an architect, or demiurge; it's just that its argumentative method is reasonable, by the ordinary canons of common cognition. But Kant's main point (again like Mill's later) is that these ordinary canons of reasoning cannot establish the existence of a perfect personal being.

Overall Kant's verdict is clear: theoretical reason has no power to prove the existence of a supreme being. Considered from the standpoint of theoretical reason alone, such a being can at most be a regulative ideal, an idea that illuminates inquiry from outside the understanding. The critique of theoretical reason endorses the *thinkability* of such a being and establishes that its noumenal existence is not ruled out; however it provides no reasonable ground for positive faith.

Yet when we turn to practical reason, things are otherwise, as we shall see (3.6).

## 2 Freedom

Some of Kant's profoundest thinking is about freedom. It has many strands, which stretch into ethics, religion, politics and the nature of knowledge itself. But we must start from his metaphysics of freedom, which provides the core of his account of freedom as autonomy. That takes us back to the third conflict of pure reason, which was left out in 1.5; its topic is freedom and determinism.

### 2.1 *The Causality of Freedom: Autonomy*

Kant makes a crucial metaphysical assumption about what it is to act freely. The assumption is that free action is the exercise of a distinctive causality by something that is not itself caused. He calls it "causality through freedom". As he observes, causality through freedom seems incompatible with the idea that everything in the world follows deterministic laws. Accordingly, the thesis of the third conflict is that there is a "causality through freedom", as well as a "causality in accordance with the laws of nature," while the antithesis states that "there is no freedom; everything in the world happens solely in accordance with laws of nature."<sup>35</sup>

Now Kant has argued that the law of universal determination by causes "in accordance with the laws of nature" must apply to all phenomena as a precondition of the unity of experience.<sup>36</sup> Thus on his own showing there is a solid case for the antithesis. As for the thesis, Kant claims that the argument for it works through a *reductio* on the antithesis. For on the one hand causality in accordance with the laws of nature mandates an infinite backward chain of antecedent causes, and on the other it also demands a "cause sufficiently determined *a priori*."<sup>37</sup> Here "determined *a priori*" means 'determined in advance;' Kant's thought then is that a causal condition could not be *determined in advance* if it came out of an actual infinity of antecedent conditions. Why not? Presumably because it can only be *determined* by something unconditional, and nothing in an actual infinity of antecedent

---

<sup>34</sup> A627/B655

<sup>35</sup> A445/B473

<sup>36</sup> In the transcendental deduction of the categories [?]

<sup>37</sup> A446/B474

causes would be unconditional. This is an appeal to the Absolute Ground principle, so it looks as though Kant is not in fact offering a *reductio* but implicitly invoking that principle as a premise.

True. But Kant's target here is the transcendental realist. His point is that the antithesis contradicts itself when it is "taken in its unlimited universality"<sup>38</sup> – in other words, when it takes the empirical world to be a *self-standing whole*. This is the way transcendental realism takes it. As argued earlier, the empirical world *conceived as an absolute and completed reality* would have to be conceived as actually infinite (or perhaps as actually finite but unbounded). Yet at the same time, Kant holds, in the world so conceived dependent causes would have to be traceable back to "an absolute causal spontaneity beginning from itself." A446/B474. Hence the contradiction.

Kant is thus assuming that *in a world that is a self-standing whole the Absolute Ground principle applies*. He resolves the contradiction not by rejecting that assumption but by denying that the empirical world is a self-standing whole – by appeal, that is, to the distinction between phenomena and noumena, and the consequent constructive view of the empirical world. Our construction of the empirical world only commits us to a *potential* infinity of causes, and while the Absolute Ground principle applies noumenally, it does not apply within the phenomenal world. Note that if Kant were taking the *pure* constructivist position, with the Absolute Ground principle and the principle of Conditionality treated as merely heuristic, there would be no contradiction to resolve. His treatment of this conflict reveals much about his thinking: that transcendental idealism is not a pure constructivism, that it envisages a world of noumena that is self-standing, and that when the world is so envisaged, not just as it appears phenomenally but as a noumenal whole, the Absolute Ground principle applies.

The argument, if sound, would indeed show that there must be a noumenal explanation of the phenomena that is distinct from 'causality in accordance with the laws of nature;' an explanation whose *explanans* is something that is itself uncaused. But it would show this at a quite generic level. Why call it 'causality through freedom'? What has this very generic argument got to do, specifically, with human beings and their freedom? Take the way noumena give rise to appearances in a knowing subject: that is an instance of noumenal 'causality,' but one cannot describe it as causality through freedom.

The relevance to human freedom is that Kant thinks free action in particular, if possible at all, must be an instance of uncaused causation. Human beings are a part of nature and in that respect, as phenomena, fall under deterministic laws. If that were all there was to them – as transcendental realism must hold, since it takes the empirical world to *be* the world – then, Kant thinks, they could not be free. But Kant has prepared the ground (in the Aesthetic and Analytic) for the claim that they are also noumenal, and that opens the possibility that as noumenal beings they can be uncaused causes – hence the possibility that their thinking and willing can be free. Wherever events in the empirical world result from free thinking and free willing we have the causality of freedom. Nonetheless, Kant claims, it is consistent to hold that such events also fall entirely under deterministic laws at the empirical level, because these laws are phenomenal, whereas the causality of freedom is noumenal.

Two steps are involved here. The first is that free action or belief is uncaused causation, the second is that uncaused causes must be noumenal.

---

<sup>38</sup> A446/B474

The insight is in the first step. It is this step that introduces the Kantian notion of autonomy. Autonomy is action from what an agent sees, for him or herself, to be sufficient reason for the action. Thus the explanation of the action is simply that the agent saw that it was the right thing to do. Where this explanation applies, we have full freedom – *full* freedom because although Kant often simply identifies freedom with autonomy, he also acknowledges that where you *could* have acted autonomously but do not do so you act freely in the sense of responsibly. Your action can be imputed to you, you are answerable for it, and potentially blameworthy.<sup>39</sup> Freedom as autonomy is thus a stronger notion than freedom as responsibility. But freedom in the weaker sense presupposes the possibility of autonomy, since in the weaker sense it is the capacity to act autonomously.

Acting autonomously – ‘giving oneself the law’ – consists in recognising that there is sufficient reason to do a particular thing and doing it *from* that recognition, making it the principle by which one acts. The ‘law’ one gives oneself is the law of practical reason. This conception of free action and belief is in some ways at least as old as Plato’s discussion of the soul. Undoubtedly however it is Kant’s formulation in terms of autonomy that has been enormously influential in the late modern period. There is something insightful and attractive about the idea that freedom is giving oneself the law. It is a governing idea of late modern ethical thought, as we shall see; though not only in the way Kant intended it, for it has been pulled in a variety of other directions, in particular in ways that ignore a crucial distinction that is presupposed in Kant’s way of understanding it, namely that between the giver and the author of the law.

We shall come back to this distinction (see 3.4). But we should first note an important sense in which Kant is right to characterise autonomous action as original or uncaused causation. A humdrum example will illustrate. Suppose I am driving to a destination and I hear on the car radio that my preferred route is blocked by an overturned lorry. I take what I recognise as the next best route among the remaining options; I do that *because* I can see that the fact that the best route is closed is sufficient reason to take that route. My recognition that there is sufficient reason to take the other route is *why* I take the other route: I act from that recognition. This is a paradigm of autonomous action.

In what way then is the cause of my action uncaused? Well, I recognise that there is sufficient reason to take route B *because there is*. Yet my act of recognition – recognising that this is what there is sufficient reason to do – is not *caused* by the fact that there is sufficient reason to take route B.<sup>40</sup> *That* is not a fact in the empirical world at all; it has no causal standing. It is a normative truth about reasons. There are, to be sure, various empirical facts in virtue of which route B is the next best: it is because I’m aware of them that I see that route B is the next best. However, *that* these empirical facts give me sufficient reason to judge that route B is next best is again not itself an empirical fact; it is a purely normative truth.

When we say that I judged that there was sufficient reason to take route B because there was, our ‘because’ is not the *because* of causality. We explain an intelligible action directly in terms of insight into a truth about reasons. Whether or not the consequent connection between judgement and action is a case of natural causality, the explanatory connection between normative truth and judgement is not.

---

<sup>39</sup> Cite.

<sup>40</sup> Point noted by Allison, 1990, p. 49, but he doesn’t explain why, that being so, Kant holds the ‘causality of freedom’ to be noumenal.



This is the heart of Kant's contrast between receptivity and spontaneity. The way in which I recognise the *empirical* facts, in virtue of which there is sufficient reason to take route B, necessarily involves some element of empirical receptivity that belongs within the causal nexus. But when I recognise that given that these facts obtain there is sufficient reason to take route B, that is not a receptivity-based judgement. Here, in judging a purely normative truth, I judge purely spontaneously. Spontaneity is inherent in autonomy.

We thus have, first, a purely normative truth which is no part of the empirical world and has no natural causality, and second, my spontaneous recognition of that truth, which consists not in an empirical relation between two distinct empirical facts, but rather in an intelligible relation between a normative truth and my recognition of that truth. In this explanation we are explaining an empirical event – my taking route B – in terms other than its accordance with the laws of nature. This is the truth in Kant's conception of the causality of freedom. But does it require any appeal to noumena?

## 2.2 Spontaneity and receptivity

We should look more closely at the notion of spontaneity (*Spontaneität*). Kant refers to the spontaneity of understanding, of will, and of reason. All these come down to the spontaneity of reason, since understanding is simply reason in its empirical role, while the will, *Wille*, is reason in its practical role, making judgements about what there is reason to do.<sup>41</sup>

Understanding is epistemically normative, practical reason is practically normative. Terms closely related to spontaneity are self-activity (*Selbsttätigkeit*) and self-determination (*Selbstbestimmung*). To say that the understanding spontaneously applies concepts is to say that I am self-active and self-determining in applying concepts. To say that *Wille* can be the spontaneous cause of my actions is to say I can spontaneously recognise and act on reasons and that I am self-active or self-determining when I do so.

More mysteriously, Kant also talks of the spontaneity of self-awareness – of the 'I think' that accompanies all my representations (as he famously puts it). In this case the salient contrast is with receptivity. Self-awareness is spontaneous in the negative sense that it is not receptive: my consciousness of thinking and willing is not acquired through some mode of receptivity that gives me access to that thinking and willing; as mentioned before, it just *is* my thinking and willing. That being so, since I am self-active in thinking and willing I am by the same token self-active in my awareness of my thinking and willing. (Beyond this there is perhaps a hint in Kant of an idea that Fichte would take forward, that in thinking and willing, and in the unity of consciousness that they presuppose, I somehow actively give rise to myself.)

Let's have another example. Suppose I am aware of a coloured light in my visual field which I judge to be green. To judge it to be *green* is to apply that concept on the basis of the visual experience spontaneously; to infer from the judgement that the light is green that it is permissible to drive on is another act that is spontaneous. Both acts, of concept introduction and inference, take place against a complex background of beliefs which are receptively grounded; but they are not themselves grounded in any receptivity – in the way that my sensible awareness of the green light in my visual field is so grounded. Whereas sensibility is receptive, applying a concept and making an inference are things I do. In the application of concepts I can be said to be self-active, self-determining: the explanation of my applying a

---

<sup>41</sup> Kant distinguishes between *Wille* and *Willkür*, where the latter refers solely to the power of choice, whether or not in line with *Wille*.

given concept is simply that I take it to be correct to apply it. Of course my active thinking occurs in the context of many beliefs and assumptions – that I’m not colour-blind, that the light is normal, and so forth. But the active thinking is ruled, or inherently constituted, by a continuing impression of correctness that comes naturally – it is spontaneous in that more ordinary sense.

When nothing goes wrong, when there is no intervention by “alien causes” (in Kant’s phrase), the chain of explanation goes from what there is sufficient reason to believe or do – to my spontaneous recognition that there is – to my believing or acting because I recognise that there is. This explanatory perspective on autonomous belief, inference, and action explains them not in terms of natural or empirical causality but as appropriate responses to normative reasons, or, to get this closer to Kant’s way of putting it, appropriate responses to *reason*. And this, Kant holds is, what freedom *is*. To believe, infer or act autonomously, fully freely, is to do so from a recognition of sufficient reason. Freedom is autonomy is rationality.

### 2.3 Empirical and intelligible

This leads us to Kant’s famous contrast between an empirical and an intelligible standpoint on human thought and action:

A rational being must view itself *as an intelligence* ... as belonging not to the world of sense, but to that of understanding; and hence it has two standpoints from which it can consider itself, and recognise laws for the use of its powers, and consequently for all its actions: *first*, in so far as it belongs to the world of sense, under laws of nature (heteronomy), *secondly*, as belonging to the intelligible world, under laws that, independent of nature, are not empirical, but have their foundation merely in reason.<sup>42</sup>

A description and explanation from the empirical standpoint of what happened when I diverted to route B would explain it in terms of a series of psychological or even just physiological events linked by natural causality. Kant assumes that an explanation of this kind is always possible and that it would be deterministic. But as he emphasises, an explanation from the intelligible standpoint is also possible. When an act or inference can be explained in the intelligible way, as proceeding from sound recognition of reasons, it is autonomous. Where it falls short of autonomy, through the intervention of alien causes, it is heteronomous.

One might think that this insightful distinction would lead Kant to conclude that freedom and empirical determinism are compatible. In a way it does, but with a tortuous complication: Allan Wood neatly described it as Kant’s thesis of ‘the compatibility of compatibilism and incompatibilism’<sup>43</sup>

The unnecessary paradox arises from Kant’s insistence that the ‘causality of freedom’ is a genuine but noumenal form of agent causality. As we saw (2.1), Kant’s suggestion, in the antinomy of freedom, is that without appeal to the noumenal the causality of freedom would be crowded out, as it were, by the causality of nature. Here Kant assumes that what he describes as the causality of freedom is not just *unavailable* in the empirical standpoint, since that deals exclusively with empirical causal explanation, but somehow *incompatible* with it. Now certainly the explanatory relation between a truth about reasons and one’s recognition of that truth is not an empirical causal relation. But that is not because it is a noumenal causal relation. It is because it is not a causal relation at all. Furthermore, if my *recognition* that there is sufficient reason to do or believe something, which is a state of mind, causes me to

<sup>42</sup> *Groundwork* III (iv 452), Timmermann translation.

<sup>43</sup> Wood, 1998, p.239

do or believe it, there is no reason to deny that that particular causal connection is empirical. So there is no competition between rival forms of causation, and it's not clear why the availability of an intelligible explanation should be incompatible with the availability of an empirical one.

A suggestion that might be made is that Kant has nowhere to put truths about reasons other than in the domain of noumena. In that case when one recognises a normative truth one is receptive to a noumenal fact, and that receptivity must involve noumenal causation. However this assumes that there needs to *be* somewhere to put normative truths. The thought is that normative truths must correspond to some facts or other, and since Kant clearly does not think that the facts that they correspond to are phenomenal he must think they are noumenal. But that is not Kant's view of truths about reasons. If these were truths about some special class of noumenal facts we could not, by his own principles, know them. Moreover, Kant's own account of how we know them is cast in terms of spontaneity, not receptivity. It thus entails that our knowledge of them is not in any way factual knowledge, since knowledge of facts requires receptivity and cannot be grounded in spontaneity alone (leaving aside the special case of self-awareness of self-activity).

The notions of sensibility, reason, self-activity and freedom hang together within the intelligible standpoint. One can ask whether that standpoint is or is not in some way 'reducible' to the empirical standpoint. That is as live a question as any in philosophy. But even if we concluded that it is not reducible, that would do nothing in itself to support the bold ontological thesis that the relation of the intelligible to the empirical is that of noumenon to appearance.

To sum up so far. Kant introduces transcendental idealism on general epistemological grounds which, he argues, call for distinctions between noumena and phenomena and between noumenal subject and noumenal object. These distinctions, as we saw, are at the level of epistemological analysis: since transcendental idealism itself rules out knowledge of the noumenal, we cannot know what specific noumenal ontology underlies them. Still, once the distinction between noumenal subject and object has been established by epistemological arguments, it becomes very tempting to treat individual human beings as noumenal subjects, and then further, as noumenal agents. So much, Kant must have thought, falls into place if we do that!

However, whereas Kant gives good reasons for thinking that a noumenal/phenomenal distinction is implicit in the very possibility of knowledge, he gives no good reasons for thinking that we need it to explain the possibility of freedom. In fact Kant himself has all the elements of a perfectly good, non-noumenal, account of the intelligible level of explanation, and thus of autonomy. Autonomy does not require transcendental idealism, and the really important Kantian insight, that normative judgements are purely spontaneous and not receptive, shows why we don't need it. Nothing is gained by interpreting the empirical/intelligible distinction in terms of the phenomenal/noumenal distinction, and much is lost.

We can note some of these costs. First, in applying the phenomenal/noumenal distinction to persons Kant puts inclinations and emotions on the phenomenal side as 'appearances'. His reason for this might be the idea that inclinations and feelings are simply things that happen to me, whereas judging and willing are things I do. But it blinds him to a very important ethical point, namely, that inclinations and emotions are reason-sensitive responses which can be assessed as reasonable or unreasonable, and hence fall under the epistemology of spontaneity. We shall come back to this (2.5). Another, very peculiar, consequence is that inclinations and emotions, as appearances, must be appearances of

noumena – specifically, Kant thinks, of the self as it is in itself. This leads him into speculations about noumenal disposition, or noumenal decisions as to empirical character, which by his own rules are out of order.

By his own rules again, if autonomy is placed in the realm of the noumenal it follows that we can never know whether we are acting freely, or even whether we have the capacity to do so. Well, it is indeed true that we sometimes don't know, and perhaps can't know that – in the way that we sometimes don't know, and perhaps can't know, what we really want, or whom we really love. It is, for sure, always easy to kid ourselves about our feelings and motives, or simply to fail naively to recognise their deep sources. Exaggerating this theme of our lack of self-knowledge is however one of the more exasperating strands in modernist thinking. In fact, in many minor and major cases, such as the example of diverting to route B, we know perfectly well why we have acted, and in particular know that we have acted autonomously. We should firmly distinguish the empirical, phenomenological question of how much in practice we know about our reasons for action from the transcendental assertion that we can in principle *never know at all* whether we have acted for reasons, since acting for reasons involves noumenal causality.

True, in the Groundwork Kant handles this implausible conclusion in an attractively elegant way, by arguing that whenever we deliberate we do so under the idea that we are free. In deliberating what to do I necessarily *postulate* that I am free to decide what to do. This gives Kant another contrast between legitimate practical postulates and knowledge proper. We cannot know we are free but – given transcendental idealism – it is legitimate for us to reason under the idea of ourselves as free. Yet elegant as this, is it unnecessary. There is no special, transcendental, problem about knowing from what reasons one acts. Nor is there *that* kind of problem about knowing whether one's reasons are sufficiently good (though there can be many others). I can often know from what reasons I am acting, and know that those reasons are sufficiently good reasons. In such cases, I know that I act freely.

#### 2.4 Negative and positive freedom

We can apply these conclusions to Kant's classic distinction between negative and positive freedom in the Groundwork. He introduces the Groundwork's final section, section III, as follows:

*Will* is a kind of causality belonging to living beings so far as they are rational.  
*Freedom* would then be the property this causality has of being able to work independently of *determination* by alien causes ...

He continues:

The above definition of freedom is *negative* and consequently unfruitful as a way of grasping its essence; but there springs from it a *positive* concept, which, as positive, is richer and more fruitful.<sup>44</sup>

Positively, he continues, freedom is the ability to originate your own action, to be its uncaused cause.

Any notion of causality, Kant thinks, analytically brings in the idea of a law under which one thing, the cause, produces another, the effect. Where an act springs from positive freedom, however, the law is not the law of nature but the law of reason – a universal normative truth determining what I should do. Positive freedom is autonomy: it consists in acting *from* that truth, in the way we have considered. It is a distinctive kind of causality in

---

<sup>44</sup> IV 446

that I don't act in *accordance* with law, as a billiard ball acts in accordance with physical laws, but *from* the law. A fully free person, in Kant's conception, is one who reliably recognises and acts from underlying universal norms, accurately recognising their implications through the wide variety of particular circumstances to which they apply.

Kant thinks this notion of positive freedom involves noumenal agent causality; I have argued that it does not. That is a major fork in the road. What remains true, however, on either view, is that freedom is not to be identified with the capacity to act otherwise. An ideally autonomous being might be simply incapable of acting against sufficient reason. This is a point Kant recognises and indeed insists on. God acts autonomously, from what is best, and thus freely, even though He *necessarily* acts that way:

The freedom of the divine will does not mean that He could have done something other than the best (for this is not even what human freedom means), but rather that He is necessarily determined by the idea of what is best; which is not so with man, and that is why his [man's] freedom is limited. (Reflexion 6078) c. 1783/84

Unlike the holy will, human beings are finite, sensuous, subject to natural impulses that limit their freedom and cause them to fall into heteronomy. They thus need – and have – the power to ‘give themselves the law’ in response to that finitude, whereas the Holy Will acts from good reasons so to speak frictionlessly: it has no alien causes to overcome.

At this point transcendental idealism takes over. Because Kant thinks of positive freedom as noumenal agent causality he treats *any* empirical cause as alien, and treats any emotion, inclination etc. as empirical. However, the notion of ‘alien cause’ that is really needed in the negative account of freedom is simply that of *any cause that constrains pure responsiveness to reasons*. An alien cause is anything that interferes with, or threatens to interfere with, recognition of normative reasons or action from that recognition. Alien causes would include causes that produce mistakes or carelessness, effects of indoctrination, wishful thinking, desires or inertia that impede the effort to do what is best, and many other such things. It does not follow that *any* empirical cause is alien.

Kant's transcendental idealism about freedom has another very important but highly questionable consequence. As he says, we are responsible for our heteronomous actions, the ones that are not autonomous, so long as we had the capacity to act autonomously. He adds that we always have that capacity. Our noumenal side always has the capacity to master our phenomenal side in the decisions we make. Indeed if the noumenal/phenomenal distinction applies to persons in the way Kant thinks it does then this conclusion follows. His transcendentalism about freedom commits him to the idea that alien causes can always be overcome, by everyone in every situation.<sup>45</sup> We all have the capacity, and the obligation, to be free. However this is not an implication of Kant's *analysis* of freedom, in negative and positive terms, as such, nor does it face up to the facts of human weakness, both cognitive and conative, which Kant so often emphasises rhetorically. In many situations, and in many people, causes that may be alien to freedom, but are in no way alien to humanity, destroy autonomy or damage it. If we accept Kant's analysis of freedom in terms of autonomy but reject his transcendentalism this is something we must be open to, and the question of how empirically important it is an utterly major issue for moral and political philosophy.

---

<sup>45</sup> “incitements from desires and impulses (and therefore from the whole sensible world of nature) cannot impair the laws which govern his will as intelligence. Indeed he does not answer for the former nor impute them to his proper self - that is, to his will; but he does impute to himself the indulgence which he would show them if he admitted their influence on his maxims to the detriment of the rational laws governing his will.” IV 458



## 2.5 Free thought and the epistemology of reason

Kant's discussion of freedom as autonomy is of first importance, both in itself and for its historical influence. Historically it is one of the great routes by which philosophy moved from the eighteenth century to its characteristic late modern preoccupations and anxieties about freedom. This is especially true of philosophy on the Continent, where the legacy of German idealism was strong, but it resonates throughout 19<sup>th</sup> century culture and moral philosophy.

None of this of course means that the Kantian view of freedom was uncontroversial. Far from it. Some of the controversy arose from the transcendental idealist setting into which Kant placed it, and which I have argued is unnecessary. But its most controversial aspect is independent of that; it arises from the central thesis that autonomy is insight into and motivation by normative truth. Obviously, then, autonomy is possible only if there is such a thing as normative truth. Denying that there is, however, has been a main theme of modernism. We shall be interested in the consequences of this denial; both for morality, where autonomy is an indispensable category, and for an important kind of philosophical liberalism in which autonomy or moral freedom plays a central role.

For the moment however we should ask whether Kant himself has epistemological resources which could answer this scepticism. The question leads back to his distinction between spontaneity and receptivity, and casts light on the importance and priority he gives to the idea of free thought.

Kant takes spontaneity to be 'agent causality' originating in the noumenal self, but we have seen no reason to agree with him. The crucial insight, that autonomy – full freedom – is spontaneity does not require any such transcendental appeal. Furthermore the appeal introduces other distortions. The most important, already noted, concerns Kant's view of emotions, feeling, inclinations. A general view of spontaneity should see it as a property of all reason-sensitive responses. If affective dispositions to feel – as well as cognitive and practical dispositions to think and to act – are reason-sensitive, then, contra Kant, spontaneity and hence freedom is a property of those dispositions too. We assess emotions such as admiration, anger, blame as reasonable or unreasonable just as we assess beliefs and actions as reasonable and unreasonable. Thus for example, you are annoyed because you have good reason to be. Or you try to restrain your anger because you see there is no good reason to be angry. As with actions and beliefs, feelings can be products of reason-recognition. Kant's blindness to this does great damage to his thinking; we shall come back to it in discussing the responses to Kant of Schiller and Hegel. (0.0.0)

Still, the fact is that for Kant the focus is firmly on thinking and willing. 'Reason,' is the power of judging purely normative claims about reasons to believe and to act. Thought is free in just the way that will is free – in fact free will is just a species of free thought. As John Rawls put it, "For Kant, there is no separate problem of the freedom of the will, as if something called "the will" posed a special problem. For him there is only the problem of the freedom of reason, both theoretical and practical."<sup>46</sup>

Reason is spontaneous in all its activity. It involves no receptivity to any special domain; it is an absolutely spontaneous active power. This spontaneity is freedom. So what is spontaneity? Further: can we maintain this notion of freedom as spontaneity while freeing it from the idea of a noumenal agent causality?

---

<sup>46</sup> John Rawls, 2000.

A spontaneous disposition is not just a naked disposition to believe or do: characteristically it comes clothed as a disposition to see a belief or action as *reasonable*. Lacking that normative dimension it can seem ‘alien,’ ‘not really mine,’ like a disposition to believe that one is being followed, or to lick a puddle on the pavement. Such dispositions are inexplicable at the intelligible level, and that undermines their status as a first-person ground for normative judgement (in this case, the judgement would be that there is *reason* to believe that one is being followed, or that one *should* lick the puddle). Their ‘alien-ness’ disqualifies them from that role. To ground a judgement about reasons a disposition has to feel to me, immediately or primitively, as one that there is *reason* for me to feel, and reason for *me* to feel. This is the phenomenological aspect of spontaneity. Then there is the negative side noted by Kant: absence of alien causes. The presence or absence of such causes may not show up at the phenomenological level. It can be hard to tell of a disposition whether it is free of alien causes; it may require a third-person perspective; it may remain uncertain, undecidable. Nonetheless the presence of alien causes undermines a disposition as a ground for belief or action.

Spontaneity, then, is what I ‘can’t help’ being disposed to see as reasonable, and thus being disposed to think or will – even after the influence of alien causes has been cleared away. In this sense it is indeed self-origination, or self-activity, not receptivity to any external given: the self-activity consists in taking on, accepting, spontaneous impressions of correctness *as* correct. This notion of spontaneity is crucial to the epistemology of reason.

There is another crucial notion; it is the *universality* of reasons, a feature of reasons that Kant takes to be basic.

Because reasons are universal we are committed to withdrawing a judgement if we think that no fault can be found, even in principle, in either the evidence or the judgement of someone who disagrees. Suppose now that evidence is not in dispute, or as in cases of purely normative judgement, that it is not at issue. We cannot then say, about a disputed judgement, ‘Well, there’s a reason for him to make the judgement but not for me.’ On the contrary, if we maintain our own judgement we are committed to viewing the other person’s judgement of reasons as faulty – as imperfectly competent in the relevant domain, or as undermined by alien causes.

From this arises an epistemological discipline. Warrant for normative judgements requires (i) genuinely serious reflection as to what one’s spontaneous dispositions are and (ii) robustly open-minded – as against dogmatic and stubborn, or weak and credulous – assessment of the responses of others. The Kantian claim is that these two pillars alone – spontaneity and dialogue – can adequately support the platform on which free thought discovers its own laws for itself, without appeal to any external authority.

This is the epistemological groundwork for the ideal of liberty of thought and discussion.<sup>47</sup> Thought that is genuinely free stands up to discussion. It is ruled solely by reason relations that it recognises through reflection on its own activity. This reflection is inherently collective: it requires unconstrained discussion with other seekers of truth, people who are genuinely responding not to dogma but to their own spontaneous normative dispositions. Of course it is possible for one person to be right, and all others wrong – but no-one can know that that they are right without engaging in dialogue with others and reflecting on others’ responses.

*Especially* in ultimate claims of reason, free debate that excludes no-one is essential – as a matter of the epistemology of the normative, not just of the ethics of democratic respect.

---

<sup>47</sup> A821/B849

It does not follow that every voice carries equal weight. In free and inclusive debate more and less authoritative voices inevitably emerge. It is important that they should – that authoritative voices should not be muffled, or hesitant in taking the lead. Putting it the other way round, one's independence or dignity is not diminished by free recognition of genuine excellence in the common pursuit of truth, wherever one finds it. On the contrary, recognising it is a mark of inward freedom.

The idea of free thought is fundamental to every aspect of Kant's philosophy. It radiates through his work and explains much about its impact. But as we have seen, it requires no theses about the transcendental status of freedom and reason. It is supported perfectly well by two other, and better, Kantian ideas: the absolute spontaneity of normative judgement and the universality of reason.

In basing a principle of liberal social philosophy – liberty of thought and discussion – on principles in epistemology together with a commitment to rational pursuit of truth, Kant and John Stuart Mill closely agree, as will emerge when we come to Mill's account of these matters. They give different grounds for this commitment, and they approach it in different ways, but for both the epistemology of free thought is the heart of their philosophical liberalism. Free thought is not so much an individual right as a public good, like free air. In his particular way Kant is quite as eloquent about this as Mill is:

Reason must subject itself to critique in all its undertakings, and cannot restrict the freedom of critique through any prohibition without damaging itself and drawing upon itself a disadvantageous suspicion. Now there is nothing so important because of its utility, nothing so holy, that it may be exempted from this searching review and inspection, which knows no respect for persons. *The very existence of reason* depends upon this freedom, which has no dictatorial authority, but whose claim is never anything more than the agreement of free citizens, each of whom must be able to express his reservations, indeed even his *veto*, without holding back.<sup>48</sup>

### 3 Morality and Religion

#### 3.1 *Morality and practical reason*

So far we have only talked about practical reason. But what about morality? It is not entirely easy to work out how Kant saw the relation between the two. Yet the question is important, not least because the nature and the very existence of morality came to be contested in the late modern period.

Famously, Kant distinguishes between hypothetical and categorical imperatives. In both cases, hypothetical and categorical, the imperative is expressed by an 'ought.' The distinction is between an 'ought' that is hypothetical in the sense that it is conditional on something else, and an 'ought' that is categorical in the sense that it is not conditional on anything. So one of Kant's most basic preoccupations re-emerges here: the contrast between the conditioned and the unconditioned.

However his remarks about hypothetical imperatives are surprisingly confusing. Are these imperatives conditional on other oughts, ultimately on categorical oughts, or are they

---

<sup>48</sup> A739/B767, CUP version. My emphasis. See also his emphasis on the duty as well as the right to think for oneself, at the end of 'What does it mean to orient oneself in thinking?' *Religion and Rational Theology*, p. 18

conditional on actual and possible purposes or ends? Kant *seems* to say the latter. Take, then, a hypothetical ought that is conditional on an actual purpose. Can it be ‘detached’? As in:

If your aim is to buy that motor bike, you ought to rob your grandmother; your aim is to buy that motorbike, therefore, you ought to rob your grandmother.

(Suppose robbing your grandmother is the only way to acquire the necessary funds). Could such a detached ought then come into conflict with a categorical ought – ‘You ought not rob your grandmother’? Could we then argue, contrapositively, that since you ought not to rob your grandmother it is not your aim to buy that motorbike? That can’t be right.

Kant says that hypothetical imperatives are analytic.<sup>49</sup> That again is hard to understand if the general form of a hypothetical imperative is ‘If you aim to secure E then you ought to adopt means that bring it about that E’ (where E is an end).<sup>50</sup> In contrast, it is plausible if we understand the general form of a hypothetical imperative as being something like

General Hypothetical Form (GHF): If you ought to make it the case that E then you ought to adopt means that bring it about that E.

GHF will combine with a posteriori information about what means bring about E to produce empirical hypothetical imperatives. Now ‘to make it the case that E’ means the same as ‘to adopt means that bring it about that E.’ Making it the case that E just is taking steps that bring it about that E. If so, then as Kant says, “the imperative which commands him who wills the end to will the means” is analytic, as is GHF. And GHF is what Kant needs – it has the right normative form to combine with the view that there are categorical imperatives: oughts which are not themselves derived by GHF. Kant can argue that if applications of GHF can ever issue in an ought, there must be categorical imperatives.

That hypothetical imperatives are analytic is important to Kant’s overall argument, because it means that no question arises about their possibility. Clearly GHF is a priori; were it *synthetic* a priori there would be, on his principles, a task of establishing its possibility. In contrast, as one would expect, Kant holds that the possibility of *categorical* imperatives is a genuine problem which requires an answer – just because categorical imperatives must be a priori and yet are not analytic.

However Kant also assumes that any categorical ‘ought’ is an ‘ought’ of moral obligation.<sup>51</sup> This is far from obvious. In talking about categorical imperatives we are talking about ‘oughts’ or requirements of practical reason. It is far from obvious that every categorical ‘ought’ of *practical reason* is or even could be an ‘ought’ of *moral obligation*. Kant does not engage with this point at all clearly. However in one passage he sketches a distinction between relative and absolute ends – ends that I set myself (“subjectively,” at my “discretion”) and ends that are set by reason itself (“objectively”).<sup>52</sup> He says little about how the notion of ‘setting myself an end’ works. But he seems to hold that when I set myself an end I take that end to have “worth” for me, to be thereby something there is reason for me to pursue. He then says that “these relative ends are only the ground only of hypothetical imperatives.” But the distinction between relative and absolute ends ought rather to produce a distinction between relative and absolute categorical oughts, where a relative ought would be

<sup>49</sup> E.g. *Groundwork*, 4: 417, 419.

<sup>50</sup> Another way to read the hypothetical imperative is: ‘You ought not *both* to make E your end *and* fail to adopt means to bring about E’. [Broome] However this too does not explain why Kant thinks the hypothetical imperative is analytic.

<sup>51</sup> *Groundwork*, 4:416, MM 6: 223

<sup>52</sup> 4:427-8.

categorical for me, and thus give rise to hypothetical imperatives of the GHF kind for me, but not for other people. In contrast, an ought that is objectively set by reason is absolutely categorical, categorical for everyone.<sup>53</sup>

So the question is whether every absolutely categorical ought is a moral ought? Kant does assume that. Thus the exact question he faces is: how are ‘absolute’ or ‘objective’ categorical imperatives possible? And he takes this to be the same question, at least for us (see 0.0), as the question how are moral obligations possible?

### 3.2 *The Categorical Imperative (i) universal law*

Kant’s famous answer, in section III of the *Groundwork*, is that the content of morality is derivable from his definition of autonomy, or full freedom (2.4). A being that is capable of autonomy is a being that can recognise and act from reasons; it follows that such a being is capable of acting from categorical or unconditional oughts. Much more contentiously, Kant thinks we can infer from the very idea of freedom what these unconditional oughts *are*. Furthermore he holds that the inference is analytic:

if freedom of the will is presupposed, morality, together with its principle, follows by mere analysis of the concept of freedom.<sup>54</sup>

The question, then, is not how it is possible for such requirements to apply to free beings, but rather *how is it possible – justifiable – for us to see ourselves as free*.

In the *Groundwork*, as we noted (2.3), Kant accepts that we cannot know that we are free beings. Instead he argues that whenever we deliberate we necessarily *presuppose* that we are free. In the 2<sup>nd</sup> Critique he takes a rather different approach, holding that we can be immediately conscious of our moral obligations. The difference is very important: it implies distinct epistemologies of practical reason and thereby different answers to the question whether freedom presupposes transcendental idealism.

A key premise, which Kant emphatically accepts, is that moral obligations apply to us *only* if we are free.<sup>55</sup> So if, as Kant holds in the *Groundwork*, 1<sup>st</sup> Critique, and elsewhere, we cannot know that we are free (because freedom is noumenal), then we cannot know that moral obligations apply to us, and so presumably can’t be immediately conscious that they do, as the 2<sup>nd</sup> Critique suggests. If in contrast the 2<sup>nd</sup> Critique is right, and we can have immediate consciousness of moral obligation, then it seems to follow that we can know that we are free. That would be incompatible with the transcendental account of freedom. This is a large issue, both in the interpretation of Kant and for the question of how Kantian ideas are best taken forward. We shall come back to it in 3.6. But we must first consider Kant’s proposed derivation of morality from autonomy.

---

<sup>53</sup> Note however that this account produces serious difficulties given Kant’s general view of practical reason. How would he combine it with his insistence on the universalisability of reasons? And what happens when a relatively categorical ought comes into conflict with an absolutely categorical ought? It’s not obvious how to get the answer that Kant would clearly want. More generally the idea that merely adopting an end has normative significance does not fit well with the notion of autonomy, which requires that one should not simply ‘accept’ or ‘choose’ anything unless one sees reason to do so. But even though Kant was plainly not interested in these questions they are in fact crucial for his ethical theory.

<sup>54</sup> *Groundwork* 4: 447. Compare the *Critique of Practical Reason*, 5: 31, where the same claim is made.

<sup>55</sup> Thus moral obligations apply to us if and only if we are free – see 3.6.



The derivation can be broken down into two stages, of which the first claims to derive the fundamental criterion of morality – *the* Categorical Imperative – while the second claims to derive morality by application of this criterion to proposed maxims of action.

Kant has several formulations of the Categorical Imperative; the one which makes the first stage of the derivation most plausible, and which he himself focuses on, is what is called the formula of Universal Law (FUL). In this formulation, the Categorical Imperative says:

Act only on that maxim through which you can at the same time will that it should become a universal law.<sup>56</sup>

How does Kant get to that from autonomy? As he rightly holds, universality belongs to the very concept of a reason: if there is reason for X to do Y just in virtue of the fact that certain circumstances apply, there is reason for anyone to do Y whenever circumstances that are relevantly the same apply. Willing to act for a reason is committing oneself to that universal principle. In this sense, acting autonomously is willing universally: acting from a maxim which one is prepared to accept has universal application.

So understood the FUL formulation simply says: act freely. Now it's not actually clear why this imperative should follow analytically from the fact that one *is* free – what is analytic (if anything) is simply that *if* one acts freely then one acts in accordance with a maxim that one can at the same time will to be a universal law.

This point might lead one to a 'voluntaristic' reading of FUL. On this reading, acting freely is a *choice*: a choice which cannot itself be described as either free or unfree. Choosing to act freely is choosing to will universally. If you are prepared to accept that your maxim should become a universal law, then it is normative for you. Others of course might have the same universalizing commitment to some maxim of theirs, in which case their maxim is normative for them. What is normative for you (if anything) is what you are prepared to prescribe universally.

The Categorical Imperative has been read in this voluntaristic way by some, as akin to existentialism's exhortation to choose freedom (0.0.0). But it is not at all what Kant intends. His view is that if you are free you are subject to a *determinate* set of categorical requirements that apply to you *because* you are free. It is not a matter of what you choose. Further, since these categorical requirements of reason are in other words simply true propositions about what every free being should do, there is no further question (granting that you are free) about whether you should follow them. It makes no sense to ask whether one should do what one should do.

So let us go back to the idea of what you can will to be universal law. Kant says that 'what you can will' is to be understood as 'what you can will without contradiction.' His interest in the idea of contradiction reflects his thesis that morality follows analytically from freedom. The thought, if pursued, would be that adopting a maxim is accepting its universalisability; but if the underlying universalised proposition is literally self-contradictory then it is false, and so you 'cannot' will it to be a universal law.

This understanding of FUL would again invite a voluntaristic reading: any consistently universalised maxim is potentially a moral principle; to make it *your* moral principle you accept the maxim and endorse its universalised form. A voluntarist of this kind might accept that morality is, in this sense, 'derivable' from freedom.

However this cannot be Kant's position either. The 'universalisable without formal contradiction' test lets in an indefinite variety of principles, many of them inconsistent with

---

<sup>56</sup> *Groundwork*, 4: 421

each other. That's fine for a voluntarist, but Kant thinks that FUL produces a unique and well-defined system of morality.

How then did he understand the test of what 'you can will'? He gives four famous examples that are meant to answer this question:

1. Maxim: 'from self-love to shorten my own life.' Kant: when universalised into a principle, this is inconsistent because the function of self-love is to maintain life, whereas this principle would permit the curtailment of life.
2. Maxim: to borrow money on a false promise to repay. Kant: this maxim could not be universalised because if everyone followed it no trust would be placed in such promises.
3. Maxim: to neglect one's natural gifts for purposes of self-gratification. Kant: a person who willed this would be willing inconsistently "for as a rational being he necessarily wills that all capacities in him be developed, because they serve him, and are given to him for all sorts of possible purposes."
4. Maxim: not to offer assistance to others who are in need. Kant: to will that as a universal maxim would be in conflict with wanting help when you yourself need it.<sup>57</sup>

These examples, and Kant's responses, have aroused a great deal of scorn (and yet, more recently, an enormous defensive literature). Indeed it is hard not to be embarrassed by them, for example when confronted by an able first-year student impatient with special pleading.

The general objection is that Kant's responses to the maxims blatantly rest on assumptions which by his own criteria he should not be making. Furthermore, even if that point is waived, they fail to identify what is really wrong with them (if anything).

Fundamental principles of morality must, as Kant rightly insists, be synthetic a priori truths.<sup>58</sup> As a priori truths entailed by freedom they apply to all free, rational beings – so they cannot appeal to any merely a posteriori property of human nature.<sup>59</sup> Kant's discussion of the examples clearly does not meet this standard. Of course he could be right about his general framework, just disastrously wrong about the examples he chooses to illustrate it. However it's not just the examples, but his method of treating them that undermines the framework.

In practice he works with a substantive, not a formal, notion of contradiction, in which a maxim is 'self-contradictory' when it is impossible for everyone to act on it, or it is in conflict with something else you can be expected to will – given some presupposed empirical assumptions, or some presupposed normative assumptions. To introduce this substantive notion of 'contradiction' is to abandon the claim that fundamental moral principles are a priori, and *a fortiori*, the project of showing that morality follows *analytically* from autonomy.

It is, for example, not strictly impossible that everyone acts on the maxim 'Break your promise when it is to your advantage to do so.' True, in the world as it is, and given human nature, it is very unlikely that disadvantageous promises could always be broken without people knowing, and so if everyone tried to act on that maxim the institution of promises might well cease to be taken seriously. Clearly however this is a matter of our empirical circumstances and nature, not of logic – whereas an absolute categorical imperative of reason is meant to hold for all rational beings in all possible worlds. And in any case, why would the

<sup>57</sup> The examples are at 4:421-423. [Kant's asperity about the first example]

<sup>58</sup> This follows by the argument that hypothetical imperatives presuppose categorical imperatives, which in the moral case must be absolute or objective.

<sup>59</sup> See the emphasis Kant places on this point in *Groundwork* 4: 425 (ll. 12-31).

fact that if everyone followed this maxim the institution of promising would break down show it to be morally wrong to follow the maxim?<sup>60</sup>

To be sure, I can ask myself whether, given my actual, contingent needs and limitations as a human being, I could ‘will’, that is, whether I could want to live in, a society in which everyone cheats whenever they can get away with it, in which there is no mutual aid, and so on. If I would not want everyone to do that kind of thing, because of the bad consequences such behaviour would have for me, or for people in general, then I am taking unfair advantage of others if I do it myself. I am unreasonably treating myself as special. If some system of cooperation is good for everyone, then I have an obligation to make a fair contribution to it, at least so long as others do. Such an appeal to fair contribution has dignity and strength. But it does not *derive* fairness from autonomy – it presupposes it. Similar points, though somewhat different in each case, can be made about the other examples.

One may well conclude that the Categorical Imperative, at least when taken in its FULL version, and so understood as to make it derivable from the idea of autonomy, has no substantive normative content. This is the ‘emptiness’ objection, which we shall consider again when we come to Hegel’s version of it.<sup>61</sup>

Certainly no determinate principles can be deduced solely from the sound formal point that reasons are universalisable. But we should recollect that Kant is not trying to derive morality from that formal point, but from autonomy.<sup>62</sup> The crucial question is what (if anything) can a free being *autonomously* will? That is, what can be willed by a will that is free of all non-rational influences on its willing?

Acting autonomously means that one accepts no aim for one’s action, and no constraint on it, unless one sees reason to pursue that aim or observe that constraint. Anyone who accepts an aim or a constraint which they see no reason to accept does not act from reason-responsiveness alone; they are being driven heteronomously by non-rational factors, “alien causes”. This rules out an instrumental, means-end conception of rationality, according to which rationality consists solely in adopting efficient means to one’s ends. Against this conception we can ask: why *should* we pursue our ends if there *is* no reason to pursue them? This or that may be your or my end, but it remains an open question whether it should be. Strictly speaking, in fact, one should deny that an instrumental conception of rationality is a conception of *rationality*, at least as Kant conceived it: as the capacity to come to a purely reason-responsive conclusion about what one should do. Rationality thus understood – free and unconstrained deliberation about reasons – requires that we should be able to pursue reasons all the way down, never accepting an end or constraint simply as given: there must be no end, no constraint, that practical reason cannot put in question, by asking and answering whether there is reason to accept it. Otherwise we are sunk in heteronomy.

The very possibility of autonomy, then, entails the existence of categorical, not merely hypothetical, imperatives. It is the crucial thing that an instrumental conception of rationality omits. The instrumentalist maxim is ‘do whatever will most efficiently advance your actual ends’. The objection to it is not that it cannot be universalised, for it can be. The objection is that this principle simply takes ends for granted, without asking whether they

---

<sup>60</sup> To cite some favourite examples that raise this question, consider ‘Always give way to others when going through a door,’ or ‘Always pay off your credit card promptly to avoid paying interest.’ Is giving way to others or paying off your credit card promptly morally wrong?

<sup>61</sup> J.S. Mill’s often cited criticism (see 0.0.0) is similar.

<sup>62</sup> This is noted by Henry Allison, 1990, pp. 204 – 210. See also Thomas Hill, 1985.

should be adopted in the first place, and so cannot be the principle of an autonomous, fully free rational agent. For this reason too, the notion that the mere fact of setting oneself an end can produce a “relative” categorical ought (as discussed in 3.1) cannot deliver an autonomous will. The question from the standpoint of Kantian autonomy is what ends do I have *reason* to adopt? The mere fact that I have an end, or have chosen an end, does not answer that question.

So far, so good. But Kant also make another claim: that all non-moral, special or particular, interests of the agent are heteronomous. They are based on feelings, inclinations that are empirical and can at most provide “subjective” starting points for hypothetical imperatives. The claim arises, as noted in 2.3, from the fact that Kant treats them as merely phenomenal, and therefore cannot see them as genuinely responsive to reason, that is, to the causality of freedom. This, in the present context, has a crucial implication. It leads to the conclusion that universal willing must be *disinterested* in the sense of taking no account of, giving no privilege to, this or that particular interest (there being no such thing as a categorically rational particular interest). Thus universal willing, if there is such a thing, is also disinterested willing. But is there any such willing? Kant’s view has to be that there is, and that it expresses a pure disposition of the will itself, independent of all empirical feelings and inclinations. The pure will contains the disposition of impartiality.

So understood, FUL is not strictly empty – at least if we allow that the will has this pure disposition of impartiality, independent of all particular interests and expressed in autonomous willing. It may still be held to be empty in another way, however – and this will bring us to Hegel’s critique of Kant’s ethics. By excluding all particular interests as non-rational it produces (if it produces anything) a merely impartialist, cosmopolitan, account of practical reason. Love of self is ruled out from autonomous willing – but so is love of family, love of country, love of home. Well, disinterested impartiality is, we may grant, one moment of autonomous willing, but is it the only one? Particular interests are not non-rational just because they are particular, nor is it clear why only a pure disposition internal to the will, a disposition that places no normative weight on any feeling, is rational.

We can ask whether practical reason can have an inherent disposition of impartiality, or *any* inherent disposition at all. We can also ask whether practical reason is inherently and *solely* impartial. Is impartial thinking the only kind of rational practical thinking? Kant assumes that it is, inasmuch as he assumes not just that the categorical imperative is impartial but that it is the *sole source* of categorical practical reasons. In the light of 3.1, one should say: the sole source of ‘absolutely’ and ‘objectively’ as against ‘relatively’ and ‘subjectively’ categorical reasons. But it was not clear what this putative contrast amounted to, and at this point the question becomes pressing. What kind of claim do these putative ‘relative’ or ‘subjective’ reasons make in ethical deliberation?

Human interests are grounded in human feelings and the desires and goals to which they give rise. Alas, Kant does not allow that it makes sense to talk about *reasons* for feelings (as against practical reasons to try to alter our feelings); hence he has no room for principles of practical reason that are grounded on such reasons. True, practical reasons can take account of the important fact that we *have* feelings. But Kant fails to recognise that feelings are themselves reason-responsive, and that reasons to feel something can in turn generate reasons to do something – reasons for action that cannot be derived from any purely impartial standpoint, yet stand as fully-fledged reasons in their own right.

If there are these reasons for feeling and for action they are not ‘alien’ to freedom, and autonomous willing can take account of them, by allowing that distinct individuals may have distinct, fully rational, interests. A revised version of FUL might then read as follows:

FUL\* Act only on a maxim through which you can at the same time will – *after taking into account everyone's rational interests in an impartial way* – that it should become a universal law.

However this revision would require a major concession from Kant: he would have to accept that the *content* of pure practical reason cannot be derived from the idea of freedom alone. We shall pursue this point in the next section.

### 3.3 (ii) *Rational nature as an end in itself*

Kant presents the Categorical Imperative as *the* criterion of the moral law. Nonetheless he gives it more than one formulation, arguing that these formulations are equivalent.<sup>63</sup> Their point is to bring out various aspects of the fundamental criterion and make it more intuitive – though he also adds that

in moral judging it is better always to proceed by the strict method, and make the foundation the universal formula of the categorical imperative: *act according to the maxim that can make itself at the same time a universal law*.<sup>64</sup>

In fact the various formulations look very different, and this has meant that distinctly different reconstructions of Kantian ethics are possible, depending on which formulation one chooses to focus on.

In the recent revival of Kantian ethics the formulation that has come to the fore is the Formula of the End in Itself (FE):

Act in such a way that you always treat humanity, whether in your own person, or in the person of any other, never simply as a means, but always at the same time as an end.<sup>65</sup>

This is indeed a powerful statement. There is something profoundly attractive in the idea that the ground of morality is the respect due to the dignity inherent in all human (or rather, all rational) beings, in virtue of their autonomy. It seems to have greater moral substance than FUL. There is intuitive force in the way Kant applies it to his four examples. If you take your own life to avoid pain you are failing to respect yourself as an end; you are not to treat yourself, as Kant puts it, as a mere means at your own disposal. Falsely promising to repay a loan is using the other as a mere means. Neglecting to develop your natural gifts shows a lack of self-respect. In each of these three examples one can see Kant's case for thinking that FE applies, whether or not one agrees with it; the argument does not look merely empty.

The last example (offering assistance to others in need) is particularly interesting in that it touches on a long-standing line of disagreement in philosophical ethics. Kant sometimes suggests that to see others *as* ends is to seek to advance *their* ends – so far as possible as though they were one's own.<sup>66</sup> Yet one might well think that not *using* others as means is rather a matter of respecting their rights and their free use of those rights, than of pursuing their ends as though they were one's own. The latter emphasis has much more the ethical atmosphere of utilitarianism, whereas many admirers of Kantian ethics admire it

---

<sup>63</sup> 4:436, 4:437

<sup>64</sup> 4:436

<sup>65</sup> *Groundwork*, 4: 429. It is sometimes called the formula of humanity; however Kant explains that it requires us to treat *all* rational beings, impartially, as ends and not merely as means – “man, and in general every rational being, exists as an end in itself ...” *Groundwork*, 4: 428. “Formula of the End in Itself” is Paton's terminology.

<sup>66</sup> MM ...



especially because it offers, in their view, a sober alternative to utilitarianism. There will be much more to say about this line of disagreement when we have set out the utilitarian tradition in ch 00, in particular in comparing the ethics of Kant and Mill.

More generally, the view that everyone is capable of freedom and that respect for that capacity is the basic value underwrites an egalitarian ethics of respect. And indeed much in Kant's substantive ethics has exactly that quality, as we shall in the next two sections. But we should keep track of Kant's own strategy. What grounds does Kant himself give for FE?

It can easily seem that with this formulation he is giving way to a teleological view: principles of morality must be derived from some final end which we simply acknowledge as having absolute worth. And indeed he says:

suppose there were something *the existence of which in itself* has an absolute worth, that as an *end in itself*, could be a ground of determinate laws, then the ground of a possible imperative, i.e. of a practical law, would lie in it, and only in it alone.

Now I say: a human being and generally every rational being *exists* as an end in itself, *not merely as a means* ...<sup>67</sup>

That certainly looks like an argument to morality from a pre-given absolute end. Moreover in the very next paragraph, which leads to the statement of FE, he seems to argue from the premise that everyone necessarily regards *themselves* as an absolute end to the conclusion that everyone just is an absolute end, i.e. an absolute end to everyone.<sup>68</sup> This bears an uncanny resemblance to Mill's notorious 'proof' of the principle of utility (in Kant's case the end is *myself*, in Mill's it is *my happiness*) and it is open to the same objections, which will be considered in 0.0.0. It does not fit Kant's official method, which is to derive the moral law from freedom.

In line with that method, Kant's considered position is that FE is just a variant formulation of the Categorical Imperative, and that the Categorical Imperative is deducible from freedom. To act freely is to act rationally; a rational agent is an agent capable of so acting, and hence subject to the Categorical Imperative. Thus if the Categorical Imperative applies to us it does so because we *have* that capacity, because we *are* rational agents—not because we 'value' rational agency.

One can get an idea of what Kant has in mind if one understands "end," as it appears in FE, to mean 'being that can act for itself from reason, autonomously'. A rational being is an *end* in the sense that it does not belong merely in the domain of *means*, as an *object* of practical reasoning, but is itself a practical reasoner, a giver of law as well as subject to law. FE then says that you should act in a way that always takes into account the law that such a being could rationally issue.<sup>69</sup> How? By taking that being's rational interests impartially into account in the law you issue by your willing. So FE is interderivable with FUL, at least when the latter is read as FUL\*, that is, as allowing that a rational being can have its own specific rational interests, and further, on an assumption we'll come to in the next section, namely, that there always *is* an impartial way of reconciling conflicts of rational interests.

On this reading the claim that a rational being is an end is not a value judgement from which the moral law is *derived*. (After all what could it be for a *person*, as against a state of affairs, to be an end in itself?) On the contrary, as Kant insists, the moral law – what we can will to hold universally – remains basic, whereas the idea of good and evil is derived:

<sup>67</sup> 4: 428 ll. 7- 8, cp ll. 34-39.

<sup>68</sup> 4: 429 ll. 39-45.

<sup>69</sup> "The formula of the universal law requires that no rational being be subject to a maxim [law] that could not arise from its own will." Sensen p. 113.

*the concept of good and evil must not be determined before the moral law (for which, as it would seem, this concept would have to be made the basis) but only ... after it and by means of it.*

... instead of the concept of the good as an object determining and making possible the moral law, it is on the contrary the moral law that first determines and makes possible the concept of the good, insofar as it deserves this name absolutely.<sup>70</sup>

Kant sees the apparent oddness of this: he describes it as a “paradox of method.” His argument is that to start with some conception of the good as the criterion of the moral law would fail to establish the aprioricity of the moral law.

His thought is that to *start* from something we value as a final end would be to start from something a posteriori – the fact that we value it (like it, etc). But why couldn’t the starting point be an a priori insight into what is absolutely good? At this point we come back to Kant’s transcendental framework. A synthetic a priori principle must be vindicated by a transcendental argument. Thus the synthetic a priori elements in theoretical reason were vindicated by appeal to forms of intuition and categories of the understanding. The same applies in the realm of practical reason. We cannot simply ground claims about value on their apparent self-evidence. We need some form of transcendental argument. In the case of practical reason that is provided by showing that in deliberating and deciding we necessarily take ourselves to be free, that in taking ourselves to be free we are not assuming something impossible, and finally, by deducing a priori practical *laws* – not *values* – from the very idea of freedom.

In Kant’s hands this strategy has an important feature to which I have already drawn attention. It reduces all particular interests to a non-rational status, leaving only the idea of universal impartial law. That is Kant’s official position. And yet, in considering the interderivability of FUL\* and FE, it has been essential to introduce the idea of specific and potentially divergent rational interests that different persons may have. These rational interests will in turn rest on the sort of beings they are. Thus, given the sort of beings *we* are, it is quite reasonable, for most of us, to want to have a principle of mutual aid generally adopted. On this expanded view a being that wills universally should take the possibility of diverging rational interests into account in its law.

However, as noted, Kant’s formal framework has no *room* for the account of reasons required to underpin such notions as ‘what a particular person reasonably wants’ or ‘the distinctive rational interests of a particular person’. Nonetheless he implicitly relies on some such account: both FUL\* and FE presuppose it. Acknowledging this explicitly would mean (a) giving up the alleged derivation of the Categorical Imperative from the *sole* idea of autonomy and (b) giving up the idea that all categorical principles of practical reason are derived from the Categorical Imperative. For if this appeal to individual rational interests, as in FUL\*, were made explicit, it would become clear that some normative principles are not derived from the Categorical Imperative but, on the contrary, presupposed by it. From Kant’s point of view this would be quite a retreat. But it might be a sensible retreat. Furthermore Kant has the essentials of an epistemology of reasons, grounded solely in spontaneity and discussion (2.2), that would allow for it. We shall come back to this in 4.3.

### 3.4 (iii) *Autonomy and the kingdom of ends: self-legislation and impartiality*

<sup>70</sup> CPR, 5:63 (Kant’s emphasis); 5: 64. Consider also 4: 436 “nothing has any worth other than that which the law determines for it.”

To act freely is to act from universal impartial law. Kant has a very striking way of putting this: it is to act from *a law we give ourselves*. That is the point of calling full human freedom *autonomy*. What then is it to ‘give oneself the law’?

Kant appeals to this notion in a third formulation of the Categorical Imperative which he calls “the principle of autonomy”. This calls on us to act under “the idea of the will of every rational being as a universally legislating will”.<sup>71</sup> The formulation returns us to the fact that Kant takes impartiality to be inherent in the Categorical Imperative, rather than to be a corollary of the formal universality of reasons. A rational will – any rational will – has the capacity to deliberate from a standpoint that is impartial, and not just from hypothetical imperatives grounded on “subjective” ends or interests. As we also noted, this entails a substantive assumption, which is that a purely disinterested will would still retain a disposition to will something determinate, rather than nothing at all. This disposition would ground the standpoint of impartiality. Since rational beings (Kant thinks) can take up the impartial standpoint just in virtue of their rationality, they can think of themselves as universal legislators.

The idea of universal legislation in turn “leads to a very fruitful concept attached to it, namely that of *a kingdom (Reich) of ends*.” A kingdom of ends is a “systematic union of several rational beings through common laws”<sup>72</sup> based on the fundamental principle that every rational being is an end – both a ‘legislator’ and a subject whose rational interests must be taken into account by every other ‘legislator’.

These two ideas – self-legislation, and the kingdom of ends – have proved profoundly suggestive, not least because they lend themselves to many readings. An affinity with Rousseau’s concept of the general will is evident, and gives them political as well as ethical significance. They structured the development of German idealist ethics. They have been powerfully revived in the thought of John Rawls.

None of this means that they are obvious or that they should be uncontroversial. Critics of the French Revolution who have found its abstract humanism dangerous have also found these Kantian conceptions dangerous. Take the notion that positive freedom is autonomy. Does it mean that a law which we *would* agree to if autonomously legislating can be said to *be our* law – whether or not we actually, empirically, agree with it? In his political philosophy Kant writes as though it does (3.8). Many philosophical liberals, through to Isaiah Berlin (0.0.0), have seen in that idea an opening to tyranny. (‘I am only acting in accordance with what is really your will,’ says the oppressor.) The question is undoubtedly important; we shall be coming back to it more than once, for appeals to, and worries about, *self-legislation* and *impartiality* are basic to late-modern political philosophy.

What did they mean to Kant? He glosses the “principle of autonomy” as follows: According to this principle, all maxims are rejected that are not consistent with the will’s own universal legislation. Thus the will is not just subject to the law, but subject in such a way that it must also be viewed *as self-legislating*, and just on account of this as subject to the law (of which it can consider itself the author (*Urheber*) in the first place.<sup>73</sup>

The parenthesis – “of which it can consider itself the author” – has sometimes been taken to show that Kant thinks the will (or, reason, as in Groundwork 4:448) *is* the author of the moral

<sup>71</sup> 4:433, 4:432.

<sup>72</sup>

<sup>73</sup> 4:431. Cp 4: 448: “Reason must look upon itself as the author (*Urheberin* – “authoress”) of its own principles independently of alien influences.” [Wood p. 161]

law. This again has led to constructivist or voluntarist interpretations of his view of morality, and (more plausibly) to interpretations that approach it to Rousseau's version of contract theory.

However to say that will or reason "can consider itself" or even "must look upon itself" (4: 448) as the author of the moral law is not to say that it *is* the author. Will (*Wille*), or equivalently, practical reason, does not literally do things in the way that an author does things, or that a law-giver does things. These are metaphors. In contrast, when Kant uses the terms 'author' and 'legislator' literally he consistently *distinguishes* between author and legislator and denies that anyone is or could be the author of the moral law:

One who commands (*imperans*) through a law is the *lawgiver* (*legislator*). He is the author (*autor*) of the obligation in accordance with the law, but not always the author of the law. In the latter case the law would be a positive (contingent) and chosen [*willkürlich*] law. A law that binds us a priori and unconditionally by our own reason can also be expressed as proceeding from the will of a supreme lawgiver . . . but this signifies only the idea of a moral being whose will is a law for everyone, without his being thought as the author of the law.

Whilst it is true that the moral laws are commands, and whilst they may be commandments of the divine will, they do not originate in the commandment. God has commanded this or that because it is a moral law, and because his will coincides with the moral law . . . No-one, not even God, can be the author of the laws of morality, since they have no origin in will, but instead a practical necessity. [God is] the lawgiver, though not the author of the laws. In the same way, God is in no sense the author of the fact that the triangle has three sides.<sup>74</sup>

Plainly, then, we are not literally the authors of the moral law, any more than God is. For that would make it "a positive (contingent) and chosen law," whereas the moral law "binds us a priori and unconditionally by our own reason." In what sense, then, do we give the moral law to ourselves?

No legislator originates the "practical necessity" of the moral law, let alone its normative content. Nonetheless a legislator, in Kant's conception, can give a law of reason a distinctive obligatoriness stemming from his legitimate will. He *discerns* the true content of an a priori requirement of practical reason and *lays it down, prescribes* it to someone as obligatory. Likewise then the self-legislator discerns the true content of the law and lays it down for himself as obligatory.

This is not a voluntarist conception. Kant's point is that we are able to discern the full content of the moral law *ourselves*, by exercise of our own reason without being told by anyone else, and can effectively 'lay it down,' 'prescribe it' to ourselves – that is, bind ourselves to act from recognition of it.

Our sense of the moral law as *law*, as *obligatory*, arises from our own finite sensuous nature. This nature can incline us *against* requirements of practical reason, yet it is also *in virtue of that nature* – through our capacity for respect when confronted with the content of practical reason – that we experience these requirements as binding. We actively impose principles of practical reason on ourselves, because in many situations we have to override personal inclination in order to act from them.

---

<sup>74</sup> *Metaphysics of Morals*, 6:227; Kant, *Lectures on Ethics*, tr. Louis Infield (NY, NY: Harper and Row, 1963) 40, 51–2) Cp 27: 283 (CUP *Lectures on Ethics*, p. 76)

The importance for Kant of this contrast between author and legislator is confirmed by the comparisons he makes between a “holy will” and the will of human beings:

For human beings and all created rational beings moral necessity is necessitation, that is, obligation, and every action based on it is to be represented as duty, not as a kind of conduct which we already favour of our own accord or could come to favour - as if we could ever bring it about that without respect for the law, which is connected with fear or at least apprehension of transgressing it, we of ourselves, like the Deity raised beyond all dependence, could come into possession of *holiness* of will by an accord of will with the pure moral law becoming, as it were, our nature, an accord never to be disturbed (in which case the law would finally cease to be a command for us, since we could never be tempted to be unfaithful to it).<sup>75</sup>

The normative *content* of the moral law is an objective requirement of reason: a holy will recognises that content as such and effortlessly acts on it. It does not experience the moral law as a law: it acts from its content without experiencing it, as we do, under the idea of obligation or duty.<sup>76</sup>

There is a supreme practical principle that is valid for all rational beings. No-one is its author; in fact Kant’s official position is that it follows analytically from the very idea of freedom. We imperfectly rational beings experience it as a categorical *imperative*: as on the one hand something that imposes on us a burden that our inclinations resist, and on the other as something that inspires in us a distinctive, humbling reverence or respect (*Achtung*). This complex emotional response plays, Kant thinks, a crucial psychological role in making it possible for us to act autonomously even in the face of recoil. Respect for what we experience as moral obligation clears the path through resisting inclinations – or brings the pain of loss of *self-respect* when it fails. It reduces the influence of self-love, and it *strikes down* self-conceit altogether, since all claims to esteem for oneself that precede accord with the moral law are null and quite unwarranted, because certainty of a disposition in accord with this law is the first condition of any worth of a person ...<sup>77</sup>

Respect for the law is thus connected with *self-respect*: in the absence of a reliably moral disposition other bases of self-esteem fall away. Furthermore, while Kant thinks that experiencing the force of our own reason, that is, its imperatival effect on us, is elevating, he also thinks that it necessarily contains an element of burdensomeness or constraint: “we stand under a *discipline* of reason” though it is “our own reason” that gives it.<sup>78</sup> To imagine that we humans could transcend the moment of constraining discipline contained in the experience of moral obligation is, according to Kant, to think that we could be beings who do not experience practical oughts as moral imperatives at all, but simply and readily act on their rational content. That is possible for a holy will, which, experiencing none of this emotional complexity, acts simply and straightforwardly from what it recognises as reason. But to think it possible for *us* is to usher in ‘*moral enthusiasm* instead of a sober but wise moral discipline’.<sup>79</sup>

This is a rather perceptive phenomenology of moral action for human beings, so far as it goes. But is it the whole truth? If – unlike a holy will – *we* have to *make* ourselves act freely can we be said to act fully freely? Is fully free action, as against mere continence or

<sup>75</sup> (CPracR, Cambridge 70) Other references.

<sup>76</sup> Cp Paton, p. 138: 66 n. 1.

<sup>77</sup> Kant, *Critique of Practical Reason*, 5:73.

<sup>78</sup> *Ibid*, 5:82.

<sup>79</sup> *Ibid*, 5:86.



self-mastery, compatible with constraint? Does self-mastery, however sublime, have the beauty of fully free action? These questions lead us to Schiller and Hegel, neither of whom (though in different ways) thought that Kant had captured what it is to be *fully* free. Kant's picture of autonomy involves what he regards as a necessary conflict between our noumenal will and our empirical feelings and inclinations: the one has to master the other, even if it does so partly by itself relying on the aid of a feeling: the feeling of respect for the moral law. *Only* an action from autonomy, thus understood, has moral worth. Kant had time to respond to Schiller on these matters.<sup>80</sup> The debate ramifies widely, and we shall be considering it again (0.0.0).

To return to the role played in Kant's theory by the notion of impartiality. Giving universal law is willing impartially. Each of us, he thinks, is capable of recognising the demands of impartial practical reason: it is this that qualifies us to take up the role of universal legislator. And, certainly, impartiality plays a vital role in ordinary ethical thinking, where it means something like seeking a solution in abstraction from one's own preferences and inclinations, as for example when one tries to come to an impartial verdict in a conflict of interest where one's own interests are not involved, or again, where one does not know what the effect on one's own interests will be.

The question we have raised, however, is whether practical reason is inherently and solely impartial. Is impartial thinking the only kind of rational practical thinking? Kant assumes that it is, inasmuch as he assumes not just that absolute categorical imperatives are impartial but that they are the *sole source* of categorical practical reasons.

If all principles of practical reason are impartial then they cannot themselves lead to conflict. Suppose, in contrast, that there are objective principles of practical reason that are not impartial, but agent-relative. We have, for example, feelings of special attachment to significant others such as family, neighbours, and indeed ourselves. Are these not perfectly justified? Do they not give rise to reasons for action in their own right? If they do, then our actions should be rationally guided by two kinds of reason – considerations of impartial right, that treat everyone equally as an end, and agent-relative reasons that stem from our perfectly justified special attachments. These are not two distinct, ordered classes: the superior class of principles of practical reason proper, and the under-class of 'reasons' based on feelings, whose status as normative reasons is at best murky. The two kinds of reason are, on the contrary, fully on a par as practical reasons, guiding rational deliberation. My *rational* pursuit of my agent-relative reasons may come into conflict with your equally *rational* pursuit of yours.

If this is so, we cannot assume that the interests of rational people will always be consistent. Nor can we assume that impartial principles can always provide a rational solution when the rational interests of different people come into conflict. They may or they may not. There is then no a priori guarantee of the possibility of *normative harmony* among human beings.

This also means that there is no guarantee of the possibility of a 'kingdom of ends,' and therefore no guarantee that that notion can provide a criterion of moral obligation. When Kant says that "morality ... consists in referring all action to the legislation by which alone a kingdom of ends is possible"<sup>81</sup> he assumes that there is a morally correct solution for every conflict between rational actors – which means, given the other things he says, a rationally

---

<sup>80</sup> In *Religion Within the Boundaries of Mere Reason*, 6: 23, note. See Allison 1990, pp 180 – 184.

<sup>81</sup> 4: 434

correct solution whose principle any rational being can see. But cannot conflicts between rational people arise in which no resolution is attainable by reason? That this is possible is part of a powerful tradition that takes the foundation of ethics to be rational self-interest. It is a tradition present in modern ethics since Hobbes. But one can put the issue more generally – not just in terms of impartiality versus self-interest. Any view that allows for irreducibly agent-relative reasons of any kind – not just self-interest – must allow that rational conflicts may not be resolvable by impartial reasons. This includes mixed views that allow for impartial, agent-neutral reasons for action as well as agent-relative ones. Take the old case of two people after a shipwreck, both heading for a floating spar which can save only one of them. If impartial reasons are the only ones involved then the two swimmers should be indifferent as to who gets the spar. But if self-interest is also rational then we have rational conflict.<sup>82</sup> Similarly if two parents can each save only their own child, in a situation where it is impossible to save the children of both. Here we are not talking about *self-interest*, but if each parent has an irreducible reason to save its own child we may still have irresolvable rational conflict.

Let us put the two issues, about self-legislation and impartiality, together. On the one hand Kant is empirically realistic about not treating human wills as holy wills. On the other hand he is philosophically unrealistic about the unlimited scope of impartiality. The viability of his ideal of a kingdom of ends is open to both an empirical and a philosophical objection.

Anyone who holds the mixed view, that there are impartial, agent-neutral reasons for action but also agent-relative ones which can and do come into conflict, may suspect that Kant's ethics provides a dangerously misleading basis for political action. Politics is pre-eminently concerned with conciliation of interests, whether by compromise or force, and this must always be so. That is in large part because we do not live in a world of perfectly rational people. But – given enough bad luck – it would remain so even if we did live in such a world. The impartialist, agent-neutral ideal of normative harmony in a possible kingdom of ends ignores that. [Connect with discussion of 'fanatical impartialism' in chapter II and with Sidgwick discussion in 0.0.0.]

### 3.5 *The honour code of humanity*

Self-legislation and impartiality are linked to another ideal that mattered greatly to Kant, the ideal of equal respect for all human beings. (In theory for all *rational* beings, but in practice this is a humanistic ideal.) The basic form of respect for Kant is, as we noted, respect for the moral law. From this arises respect as the distinctive attitude that is proper towards someone who does his or her duty, observes the imperatives of morality. The latter attitude is implicit in Kant's phenomenology of moral experience: to experience a ground for action as a moral obligation is to experience it as something one cannot violate on pain of loss of *self-respect*.

What should we call this kind of respect – respect for conscientiousness, dutifulness, righteousness? Nowadays these words have connotations of self-satisfaction, of mere rule-following and so forth. We might prefer to talk, for example, of respect for moral integrity. This is interesting in itself: 'integrity' – like other words such as 'authenticity' – is a concept that reveals how our ethical culture has absorbed criticisms of Kant and directions taken away from him. Yet the older words pick out the exact quality that in Kant's opinion deserves respect – namely acting from duty.

Most of us would readily grant that people who can be counted on to do the morally right thing, for the reasons that make it morally right, deserve a kind of admiration that is

---

<sup>82</sup> For Kant's brief and not very helpful remarks about this case see MM 6: 235-6

distinctive, and that exactly fits the name ‘respect.’ Whether this is the only kind of virtue or human excellence is another question, to be pursued when we come to the Schillerian ideas that were mentioned in 3.4. At any rate ordinary moral attitudes certainly expect respect for the moral law: a disposition to conform to it, even, in a sense, to fear it (as a sailor’s respect for the sea involves a fear of its power). Good people and not so good people can feel something like this attitude towards a moral obligation, a feeling of its uncompromising importance, unyieldingness, necessity. The two kinds of respect, respect for the moral law and respect for moral worth, go together and strengthen each other.

Kant also thinks of respect in a third way, as the distinctive attitude that is appropriate towards any being that has the *capacity* to make and act on moral valuations – a capacity which Kant thinks all human beings have:

a human being regarded as a *person*, that is, as the subject of a morally practical reason, is exalted above any price; for as a person (*homo noumenon*) he is not to be valued merely as a means to the ends of others or even to his own ends, but as an end in itself, that is, he possesses a *dignity* (an absolute inner worth) by which he exacts *respect* for himself from all other rational beings in the world. He can measure himself with every other being of this kind and value himself on a footing of equality with them.<sup>83</sup>

This is respect for persons as moral subjects – it is respect for human beings as each capable of giving themselves the moral law.

How much does this third idea of respect coincide with moral common sense? It may seem to overlap with an ideal of civic equality, of treating everyone with dignity, which most people would accept. We shall be returning to this ideal more than once. However, civic equality does not require the noumenal and absolute basis that Kant gives to respect for persons, as in the passage just quoted. If that noumenal basis is confused, then the ideal of civic equality is better off without it.

More abstractly, Kant’s notion of respect for persons picks out rationality as the sole basis of moral standing. And we can agree that some distinctive moral rules govern our behaviour towards rational agents precisely in virtue of their rational agency. Most obviously, rational competence underpins the right to decide for oneself over a large sphere of action that does not affect the interests of others. This is respect for the *rights* of persons. To make rationality the sole basis of moral *standing* is to go much further. Is the capacity to act morally the only thing on which the status of human beings as ends in themselves depends? Don’t animals have any moral claims in their own right as animals? Doesn’t the natural world deserve respect in its own right?<sup>84</sup> Or, to raise a different kind of question about Kant’s individualistic humanism: while human beings have a legitimate importance to themselves, from their own point of view, are they not part of a larger whole in which they have their place, and an importance relative to that place, but in which only the whole has *absolute* importance? This would be Hegel’s view.

Cogent questions can also be asked about Kant’s egalitarianism of respect. Is autonomy an all or nothing quality? Does every human being have it in equal measure? Kant has two ways to go here. First he can take refuge in his transcendentalism: transcendently speaking, every human being is an absolute moral subject, even if empirically human beings vary greatly in their moral worth. This is the abstract route that makes transcendental rational agency the only end in itself. Second he can emphasise, as he also does, that awareness of the

---

<sup>83</sup>Kant, *Metaphysics of Morals*, 6: 434-5.

<sup>84</sup> Kant on animals.

moral law is not a matter reserved to outstanding intellects, and that acting from it does not require outstanding will power. Moral worth is a quality that the simplest human being can have, and in a Rousseauesque twist, may be more likely than a sophisticate to have. This is the humanistic route; it puts faith in the simple human goodness that it takes to be immanent in every human being.

It is illuminating to see how Kant's notion of respect looks if we start not from common-sense morality but from a certain enlightenment ideal: call it *the honour code of humanity*. An honour code associates fine or noble action with honour, esteem and pride, ignoble action with dishonour, contempt and shame. Much of what Kant says about servility and self-respect mobilises these feelings, while it at the same time detaches nobility of character from the idea of social rank with which it is ideologically fused in real-life aristocratic codes of *noblesse oblige*. Kant invokes an honour code for Everyman – every person, moral subject, representative of humanity:

A human being can ... be an object of my love, fear or admiration even to amazement and yet not be an object of respect. His jocular humour, his courage and strength, the power he has by his rank among others, could inspire me with feelings of this kind even though inner respect toward him is lacking. Fontenelle says "*I bow before an eminent man, but my spirit does not bow.*" I can add: before a humble common man in whom I perceive uprightness of character in a higher degree than I am aware of in myself *my spirit bows*, whether I want it or whether I do not and hold my head ever so high, that he may not overlook my superior position... *Respect is a tribute that we cannot refuse to pay to merit, whether we want to or not; we may indeed withhold it outwardly but we still cannot help feeling it inwardly.*<sup>85</sup>

This is moving and true. The respect that is due to moral merit, uprightness of character, is distinguishable from all questions of rank, and from all other excellence. Respect of this most unchallengeable kind is due to *any* person who grasps their moral obligations and acts, even in adverse circumstances, on what they grasp.

Honour codes confer self-respect on the basis of one's standing and one's ability to live up to it. Respect for the standing and respect for the individual is connected through the code associated with the standing, by the idea that someone whose standing places them under a code that deserves respect deserves respect. In the ideal type of an honour code the self-respect that distinguished standing confers then arises not just from *having* that standing but from *living up* to its obligations. Essential to the point is the contrast to those who don't have the standing and thus don't have the obligations.

For Kant the capacity to respond to practical reason puts you on a footing with every other being that has that capacity. Our standing is our standing as moral beings *in contrast to the rest of nature* – as distinguishing us noumenally from nature. This is the romantic heroism of noumenal freedom. Respect for noumenal freedom is related to respect for personal conscientiousness in exactly the way that an honour code links a standing to success in living up to that standing. Honour, or respect, results from living up to one's standing, shame and disgrace attend failure.

Kant connects the two forms of respect – respect for moral capacity and respect for moral conscientiousness – in this way in his fascinating discussion of servility in *Metaphysics of Morals*:

This duty with reference to the dignity of humanity within us, and so to ourselves, can be recognised, more or less, in the following examples.

---

<sup>85</sup> CPracR, 5:77.

Be no man's lackey. – Do not let others tread with impunity on your rights. – Contract no debt for which you cannot give full security. – Do not accept favours you could do without, and do not be a parasite or a flatterer or (what really differs from these only in degree) a beggar. Be thrifty, then, so that you will not become destitute. – Complaining and whining, even crying out in bodily pain, is unworthy of you, especially if you are aware of having deserved it; thus a criminal's death may be ennobled (its disgrace averted) by the resoluteness with which he dies. – Kneeling down or prostrating oneself on the ground, even to show your veneration for heavenly objects, is contrary to the dignity of humanity ... one who makes himself a worm cannot complain afterwards if people step on him.<sup>86</sup>

The distinctions Kant makes here, between humility and servility, arrogance and self-respect, appeal to a discipline and ideal of the noble. In a group regulated by an honour code the code is the fundamental object of reverence. Self-respect is founded on the degree to which one knows oneself to have lived up to it, humility on the fact that one can never do so fully. One measures oneself by the code to which one is bound and by no other standard.

Kant writes about morality in this vein. Humility, he says, as against servility, arises not from comparing ourselves to other people but from judging ourselves by the moral law. Arrogance involves inappropriate comparison of oneself with others, self-respect, the judgement that one has succeeded in following the code.<sup>87</sup> His strategy of persuasion is to put every person, every representative of humanity, on their mettle by appeal to an honour code that gives to all the “sublime vocation” of rational beings. Mere nobility of birth or social power become utterly irrelevant. *All* of us are placed distantly above – incommensurably above – the merely natural domain. This is Kant’s ideal, or honour-code, of humanity.<sup>88</sup> It is effective against honour codes that are based on social privilege precisely because it drives them off the ethical ground to which they lay claim.

There is obviously a good deal more to Kant’s ethics than this, and which does not depend on this – as we have seen. Yet the honour code of humanity is very important to its appeal. Within late modern ethics there is a current of individualist humanism, a commitment to the dignity of all human beings simply as human beings. Kantian transcendentalism provides this humanism with a cosmic setting which can have genuinely inspiring effects. It can also seem hollow or quixotic. Can this ideal, with or without transcendentalism, or any

---

<sup>86</sup> Kant, *Metaphysics of Morals*, 6:436-37. Compare 6:463: “I cannot deny all respect to even a vicious man as a man; I cannot withdraw at least the respect that belongs to him in his quality as a man, even though by his deeds *he makes himself unworthy* of it” [my emphasis]. Kant proceeds to condemn certain punishments as disgraceful because they “dishonour humanity itself.” Similarly, an aristocratic society will hold that an aristocrat deserves a certain respect just in virtue of his status as an aristocrat, despite his wrongdoing, and shield him from certain punishments on the grounds that they would dishonour the aristocratic status itself.

<sup>87</sup> *Ibid*, 6:435-6.

<sup>88</sup> Manfred Kuehn, (Kuehn 2001, p.41-43), rightly highlights Kant’s hostility to *aristocratic* notions of honour. Nonetheless, as we see, Kant himself appeals to honour categories, such as that of ‘distance’ and ‘standing’, in setting out his ideal of humanity. For example in treating lying, or suicide, as violations of duties to *oneself* (rather than as wrong if and when they let *others* down), he sees them as incompatible with one’s standing as a noumenal law-giver. This draws on a disposition we already have to see these acts as dishonourable (‘I am not the sort of person who does that sort of thing’).



honour code, however universal, accomplish the continuous humdrum work of constituting and reconstituting moral life? Or does that require something closer to human bedrock than the late modern ideal of the dignity of humans simply as humans? Something that is closer to the moral sentiments and the reciprocal relations of recognition, ethical punishment and reconciliation which these sentiments mobilise and direct? We shall come back to these questions.

[Send this para to 0.0.0? Eg Hegel, or Nietzsche? Taking the honour code out of its social context, in which *status* – hierarchy and membership – are essential, abstracting it and universalising it but still appealing to its ethical force. H&N: put it back. ]

### 3.6 *The postulates of freedom, immortality and God*

Kant's ethics combines an objective account of the content of practical reason with a subjective account of the impression that it makes on us, that of imperative moral law. The impression arises from the impact of pure reason on our feelings, and through them on our will to obey. It does not require an explanatory appeal to God as law-giver. Furthermore, Kant frequently insists that the practical *content* of reason does not require an explanatory appeal to God as author. It is true that thinking of the moral law as the law of God can provide incentives to moral life, some admirable, some less so. But the normative authority of the moral law does not depend on thinking of it that way. Add to that Kant's critical analysis of theoretical arguments for the existence of God, which concludes that there can be no such argument: theoretical reason cannot tell us that there is a God. Overall it follows that neither our knowledge of the world nor the authority over us of moral law presuppose that there is a God.

These conclusions are central to Kant's Critical philosophy, arising from its most fundamental tenets. However they tell only one side of the story. For Kant does think that we have reason to believe in God. The basic thought throughout is that faith in God does not rest on theoretical reason but is a "need" of practical reason. A very large part of his writing is devoted to defending the reasonableness of religious faith in these terms – in all three Critiques, in *Religion within the Boundaries of Mere Reason*, and in other writings. The contrast, repeatedly, is between a reasonable religious faith reached through the moral law on the one hand, and, on the other, doctrinaire religion, whether based on bad metaphysics or dogmas derived from the bible or church tradition.

Within that framework Kant has many serious, one could say devout, things to say about Christian spirituality. We must take the attitude and content of his religious faith seriously – even as we recognise that to many, both among his contemporaries and in subsequent decades, his analysis of religion seemed to destroy faith rather than save it.

Two important notions in Kant's thinking have already been noted: the notion of a *regulative principle* and the notion of a *postulate*. A regulative principle provides heuristic guidance and inspiration for theoretical inquiry; but it is neither constitutive of, nor a conclusion of, such inquiry, nor is it a substantive assumption presupposed by the inquiry. Unlike the inquiry it regulates, it makes no ontological claims. Although Kant does suggest that the Absolute Ground principle and the principle of Conditionality can be seen as regulative principles, in practice, as we saw, he handles them differently. He takes Conditionality to be a constitutive (synthetic a priori) principle in our construction of the empirical world, and he takes the Absolute principle as a pure principle of reason which applies at the noumenal though not the empirical level. This is implicit in one of his

arguments to things in themselves, and appears in his discussion of the antinomy of pure reason; it is also in play in his argument from conditioned to unconditioned value. [5:107 – 8]

Where Kant does treat the Absolute principle as a regulative principle for empirical inquiry he takes its function to be that of animating a continuing search for broader and simpler explanations. In the 3<sup>rd</sup> Critique the idea of a Divine designer is invoked in this way: the hypotheses we make about nature can be usefully pulled together by a regulative principle that directs us to think of the laws of nature as products of an intelligence, an intelligence that works by methods to which our own thinking has an affinity. In this role, the idea of a Divine designer has, according to Kant, heuristic value; nonetheless it is not constitutive in our construction of the empirical world of nature, nor is it a conclusion of our inquiry into nature, nor is it a *logically* indispensable presupposition of our inquiry. Such an ‘as if’ status does little for the reasonableness of literal faith and is not invoked by Kant in that role.

Contrast the idea of freedom as it appears in the *Groundwork*. It is not a merely regulative principle: it is an ontological claim. It is not that we think of ourselves *as if* we were free. No: whenever we deliberate about what to do we necessarily take ourselves to *be* free. This is a logically indispensable presupposition; transcendental idealism then establishes the possibility that we actually are what we take ourselves to be.

This argument can get to the indispensability and coherence of the assumption of freedom – but it cannot give us *knowledge* of its truth. If a postulate is to be distinguished from a regulative principle on the one hand and a synthetic a priori truth on the other, the idea of freedom is exactly that. Practically we have to think of ourselves as free, but since freedom is noumenal we cannot know that we are free, although we can know that it is not incoherent to think of ourselves as free.

Now if our freedom is a practical postulate in this sense, and if it is analytic that the content of the moral law applies to a being if it is free, then Kant can legitimately hold that the proposition that the moral law applies to us is a practical postulate in the same sense. We considered this line of thought, in the strong form Kant gives it, in 3.2. Its effect is to take not just our freedom but our moral obligations to be practical postulates that we are committed to whenever we deliberate about what to do.

Kant also holds that if the moral law applies to a being then that being is free. (This produces his “reciprocity thesis:” a being is bound by the moral law if and *only* if it is free.) So if the moral law applies to *us* then *we* are free. The implication depends on the thesis that ‘ought implies can.’ If there is something that you ought to do then it must be in your power to do it – a plausible principle at least for uses of ‘ought’ in which ‘ought’ expresses a duty.<sup>89</sup>

This gives Kant another possible argument to our freedom, which some commentators have highlighted. By ‘ought implies can’, the proposition that we are bound by the moral law entails that we have the positive freedom to act from it. So if we know that we have moral obligations then we know that we are free.

But do we *know* that we have moral obligations? In the *Groundwork* the answer is clearly ‘no’. That we are bound by the moral law is a conclusion we infer from the postulate of our freedom; hence it has the same status as that postulate – we are committed to it whenever we deliberate but we cannot know it.

In the second Critique however, and elsewhere, Kant’s answer may be ‘yes.’ This reading centres on what Kant there refers to as “the fact of reason”:

The moral law is given, as it were, as a fact of pure reason of which we are a priori conscious and which is apodictically certain...

---

<sup>89</sup> Stern 2004.

and

The consciousness of this fundamental law may be called a fact of reason, since one cannot ferret it out from antecedent data of reason, such as the consciousness of freedom.<sup>90</sup>

It is not clear whether the “fact” is that certain moral obligations apply to us or that we are “a priori conscious” that they apply to us, or both.<sup>91</sup> At any rate the point of talking of “fact” here is not to invoke the very unKantian idea that there is a domain of moral fact to which we are receptive by some faculty of intuition. Rather, it is to make a contrast with the *Groundwork* position, where the moral law, or rather the fact that it binds us, is *not* “given,” where *all* we know is that *if* we are free the moral law binds us. Kant seems to move away from that indirect route, which does not give us knowledge that the moral law binds us, and instead now takes us to be directly conscious of our moral obligations.

One might reasonably think that Kant takes ‘consciousness’ of the moral law to be knowledge of it,’ so that in the 2<sup>nd</sup> Critique he takes us to *know* that we have certain moral obligations.<sup>92</sup> Hence, since we know that ‘ought implies can,’ we know we have the power to act from them, thus, that we are free to do so. For example, we know that we ought not lie. So we ought not to lie and hence, by ‘ought’ implies ‘can’ we have the positive freedom not to lie. If this reading is correct then ‘postulate’ no longer means something we *do not know* but in our practical deliberation have to assume. It is rather a presupposition of something we know to be true – that we have some obligations. If we assume that what is presupposed by what we know is something we can know, it will follow that we *can know* that we are free.

These two approaches, in *Groundwork* and in the 2<sup>nd</sup> Critique are importantly different; we shall discuss them, along with the question of practical ‘knowledge,’ further in 4.2-3. But first we must consider the immortality of the soul and the existence of God, for Kant also treats these, like freedom, as postulates of pure practical reason.

The argument to the immortality of the soul starts from an ethical proposition: we must strive to realise virtue in ourselves to the highest degree, that is, strive to achieve holiness – “complete conformity of [our] dispositions with the moral law.” Since ought implies can it follows that we can do that. However, given our imperfect nature holiness is not something we could achieve in any finite time, at any moment of our empirical existence. It follows then that self-perfection will take an infinite amount of time, or rather an eternity, and since it can be achieved, it follows that we are immortal.<sup>93</sup>

The postulate of the existence of God again starts from an ethical proposition. Virtue is unconditionally good: its goodness does not depend on anything else. But it is not the highest good – whole and complete – for that also requires that the virtuous should be happy: not because happiness is itself unconditionally good, as virtue is, but because a virtuous person *deserves* happiness. The highest good would be achieved in a world consisting of perfectly virtuous people who were perfectly happy.<sup>94</sup> Now this highest good *ought* to be

<sup>90</sup> CPracR 5: 47, 5: 31. Kant is denying here that we are directly *conscious* of being free, not that morality follows from freedom. [5: 121. See also 6:49, note, 6:50]

<sup>91</sup> See Allison 1990. ch. 13 for discussion. Also Ameriks 1981

<sup>92</sup> The Preface speaks of “the fact that there really is freedom, for this idea is revealed by the moral law” (5:4).

<sup>93</sup> Kant holds that God sees this eternity of effort extra-temporally. The quotation is from 5:122.

<sup>94</sup> 5:110. (As Broad notes (1978, p. 265) Kant here anticipates Moore’s principle of organic unities.)

realised, therefore it *can* be realised. But it cannot be realised by our own efforts alone. Therefore there must be a divine being that has the power to distribute happiness according to desert.<sup>95</sup>

In both arguments we could question the premises; we shall come back to them in a moment. But the second argument also has two rather notable flaws.

First: granting that some state of affairs would constitute the highest complete good, or even that it would be a very great good, it does not follow that anyone can realise it. For example a state of perfect peace, or the elimination of all disease, would be very great goods, but it does not follow that anyone can achieve them. In such cases to say ‘it ought to exist,’ would mean no more than that it would be best if it did. It does not follow that it ought to exist in the sense in which *ought* implies *can*: i.e. that someone ought to bring it about. The most one could say about the supreme good is that *if* anyone could bring it about, they ought to do so.

Second: if some being has the obligation to produce the complete good, it only follows by ‘ought implies can’ that it *can* produce it. Unlike God, who is by definition perfectly good, this being might be indifferent to the good or positively hostile to it. So this argument could at most entail the existence of a being capable of producing the complete highest good (or of beings jointly capable), not the existence of God.

Still, we must assume that Kant thought these arguments to be sound. So we can ask what their premises show of Kant’s own ethical attitude. Consider then the ethical premises from which the immortality of the soul and the existence of God are derived. These are that we each have a duty to perfect ourselves in holiness, and that the supreme and complete good ought be realised. What is striking is how distinctively they belong to a Christian, Protestant spirituality.

First of all there is the emphasis on perfection, as against improvement, and on *moral* perfection. An ideal of self-improvement is common to many ethical outlooks (not least in the 19<sup>th</sup> century). But such an ideal does not entail a *duty to perfect* oneself, or any possibility of doing so. Moreover ideals other than the Christian ideal of holiness can give content to the aim of self-improvement – as Kant knew:

if I consider *Christian morals* on their philosophic side, then, compared with the ideas of the Greek schools they would appear as follows: the ideas of the *Cynics*, the *Epicureans*, the *Stoics*, and the *Christians* are *natural simplicity, prudence, wisdom, and holiness*. With respect to the path for attaining them, what distinguished the Greek schools from one another was that the Cynics found *common human*

---

<sup>95</sup> “It is not to be understood by this that it is necessary to assume the existence of God *as a ground of all obligation in general* (for this rests, as has been sufficiently shown, solely on the autonomy of reason itself). What belongs to duty here is only the striving to produce and promote the highest good in the world, the possibility of which can therefore be postulated, while our reason finds this thinkable only on the presupposition of a supreme intelligence; to assume the existence of this supreme intelligence is thus connected with the consciousness of our duty, although this assumption itself belongs to theoretical reason; with respect to theoretical reason alone, as a ground of explanation, it can be called a *hypothesis*; but in relation to the intelligibility of an object given us by the moral law (the highest good), and consequently if a need for practical purposes, it can be called *belief* (or faith: *Glaube*) and, indeed, a pure *rational belief* since pure reason alone (in its theoretical as well as in its practical use) is the source from which it springs.” 5:126

*understanding* sufficient, the others the path of *science* alone; but both found the mere *use of natural powers* sufficient for it. Christian morals, because it frames its precept so purely and inflexibly (as must be done), deprives the human being of confidence that he can be fully adequate to it, at least in this life, but again sets it up by enabling us to hope that if we act as well as is within our *power*, then what is not within our power will come to our aid from another source, whether or not we know in what way.<sup>96</sup>

This captures an ethically important Christian attitude to the world. It is for us to do our Christian duty, confident that God, whether or not we understand His ways, will do whatever more is needed. This faith gives Christians hope that it is neither presumptuous nor self-centred to aim for holiness, and that the highest good *will* be achieved. It is also Christian faith that takes this highest good to be blessedness merited by virtue. From a non-Christian ethical standpoint these elements of Christian faith may well seem unworldly, unrealistic, fantastic even.

So the premises of Kant's arguments for immortality and God draw on Christian faith and a particular ethical tradition within it.<sup>97</sup> That does not make them circular in any strict sense. Nonetheless a Christian ethical faith is the source of their power, and we can ask whether such a faith is possible without a prior belief in a loving and all-powerful God. That in turn throws light on Kant's own spiritual stance.

### 3.7 *Christianity as a moral faith*

It is possible that Kant's philosophy of religion has done more to undermine religious faith than to vindicate it. Yet vindicating it was Kant's declared aim.

Some have felt this declaration to be disingenuous. Did Kant have any personal religious belief at all – let alone Christian belief? Several contemporaries who knew him well thought not: "Having postulated God and immortality, he himself [one of them wrote] did not believe in either."<sup>98</sup>

Kant's voluminous writings on religion do, in my view, witness his faith in God, and in a certain ethical sense, his Christianity. He did not have to write as he did, or at such length, unless he meant what he said – and honesty was important to him.<sup>99</sup> As is

<sup>96</sup> 5:127, note

<sup>97</sup> "Be ye therefore perfect, even as your Father in Heaven is perfect." Matthew 5, 48. Christians have been divided about what this injunction means. Does it apply to everyone, and if so how? Catholics taught a "double standard of perfection." Is perfection in this life possible without the assistance of divine grace? (commonly held not.) Kant on grace: 6: 53, 6: 174, 191-2, 7:43 Pelagianism + immortality = Kant. Methodism and Wesley. On Kant's view of Augustine and Pelagius see Wood, 'Kant's History of Ethics.'

<sup>98</sup> Kuehn, p.3; see also pp. 5, 11, 138, 391-2. Compare Byrne 2007, Pasternack and Rossi 2014.

<sup>99</sup> In a letter to Moses Mendelssohn of 6 April 1766, quoted in Kuehn 2001, p. 172, Kant declared "Although I am personally convinced with the greatest clarity and satisfaction of many things which I will never have the courage to say, I will never say anything that I do not mean (*dencke*)." And "while all that one says must be true, this does not mean that it is one's duty to speak out the whole truth in public" quoted Reiss, p. 2. That is a quite understandable policy, but it can be hard to follow it without seeming to lack candour.



clear from his ethical proof of the existence of God, the God in whom he had faith was an objective, personal, and perfect being.

The premises of his ethical proofs – that we have a duty to strive for perfect Christian holiness, and to work in partnership with God for the highest good – bear the distinctive stamp of a Christian and Protestant conception of our life-task.<sup>100</sup> They may well originate in his Pietist upbringing. However they remained with him, even though his mature view of Pietism was by no means uncritical: he developed serious reservations about its moral quietism and its anti-intellectualism.<sup>101</sup> Nevertheless he always agreed with its emphasis on inner religious faith and righteous conduct in opposition to church institutions, rites, and doctrine – what he called “priestcraft:” “the dominion which the clergy has usurped over minds by pretending to have exclusive possession of the means of grace.”<sup>102</sup> His comment about the pietism which he knew in his youth, at home and in the Collegium Fridericianum, rings true:

Even if the religious consciousness of that time and the conceptions of what is called virtue and piety were by no means clear or satisfactory, it yet contained the root of the matter. One may say about pietism what one will. Enough! The people who took it seriously were distinguished in a way which is worthy of greatest honour.<sup>103</sup>

But Kant’s rejection of any distinctively Christian rites or doctrinal and historical claims went well beyond Pietism. Christianity as he conceived it was not just primarily, but exclusively a moral faith. The regretful conclusion of one of his more orthodox friends, that he “viewed Jesus as the personified ideal of perfection,” not “as the sufficiently proven messenger and son of God, the savior of mankind,” is entirely plausible.<sup>104</sup> This view of Christ would become common in nineteenth-century moral philosophy. Kant hoped for and expected an eventual free moral community of the faithful, with no “degrading” distinction between laity and clergy, liberated from all compulsory historical and supernatural claims, dogmatic catechisms, and rites – a kingdom of ends in which “God may be all in all.”<sup>105</sup>

So what did he mean by faith, *Glaube*? Here it is of first importance for Kant that the ontological content of faith is derived from, and *can* only be derived from, morality, rather than vice versa. This stance, though deeply puzzling, explains his disregard of historical tradition or Church dogma. The content of religion must be derived within the limits of reason alone – and since the theoretical proofs have been shown to be nugatory,

---

<sup>100</sup> There is a question about how his thought was developing at the very end of his life. See Guyer, who quotes the following from late notes published as the *Opus Postumum*: “God and the world are “not substances outside my thought, but rather the thought through which we ourselves make these objects.” But on these late notes cp Louden 2015.

<sup>101</sup> Kuehn 2001. Wood, intro to *Religion and Rational Theology*, xii, xxiv

<sup>102</sup> 6: 200. Cp Kuehn 2001, p. 371.

<sup>103</sup> Wood intro, p.xii. Collegium Fredericianum: Pietist school attended by Kant...

<sup>104</sup> Quoted in Kuehn 2001, p. 11, from Borowski’s biography of Kant (1804). See also Kant’s reply to Johann Kaspar Lavater, described in Kuehn 2001, p. 224 and quoted by George di Giovanni, Kant Cambridge Edition, *Religion and Rational Theology*, p. 50: “I ... seek in the gospels not the ground of my faith but its fortification.”

<sup>105</sup> *Religion Within the Boundaries of Mere Reason*, 6: 121. Kant is quoting “God may be all in all” from I Corinthians 15:28.

that leaves only practical reason. Theoretical reason can only establish the *possibility* of God.

To be blunt, Kant's working out of this view is extraordinarily unconvincing. His ethical proofs are no better than the theoretical proofs he rejects. Furthermore he has to hold that their premises are deliverances of practical reason – within his philosophical framework this would have to mean that they can be deduced from the Categorical Imperative. In reality, they quite obviously belong to a particular Christian tradition of spirituality. A deeply impressive tradition, but in no way the voice of Pure Practical Reason.

The failure of the ethical proof of the existence of God is disastrous for Kant's defence of a theistic religion. Even if for the sake of argument we take it as sound, however, there remains the perplexing question of how to understand the elusive contrast between knowledge and faith. He himself is clearly bothered about how exactly to state it and devotes quite a few pages to trying to clarify it – it cannot be said successfully. It is one of his more troublesome legacies to late modern ethics.

What is the problem? The arguments for immortality and God differ from the argument for freedom, in that they require rather specific ethical premises, whereas the argument for freedom can work on any true assertion that we have a duty. In all three cases however the premise is ethical. So can we *know* these premises to be true, as mooted in 3.6? If we can, then granting Kant's arguments to be valid, we can know their conclusions. We can know that we are free, that we are immortal, and that God exists. Yet these are ontological conclusions about noumena, which Kant holds cannot be known. Not merely does he hold in general that we cannot have substantive knowledge of the noumena, he definitely and repeatedly says that we cannot know that there is eternal life or that God exists.<sup>106</sup>

Consistency would be restored if we took it that Kant thought that only *theoretical* knowledge is knowledge. Perhaps he implicitly takes that view in the case of purely normative propositions: we can be legitimately confident in them, warranted in acting on them, but cannot be said to *know* them. This would not be an entirely surprising thing for him to think, given how sharply he contrasts the epistemology of the factual, which involves receptivity, and the epistemology of the normative, which does not. But he never does say this.

He does say that an ontological conclusion derived (via ought implies can) from a truth of practical reason has only practical significance. What can we make of that? He does not mean that there is *practical* reason to *cause* oneself to believe it. This might be the view of someone who advanced a 'pragmatic argument' for the existence of God, to the effect that one should believe in God because such a belief contributes to one's peace of mind. That is not how Kant's argument for the existence of God works. Rather, were his argument valid, the ethical proposition that I should devote myself to promoting the highest good would entail the ontological proposition that God exists. In contrast, the proposition that believing in God gives one peace of mind does not entail that God exists. In itself it gives one *no* reason to believe that God exists (as against, for example, reason to take a course of indoctrination that will cause one to believe that God exists).

If Kant's argument to the existence of God were valid, it would give us exactly the same grounds for believing in God as we have for believing the premise of the argument. But, if these seemingly warranted ethical premises turn out to presuppose ontological

---

<sup>106</sup> A829/B857. Cite later texts.

propositions that we have no theoretical warrant to believe, how *can* we be confident them? Why doesn't the argument go into reverse – from lack of warrant for the ontological claim to lack of warrant for the ethical claim? It seems that Kant has no answer to that.

Nowadays it is common to hold that 'faith' is a matter of 'choice.' Kant make no such claim. Nevertheless something like it may be helpful. Imagine someone who makes it his life-task to advance the cause of a perfect socialist society. Allow also that that the socialist ideal has ethical merit, whether or not we agree with it. This man believes that such a society is possible; he realises that his own work for it only makes sense if others will play their part in bringing it about. If there is very little or no reason to believe that others *will* play their part, is his commitment nullified? Suppose that while there is no reason to suppose they will, there is no reason to suppose they won't. It is *possible* that they will, not just as an abstract logical possibility, but as a bet on human nature.

One might see this ideal as an integral 'ethical/factual package:' socialism is the highest form of society, it is attainable if people co-operate in working for it, there is no decisive reason to think they won't, so I make it my task to strive for it. Is it irrational to 'believe' in, have 'faith' in, 'commit' to, the whole package – to see working for it as life-defining task? In some conditions it may be: for example if the ideal of socialism can by now be seen to be utterly flawed. But in other conditions faith in the ideal may not be irrational, and it may not be irrational to believe that others will join in bringing it about. There is a balance between the power and feasibility of the ideal and the meaning one's life gains from devotion to it.

Kant's ideal may be thought to draw implicitly on a similar total ethical/factual package, in this case the Christian one, even though that is not how he presents it. The Christian ideal is realisation of personal holiness and of the highest good; to make it one's task gives life meaning, yet only makes sense if one can believe that there is a God who will make its achievement possible. The epistemic circumstances are that there is no theoretical reason to think that God does not exist, even though there is no theoretical reason to think he does; perhaps in these circumstances taking up the ideal as a life task, and thus the faith in God that is essential to it, is a viable project?

The idea that working for this ideal, and hence believing in God, in a way makes sense, even though it is also in a way 'absurd,' takes us forward to Kierkegaard. We shall come back to it there. It is not, however, a view that sounds like Kant. Kant's claim that belief in God and immortality are a "practical need" may open up these questions, but he does not answer them in this way, and in the end his own position remains enigmatic.

What is certain is that after the death of Frederick the II ('Frederick the Great') Kant's new political masters saw his view of religion as a threat to the moral order and stability of the state (nor would they have changed their minds if they had had a full understanding of it). A shift took place from the days of Frederick's enlightenment despotism. He, himself a free thinker, practiced tolerance in religious matters (though according to Lessing this was the only freedom his subjects were allowed). When Frederick William II acceded to the throne in 1786, however, he decided that orthodox religion, a pillar of the state, needed to be defended by the state. Since Kant was by now accepted as Prussia's leading thinker – a fact recognised by the king and his advisers – the campaign to impose religious orthodoxy did not affect him for some time. However the publication of *Religion Within the Limits of Mere Reason* triggered action. Kant received a shot across the bows: it took the form of a letter dated 1<sup>st</sup> October 1794 in the king's name:

Our most high person has long observed with great displeasure how you misuse your philosophy to distort and disparage many of the cardinal and basic teachings of the Holy Scriptures and of Christianity; how you have done this particularly in your book *Religion within the Boundaries of Mere Reason*, as well as in other shorter treatises. We expected better things of you, as you yourself must realize how irresponsibly you have acted against your duty as a teacher of youth and against our paternal purpose, which you know very well. We demand that you give at once a most conscientious account of yourself, and expect that in the future, to avoid our highest disfavour, you will be guilty of no such fault, but rather, in keeping with your duty, apply your authority and your talents to the progressive realization of our paternal purpose.

Failing this, you must expect unpleasant measures for your continued obstinacy.<sup>107</sup>

This has the true note of autocracy: both personally intimate and menacing. The measures to be expected might indeed have been unpleasant. They could have included dismissal from his University post, loss of pension, and banishment.

According to friends Kant had acquired the material means to look after himself and was not worried by these dangers.<sup>108</sup> So it was not from cowardice that the letter he sent back in reply on October 12<sup>th</sup> complied with the King's demands. In all his teaching and writing, he wrote, his concern was solely the *philosophy* of religion, rather than biblical or theological studies; he therefore made no appraisal of the truth of scriptural revelation or of Christianity as such. Nonetheless he had, he claimed, shown their consistency with the moral content of a philosophically defensible religion, and had always respected them as the best "means of public instruction for establishing indefinitely a state religion that is truly conducive to the soul's improvement." Finally, he promised "as your Majesty's most loyal subject" that he would "hereafter refrain altogether from discoursing publicly, in lectures or writings, on religion, whether natural or revealed."<sup>109</sup>

Given Kant's repeated objection to saying anything he did not believe to be true (see footnote 00) – though not to being economical with the truth – what he does say in his letter casts interesting light on his position. In the first place he accepts that "the public *religion of the land*" is a legitimate matter of state policy, since public religion is implicated in the maintenance of civil order. Thus the state does not act beyond its powers in prescribing what religious content should be taught. He also *appears* to accept that the King's "paternal purpose" – the improvement of his subjects' souls – is legitimate. However in this case all that he strictly *says* is that he knows that the King has this "highest paternal purpose," and that he himself has done nothing inconsistent with it. Elsewhere he says that a "paternalistic" government "is the most despotic of all (since it treats its citizens as children)."<sup>110</sup> It does not follow, of course, that Kant would have thought any effort by a government to improve its subjects' souls, or their virtue, illegitimate. There might be non-paternalistic ways of trying to do that.

<sup>107</sup> 7:6. Kant published this letter, together with his reply to it, in 1798 after Frederick William II's death, as the preface to a collection of essays entitled *The Conflict of the Faculties*.

<sup>108</sup> Kuehn, 379

<sup>109</sup> 7:10. After Friedrich Wilhelm II's death the qualification "as your Majesty's most loyal subject" was used by Kant as an excuse for further public discourse on religion. He contended, not very plausibly, that his use of the phrase in his reply had restricted his promise to the duration of the king's reign.

<sup>110</sup> *Metaphysics of Morals* 6: 317; cp *Theory and Practice*, 8:290-1.

Acknowledging that the state has a legitimate role in determining the content of religious education is not the same as accepting that it is discharging it correctly, or even that it should necessarily be doing much on this front at all. What Kant actually thought about the activities of Frederick William II's "Commission of Investigation," as he made clear after the King's death, was that it had instituted a "drive towards a faith ever more estranged from reason," a "nonsense" which had "now been brought under control." 7:10 – 11.

Nonetheless, given the common present-day assumption that liberalism requires state neutrality on matters of religion Kant's admission of a religious role for the state may come as a surprise. (Equally, of course, our present-day tendency to make the opposite assumption calls for justification.) Kant underpinned his view by a distinction between the public and the private use of reason. It appears in 'What is Enlightenment?' (1784) and is used again in "The Conflict of the Faculties:"

by the public use of one's own reason I understand that use which someone makes of it *as a scholar* before the entire public of the *world of readers*. What I call the private use of reason is that which one may make of it in a certain *civil* post or office with which he is entrusted. 8: 37

Thus for example a clergyman must in his capacity as a clergyman teach "the creed of the church he serves." 8:38. His use of reason in this role is "private" in that it is used to carry out the tasks of his institutional role in the most effective manner. But as a scholar he should have complete freedom of speech, including freedom to criticise what he takes to be errors in that creed.

In this respect, however, the Faculty of Philosophy has a special role. Kant holds that philosophy is by its very nature free thinking. In effect, then, in instituting a Faculty of Philosophy the state itself institutes a body tasked with the public use of reason.

This is not an unreasonable overall position, especially if one is convinced that Christianity has a uniquely valuable moral content that should inform the culture of state and civil society. In its time it was the view of a moderate enlightenment thinker. It is not the more radical view that already existed then and has come to prevail in much liberal thinking, according to which the state should not meddle in matters of religion, and religion should not meddle in affairs of state, at all.

### 3.8 *Politics and freedom*

Taken as whole, Kant's political thought is a version of what we might now call liberal constitutionalism. At that time (as noted in II.0) the word 'liberal' had not yet acquired the meaning, or rather meanings, it has in politics today. Kant terms the kind of constitution he argues for "republican."

The first thing he means by that is a constitution that separates the executive office of the state from its legislature. A state that does not make that separation is, he says, "despotic." Separation is essential for the rule of law itself – essential to the idea of a *Rechtsstaat* or 'law-state' that became an important Kantian legacy to German jurisprudence. Substantively, Kant's constitutional ideal entrenches legal protection of the equal rights of all citizens, protection of liberty of thought and discussion, and a legislator or legislature that – in some sense – represents the people. But none of this makes Kant a democrat:

So that a republican constitution will not be confused with a democratic constitution (as usually happens) the following must be noted. ... Of the three forms of state [monarchy, aristocracy and democracy], that of democracy in the strict sense of the word is necessarily a *despotism* because it establishes an executive power in which all



decide for and, if need be, against one (who thus does not agree), so that all, who are nevertheless not all, decide: and this is a contradiction of the general will with itself and with freedom ...

Even if the other two state constitutions are always defective insofar as they leave room for this [despotic] kind of government, in their case it is at least possible for them to adopt a kind of government in conformity with the *spirit* of a representative system ... whereas a democratic constitution makes this impossible because there everyone wants to be a ruler.<sup>111</sup>

This tripartite division of forms of government is traditional, but not very helpful for understanding Kant's view. The best form of government, he says, at least in this essay, is monarchy (8: 353). In saying this he is thinking of monarchy as a form of 'republican' constitution, i.e. with a division between executive and legislature. It is consequently not entirely clear whether he thinks of the monarch as ruler, that is, the head of the government's executive branch, or as a sovereign legislator. On the former view, the monarch would be the chief executive, presumably acting in accordance with laws originating in in some kind of representative assembly.<sup>112</sup> On the latter view the monarch would be sovereign legislator, making law in the name of the people.

This would not be an inconsistent view. Kant takes the ultimate *normative* source of sovereignty to be 'the general will.' We have already noted the multiple connotations of that phrase at that time (II.0). In Kant, as in Rousseau, the general will is not the will of most, or even all. Bearing in mind that will, for Kant, is practical reason, the general will becomes in his hands the Categorical Imperative. It is *general* in that it is the will of each one of us considered transcendently as a rational being. At least at this level of theory, it does not much matter what kind of sovereign realises the general will, so long as some sovereign realises it – acting within a republican constitution. The idea of a social contract, which Kant also invokes, becomes a purely hypothetical representation of the Categorical Imperative. In fact, since – as we saw in 3.4 – Kant makes no allowance at the fundamental level of practical reason for divergent rational interests – that is, for the possibility of fundamental conflicts warranted on both sides by irreducibly agent-relative reasons – the idea of a social contract becomes otiose. One could simply ask what principles a purely rational legislator would lay down, that is, what moral requirements the Categorical Imperative imposes, and how these should be applied to empirical political questions of constitution and policy. If a prince legislated in accordance with and from the Categorical Imperative he would represent 'the people giving itself the law' as much as an assembly of representatives would.

This is understandable as a consequence of Kant's abstract ethical theory. At the same time it is one of the points at which his theorising most flagrantly misses the reality of politics – in this case the crucial role of a representative assembly in conciliating rational but divergent interests.

Whether the monarch is cast as legislator or as chief executive, it is not clear why monarchy would be any less likely to transgress the distinction between the two than democracy would be. In *Metaphysics of Morals* 6:339 (two years after *Perpetual Peace*) Kant recognises the danger. He now calls the form of government in which "only one is legislator" "autocracy." Autocracy, he says, can be defended as the simplest form of government, but with regard "to right itself this form of state is the most dangerous for a people, in view of

---

<sup>111</sup> *Perpetual Peace* 8: 352/3.

<sup>112</sup> 8:352-3; 6: 352 – 3. Frederick II's description of himself as only the highest servant of the state, which Kant quotes with approval, would seem to cast him in the role of chief executive.

how conducive it is to despotism” 6: 339. So he distinguishes here between autocracy and monarchy: it looks as though his favoured constitution has a legislative assembly elected by the citizens, with the monarch ruling in accordance with its laws.<sup>113</sup>

Kant’s objection to democracy in the “strict sense,” direct rule by the people, is that it would permit a majority to oppress a minority. It was, then, the same worry as that of the French liberal constitutionalists who drafted the monarchical French constitution of 1791 (II.0.0), and indeed of liberals throughout the nineteenth century.<sup>114</sup>

Present-day democracies have a universal franchise, and they take possession of a vote to be a universal *right* (over a given age [cp issue of prisoners]) rather than a distinct *standing* for which one has to qualify by meeting certain conditions. That was not Kant’s conception. Kant took the latter view. Every *citizen* has a vote, but not every member of the state is a citizen.<sup>115</sup> Citizenship for Kant is a standing that consisted of three properties: freedom under the law, civic equality, and civil independence. The difficulty lurks in the third of these: civil independence, or “civil personality.” It requires that one is not “under the direction or protection of other individuals.” That excludes

An apprentice in the service of a merchant or artisan; a domestic servant (as distinguished from a civil servant); a minor (*naturaliter vel civiliter*); all women ...<sup>116</sup> The pettifogging examples Kant gives in various places<sup>117</sup> of who does or does not qualify for citizenship by this criterion of “independence” do not cover him in glory.<sup>118</sup> For us of course what especially stands out is his unqualified exclusion of all women. This was an exclusion that another great liberal, John Stuart Mill would vehemently oppose. That was later; but we have seen that prior to 1793 the enfranchisement of women was being strongly canvassed in the French Revolution, even though the 1791 constitution did not implement it (0.0.0). Kant’s complete failure to engage with the issue, especially in light of his deep interest in and support for that early phase of the French Revolution – we’ll come to it in a moment – is at best thoughtless. However the disagreement among liberals about the enfranchisement of women should not obscure the point that there was an underlying liberal philosophical *agreement*, right through the nineteenth century, that the vote was not a universal right but a responsibility or trust which required, for the safety of all, appropriate qualifications on the part of those to whom it was entrusted. The significant disagreement was about what qualifications were appropriate.

Kant’s notions of citizenship and civil personality were narrow even by the standards of his time. However he did not hold that the laws of justice applied only to citizens (or as the drafters of the 1791 constitution would have said, active citizens). On the contrary, he emphasises – as indeed one would expect from his underlying ethics – that they apply impartially to everyone with respect to everyone – citizen or non-citizen. What then are these laws?

He gives a systematic answer in *The Metaphysics of Morals*, where he divides the field of morals into what he calls the Doctrine of Right (*Rechtslehre*) and the Doctrine of

<sup>113</sup> Maliks 2014 says that Kant takes the monarch’s role to be executive. [Reiss, p. 30. VI: 338-9. *Freedom and the Construction of Europe*, vol 2 eds Skinner and van Gelderen.]

<sup>114</sup> II.0.0 and 0.0.0; Sieyès reply to Tom Paine on democracy – despotic.

<sup>115</sup> Compare Sieyès distinction, adopted in the French constitution of 1791, between active and passive citizens.

<sup>116</sup> 6: 314 - 5

<sup>117</sup> E.g. *Theory and Practice*, 8: 295: a wigmaker does, a barber does not.

<sup>118</sup> Cite a discussion of Kant on civil independence.

Virtue (*Tugendlehre*). The former deals with duties to others (individuals, nations and states) that correspond to *rights* on their part. Virtue, in contrast, covers duties to oneself – for Kant certainly thinks that we have duties to ourselves (0.0) – and “duties of virtue to others,” that is, duties which do not arise from others’ rights. The significance of a right is that it can be *demande*d and *enforce*d – a conception of a right that goes back at least to Grotius. Justice is the equitable enforcement of rights (including compensation for the violation of rights). Kant’s Doctrine or Right is thus his theory of justice.

Individuals have natural rights [defined at 6:237]; these are constitutive of natural-legal notions of property, contract, person and state, and must therefore guide and constrain the legislature in its institution of positive laws.<sup>119</sup> There is an “innate right” of “outer freedom” 6:237; acquisition of external objects takes place under the principle “What I bring under my control in accordance with the laws of outer freedom, and will to become mine becomes mine.” 6:264. However Kant thinks that this principle itself issues from the general will, and adds that private property can exist only “provisionally,” not “conclusively” in the state of nature.

One gets a clearer sense of what he means if one compares his view of punishment (6:331) with John Locke’s view in the 2<sup>nd</sup> treatise. For Kant, punishment is the original and exclusive right of the state. For Locke, the right to punish and enforce compensation is an original right of individuals, which they hand over to the state only by agreement amongst themselves. In effect Locke and Kant agree on the underlying idea, namely that “conclusive” rights presuppose a right of punishment. But Locke thinks such a right is held by individuals in the state of nature, whereas Kant thinks it exists only in, and as the exclusive right of, a state. Hence for Kant individuals cannot have any “conclusive” rights (other than that of outer freedom) except in the civil condition.

This may seem a slight distinction, especially if one compares what Locke and Kant say about freedom, understood as liberty in a civil state:

Freedom is not, as we are told, *A Liberty for every Man to do what he lists* ... But a *Liberty* to dispose, and order, as he lists, his Person, Actions, Possessions, and his whole Property, within the Allowance of those Laws under which he is; and therein not to be subject to the arbitrary Will of another, but freely follow his own.

One cannot say: the human being in a state has sacrificed a *part* of his innate outer freedom for the sake of an end, but rather, he has relinquished entirely his wild, lawless freedom in order to find his freedom as such undiminished, in a dependence on laws, that is, in a rightful condition, since this dependence arises from his own lawgiving will.<sup>120</sup>

For both of them the civil liberty that is really worth having is freedom from illegitimate interference. It is *not* freedom from any normative constraint: it is, rather, freedom from any constraint other than a law that is just, that is, in accord with the system of natural rights. Nonetheless there remains a difference, epitomising English and German thinking. Locke is a robust individualist; he thinks that the state has no rights that do not derive from individuals who have freely transferred their rights to it. His political point is simply that the rights of individuals cannot be *effectively* – as against legitimately – enforced without a transition to

<sup>119</sup> States, nations, or peoples also have rights against each other, even in the absence of a world state; these constitute “cosmopolitan right”, that is, international law.

<sup>120</sup> Locke, *Second Treatise of Government*, §57. Kant, 6: 316

civil society. The state is not ultimately more than a collective instrument, on which individuals agree for the security of themselves and their property.

For Kant, in contrast, the state has an ethical significance that is not merely instrumental, but that it possesses as such. The state (at least in its idea) embodies the general will: this – it seems – is why only it has the original right to punish. Yet at the same time its laws are ‘my laws’ in the sense that they arise from my “own lawgiving will.”

This momentous line of thought would become a fixture in post-Kantian idealist political philosophy. We have already noted the anxiety it arouses in many less metaphysical liberals (3.4). Given this conception of the state, does Kant really endorse the notion of natural right? Not in the sense of a right that conclusively exists in the state of nature. But yes, in the sense of right whose justification as such can be established by reason alone, irrespective of revelation, empirical arguments, or actually existing positive laws.

Kant’s famous refusal to allow a right of rebellion, or even of resistance, against the state can be seen, against this background, as flowing in particular from the thought that *any* working political state expresses, however imperfectly, the general will – whereas no state of nature does. As so often with Kant, if we leave out the background thought the explicit arguments look like empty logic-chopping, generating specious ‘contradictions’ to which there are obvious replies. So if his argument (0.0) is that you can’t lawfully rebel against the legislator, because the legislator is the source of all law, the answer is that the legislator gives all positive law but is not the author of natural law. Natural law may allow a rebellion to strike down positive laws that infringe it, for example by infringing the rights of minorities. Or again, if his claim is that any constitution that explicitly envisages conditions for overthrow of a supreme ruler is contradictory, the answer is that under such a constitution there is by definition no *supreme* ruler, and the question is, why is that a problem? 6: 319. Why can’t there be a state (a civil condition) without a supreme ruler? A posteriori arguments are needed to show that such a state could lead to disaster. But there is no sign, at least explicitly, that such arguments are at work in Kant’s thinking. What seems to be at work is the conception of the state as the *unique* embodiment of the indivisible sovereignty of the general will.<sup>121</sup> To rebel against such a state would be contradictory, because I would be putting my will against my will. But from a less metaphysical standpoint the obvious answer is that not *every* state is the embodiment of the general will or alternatively, that not *every* general will expresses my will.

The general will would only be my will if there was some intelligible sense in which I or my will could be said to *be* reason. Something like that is hinted at in the idea that as a rational being I am noumenal. This Kantian doctrine is an important stage in the rise of some famous idealist identities: my will as a rational being and the general will, the general will and the state. We shall come to a further stage when come to Hegel. And we shall come back to some further aspects of Kant’s doctrine of right – punishment, marriage, children, civil society and the state – when we discuss Hegel’s critique of Kant and his very different treatment of these topics in the *Philosophy of Right*.

Meanwhile we return to Kant’s general account of the Doctrine of Right. It raises a question of method which became a recurrent theme in late-modern ethics: the question of ‘moral common sense’ versus ‘theory.’

Kant’s substantive account of natural rights shows a notable degree of convergence with other such accounts. Is that because the particular rights on which they converge are self-evident to moral cognition? Here I use the term ‘self evident’ in the philosophical sense

---

<sup>121</sup> Cp Louis XV’s reply in 1766 to the pretensions of the Paris *parlement*, p.00 (ch II).

in which a proposition, or a set of propositions, is self-evident if it can be seen to be justified 'on its own evidence', i.e. just by careful reflection on and insight into its content, without any need to infer it from anything else. This claim about a proposition is compatible with holding that its justification can be enhanced or decreased when it is considered in a corpus of related propositions, as when one claim about a natural right is supported by or in tune with others, and the whole system is argued to be self-evident. Can self-evidence, in this sense, be justification enough? Or must a theory of natural rights, to carry philosophical legitimacy, be derived from or in some way systematised by some higher, more general principle?

The question carries special force when we are considering a natural-right theory of justice. Utilitarian approaches to justice dispose of natural rights altogether, Marxists argue that natural rights are bourgeois ideology, while many in the 20<sup>th</sup> century have theorised justice not in terms of natural right but in terms of a new, and seemingly distinct and separate, idea of 'social justice.' These responses to the notion of a natural right are sceptical inasmuch as they either reject the notion head on, or at most provide what may be called a sceptical reconstruction of it, as we shall see. Can a simple appeal to common moral cognition of rights stand up against this tide?

The Kantian position on this, as ever, is not sceptical but Critical. That is, here as elsewhere the question to be asked is: how is common cognition of rights possible? It might be possible to answer this Critical question without appealing to a deduction from some higher-order principle. Nevertheless Kant does propose a deduction. The Doctrine of Right, and indeed the very distinction between Right and Virtue, is derivable, he holds, from the Categorical Imperative. He does not in fact do very much to illustrate that claim. However he introduces an important intermediate principle which he claims can be deduced from the Categorical Imperative, and which at least gives shape to the system of natural rights. It is the Universal Principle of Right:

Any action is right [i.e. consistent with principles of right] if it can coexist with everyone's freedom in accordance with a universal law, or if on its maxim the freedom of choice can coexist with everyone's freedom in accordance with a universal law." 6:230

As Kant notes, this can be read as a principle of maximum equal liberty.<sup>122</sup> So in order to deduce individual property rights from it one would need to show that some structure of property rights, complete with original acquisition, maximises liberty. Though that looks interesting, it also looks challenging. Perhaps establishing a reasonable give-and-take between particular convictions about rights and the maximum equal liberty principle would be enough to satisfy the Critical question.

How the principle of maximum equal liberty relates to the Categorical Imperative is another important issue. The principle is a principle of freedom, so it might be supposed to follow from the Categorical Imperative through some linking idea of freedom. Or, if it is not possible to derive it in any serious sense from the Categorical Imperative proper, we could ask how it might be derived from premises involving an appeal to autonomy. The problem is that autonomy is one idea, civil liberty is quite another. The fact that we are equally capable of autonomy does not obviously entail the normative claim that we should have maximum equal liberty. What could the linking thought be?

---

<sup>122</sup> "the greatest feasible right under the law to act as one chooses, subject to the equal rights of others – negative liberty."



There seem to be two possible links. One starts from the claim that we do in fact strongly want to be left to make our own decisions plus the thesis that we are all capable of doing so. Here the basic – undoubtedly powerful – ethical thought is: who are you to tell me what to think or do? – I’m a rational being, perfectly capable of making my own decisions, and that is what I mean to do. Obviously this rests on an empirical claim about people’s capabilities and desires. The other line of thought starts from an ethical claim, which as we have seen Kant endorses, namely, that our task in life is to perfect our autonomy, and then again an empirical premise, that the way to perfect our autonomy is by practising it in the widest circumstances, without assistance. Kant makes a claim of this kind: “we do not ripen to freedom otherwise than through our *own* attempts.”<sup>123</sup>

Unsurprisingly, then, getting to liberalism about politics from the idea of autonomy requires some ethical principles about what the value of autonomy is, and some empirical claims about what human beings are like. From here there are two ways to go (they could be combined). One is simply to re-assert the theory of natural rights as basic, without trying to derive it from something else. The other is to develop, and robustly defend, the ethical principles and empirical claims that are required to get from autonomy to a liberal political constitution. Given the way the ethical ideas that underlie political thought would develop in the nineteenth century, the second, more difficult and uncertain route, would be the way for liberals to go.

To complete this sketch of Kant’s political thought and attitudes we should take into account his lifelong enthusiasm for the French Revolution. Contemporaries were evidently struck by it. At a time when hostility to what was happening in France was in full swing in Prussia, he went out of his way, as one acquaintance said, “to endorse the principles of the French Revolution, defending them even at dinners in the noblest houses.”<sup>124</sup> He followed all its details, as reported in the newspapers, avidly: as another acquaintance said,

He lived and moved in it and, in spite of all the terror, he held on to his hopes so much that when he heard of the declaration of the republic he called out with excitement: ‘Now let your servant go in peace to his grave, for I have seen the glory of the world.’<sup>125</sup>

In *The Conflict of the Faculties* he writes that the revolution may succeed or miscarry; it may be filled with misery and atrocities to the point that a right-thinking human being, were he boldly to hope to execute it successfully the second time, would never resolve to make the experiment at such cost.

But he draws attention to the participation in spirit, and the enthusiasm, it has inspired among onlookers in other nations. Enthusiasm, he continues, “as such deserves censure,”<sup>126</sup> yet it arises from a belief in the revolution’s essential justice and thus shows that the human heart can be inspired by a struggle for the right, in a way that it can never be inspired by self-interest or monetary reward (or one might add the other blandishments of a despotism). This disposition of the heart is his best hope for human progress.

It is a somewhat more cautious line on the Revolution than what we are told were his private views. Nonetheless it seems that even in 1798, when *The Conflict of the Faculties* was

<sup>123</sup> 6: 188, note.

<sup>124</sup> Quoted from Metzger in Kuehn p. 4.

<sup>125</sup> Quoted from Malter, Kuehn, pp. 341-2. (See the whole account on pp340 – 43.)

<sup>126</sup> 7: 85-6. An interesting remark, in view of the descriptions of his private attitude just quoted! But Kant is using “enthusiasm” in an older sense, to mean something like ‘fanaticism.’

published, his private support for it remained very strong<sup>127</sup> – this was after the execution of the king in 1793, the execution of Robespierre and St Just and the takeover by the Directory in 1795, and just before Napoleon’s *coup d’état* in 1799 (the ‘eighteenth Brumaire.’)

As we saw in the last chapter, the meaning of the terror had by then become, to many outside observers, a major intellectual issue. Many thinkers who welcomed the Revolution’s initial principles had become horrified by its terror; for them it became a fundamental question whether the principles and the terror were linked. Answers to this question developed along conservative, liberal and radical or ‘Jacobin’ lines. Was terror and descent into anarchy or dictatorship the inevitable result of rebellion? Was it that progressive constitutional reform had been tragically taken over by the irresoluteness of the king and the politicians’ manipulation of, or submission to, the Paris mobs? Or was there a harder message, that real progress to a new and better society could only be achieved by ruthless repression of the revolution’s enemies?

Kant was no Jacobin follower of Robespierre. Given his cautious monarchical republicanism, his criticism of democracy, his denial of a right of rebellion, and his outright opposition to regicide (6: 320), one might expect him to oppose the coup of June 1793. There is probably too little evidence to assess his considered personal view. As far as the issue of rebellion is concerned, his argument was that what had occurred in 1789 was not a rebellion, because the king’s summoning of the estates general implied his abdication of sovereignty to them –

they were representatives of the entire nation after the king allowed them to pass decrees in accordance with indeterminate powers. Before that, the king represented the nation ...<sup>128</sup>

One must remember Kant’s objection to the idea of divided sovereignty. Presumably what happened on his view of things was that the king freely ceased to be the autocratic legislator, and became the chief executive – “roi des Français.” The National Assembly became the sovereign legislator, at which point its Declaration of the Rights of Man and of the Citizen became legitimate and rebellion against it illegitimate.

In its own terms this is not unpersuasive – though it overemphasises the king’s freedom of choice. Alas, it also exemplifies all too well the legalism that Kant brings to practical political thinking. We do not find in it wise reflection on the big, indeed the fateful, questions about the Revolution that divided conservatives, liberals and Jacobins. We find in it no assessment of the significance of terror at all. Given the kind of thing that Edmund Burke had to say about Richard Price (0.0.0), one can imagine what he would have had to say about Kant.

#### 4 *Kant: reflection and assessment*

There have been many efforts to reinterpret Kant or revise him. This is entirely understandable. His philosophy is packed with interesting ideas that cry out for development, pregnant with new possibilities, deeply resonant to modern sensibilities. At the same time it has clear fracture lines, elements that point in different directions. The historical Kant is a complex figure; there is every reason for anyone who wants to develop Kantian resources in

<sup>127</sup> Kuehn, p. 343.

<sup>128</sup> quoted Kuehn, p. 375, Cp MM 6:341-42. [17<sup>th</sup> July Louis XVI drives from Versailles to Paris and dons Tricolour Cockade.]

their own philosophical way to do so. However, since our concern is with the historical development of ethical ideas it is relevant to ask how *Kant* took his transcendental idealism, his ethics, his account of religion – what “Critique” or “Critical philosophy” meant to him.

That is not an easy question, not least because there seems to be no *one* way in which he took it. He may, implicitly at least, have had some variant possibilities in mind. There is much truth in Wilhelm Windelband’s comment that to understand Kant is to go beyond him.<sup>129</sup> Even to place his philosophy historically we need some idea of the space of possibilities around it. This final section is given over to some reflection and assessment.

The core of Kant’s influence is his treatment of knowledge. The first Critique stands firm, a development no-one can simply reverse. In contrast Kant’s account of ethics and religion is less masterly but also more personal. One hears in it a particular human voice. Here Kant comes across as a thinker of penetrating insights into freedom, a mind and character of moral grandeur, yet of some rigidity and many forlorn commitments and hopes. In many ways it is just this quality that has kept people interested in what he had to say. His very idiosyncrasy has left a permanent mark on late modern ethics.

#### 4.1 *Idealism and Critical philosophy*

Let us begin our reflection, however, with the Critical philosophy itself. Here there has always been a great, tantalising, question. Can anything like Kant’s Critique of knowledge be preserved, if the much-maligned distinction between phenomena and noumena is discarded? The essential Critical claim is that knowledge is possible only if it includes elements that are synthetic a priori. Can this analysis of the possibility of empirical knowledge be maintained without introducing noumena? More broadly, can a Critical defence of that possibility be given without acknowledging limits to our knowledge of the world?

If, with Kant, we answer in the negative, we face questions about the knowing subject and the phenomena. Who or what is the subject and what are the phenomena?

One approach to Kant takes “phenomena” to refer to the subjective experience of an individual mind. The subject is the subject of individual consciousness. Kant often talks in that way, and was often read in this way in his time. It would be wrong to deny that this is a strain in his philosophical thinking. Call it the *subjective reading*.

The basic question is then, for me, about *my* individual sensory experience. How do I know that my seeming perceptions, of objects that are external to my subjective experience, are genuine? There is a temptation, on this reading, to take it that “noumena” refers to the non-phenomenal ‘external’ causes or grounds of that experience. Kant’s analysis then becomes a proposal in the epistemology of perception, which holds that these external causes exist but cannot be known. One standard nineteenth century view, developed in independent ways by Mill and then by Mach, Avenarius and others, agrees that they cannot be known – because they do not exist. There are no objects that exist independently of subjective experience and cause that experience. We should instead treat talk of empirical objects reductively, as nothing but a hypothetical or ‘thought-economical’ way of talking about subjective experience.<sup>130</sup>

So does Kant take “noumena” to refer to external causes of subjective experience? We must apply his all-important distinction between empirical and transcendental idealism.

<sup>129</sup> "Kant verstehen, heißt über ihn hinausgehen."

<sup>130</sup> In Kant’s time some claimed that his view differed little from so-called “subjective idealism” of Berkeley. He strenuously reject that – B44-6/A28-30, Beiser. Broad p. 23. (This term would be better applied to Mill’s later view – see 0.0.)

At the empirical level, there is no philosophical obstacle to knowledge, whether of objects of perception – trees, mountains, planets – or of theory – electrons, waves, etc. It is for science to find out how much we can know about them, and *how* we know about them. At this level the limits (if any) to our knowledge are empirical, not transcendental. Thus the claim that there are noumena, whose intrinsic properties are *in principle* unknowable, belongs to a different, ‘transcendental,’ level of thinking. As Kant makes clear, it is supposed to follow from an a priori analysis of the preconditions of the possibility of a posteriori knowledge. So we cannot identify noumena with the empirical objects of perception and scientific theory.<sup>131</sup>

Consequences follow for Kant’s conception of the knowing subject: (1) *it has knowledge of empirical objects by empirical receptivity*, and (2) *it has this knowledge only in virtue of being transcendently receptive to noumena*. So we must apply the empirical/transcendental distinction to the subject, as Kant does. There is the knowing subject as phenomenon – as it appears to itself – and the knowing subject as noumenon – as it is in itself.

We think of empirical knowers as plural: you, me and many others. Does anything force us to think of the noumenal subject as plural? Remember that we cannot *know* that the categorial distinction between unity and plurality applies at the noumenal, as against the empirical, level. So we cannot know, and perhaps cannot even think of, ‘the’ knowing subject, self or ego as *either one or many*. It may be that absolute subjective spirit has to be conceived in terms that go beyond the category of unity versus plurality. And yet the *empirical* self still has to be thought of as the empirical appearance – presumably to itself? – of that absolute subjective spirit. It is easy to see how this puzzling outcome could inspire a development of idealism after Kant (see 0.0).

There is another interpretative option (already followed in 1.2): the *objective reading*. This takes seriously Kant’s claim to start from empirical knowledge. His aim, that is, is not to dispute with the sceptic *whether* empirical knowledge is possible, but simply to lay bare what must be the case given that it is.<sup>132</sup> On this reading, ‘phenomena’ refers to the empirical world – the world that is the object of ordinary observation and scientific theory. The knowing subject is collective: humanity, or rather, any being that can take part, with *us* – the self-reference is essential – in the collective activity of empirical inquiry.

Kant’s analysis, on this view, starts from the empirical world that we know through observation and ordinary canons of reasoning. It also takes empirical subjects and their psychological experience (the experience known to them by “inner sense”) to be objects in that world. It postulates no ontologically more basic level of pure subjective experience that is ‘outside’ it.

The Critical argument fits the objective reading well. Our collective knowledge of the world is possible only if it includes some elements that say something about the way the world is, yet are at the same time a priori. These elements have that status because *we* contribute, collectively, to the social construction of empirical reality. Our contribution is fixed by norms we share, and these in turn are epistemically grounded on spontaneous sensory and cognitive dispositions that we also share. This common structure of sensibility

---

<sup>131</sup> This is not the categorical assertion that they are *not* identical with empirical objects. But the more the constructivism of Kant’s view of empirical objects is stressed, the more the latter assertion seems to follow.

<sup>132</sup> That this is Kant’s aim has been emphasised by Karl Ameriks. See Ameriks 2000, e.g. p. 55 ff.

and understanding legitimates a priori default principles for thinking about the world, and canons for theorising about it.

Why though, on this objective reading, should we need a further step, from the possibility of synthetic a priori knowledge to transcendental idealism – with its perplexing distinction between phenomena and noumena? The world is our world: our knowledge of it is grounded in our sensibility and understanding. How would that show that there must be transcendental limits to our knowledge of it?

A way to approach this question is ask whether there might be other theories of the world, grounded in other kinds of sensibility and understanding, differing from ours. Suppose these other theories were as good as ours in predictive terms; that they went on being neither better nor worse than ours as both they and ours were improved; and that our theories and these other theories never converged as they improved. Then it seems that we simply could not know which theory, if any, was true of the world as it really is.

Someone might doubt whether we can understand the possibility of a sensibility and understanding that differs from ours. We don't have to argue this point, however, because we don't have to assume such differences. Even within our own sensibility and understanding an under-determination of theories of the world could arise. Rival theories could match each other in predictive terms, go on matching each other as they were improved, and yet never converge. The Kantian picture, in which we apply our understanding and its categories to the input of shared sensibility, contains no philosophical feature that could rule this out.

Suppose then that such an under-determination prevailed over a long period. How would we react? I suspect that we would indeed eventually conclude that we cannot know which theory is true. In that case we would be agreeing with the Kantian view: there is a way things absolutely are but we cannot know it. The very possibility of such an outcome shows us a philosophical limit on our knowledge – at least if the Kantian reaction to it is correct.

That reaction (we cannot know which of the optimal theories is true) assumes that it makes sense to ask not just what is true *within* a given theory, but also which *overall theory* (if any) is true. Is this not a kind of transcendental realism? If so, isn't Kant getting to his distinction between noumena and phenomena only by assuming the very thing he means to reject? Without transcendental realism, it could be urged, we can make sense only of a conception of truth that is *internal* to fully developed total theories of the world, in which case no distinction between phenomena and noumena follows. However Kant could reply that what he rejects is not the idea that truth is absolute – which is arguably inherent in the very idea of truth – but the idea that there are no philosophical limits on our knowledge of truth.

The possible reactions to our thought-experiment seem to be scepticism, relativism, or transcendental idealism. If scepticism or relativism are the only alternatives to it, transcendental idealism has considerable attraction.

But what if it could be shown philosophically – by a priori reasoning – that the greater the comprehensiveness and predictive power of our theories of the world, the more they must converge to a single account? It seems unlikely that that could be done. But it is in fact Hegel's speculative project. He argues that convergence on a single total view is the necessary result of the dialectic through which theories develop. This is neither scepticism nor relativism – but neither does it allow for any distinction between noumena and phenomena. We can know the world as it absolutely is. The attractiveness of this line of thought is obvious – but, of course, only if the argument that dialectical reasoning necessarily converges in the long run is sound. If it is not, we remain with transcendental idealism.



Both the subjective and the objective reading of Kant set up a momentum beyond themselves, to some variety of idealism. To put this in Hegel's terms: in the subjective reading, the momentum is towards thinking of the knowing subject as Absolute 'Spirit' or 'Subject'; in the objective reading, it is towards thinking of the world as the Absolute 'Concept' or 'Notion'. Subjectively, the momentum is towards the conclusion that there is only Spirit and its activity. Objectively, it is towards the conclusion that there is only Concept – the world as full 'realisation' of the final comprehensive conceptual structure to which dialectic leads.

The lines of thought we have sketched do not amount to what came to be called *absolute* idealism, so long as they have not taken the radical step of dispensing with the distinction between subject and object. In absolute idealism proper, Nature, Concept and Spirit have to become one. To take that step is to undermine or 'overcome' fundamental Kantian distinctions: between sensibility and understanding, understanding and willing – most generally, between *receptivity* and *spontaneity*. And that is indeed the way in which philosophical discussion in Germany after Kant went.

In following that discussion we shall find that 'overcoming' these distinctions is a leap into very strange territory. These distinctions are not just Kantian. They are distinctions which it is very hard, if not impossible, to think away. Moreover Kant's 'dual contribution' picture builds very solidly on them. The world-as-we-know-it is the product of an active impulse from us but also of an input that we passively receive, from the world as it is. In the objective reading, the knowing subject is simply an idealisation of the process of discussion among empirical persons – the collective 'we'. We empirical subjects engage in an activity of reason, interpretation, on common material which arises from our shared cognitive dispositions. This basic picture does not entail that there is a *pure given*, into which reason has not already penetrated in some way. Nor does it entail that *specific* elements in our conceptual scheme can be picked out as the synthetic a priori: it may be that a posteriori and a priori strands permeate the whole. Even so the central point remains: some distinction between an *object* given in sensibility and a thinking *subject* seems impossible to avoid.<sup>133</sup> Of course we can still ask what the significance of the distinction is. Does it force itself on us solely for the analysis of knowledge, or does it express an inescapable ontological divide? Further discussion of this issue and its ethical significance is matter for the next chapter.

#### 4.2 *Two heroic illusions about freedom*

Meanwhile, since our concern is with what Kant himself thought, we cannot accept either of the readings considered so far as accurate. Each of them fits many aspects of what Kant said, and shows how those aspects could be developed. But neither of them fits what he says about will, freedom, morality and religious faith. And for him these were the subjects of greatest importance.

The way Kant deals with them assumes that each individual rational agent – each person – is noumenal. Both as subjects and as agents persons stand in a responsive relation to reason. This responsiveness to reason, Kant rightly says, is freedom: freedom and rational agency are one and the same. Just because freedom is rational agency, Kant holds that it is noumenal. If we attempt to understand freedom empirically, he thinks, we find it to be incompatible with empirical determinism; hence, since empirical determinism holds, the

---

<sup>133</sup> Intellectual intuition.

concept of freedom cannot have empirical application. The ‘causality of freedom’ – i.e., responsiveness to reasons – must be noumenal, and so freedom is noumenal (2.1).

Thus Kant argues that *both* our knowledge of nature *and* our standing as free agents presuppose transcendental idealism: “Were we to yield to the illusion of transcendental realism, neither nature nor freedom would be possible.”<sup>134</sup> And since it is individual persons, persons among other persons, who are rational and free, it follows that each individual person is a noumenal being. This is of utmost importance to Kant’s account of acting from duty, individual moral worth and personal perfection. But it is hard to reconcile with either the subjective or the objective reading. Both readings tend in the direction of a *unique* knowing subject. In the subjective reading, it is Spirit, the transcendental self, in the objective reading the knowing subject is not personal at all: it is better thought of as an idealisation of *common cognition*, the *cognitive culture*, the *collective we*.

The argument from (knowledge of) nature to transcendental idealism may be sound, but we have argued in section 2 that the argument from freedom is not. Kant’s very real insights into ‘positive freedom’ force no transcendental conclusions. He has not shown, in the Transcendental Dialectic of the first Critique, that the antinomy of freedom and determinism can only be resolved by transcendental idealism. Not has he shown that the causality of freedom must be noumenal. He is right to hold that it is not empirical causality, the causality of nature. However that does not show that it is a different, non-empirical causality. The relation between reasons for action or belief, and a subject’s recognition of them, is not a causal relation at all. It is compatible with that point that each particular recognition of reasons will have its empirical causes (for example that one is suddenly caused by some experience to reflect about what to do). Furthermore the power to act from reasons for action that one recognises as such is an empirical power, a power one may have in greater or less degree. There is no philosophical conflict between freedom and determinism here, although certainly one’s freedom may be *empirically* limited by lack of competence in recognising reasons, or lack of psychological power to act from them.

We can thus reject Kant’s transcendentalism about freedom while accepting both his identification of freedom with rationality and the objective reading of his argument from the possibility of knowledge. What then does transcendental idealism do for Kant’s ethics?

It dramatizes a heroic view of ethics; it fits with great Christian ideals. It gives human beings a dignity based on their elevated standing above mere nature. It allows for an assumption of transcendental equality among them. It plays a role in his practical proofs of immortality and God, which picture us as a kind of transcendental community on the march to individual holiness and the highest good. This ethical ideal and eschatological picture was of great importance to Kant and to many of his readers. For all that, Kant does not establish the transcendental character of freedom. That freedom is transcendental is the first heroic Kantian illusion.

The other heroic illusion is the idea that morality can be deduced from freedom. Kant argues, rightly, that freedom is acting from laws, universal norms, of reason. He then argues that from very *idea* of a universal norm of reason one can derive a normative *content*: a substantive set of pure practical principles. No extraneous empirical premise can be allowed, for if it were, the derived practical principles would not be spontaneously accessible to any rational being. Nor is any substantive normative premise required.

This last claim aroused justified scorn. Among the philosophers who dismissed it were Hegel (0.0.0) and Mill (0.0.0). Both made the by now very familiar point that a

---

<sup>134</sup> A543/B571

normative conclusion could be derived from the Categorical Imperative only if some normative content had already been put into it. Without normative content it was, in Hegel's phrase, an "empty formalism" (0.0.0) This aspect of Kant's ethics, more than anything else, lowered its reputation for well over a hundred years. The only defence against it that we have found is to interpret the Categorical Imperative as implicitly containing a positive, and hence not purely formal, requirement of impartiality. But this is not in the end a full defence, because it still involves a non-formal move from the negative idea of disinterestedness to the positive idea of impartiality.

#### 4.3 *The epistemology of morality*

The two illusions about freedom give Kant's vision of human being much of its inspirational romantic quality. They generate the distinctive Kantian picture of the heroism of moral life and they make possible the elegance of the argument in Groundwork III. It is indeed disappointing to give them up. All the same, much of Kant's moral philosophy can survive if we remove them. Much can still be made of distinctive and important Kantian ideas, both as regards the epistemology of moral life and as regards its substantive content.

Take first the epistemology. On the heroic reading Kant has no need of an epistemology of moral judgements. Morality follows analytically from freedom, while our notion of ourselves as free is a postulate in terms of which we necessarily think, not a judgement whose truth can be, or requires, to be justified. Without the two illusions, that whole argument goes. What is still true is that we are often perfectly capable of recognising reasons to act, acting for such reasons and knowing what reasons we have acted for. Thus the possibility of autonomy remains, and with it the possibility of thinking of human beings in the intelligible perspective. 'Empirical versus intelligible' does not require 'empirical v transcendental.'

What we do now need is an epistemology of reasons, and that can be an epistemology that treats our knowledge of reasons to believe and our knowledge of reasons to act uniformly. We can get such an account from Kant by drawing on the epistemology of spontaneity and free collective discussion that was discussed in 2.5. That is in effect an epistemology of 'reflective equilibrium', in John Rawls's illuminating phrase. We bring our spontaneous normative responses into play, testing, correcting and refining them against each other and against those of other people until we reach an account that seems steadily right on reflection. 'Reflective equilibrium' is by no means unKantian; it simply brings us back from the heroic transcendental argument of Groundwork III to an equally Kantian appeal to common cognition, where common cognition now includes cognition of the intelligible domain of reasons as well as cognition of nature.

This epistemology of the normative – the epistemology of spontaneity and discussion – combines naturally with the objective reading. There is collective knowledge of the empirical world, and there is collective normative insight into reasons. The first, just because it is knowledge of the empirical world and thus calls on receptivity, requires transcendental idealism. The second, just because it is normative knowledge and normative knowledge is *not knowledge of a world*, does not call on any receptivity, rests epistemically on spontaneity and dialogue alone, and thus does not require transcendental idealism. Purely normative propositions, on this development of Kant's view, make no claims that are *either* empirical or noumenal. So Kant can rightly claim that we are "a priori conscious" of them, and that they can be "apodictically certain" (3.6).

#### 4.4 *Kantian ethics without the illusions*

Rejecting the heroic illusions need not, in itself, lead us to reject foundations for morality that are recognisably Kantian. Kant's various formulations of the Categorical Imperative retain an ethical content that remains promising, even if those formulations do not follow, as he thinks, from the very idea that we are free.

A strict Kantian might object that to accept a foundational ethical principle simply on the grounds of its plausibility and promise, without providing for it any Kantian "deduction," is to fall back on "intuitionism" about ethics – a position that Kant was concerned to avoid.<sup>135</sup> However it need not be intuitionism in Kant's sense of intuition, that is, it need not involve the idea that normative knowledge requires receptivity to some special domain of normative facts. Kant has a perfectly good epistemology of the normative, involving the notion of spontaneity alone. He is committed to it by his conception of the spontaneity of both theoretical and practical reason, and that conception is in turn fundamental to his Critical philosophy as a whole. Furthermore, in the passages we have noted, Kant takes it to be a fact that we are conscious of moral obligations. Thus both approaches to ethics – the 'heroic' route of transcendental freedom and the route that starts with common cognition and seeks 'reflective equilibrium,' can be grounded in things Kant says. And the route of reflective equilibrium, though less metaphysically ambitious than the heroic route, avoids the objections of empty formalism that the heroic route rightly incurs.

So we should take a second look at what can be done with the Categorical Imperative if we give up Kant's heroic ambitions for it. Take first the formulations of Universal Law, Autonomy, and the Kingdom of Ends, which we can treat together as a group.

Consider the maxim 'Do not pay for your travel on public transport whenever you can do so without being caught and without being noticed by fellow passengers.' That is universalisable without contradiction. If everyone followed it it would not lead to the collapse of public transport, though the salary bill for ticket inspectors, and so the cost of tickets, or of public subsidy, would rise.

At first sight, adding a requirement that one should be prepared to make one's maxims public does not help. If some attention-seeking non-conformist is prepared to go about being quite open about his tricks for non-payment, that doesn't make his non-payment right. He is still taking unfair advantage of a system of collective co-operation to which he ought to contribute his fair share.

Since we are not now trying to find a version of the Categorical Imperative that is free of pre-given moral content, we can openly appeal to the notion of fairness. And here the idea of publicity starts to play a role. We could say that an acceptable moral rule is a requirement that any reasonable, fair-minded person could agree to – one that could be proposed and accepted in a public debate among such people.

Fair-mindedness, however, is not about ignoring one's own legitimate personal (agent-relative) interests. It is a matter of not imposing them on other people in a way that those others could reasonably reject as treating *their* interests unfairly. This approach to morality gives an essential role to the idea of reasonable personal interests, which therefore needs further analysis. We noted in 3.4 that Kant has no satisfactory basis for this idea. The problem lies in the inadequacy of his account of practical reasons (3.1); he never thinks through the normative status of reasonable agent-relative aims, for this involves the idea that some agent-relative reasons are just as basic as agent-neutral ones. But since we are now

---

<sup>135</sup> cite

considering how to go beyond Kant, we can say that this only shows something about the direction in which any rethinking of those formulations of the Categorical Imperative must go.

And indeed much can be got out of this line of thought, as the last forty or fifty years of moral philosophy have shown.<sup>136</sup> But fertile as it is, can it cover everything that we or indeed Kant would think of as moral? Or is such completeness another Kantian ambition that has to be given up?

One might expect the formula of universal law, as developed in this way, to cover at most those areas of morality that involve rights, and not those areas which Kant covers in the *Metaphysics of Morals* under the Doctrine of Virtue (3.8). It would cover ‘what we owe to each other,’ to quote the title of Scanlon 1998<sup>137</sup> – so not duties of virtue to others or to oneself, in which Kant was very much interested. It would not, for example, cover obligations of gratitude, which we ‘owe’ to others in a weaker sense, nor would it cover requirements of self-respect.<sup>138</sup>

Turn then to the formula of ends. That has seemed, if anything, more fruitful. One could develop the idea of treating everyone as an end in themselves in the same direction as the one we have just been considering – in the direction of a rights theory, or some form of ideal contractualism. But the attractive thing about the formula of ends is that it is more capacious, because it allows for the idea that we should always treat *ourselves* as ends, that is, not merely use our humanity and its potentials simply as means to the ends we happen to have, but respect it and improve it for its own sake, or rather for the sake of ourselves as persons.

These are all ways of developing Kant’s ethics that can take inspiration from things he said. They have become a major activity in the post-modern period, under the influence, most importantly, of John Rawls. But they were not pursued in the late modern period, when neither rights nor social contract were in the ascendant. It is also important that they have come to be seen as a principled reaction to utilitarianism, whose force and historical impact we have yet to consider. It is in that context that their ethical significance really stands out.

Politically however the revival of Kant-inspired ethics goes beyond its promise of providing an alternative to utilitarianism. It shows something even more significant about the development of our democratic culture. Equal respect, not treating people in a way that ‘disrespects’ them, has become a key notion – and this, at least, in the abstract, sounds like a Kantian ideal. We shall come back to that as well.

#### 4.5 *Beyond the enlightenment*

The motto of enlightenment, according to Kant,<sup>139</sup> is ‘*Sapere aude*’ – Dare to Know. Some critics of enlightenment who came after him might have said ‘Dare to Believe,’ ‘Dare to Will,’ or ‘Dare to Feel.’ Others might have said that romantic ‘daring’ was not the issue – it

---

<sup>136</sup> Rawls, Scanlon ...

<sup>137</sup> Scanlon does not think that the idea of ‘what we owe to each other’ can cover the whole of morality.

<sup>138</sup> And it does not cover the notion of desert, which is important to Kant for example in his conception that the highest good is *merited* happiness, or in his retributive account of punishment.

<sup>139</sup> ‘An answer to the question: What is enlightenment?’ (1784): “*Sapere aude!* Have courage to make use of your *own* understanding! is ... the motto of enlightenment” 8: 35.



was a question of understanding, with enough humility, what human beings are really like, and finding ways to reconcile them with themselves. These critics were at one in seeing Kant's ethical theory as a one-sided rationalism, however much they admired other parts of his thought. Their view is comparable to views of French revolutionary violence that saw it as one-sided political rationalism acting on a human proneness to simplicities – *too much* 'daring to know.' It was a reaction in favour of historical sense, scepticism about reason, and reflection on the weakness of human nature. Of course Kant's *substantive* ethical, political and religious outlook was not lacking in these qualities. If his ethical theory had elements of Jacobin rhetoric and over-reaching, his substantive outlook was that of a moderate-enlightenment thinker, conscious of humanity's "crooked timber,"<sup>140</sup> with a very strong commitment to stability under law in politics, and a strongly ascetic-Christian ideal of life.

In epistemology two lines of Kantian analysis stand firm: his conception of free thinking and his transcendental idealism about knowledge. They are closely connected with each other and with the Critical approach to philosophy.

A powerful response to Kant's epistemology of free thought is that reason itself is not and cannot be free. How could a reason that works from spontaneity alone be objective? If reason is not to collapse into subjectivism, scepticism, nihilism and so on, it needs a foundation of *non*-rational belief, received from revelation, or authoritative tradition, or both.

This was the basis of Catholic reaction against enlightenment or religious 'liberalism.' In another way it was also the basis of counter-enlightenment reaction in Germany. The question was how to find a solid basis for faith, belief – *Glaube*. Was it to be found in the tradition of Catholic Christian doctrine? In the bible and private religious conscience? Or did Kant's epistemology undermine all this – whether or not he intended it too – leaving only voluntaristic options for belief, faith, value? More and more the latter became the late-modern road. Two distinctive Kantian notions – self-legislation and moral faith – had an irrationalist future in the late-modern period of which Kant would not have approved. Especially in the twentieth century, they both received a strongly voluntaristic turn. (0.0.0)

This is not Kant's view. What he thought the Critical philosophy had shown, on the contrary, was how spontaneous human reason *is* able to achieve objectivity: how objectivity and autonomy, true freedom, turn out to be two sides of the same thing. To achieve this result was, if anything, even more clearly conceived as a basic aim in idealism after Kant than it was in Kant. Inevitably this raises a question: does recognising the identity of free and objective thinking require one to accept the truth of some kind of idealism?

Critical philosophy does not need an account of the a priori that is anything like as bold as Kant's account, which picks out distinct elements in our knowledge and takes them to be not only a priori but indefeasible. Yet even a more 'liberal', holistic and fallibilist view must, to achieve its aim, retain Kant's essential philosophical argument, the skeletal structure of his philosophy. It is the argument that knowledge is possible only if it contains elements that are synthetic a priori – a rejection of empiricism – and that such elements are possible only if what Kant calls transcendental realism – the deep metaphysical assumption that underlies empiricism – is false.

Later idealists agreed with this revolutionary argument. They took it (rightly or wrongly) that rejecting transcendental realism means accepting some form of idealism. But what kind? Did it have to be Kant's transcendental idealism, with the uncomfortable dualisms and the limits on knowledge that Kant thought it required? Or could it be a more thoroughgoing, more unifying and harmonious idealism?

---

<sup>140</sup> 'Idea for a Universal History...'

One type of criticism of Kant says that his ‘revolutionary argument’ quite simply fails. It yields no intellectual answer to scepticism. Scepticism can be parried only by a non-rational attitude or act: submission to teaching or a willed act of faith. The other type, that of later idealists, says that the revolutionary argument is sound. Rejecting transcendental realism is indeed the key; but Kant does not see far enough into the kind of idealism that must replace it.

These are purely philosophical criticisms of Kant’s epistemology, and can be assessed as such. Yet they were animated by concerns that are spiritual. If, after Kant, pre-Kantian ‘metaphysical’ arguments for theism were dead – if his ‘ethical’ arguments failed – where could one find spiritual comfort? A return to Catholic teaching, taken by some,<sup>141</sup> was one possibility. Another was the voluntarist route, which we shall come to in considering Kierkegaard and Nietzsche. This sought it through some form of personally willed resignation or leap – to God, or to a new hierarchy of values. But this came later. Absolute idealism, the topic of the next chapter, sought to go further down Kant’s road, rather than abandoning it. *Kant’s* idealism seemed to imply a bifurcation, alienation, from the world and within the self. On its ethical and religious side German idealism after Kant was a major effort to deal with that.

---

<sup>141</sup> Friedrich Schlegel ... Newman.