CHAPTER 2

Kant's Antinomies of Freedom and Teleology

The chapter on 'Teleology' in *The Science of Logic* starts equating the antinomy of freedom and necessity to the antinomy of teleology and mechanism. These antinomies were introduced by Kant in the *Critique of Pure Reason* and the *Critique of the Power of Judgement*, respectively. Hegel recalls the Kantian presentation of both antinomies in the introduction to 'Teleology'. Despite their obvious differences, he claims that these two conflicts are essentially the same—if anything, the former is a broader formulation of their common logical ground.[1] In Kant's work, however, it is not apparent that they are different forms of one single opposition. It will, therefore, take some elaboration to show that Hegel is right that the concepts opposed in the two conflicts are basically the same. However, I will also argue that Kant would be willing to admit this, despite initial appearances. Understanding the logical identity of these antinomies will prove crucial to commencing the interpretation of 'Teleology' on the right foot.[2]

Hegel also believes that Kant has not resolved the conflict underlying the antinomies, despite his many assurances about having found a 'solution' to them. I will defend the claim that Hegel is also right in this regard, given the idea he has for a proper solution, as for Kant the concepts remain opposed after the sublation of their antinomy. I will argue that Kant would not contradict Hegel on this point either. Kant's solution does not imply

[1] The text asserts both the difference merely in generality and the essential sameness (GW 12:157, 158).

[2] It should help to forgo, for instance, Karl Ameriks's approach (in Ameriks 1985). According to Ameriks, Kant's argument in the first *Critique* may be 'faulty' and is no theoretical proof that *we* are uncaused causes, absolutely spontaneous, but at least he addresses the issue, whereas Hegel would have kept a 'strategic silence' about this condition (1985, 31). However, taking Kant's failure seriously and realising the essential identity of those antinomies, as Hegel does, is actually—I aim to show—a first step towards making some philosophical progress. Ameriks recognises that, in preparatory work for *The Science of Logic*, Hegel suggests that the category of 'interaction' provides a 'true solution' to the antinomy of freedom and the causality of nature, but unfortunately, he cannot find a development of this suggestion in the published version (1985, 32). Later, in Chapter 7, we will see how Hegel indeed connects interaction to purposiveness in 'Teleology'.

a unification of the opposing ideas, nor that one of them is true and not the other. He instead claims to have demonstrated their mere 'compatibility' in the first *Critique*. Hegel, indeed, considers this outcome a failure. He interprets Kant's accommodation to consist only in internalising the antinomies, thus leaving the actual conflict between the involved concepts of causality unresolved.[3]

As a matter of fact, although Kant appears satisfied with his proof of compatibility, he admits important limitations to its conclusion. In the *Critique of Pure Reason*, he acknowledges that he has not really proved the 'possibility' of a free causality. In turn, in the *Critique of Practical Reason*, he concedes that the objective reality of freedom he is able to demonstrate on the basis of the first *Critique*'s proof is only 'practical'—freedom that we cannot help but attribute to ourselves while deliberating and acting and, besides, freedom with regard to wanting, not to acting. Moreover, at the beginning of the *Critique of the Power of Judgement*, Kant accepts that the compatibility between the causality of absolute spontaneity and natural causality might merely mean that opposite principles can coexist but are perhaps separated by an unbridgeable hiatus. One of the purposes of the third *Critique* is precisely to argue that these principles can not only coexist but that they can be connected.

I shall dedicate this chapter to justifying Hegel's assumption that the antinomies of freedom (and natural causality) and teleology (and mechanical causality) are essentially the same (Section 2.1); to arguing that, despite Kant's 'solution' to the antinomy in the *Critique of Pure Reason*, the concepts remain at odds and their opposition simply interiorised, as Hegel claims (Section 2.2); and to showing that Kant's reconciliation in the first two *Critiques* is of limited value for Kant himself (Section 2.3). In the next chapter, I shall discuss the expectations raised later by the third *Critique* in this regard—which also remained unfulfilled, according to Hegel.

2.1 Two Antinomies—But One Single Conflict

In the *Critique of Pure Reason*, Kant contrasts the idea that events have an adequate explanation, or their cause, in previous events to the concept of a principle of explanation that determines strictly by itself a whole series of effects. This contrast is analysed in a chapter dedicated to the phenomenon, which Kant believes he has discovered, that reason argues contradictorily when inquiring into the ultimate, unconditioned conditions of

[3] See Wood 2010, 261–4 on Kant's alleged failure.

finite things. Kant is convinced that its natural interest in complete explanations, not relative to further conditions, leads reason to conceive in principle incompatible causal grounds of ordinary events: namely, causes that are not effects and equally infinite series of causes that are effects.

In support of each of these opposing ideas, Kant famously constructs an 'apagogical proof'. Apagogical proofs start by negating what will eventually become their conclusion (i.e., affirming the contradiction of their conclusion) and then show, or so they claim, that this negation (or contradiction) is unacceptable on logical grounds. For instance, the apagogical proof that intends to demonstrate that there must be a causality that is not the effect of another starts by assuming that there is no such thing; it then aims to show that the causality that is the effect of another supposes a prior causality that is not alike, so the initial assumption must be rejected. Similarly, the apagogical proof that intends to demonstrate that all causation is the effect of a previous causation starts by assuming that there is a causality that is not such an effect and then aims to show that the admission of a causality that is not the effect of some other precedent does not accord with the representation of a unitary, coherent experience of objects whose changes are governed by laws and is therefore unavailable to us.

Kant thus purports to prove the 'thesis' that there is a causality that operates spontaneously by strategically supposing that causes without exception have their explanation in antecedent causes. The difficulty that this assumption must face is obvious. The very idea that an actual event has its explanation in something that has previously occurred implies that the occurrence of the present event is determined in advance by what happened before. However, if this is so, the argument goes, then it cannot be admitted that the determination is provided by an endless series of past events. If the series has no beginning, then it was not determined beforehand that the current event necessarily had to occur, in accordance with the general law that connects events to other events, theoretically without exception. Kant therefore concludes that the very idea of an occurrence determined *a priori* by a law apparently supports the admission of a causation that is not the effect of a preceding causation.

The argument in support of the contradictory thesis, which Kant calls the 'antithesis', is slightly less simple. It similarly starts by denying the argumentative target: here, by affirming that there is a causality that works spontaneously. Arguably, this means—as explaining events is the aim here—that the spontaneous cause happens at some point to cause the said series. It follows that there is a change in the cause from a state that does not give rise to the series to one that does. These states that the cause goes

through therefore build a sequence. The question is, what prevents us from causally relating the states of that sequence to each other so that the posterior states are effects of the previous states? The argument runs that the rational motivation to connect some events to others in the series of effects equally requires us to associate some states with others in the presumed spontaneous cause of the entire series. As the advent of any state of the cause that could be related to the beginning of the series of effects has to be represented as occurring at a certain time, it turns out to be impossible for us to conceive of a causality that does not itself have a beginning and, therefore, we cannot truly think of a 'being a cause' which is not a 'being an effect'.

For Kant, the problem is that the two representations of a full explanation that are opposed here to each other are equally suspect. In one case, the final explanation is not given, and in the other case, the explanation that is affirmed as an endpoint does not seem to be apt for the task. Kant says that one is too large and the other too small (KrV A 422/B 450). He means by this that the infinite series of antecedent causes is too long and the singular cause that pretends to suffice is too insignificant. It could also be said, inversely, that the representation of a cause that terminates the explanation is too 'large' and that of a real cause that is just another effect is too 'small'. The important thing is, however, that both notions of a full explanation attract and repel us, to an equal extent.

The two representations and their alleged proofs make up the 'third conflict of transcendental ideas of the antinomy of pure reason', which I will abridge by calling the third antinomy of reason (or simply the third antinomy), or, as Hegel does, the antinomy of freedom and necessity. In the form it takes in Kant's text, the antinomy confronts the concept of an absolutely spontaneous causality (a 'first mover' or, in other Kantian terms, 'transcendental freedom') with that of the causality in time of antecedent events. In traditional terms, the concept of efficient cause (which would be the second) opposes the concept of a causality that is *somehow* efficient, but which, unlike ordinary efficient causes, because it is not moved to cause, exhibits rather a completely different, unique causal power (in fact, Kant speaks of legality in the former case and of 'absence of legality' in the latter).

In this sense, we might think that there is no direct relation between the third antinomy and the concept of final cause, contrary to Hegel's assumption in 'Teleology'. To be sure, the 'first mover' mentioned is defined in the first *Critique* only negatively: being *first* means not being moved, not being moved to move, having no cause and, furthermore, not being itself

in motion while moving some other thing. It does not appear to mean, therefore, to move *in order to* achieve an outcome, that is, to move for the sake of something (and hence to have a purpose).

However, with the intention of making the idea that such causes or movers exist—or that they may exist, at least—more plausible, Kant associates the abstract and difficult concept of uncaused cause with that of spontaneous (human) action:

> We have really established this necessity of a first beginning of a series of appearances from freedom only to the extent that this is required to make comprehensible an origin of the world … But because the faculty of beginning a series in time entirely on its own is thereby proved … now we are permitted also to allow that in the course of the world different series may begin on their own as far as their causality is concerned, and to ascribe to the substances in those series the faculty of acting from freedom. … If (for example) I am now entirely free, and get up from my chair without the necessarily determining influence of natural causes, then in this occurrence, along with its natural consequences to infinity, there begins an absolutely new series … For this decision and deed do not lie within the succession of merely natural effects and are not a mere continuation of them … (KrV A 449–50/B 477–8; my emphasis)[4]

The strange idea of a principle that moves without being in motion is thus related in this remark to the supposedly less strange idea of a 'completely free' action that is carried out 'without the necessarily determining influence of nature'. Kant also provides a primarily negative characterisation of the freedom of this kind of action (or, more precisely, of the freedom of the decision to which it obeys)—it is, we read, independence from coercion, not being subject to a determining natural influx. Decision and action, therefore, in Kant's eyes, 'are not part of the sequence of mere natural effects in any way'. However, the examples Kant chooses make him at once move from talking about actions that are simply not based on empirical conditions to talking positively about actions based on 'grounds of the understanding' (see KrV A 545/B 573).

[4] In this paragraph, Kant assumes for once that the need for a first start has indeed been demonstrated, although he actually denies that the proof is feasible. The demonstration is only apparent: an illusion and a sophistry. Nevertheless, for Kant the sophistry is due only to an illegitimate premise. For Hegel, on the contrary, the arguments in the antinomy of the first *Critique* prove nothing, because they take for granted what they are meant to demonstrate (GW 11:115–20; 12:157–8; cf. *Enz.* § 48 R). For a defence of the soundness of Hegel's charge, see Houlgate 2016, 47–8, and the extensive treatment in Sedgwick 2012, 169–80. However, for my purposes, the only thing that matters is that Kant assumes that if it had been possible to demonstrate that uncaused causes are intelligible, we would have to admit that there may be more than one such principle, and we would be allowed to think that human actions have a causal principle of this sort.

The idea of the freedom of a motionless motive principle thus turns out to be the foundation of a 'practical concept' of freedom for Kant, which is the concept of the freedom of those who act in light of reasons (see KrV A 533/B 561). According to Kant, such a freedom, which we attribute to human agents, 'presupposes that although something has not happened, it nevertheless *ought* to have happened' (A 534/B 562). Freedom in a practical sense is therefore conceived as *independence from natural causes to act* in accordance with the grounds that the understanding may represent. Our understanding conceives something that ought to be and, if we are practically free, it is because there is a causality that can turn it into reality (that can comply with that 'duty'). Through this means, the idea of an explanation of action in teleological terms is eventually introduced in the *Critique of Pure Reason*—without ceremony and without any proper thematisation, but in a way that will gain importance in later works.[5]

The link between the absolute spontaneity of the third antinomy and the concept of final cause is made explicit later, in the *Critique of Practical Reason*. Kant considers the possibility opened up by the 'solution' to the antinomy of freedom and necessity precisely as a preparation for the introduction of the concept of an action carried out for the sake of an end, an action that has an end for its explanation and hence a final cause. He devotes his second *Critique* to studying whether reason has what he calls a 'practical use'. The question is whether reason can be the cause of actions. Kant believes that such a capacity will be demonstrated only if it can be shown that the considerations that reason can entertain regarding what must be done are capable of causing it to be done, that is, to occur in the absence of empirical causes that produce it by themselves.

The issue is therefore whether something that reason can represent can lead to action by itself—whether a mere representation of a possible action, by virtue of some singularity it may exhibit, by virtue of some singularity of what is represented in it, can cause the action. From Kant's point of view, the relation between the causality of reason and final causality derives from the fact that the power to elicit a course of action for the sake of some represented duty is attributed to the rational faculties of agents. When it is not passions, desires or felt needs that move us to act, we arguably act, exceptionally, precisely *to* see a certain action carried out and not because this kind of action is required to produce the outcomes that we want it to have. Of course, any action we can represent may be the object of a passion or desire. The point is that for Kant, we can only speak

[5] See Allison 1990, 33.

of the causality of reason and, specifically, say that an action has been carried out purely because of rational considerations of our duty, when there are no such passions or desires intervening in its realisation.

In my opinion, it is this account of the practical-causal role of reason that properly establishes the concept of final cause in Kant's critical system.[6] The action caused by those considerations is defined as an action that occurs so that what ought to be done is actually done. Only in this case does Kant think that the cause of human action is not an antecedent event, strictly speaking. Only then does he think that reason itself, instead of something that happens to reason, is the cause of action. Admittedly, we could retort that it is the *representation* of what ought to be that causes the action, but such a representation, from Kant's point of view, is not an empirical event. Even becoming aware of that representation is not treated here as an empirical event. If there are such actions, then according to Kant it is the case that there are actions that are carried out according to the representation of duty *because* duty is a self-sufficient basis of practical determination for reason. In his eyes, only when this occurs can we properly say that an action has been done to fulfil a duty, because something ought to be done, regardless of the empirical appeal that the expected consequences might exert on our faculties. Only in such circumstances can we rightly say that an action has been taken not to satisfy a desire, not because a need has been felt, not because something has happened that has led naturally and necessarily to action, but because of—and not merely in accordance with—moral duty.

The *Critique of Practical Reason* thus contrasts natural causality with final causality, at least when speaking of human actions. Kant's reflections on the third antinomy lay the ground for understanding this relation. They show what it takes for an event—or series of events—to have a cause that is not efficient. They show therefore what it takes for it not to have a cause in an antecedent event related according to laws to subsequent events. We need indeed, Kant argues, an 'absence of legality', which is in itself not an insignificant demand. In the second *Critique*, the absence of legality is interpreted as a condition that makes it possible for something to happen *in order to* make something happen, in the final-causal sense of 'in order to'. Kant's point of view is that the absence of legality allows causes that are ends to intervene.

[6] This consequence is made particularly clear in Kant's introduction to the Doctrine of Virtue in *The Metaphysics of Morals*, in which there is talk of 'the end that is also a duty' ('den Zweck … der zugleich Pflicht ist', AA 6:382). However, herein I will not delve into the concepts that this later work articulates.

Strictly speaking, the third antinomy confronts a familiar concept of causality, that of natural events according to laws, to a causality of another kind, one without positive determination. However, Kant treats his discussion as preparing the ground for his admission, in his system of idealism, of the concept of final causality. Centuries before, Aristotle introduced in the *Metaphysics* the concept of the first mover—which Kant invokes in his remark on the thesis of the third antinomy—as a mover by attraction, which moves as a purpose does.[7] Similarly, Kant deals with the absolute spontaneity of the third antinomy as part of an idea of freedom as the power to act as exclusively attracted by an unconditional duty. It is this conceptual connection that I think justifies Hegel in his treatment of the third antinomy and the antinomy of teleology and mechanism in the *Critique of the Power of Judgement* as posing the same fundamental logical problem.[8]

2.2 Conciliation without Unification

Hegel considers the arguments of the antinomy of reason to be convoluted, superfluous and ultimately invalid. Even worse is, in Hegel's eyes, that Kant's solutions to the different conflicts do not overcome the opposition between the main concepts involved. However, I aim to show that Kant might as well admit that the opposition persists. The third antinomy, for example, confronts the concept of causality in accordance with natural laws with a concept of causality alien to that lawfulness. Given his solution to the antinomy, Kant infers that both concepts have application, that

[7] This happens in *Met.* Λ 7: 'there is a mover which causes motion without being moved, being eternal, and substance, and activity./This is how the objects of desire and of intellect cause motion; they cause motion without being moved' (1072a24–6). There is an ambiguity in *kinoúmenon* ('without being moved'), as in 'move', that has both a transitive and an intransitive meaning, so the Greek could also be translated as 'without being in motion'. I agree with Ferrarin that in the context of *Met.* Λ 7 'unmoved' means that motion is denied in the first mover (2001, 118–20). Commenting on Book Λ of *Metaphysics* in his lessons, Hegel says that '*Die absolute Substanz, das wahrhaft an und für sich Seiende ist danach das Unbewegte, Unbewegliche und Ewige*': that is, the unmoved, unmovable and eternal (VANM 8:71). On the first mover as a final cause, see Daniel W. Graham's remarks on Aristotle's *Phys.* Θ (1999, 179–80; on the ambiguity I have mentioned, see 199).

[8] For a contrasting view, see Willaschek 1991, 54–5. Willaschek holds that there is no remainder of '*finalistichen Denkens*' in Kant's talk about ends. In my view, however, the very idea of an unmovable cause is already a remainder, to say the least. Another remainder is, arguably, Kant's account of respect (*Achtung*) for the moral law in his *Groundwork of the Metaphysics of Morals*, as awareness of a *subordination*, the *effect* of the law rather than its cause, and a feeling not acquired by means of influence, but self-wrought (AA 4:401). Respect is, subjectively, the cause of moral action, although, objectively, the moral law is its cause, and respect rather an effect of the law. The (subjective) cause, that is, respect, is the (objective) effect of its own (eventual) effect, that is, the (observed) law.

their respective applications are not reciprocally hindered, but that the concepts remain the same, without transformation or redefinition. They therefore remain opposed.

This claim might be innocuous as the seemingly inconsequential fact that we have different concepts for different phenomena. Yet, in Kant's work it implies that two *opposing* explanations of the same empirical change can be combined, and Hegel considers this prospect outright inadmissible. To account for this Kantian implication that arouses Hegel's protests, I shall first present the solution in the *Critique of Pure Reason* to the third antinomy.

Kant grounds his solution on the doctrine that he calls 'transcendental idealism'—transcendental idealism is the 'key' to the solution, as he puts it. The *Critique* is supposed to have 'sufficiently proved' this doctrine beforehand, in its earlier stages. Recapitulating this philosophical achievement, Kant defines transcendental idealism, in the chapter on the antinomy, as follows:

> … everything intuited in space or in time, hence all objects of an experience possible for us, are nothing but appearances, i.e., mere representations, which, as they are represented, as extended beings or series of alterations, have outside our thoughts no existence grounded in itself. (KrV A 490–1/B 518–9)

Transcendental idealism is a doctrine that can be formulated in a few words, as we can see, and it is analysed by Kant in two components. The first component is that the objects that constitute our experience are mere 'representations' (or 'perceptions') and as such 'cannot exist at all outside our mind' (KrV A 492/B 520). They are actually very qualified representations, as they are determined according to laws—the laws of what Kant calls 'the unity of experience'—by reference to all other perceptual representations. However, they are merely representations and therefore cannot be considered 'things in themselves', whose existence and determination would be independent of their being represented in one way or another. Of course, they are representations of things, and hence not pure appearances (see KrV B 69), but as they are represented, they are not things, nor are they comparable to things.

The second component of the doctrine is the view that we represent these objects in our minds in a certain way by virtue of certain faculties of sentience and thinking. This ingredient, in turn, is complex and has several elements itself, which we will not further analyse or discuss here. It should be noted, however, as the summary points out, that the first thing the component means is that the representation of objects as 'extensive beings' (i.e., as beings that occupy a space) and as a 'series of alterations' (i.e., as occurring at

and for a certain time) depends on the determined way we represent them: the objects of experience are objects that do not exist by themselves ('have … no existence grounded in itself'), Kant claims, as extensive beings or as series of events. Space and time themselves become representations too for Kant, and not 'things', representations with respect to which the representations of particular objects are determined, that is, defined. For this reason, space and time are both representations and 'ways' of representing.

These essential components of transcendental idealism provide the basis for Kant's solution to the third antinomy. The antinomy emerges in the attempt to give full explanations for changes in empirical objects. Reason is guided in that attempt, according to Kant, by an elementary rule: if the conditioned is given, then its condition must be given too. This rule leads us either to the idea of an unconditional or spontaneous cause, or to the idea of a series of conditioned causes that has no end and thus becomes, as a series, unconditional. However, the representation of these two kinds of purportedly unconditional explainers is problematic, as we have seen. Kant thus proposes to solve the antinomy, duly circumventing these difficulties, by relativising the rule. Insofar as the objects of experience are mere representations of conditioned objects, the rule commands that those objects be referred to other objects and these further objects to others, and so on indefinitely, inasmuch as we need to determine them, account for them, individuate them or the like. However, as these objects are not 'things in themselves', and, as a consequence, they are not given as things, we are not obliged to suppose that their condition is given as a thing. The rule does not, therefore, commit us to affirming unconditional causes of one kind or another, but simply to looking in a certain direction for the necessary determination of the objects under scrutiny—namely, in previous events.

The outcomes of that search are inevitably always limited. However, the search always makes sense, as Kant hastens to emphasise. Its sense is guaranteed by the representation of space and time as *infinite* given magnitudes (KrV A 25/B 40, A 32/B 47–8). We are supposed to represent space as having no end and time as having neither beginning nor end because, according to Kant, we cannot conceive of the lack of space or a multitude of independent spaces and neither can we conceive of the elimination or suspension of time or the multidimensionality of time. However, none of these points imply that things exist in infinite, unbounded containers, as space and time are nothing but representations that act as the basis of other representations, as I have just pointed out.

We can therefore speak of the conditions of the conditioned (or finite) in several ways. We speak of *empirical* conditions when we consider

representations of objects and events that help us determine the representations of other objects and events. However, we can also speak of *transcendental* (or epistemic) conditions if we refer to the representations of space and time that, according to Kant, serve as the unconditional basis for the representation of particular objects. Finally, we can speak of *meta-transcendental* (or meta-epistemic) conditions in order to thematise the assumption, characteristic of transcendental idealism, that our representations have 'non-sensible' causes, as Kant puts it, which are *ex hypothesi* and inevitably unknown.

It is by virtue of these multiple meanings of 'being a condition' that Kant finds some use for both the representation of spontaneous causes and the representation of causes that are also effects. The causes that are also effects are empirical conditions, in the terms of the previous paragraph. An unconditional, spontaneous cause, on the other hand, cannot be an empirical condition. Can it not, however, be a thing: a 'thing in itself', or a 'non-sensible cause' of our representations? Not only may it be a thing, according to Kant, it is rather that, if time is the form of our representations, then it is not a thing by reference to which those non-sensible causes are determined as phenomena. Thus, if things themselves are causes, as Kant suggests, we have no reason to think that they are causes in the sense that (empirical) events are empirical or natural causes. Things, unlike our representations of them, do not exist in space and time. Therefore, we cannot look for a ground of their being caused going back to a previous time. Consequently, transcendental idealism makes conceivable, in Kant's eyes, a *conciliation* of the theses that are opposed in the third antinomy. The two stances are defensible from different *points of view*, because that which, as a represented thing, has a cause itself caused or conditioned, can have a cause not itself caused in the thing of which it is a representation.

Applied to the investigation of decisions and actions, such a conciliation means for Kant that the same action, as an observable phenomenon, can have two kinds of causes and be subject to two independent causal analyses. As an observable phenomenon, an action is an event. As such, it has its determination, necessity and *raison d'être* in previous events: 'If we could thoroughly investigate all the phenomena of the human will, there would be no action of man that was not predictable with certainty and that was not known as necessary taking into account their preconditions' (KrV A 549–50/B 577–8). However, as a phenomenon it also has a non-sensible cause, since it is the phenomenon of an as such non-phenomenal thing. According to Kant, the action is a phenomenon of the agent or, more specifically, of the character of the agent that Kant calls 'intelligible' (because

we can think of it, but it cannot be perceived). We may speak from the point of view of transcendental idealism both of the observable (or empirical) character of the agent and of the character of the thing, neither represented nor representable, that the agent is itself. The observable character should be understood as a complex empirical state in which the agent happens to be, and therefore interpreted as an antecedent event or state of affairs, which decision and action succeed and from which they follow. Instead, the unobservable character should be conceived of as an unconditional cause of action. In Kant's opinion, it is this second character which acts as an unconditional cause that justifies the imputation of the actions that agents carry out 'as if their author began, with total spontaneity, a series of consequences' (KrV A 555/B 583) and it is therefore irrelevant for this imputation how many empirical factors seem relevant (even decisive or necessary) to setting in motion and developing the action.

The combination of perspectives provided by transcendental idealism allows us to admit, in this sense, two opposing explanations (i.e., two *truths*) for the aetiology of action. As I have said, however, for Hegel this only internalises the formal opposition. We are told that we can explain action in two different ways. According to one of these explanations, actions are absolutely spontaneous, can only respond to grounds of the understanding and can be carried out so that something wanted (in a certain sense) becomes true. According to the other explanation, actions are events that have in previous events—which are the effects of even earlier events—their necessary trigger, spontaneity is entirely missing in them and there is no point in saying that one acts to achieve an end, not even for Kant. However, providing these explanations cannot be compared to pointing out various conditions or circumstances that make an event possible. Obviously, if an action is necessary, in the sense in which Kant understands the word, it is not spontaneous, in the sense in which Kant understands the word. Hegel therefore takes it that, on Kant's account, explanations can be combined at the price of not being able to unify views of action. In his opinion, Kant's strategy to combine those explanations relocates within reason, that is, internalises or interiorises, yet does not solve, the problem that reason cannot leave unsolved, according to the first *Critique*.[9]

[9] Some attempts at defending Kant's 'solution' as a respectable form of compatibilism are based on an interpretation of transcendental idealism as a doctrine concerning the legitimacy and compatibility of two kinds of descriptions of one single object; see, for example, Meerbote 1984 and Hudson 1994. It is therefore significant that Hegel rejects Kant's alleged solution for proposing precisely a mere compatibility of irreducible descriptions or interpretive principles (GW 12:158).

2.3 The Shortcomings of Kant's Solution

Kant himself explicitly admits other important limitations of his solutions, even though he does not recognise the tension observed above. Regarding the antinomy of reason, what the first *Critique* proves is that the two ideas, conceptually opposed, of the causality of freedom and the causality of nature are not in practice at odds. Empirical causes can thus be considered real or true, and yet there may be a (non-sensible) causality that starts 'without starting' (without being moved nor itself moving) a potentially endless series of effects. However, there *may* be such a causality. Kant does not assert that there *is*, because he believes that we cannot know objects beyond the reach of our experience. What cannot be known, however, can at least be thought, but to think about it, or to conceive of it, does not entail in this case that its possibility is properly cognised, as Kant emphasises. Neither is knowledge of this kind implied by the fact that the conciliation between an unconditional causality and the conditional causality of empirical causes has arguably been reasoned. The possibility of a non-moving cause, as Kant admits, *has not yet been proved*.

I am persuaded that this admission means that the Kantian solution to the antinomy proves only that natural necessity is not a hindrance to the actuality of the causality of freedom. There could be other such hindrances, nonetheless. Demonstrating the possibility of this causality would be like removing all potential obstacles or, failing that, straightforwardly demonstrating its reality (its 'objective reality'). In the *Critique of Pure Reason*, however, Kant is apparently satisfied with far less than that:

> It should be noted that here we have not been trying to establish the *reality* of freedom, as a faculty that contains the causes of appearance in our world of sense. For … it could not have succeeded, since from experience we can never infer something that does not have to be thought in accord with the laws of experience. Further, we have not even tried to prove the *possibility* of freedom; for this would not have succeeded either, because from mere concepts *a priori* we cannot cognize anything about the possibility of any real ground or any causality. … [To show] that this antinomy rests on a mere illusion, and that nature at least does not conflict with causality through freedom—that was the one single thing we could accomplish, and it alone was our sole concern. (KrV A 557–8/B 585–6)

This remark, with which the section on the third antinomy ends, in my opinion constitutes *velis nolis* the concession that the achieved conciliation is, after all, merely relative. If, on the one hand, we are told that the phenomena of the sensible world must have non-sensible causes, as they are appearances of something that appears, on the other, we are told that the

possibility that this causality is the unconditional causality, through freedom, that we were looking for all along has not been demonstrated nor *can* be demonstrated. It cannot even be demonstrated, I insist, that this merely intelligible causality can be spontaneous. So, Kant's arguments regarding the antinomy do not rule out that the idea of spontaneous causality does not apply to anything at all. Besides, Kant's arguments do reject the notion that an object of experience, as we experience it, could act as a spontaneous cause. Thus, bearing in mind the obligatory nature of the empirical explanation of actions, it is fundamentally misleading to say that the Kantian solution refutes the scepticism of the supporters of the antithesis.

In this sense, we might think regarding the failure that Hegel attributes to Kant—that his solution simply internalises the opposition or the original problem—that, if Kant himself did not have to admit this failure, it was thanks to his willingness to radically redefine the problem he was addressing. Kant does in fact not suggest any way in the first *Critique* to make room in nature for causality through freedom.[10] Rather than reconciling two causalities, the *Critique of Pure Reason* renders one world, under the regime of the laws of nature, compatible with another world, which is possibly under the regime of other laws.

Indeed, the other two *Critiques* are designed, at least in part, to compensate for the modest achievements of the first in this regard. However, they suffer from their own limitations, as I aim to show. To begin with, what does Kant's argument about the objective reality of freedom prove in the *Critique of Practical Reason*? Kant claims that his proof of objective reality makes the demonstration of the possibility of freedom unnecessary, as what exists must obviously be possible (AA 5:3–6). The proof starts, notably, by arguing that if someone ought to do something, they must be able to act accordingly. In turn, that we are obliged to act in a certain way is known to us immediately on Kant's account, because the 'voice' of moral conscience cannot fail to be heard. However, what can we do if our moral duty is to act in a particular way? Kant's answer is that we can at least *want* to do it. It is not that we can *do* it. If we were actually to do it, our efforts to that end would have to be successful, yet the relevant success depends, according to Kant, on circumstances and powers that

[10] Regardless of the effect that a reconstruction of the solution may have in making it philosophically respectable, the interpretation cannot overlook the way in which Kant limits the scope of his argument. In light of this limitation, it appears natural that Kant tasks the *Critique of Practical Reason* with the proof of the possibility of freedom. Those who claim that the argument of the first *Critique* implies a form of compatibilism are forced, I believe, to question the necessity of the second *Critique* for this purpose.

are beyond our immediate control. Yet, if something is not in our hands now (i.e., if we are not able to do it now), how could we now have the duty to do it?[11]

Consequently, the moral law as Kant conceives it is a law that refers purely to the will (AA 5:15, 21). It is not related to the action that is a phenomenon and, in fact, it is not even related to trying, that is, to mobilising and employing means. Anything that implies that certain things actually occur as an effect of my will depends, in Kant's eyes, on certain faculties producing those things and on suitable circumstances allowing the faculties to do so. The only thing that the moral law commands, strictly speaking, is a 'movement' of sorts of the will. The will, which Kant calls in the first *Critique* '*arbitrium liberum*', is conceived in the second *Critique* as the ability to bend to the law, abide by the law and to listen to and heed the call of duty. As such, as a capacity to accept the law or, better, to accept the law *as such*, given that its fulfilment is unconditionally obligatory, the will is the capacity to work so that something is fulfilled. We may therefore say that the law is the end of the will and that the will is the capacity to act to achieve that end. The law would thus be the cause that, without being in motion, moves the will. However, does the representation of the law thus start an endless series of effects, as the causality of freedom is meant to do according to the *Critique of Pure Reason*? It certainly does not initiate an endless series of actions. It does not even seem likely that it initiates a single action that could then have an endless series of consequences.

This realisation poses several difficulties. First, when comparing the arguments of the first two *Critiques*, the impression is produced that freedom, both in a 'cosmological' (and metaphysical) and in a practical sense, cannot be the cause of a phenomenon and, therefore, nor can it be the cause of an action that is a phenomenon, that is, an action that can be perceived in some way. This causal impotence becomes even more evident in the argument of the *Critique of Practical Reason*. Of action as an observable event, Kant admits only one kind of cause: 'empirical' faculties in empirical circumstances, acting according to their own 'empirical character'. Kant has not developed an argument proving that actions could have another type of cause. It is because Kant has not developed such an

[11] The reader should notice the *reversibility* of Kant's argument. It not only leads from a concept of duty to a concept of power, but also in reverse: from a concept of power (namely, freedom) to a concept of duty (AA 5:29). In fact, we obviously reason that because children cannot fulfil the duties of adults, they do not have to—therefore they are not required to do so.

argument that his moral psychology is necessarily complicated in order to account for how the representation of a duty can be related to actions through interests (that it could awaken) and already-acquired virtues that predispose us to act according to duty.[12] Kant's arguments prove only that the will *may* want those actions; yet, willing does not *per se* lead to action.

Another problem that arises is the following: what does 'want' mean if it does not even mean *trying*? Or, to put it differently, what is proven as happening when what is proven is something essentially imperceptible and that may not have perceptible effects? It seems obvious to me that Kant, throughout his reflections on these matters, exploits the fact that we commonly understand willing as a causal factor, even as an initial phase when carrying out an action. This would make sense of the conclusion of the argument that leads from duty to power. Kant also expects the reader to understand power as the capacity for at least minimal action, that is, as the capacity to *undertake* something. However, what Kant gives with the hand of implications, he takes away with the hand of precisions. When specifying what it is that has been shown to be possible, he can only assert that a duty is shown to be possibly accepted or assumed. Only that possibility of accepting or assuming is possible with certainty. That which can be done is not therefore an action, strictly speaking. It is a motion of the will that may have no consequences outside the will—if it has any consequences 'inside'. That there is compliance with the law is as much possible as non-compliance and equally irrelevant in practice. In sum, if the only thing that has been demonstrated is that it is possible to embrace the law, but not to carry out the law, then nothing determined about the fulfilment of duty has been demonstrated.

In my opinion, there is indeed evidence in Kant's work that these difficulties are perceptible and that his argument requires that the proof of freedom—of the power to comply with the law—be more than proof of acceptance of the law. I think it is this need which drives Kant to be consciously ambiguous when talking about the conclusion of his demonstration. When he gives examples of that power, it is actually very clear that the argument is conceived to demonstrate that it is possible to *act* in a certain way.

[12] The essence of Kant's moral psychology is found in the section 'On the incentives [*Triebfedern*] of pure practical reason' in the *Critique of Practical Reason* (AA 5:71–89). The core of the doctrine is that the moral law is, as represented by the agent, a subjective ground for the determination of action, 'that is, an incentive—to this action inasmuch as it has influence on the sensibility of the subject and effects a feeling conducive to the influence of the law upon the will' (AA 5:75). The requirement that the moral law be the (only) subjective basis for determining moral action is what is known as Kant's 'rigorism'. A penetrating discussion of this aspect of Kant's moral philosophy, on which we cannot dwell here, is found in Pippin 2001.

One such example of the power to comply is that of the person who is supposedly irresistibly inclined to pleasure. What will happen if they are threatened with death if they do not resist that inclination? They will surely resist it, but perhaps because they are willing to exchange the present satisfaction of their inclinations and appetites for many future satisfactions of the same kind. However, what will happen if they are forced to testify falsely against an innocent party under the same threat? Will they be able to give up those many supposedly irresistible satisfactions or not? Kant's answer is resolute: we do not know what they will do (even they surely do not know), but that it *could* be that they renounce these satisfactions and do not falsely testify, they will ultimately have to admit *with certainty* (AA 5:30). They might accept the risk, and not testify mendaciously, and thus jeopardise the safety of their lives and the probability of future pleasures. According to Kant, this is the conclusion to be drawn from the test. The recalcitrant hedonist will not be able to simply say: 'I acknowledge that I must not lie before the judge, since I am aware of the innocence of the accused, therefore I recognise the law, I make the law my own, I represent the law as the only thing that I myself can want to be done in situations like this, but I do not see myself able to testify on their behalf, given how true the accusers' threats and the fear they instil in me are'. Kant wants us to admit that they are capable, that is, that they *can* do, whether they like it or not, what they *ought* to do and that even if they are unlikely to do so, in view of their fondness for pleasure in the past, even if everything indicates that they will not, it is still *possible* that they will. Kant's demonstration of freedom, therefore, is not supposed to be proof that we can simply embrace the law. It is supposed to be proof that we can *do* what the law says we ought to do.

This illustration could convince us that, at least in the examples, Kant is not satisfied with demonstrating that a motion of the will for the sake of duty is always possible. It could therefore convince us that the examples contradict the limitation of the test to that motion. Yet, Kant manoeuvres opportunely to avoid this contradiction by means of a new general qualification of the meaning of the demonstration. This manoeuvre is a kind of *definitive* general measure to overcome the main difficulties that I have pointed out. Kant's reasoning seems to imply that we must choose between rejecting the rule of moral law or denying our ability to act to carry it out. However, we are not actually confronted with this dilemma, as we are about to see.

The general qualification of the proof is as follows. The demonstration of the 'objective reality' of freedom, in the sense of absolute spontaneity,

deals from the beginning with what in practice we cannot avoid thinking, accepting or supposing. We are not supposed to avoid wondering what is reasonable to want, regardless of what we actually want. We could want other things, apparently. We can, it seems, *want* to want other things. We can therefore ask ourselves what we should want. According to Kant, this question immediately subjects us to a law. More precisely, this question *in practice* subjects us immediately to a law. Kant does not claim that there is a law to which we are bound whether we ask those questions or not. If he did make such a claim, he would owe us an explanation regarding that law: who has enacted it, where has it been enacted, how can we make sure of its content and, above all, what obliges us to comply with it? For Kant, however, reason itself is the source of the law. This means that the operation of reason, its willingness to ask for explanations of what is, be they states of affairs, actions or dispositions for action, generates by itself a law and a corresponding subjection. The source of the law is a fact, but, as Kant says, it is a fact (*Faktum*) *of reason*. Consequently, Kant, who considers reason a faculty of cognition in human beings, has to see the law, inasmuch as it actualises itself against other faculties of ours (against the 'lower' faculty of desire, specifically), as a *phenomenon* of reason. Under the operation of reason, we *appear to ourselves* as subject to a law that we are capable of complying with. We cannot consider ourselves *but* subject to such a law and, correspondingly, we cannot consider ourselves *but* capable of complying with it. In the *Groundwork of the Metaphysics of Morals*, Kant explicitly introduces this view: we cannot act except under the law of freedom (AA 4:448). He never tires of pointing out that the objective reality of transcendental freedom is indubitable, but 'only practical':

> … the concept of an empirically unconditioned causality, although theoretically empty (without an intuition that fits it), is nonetheless always possible and refers to an undetermined object; … the concept is nonetheless given signification in the moral law and consequently in a practical reference. Therefore the concept, even though I do not have an intuition that would determine its objective theoretical reality for it, does nonetheless have actual application that can be exhibited *in concreto* in attitudes or maxims, i.e., it has practical reality that can be indicated; and this is indeed sufficient to justify it even with regard to noumena. (AA 5:56; cf. 48, 49)

It is by means of this qualification that Kant can say that his critique of the practical use of reason does not in any way correct the conclusions of his critique of the theoretical use. The 'objective reality' of unconditional freedom turns out to be objective validity for reason in its practical use. It is a reality, therefore, *analogous* to the empirical reality of space, time and

the objects of experience. What the second *Critique* provides is an analysis of the obligations of reason when guiding our action, just as the first *Critique* provides an analysis of the obligations of reason when guiding our beliefs (or doxastic commitments). The main difference between the results of the first and the second is that the validity of the pure forms of cognition refers to a matter that is sensed (i.e., 'given' to the senses) by virtue of the fact that when we speak of cognition we speak of how things are and what there is (in German: what things *es gibt*, what things 'are given'), not of what things can be or how actual things can be. On the other hand, the validity of practical imperatives, as they refer to a (purely) possible will, is understood to be rational and self-sufficient, independent of experience.

It is due to this limitation that it is not as important as it seems whether the argument by which that validity is affirmed speaks of actions or of elementary and imperceptible motions of the will. As the argument concerns how we consider ourselves *before* the law, we may well have to see ourselves as capable of acting in a certain way. The argument does not prove that we *are* indeed capable—it simply proves that before the law we have to consider ourselves as having the power to fulfil our duty. What is thus established is what kind of representation can be thought to play the role of a practical end of reason. However, the proof that such a representation becomes something for the sake of which reason actually acts has objective validity in practice only. As agents, we are forced to recognise the causality of reason.

The inference I draw is that the theoretical application of the concept of the final cause of action, which is not guaranteed in the *Critique of Pure Reason*, is not guaranteed in the *Critique of Practical Reason* either. This other *Critique* simply adds to the realm of application of empirical concepts (such as that of empirical cause or an event that causes) a realm of application of practical concepts (such as that of an end that is a cause), but it does so without theoretical consequences. We must not be deceived by the declarations with which the second *Critique* begins concerning a freedom that corresponds 'in reality' (*in der That*) to the human will, the demonstration of the 'reality' (*Realität*) of that freedom, a freedom that would manifest itself as 'effective' (*wirklich*), the *a priori* 'knowledge' (*wissen*) of the possibility of the idea of freedom and the possibility 'which previously was only a *problem* and here becomes an *assertion*' (AA 5:5). The final word of the *Critique of Practical Reason* is that freedom has practical objective reality—no more, no less. Our deliberations and decisions are necessarily oriented by the idea of an independence from the sensible world that the will is supposed to exhibit. This conclusion establishes that the point of

view of final causality cannot be cancelled or suspended. However, it does not take us one iota further or amend at all the conciliatory projects of the first *Critique*.[13]

Kant recognises these argumentative shortcomings in the *Critique of the Power of Judgement*. In the published introduction (the so-called 'second introduction'), he underlines that the causality of freedom 'is to take effect in the world' (AA 5:195), that there must be its 'manifestation … in the sensible world' (5:196) and that the concept of freedom 'is meant to actualise' the end proposed by its laws in the world we see and feel (5:176). Kant thus clarifies once and for all that what commands the moral law is more than a motion (of which we can 'think') of the will. Furthermore, however, he adds that the possibility of this concept becoming real must be assumed to be contained in nature itself, that is, in the world of objects.

The first argument that, according to Kant, needs to be developed to admit that freedom becomes true is still the argument of the first *Critique* that purportedly solves the antinomy of freedom and necessity. In fact, Kant refers explicitly to his original transcendental-idealist solution in two places in that introduction, as having proved that the legislation of the understanding does not imply that it is impossible, inadmissible, a causation of a peculiar kind and different from that of events (AA 5:175, 196). In his third *Critique*, however, it seems necessary—as it did not seem before—to ensure the power of nature to be determined in accordance with the principles of that other causality (the causality through freedom) laid out in the *Critique of Practical Reason*. Nature must indeed be able to be conceived as *allowing* the realisation of freedom, for, as Kant argues, 'through that concept we cognize the possibility of the final end [of practical reason] that can only be actualized in nature and in harmony with its laws' (5:196).

The *Critique of the Power of Judgement* is therefore meant to reason not a simple 'conciliation' (*Vereinbarung*) between the order of nature and the order of freedom or, as Kant also describes them, between the legislation of the understanding and the legislation of reason, but instead

[13] Allison infers that the argument on freedom in the *Critique of Practical Reason* goes further than that of earlier works in only two fundamental ways: on the one hand, it makes clear that transcendental freedom is the *raison d'être* of the authority of the moral law, that is, that a pure practical freedom is not sufficient for morality (which would solve coherence problems of the first *Critique* that cannot be entered here); secondly, because the first *Critique* and the *Groundwork* show how important the idea of transcendental freedom is for our conception of ourselves as rational agents, whereas the second *Critique* provides a proof of the reality ('actuality') of that freedom, although a proof 'merely from a practical point of view' (1990, 248). The conclusions of his study as I understand them, therefore, are at the end of the day consistent with mine.

a 'connection' (*Verknüpfung*) between the two. In the introduction, Kant grants that a simple conciliation can exist between different realms, separated by a 'great gulf', 'so that it is not possible to pass from the former [i.e., the realm of the natural concept] to the latter [i.e., the realm of the concept of freedom]' (AA 5:175; see 195). The conciliation might consist in the proof of a mere coexistence of one next to the other. The task of Kant's new *Critique* is, therefore, to build a bridge between those orders, realms and laws. It is a mediation unforeseen in the arguments of previous works, which Kant later regarded as crucial for the concept of the practical principle of unconditional validity to make sense and be applicable.[14] However, as the next chapter will make clear, the arguments of the third *Critique*, to Hegel's despair, are themselves of limited import. They do not cut across the veil of transcendental idealism, they do not transcend the opposition that separates the free and end-oriented from what is necessary or mechanical and they leave the concept of purposiveness as nothing else but a problematic concept.

[14] Paul Guyer has played down the task, to some extent, by arguing that the third *Critique* only has to provide a 'sensible' confirmation of an already proven efficacy of freedom as governed by moral principles (2006, 425). It is clear, though, that in my reading the aporiae of the previous works, which Guyer ignores in his paper, make something else necessary.