



The General Will: Rousseau, Kant, and Hegel

Citation

Hatch, Chandler Abram. 2022. The General Will: Rousseau, Kant, and Hegel. Doctoral dissertation, Harvard University Graduate School of Arts and Sciences.

Link

<https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37371141>

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

<https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFIOWI2M2EwYTg>

Accessibility

<https://accessibility.huit.harvard.edu/digital-accessibility-policy>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#)

HARVARD UNIVERSITY
Graduate School of Arts and Sciences



DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the
Department of Philosophy
have examined a dissertation entitled

The General Will: Rousseau, Kant, and Hegel

presented by Chandler Abram Hatch

candidate for the degree of Doctor of Philosophy and hereby
certify that it is worthy of acceptance.

Signature Christine A. Korsgaard

Typed name: Prof. Christine Korsgaard

Signature Lucas Stanczyk

Typed name: Prof. Lucas Stanczyk

Signature Matthew B Boyle

Typed name: Prof. Matthew Boyle

Date: 12/01/2021

The General Will: Rousseau, Kant, and Hegel

Chandler Abram Hatch

December 2021

© 2022 Chandler Abram Hatch All rights reserved.

The General Will: Rousseau, Kant, and Hegel

This dissertation argues that a central, guiding aim in the political philosophy of Rousseau, Kant, and Hegel was to produce an account of the general will, and that the accounts of the general will in Kant and Hegel are each aimed at improving upon their predecessors' accounts. Moreover, reading Kant's and Hegel this way yields compelling answers to thorny interpretive questions.

In the first chapter, I argue for a novel interpretation of Rousseau's general will as combining two elements often separated in the Rousseau literature: a publicly shared conception of the common good and a collective procedure whereby all participate in determining the general will. Together these constitute a collectively shared capacity to act in pursuit of the common good.

Kant's version of the general will, I argue, aims to overcome the difficulty that it seems almost impossible to realize a publicly shared conception of the common good and Rousseau's prescribed collective procedure in the modern state. Kant's solution is that each is to be found not empirically, but as part of an idea of reason. I argue that reading Kant's argument as a refinement of Rousseau's offers an answer to the vexed question of the derivation of the Universal Principle of Right (UPR): The argument for UPR is simply a specified form of the argument for the Categorical Imperative from the

Groundwork and the 2nd *Critique* in which freedom has been specified as external freedom.

But Kant's idea of reason cannot be exhibited empirically, and hence, our participation in legislation is discovered in philosophical reflection and not in experience. This is part of what Hegel is referring to in describing Kantian morality as a mere ought. Hegel thinks that construing morality as a mere ought leads to several tensions within morality. I provide an interpretation of several of the specific tensions Hegel outlines in the *Encyclopedia*. A proper understanding these tensions, I argue, helps us better understand Hegel's famous formalism objection. Hegel's point, I argue, is not that Kant's formal principle cannot tell us what to do in any particular circumstances, but that it requires the input of (from a rational perspective) arbitrary particular circumstances to yield any verdicts. Such verdicts, he claims, fall short of complete rationality, and the actions they endorse fall short of complete freedom.

I offer an interpretation of Hegel's own version of the general will that responds to Kant's by attempting to show how an idea of reason can, and indeed must be an object of experience. Such an idea, like Anselm's concept of God, necessitates its own existence. Hegel illustrates what such an idea would be by describing various kinds of being that approximate it. I examine in particular organisms; self-conscious, mutually recognizing individuals; and persons. I argue that the state achieves a kind of self-necessitation that these previous kinds of being lacked.

Having shown how Hegel conceives of the state as an empirically existent idea of reason, I turn to the question of why participation in the life of the state or ethical life is

rational for us or from our practical perspective. I argue that Hegel returns to Rousseau's thought that the general will is mine less because it belongs to me than because I belong to it. However, in returning to this thought, Hegel rejects Rousseau's commitment to direct participation in legislation. I offer an interpretation of Hegel's argument for an organic division of responsibilities as more rational than equal participation.

Table of Contents

Title Page	i
Copyright	ii
Abstract	iii
Table of Contents	vi
Acknowledgements	vii
Introduction	1
1. A Shared Capacity Account of Rousseau's General Will	23
2. Kant's General Will	66
3. Hegel's Criticisms of Kant's Morality Part 1	119
4. Hegel's Criticisms of Kant's Morality Part 2	164
5. Hegel's General Will	200
Conclusion	259
Bibliography	266

Acknowledgements

The first seed of this dissertation traces to Jeremy David Fix's suggestion that I work with Matthew Boyle on a second-year paper about Hegel. I had scarcely studied Hegel at the time, but I was curious. Matt is to be credited both for warning me that Hegel is impossible and for supervising me anyway, which included reading much of the *Philosophy of Right* and some secondary literature with me. My original plan was to study Hegel's philosophy of action, but I was distracted *en route* by Hegel's claim that punishment is the will of the person punished, which became the topic of my paper.

It was in my examination that Christine M. Korsgaard pointed me to the crucial background to Hegel's arguments (she suggested I start with Hobbes and Kant) that proved essential to this dissertation. Matt once again graciously offered to read with me, this time most of Kant's *Doctrine of Right*.

Matt also impressed upon me the importance of Fichte to the story I am telling. Although I have not explicitly discussed Fichte in this dissertation, my study of Fichte's *Foundations of Natural Right* greatly deepened my understanding of Rousseau, Kant, and Hegel. Stimulating discussion with the Fichte Reading Group, including Thomas Pendlebury, Sanford Diehl, and Byron Davies, was most helpful in that regard.

Lucas Stanczyk showed up at Harvard just at the right moment (my return to studying Hegel) and was extremely helpful and generous in giving comments on my papers and suggesting and discussing secondary literature with me.

I received regular and valuable feedback throughout from Christine Korsgaard, particularly in a remarkable stretch when I was under a deadline to polish the fourth chapter: For weeks she returned draft after draft with comments with very little lag time.

Matthew Boyle was likewise a regular source of valuable feedback, right up to our last meeting before my defense on the bank of the Charles.

I also received much helpful feedback from Harvard's Moral and Political Philosophy Workshop and from Rachel Achs, Sanford Diehl, Lidal Dror, and Thomas Pendlebury in particular.

Introduction

It is a common thought in contemporary liberalism that just rule must be in accordance with principles justifiable to the ruled. That thought may perhaps be traced to Rawls, who argued for a conception of justice characterized by its justifiability to individuals behind a veil of ignorance: Justice is what we would agree to if we didn't know various biasing facts about ourselves. Rawls in turn acknowledged that his conception of justice was inspired by the contractarian tradition, and in particular, to “the high point of the contractarian tradition in Kant and Rousseau.”¹

But there is an important point of discontinuity between the liberals of today and their contractarian forebears. For the great contractarians, it was not enough that one be ruled merely according to principles expressive of one's nature as a free and equal rational being; one must in fact rule oneself. Indeed, for the great contractarians, the latter claim was the explanation for the former: The reason why each must be ruled according to principles expressive of her nature as a free and equal rational being is that only in that way can each rule herself. This thought has largely vanished from contemporary liberalism.²

There's an obvious appeal to the thought that freedom requires that I rule myself. A slave whose master rules him in accordance with principles justifiable to the slave does not seem fully free. But how could there be an orderly political society in which each

¹ Rawls, *A Theory of Justice* p. 222

² Though see my remarks on Rawls in the conclusion.

individual rules herself? How could there be “a form of association that [would] defend and protect the person and the goods of each associate with the full common force, and by means of each, uniting with all, [would] nevertheless obey only himself” (*SC* 1.6.4)? This was Rousseau’s great question, and the central question of this dissertation. How can there be anything recognizable as a law-enforcing state where each citizen obeys only himself?

This dissertation will explore three contrasting answers to this question, the two that constitute “the high point of the contractarian tradition,” Rousseau and Kant, and one who left the contractarian tradition behind but incorporates elements from each of their answers, Hegel. Each of their answers centers on the idea of a general will, a will that is the will of all the people and which legislates the laws by which they are governed. Insofar as the general legislative will is the individual’s will, the laws which govern the individual are products of his own will. Hence, in obeying the law, the individual obeys only himself.

Rousseau, Kant, and Hegel were not the first to articulate political theories on which the sovereign’s will could be said to be the will of the people. But three things set these three philosophers apart from those that preceded them: first, the centrality in their political theory of the principle that each must obey only himself; second, their creativity in working out a theory of the general will that allows each to obey only himself; and third, a chain of influence by which Kant and Hegel each borrowed and transformed the ideas of their predecessors.

The first two points warrant a detailed investigation of their views. As I remarked

earlier, there is an obvious appeal to the thought that each must rule herself. Kant expressed this thought by saying that rational nature is an end in itself, that it should never be treated as mere means. To make decisions about another without consent is to treat a rational being as a mere means, to show disregard for the ends that it gives itself.

Attractive as this thought may be, it can seem difficult to imagine how a satisfactory political theory can be articulated on its basis. Fortunately, Rousseau, Kant, and Hegel have done the work of elaborating three such theories. Even if each of these theories is found wanting, each offers brilliant ideas about how the sovereign's will can be the will of the citizens. A clear understanding of the successive contributions of these philosophers to the topic may inspire future contributions.

If the first two points of unity justify an investigation of these philosophers' views, the third point of unity, the chain of influence from Rousseau to Kant to Hegel, is what justifies the treatment of their views together in one monograph. The deep interpretive challenges regarding Kant's and Hegel's political philosophy are well known. As I will try to show in this dissertation, some of these central interpretive questions can be resolved by reading Kant and Hegel as responding to their predecessors' thoughts on the general will. Most saliently, using Rousseau's argument that the general will must issue laws with universal form as a guide, we can discern a parallel argument (otherwise mysteriously lacking) for the derivation of Kant's Universal Principle of Right in the *Groundwork of the Metaphysics of Morals* and the *Critique of Practical Reason*. And using Kant's description of an idea of reason and his thought that the civil constitution (including the general will) is

such an idea, we can discern an argument for the state as an empirical idea of reason in Hegel's always difficult to interpret *Encyclopedia of the Philosophical Sciences* and *Philosophy of Right*.

This interpretive payoff is in turn part of the argument for the first point of unity that I suggested above: The principle that each must obey only himself is the central principle of the political theory of each of these philosophers. It is clear that Kant and Hegel put freedom at the core of their political philosophy, but perhaps less clear that they intend to derive their whole system from the thought that each must obey only himself. And although Kant's and Hegel's various remarks on the general will confirm that each adopted the idea from Rousseau, it is not obvious that the general will has as important a role to play in their views as it did in Rousseau's. I shall give an interpretation of Kant's and Hegel's political philosophy on which the general will that enables each to obey only himself is the central thought. One of the chief virtues of my interpretation will be that such an interpretation resolves some of the vexing interpretive puzzles regarding their views.

It is no mistake, I think, that contemporary philosophy has failed to understand Rousseau, Kant, and Hegel in the way I shall be reading them. As I remarked before, a point on which contemporary liberalism is distinct from its contractarian forebears is that it does not concern itself with the question of how individuals can be said to rule themselves. It is enough for contemporary liberalism that individuals be ruled in accordance with principles justifiable to themselves. Rousseau's, Kant's, and Hegel's

concern that the sovereign will be the people's will is easily reduced in contemporary eyes to a concern that the sovereign will conform to principles the people can will. But this reduction makes it difficult to read their arguments.

I will not seek in this dissertation to militate between the great contractarians and contemporary liberals on whether I must be ruled by my own will or whether it is enough that I be ruled in accordance with principles justifiable to me. But this dissertation may serve as a helpful preliminary to such a debate. As I have said, I take the thought that I am to be ruled by my own will to be attractive, if only it can be shown to be possible. But it can be difficult to imagine how the demand that I rule myself could be met without anarchy. Rousseau, Kant, and Hegel offer three distinct answers to this demand. Whether or not one finds their answers are fully satisfying, they illuminate possibilities that may at the very least generate hope for a solution to a problem that might at first glance seem insoluble.

To appreciate the richness of their answers, and to understand the background against which they are working, we would do well to glance at some of their predecessors' explanations of how the ruler's will (qua capacity) is the citizen's or subject's will. As I mentioned before, this thought spans the contractarian tradition. Though Rousseau was the first to shape it into the answer to the fundamental question of political philosophy, his version is in part an incorporation of and a reaction against his predecessors' ideas on the subject. We will see two ideas that Rousseau picked up from his predecessors: the thought that something is my will only if it involves my capacity of willing and the

thought that the common good is constitutive of the public will. We will see one idea in particular against which Rousseau reacted: the idea that the sovereign's will becomes the individual's by a process of authorization. But before we get to any of these influences, we must see the early theories that led to authorization-based theories.

Theologically motivated theories of representation

In his excellent study of the roots of the view that the state represents the people,³ Eric Nelson traces the idea that the state's will is the will of the citizens to theological questions. Early modern theologians were concerned about how the consequences of Adam's transgression could justly be applied to Adam's descendants. They were also concerned about how Christ could atone for the sins of Adam's descendants. In other words, how could I incur guilt or be exonerated by another's action? One possible answer is that Adam and Christ each represent me, such that whatever they do is attributable to me as well.

In virtue of what do they represent me? Various answers were proposed. One view was that God simply appointed Adam and Christ to represent me. Since God created me, he has the right to appoint representatives for me. But it's difficult to reconcile this view with God's goodness. Surely an omnibenevolent God would not hold me accountable for Adam's transgression, even if it is his right.

One more subtle variant of this view was that Adam represents me because he bears a certain sort of resemblance to me (created by God), such that I can be said to be present

³ Nelson 2019 ch. 1.

in him. Advocates of this view differentiated between accidental and essential resemblances. That someone happens to look like me does not entitle him to act on my behalf, for the resemblance is merely accidental. But someone whose essential nature resembles mine may represent me. Adam, being the father of humanity, bears this essential resemblance. This view also partly explains why Christ needed to take on flesh to atone for our sins: Without human form, he would not bear to us the essential resemblance necessary to represent us.

This resemblance account of representation was agreeable to defenders of Parliament's authority, since Parliament, consisting of members from the various constituencies, could be said to resemble the whole Kingdom much better than a single monarch. But even this more complex account is hard to square with the conviction that I cannot be responsible for others' actions without my consent. Moreover, my consent seems sufficient to designate anyone to represent me, however little we resemble one another. Hobbes takes advantage of these claims to turn the Parliamentarians' appeal to representation against them.

Hobbes: Authorization by Consent

According to Hobbes, legitimate representation happens when representatives "have their words and actions *Owned* by those whom they represent. And then the Person is the *Actor*; and he that owneth his words and actions, is the **AUTHOR**: In which case the Actor acteth by Authority" (*Leviathan* 1.16.4 p. 81). In the case of legitimate sovereignty, the subjects are the authors of the acts of the sovereign, and the sovereign is the actor

whose acts are attributable to or owned by the subjects.

As Quentin Skinner has pointed out, in connecting the sovereign's authority with the people's authorship Hobbes is borrowing language from the Parliamentarians of the 1640s. For example, Henry Parker writes in 1642 that "Power is originally inherent in the people... and so man is the free and voluntary Author, the Law is the Instrument, and God is the establisher of both" (1-2). But as we shall see, the doctrine Hobbes derives from this notion of authority is very different from that of Parker.

For Hobbes, the process by which the sovereign is given authority to act on behalf of the people is authorization. Authorization and a transfer of the right of self-governance are the two coercion-legitimizing acts in Hobbes's social compact: "I Authorise and give up my Right of Governing my selfe, to this Man, or to this Assembly of men, on this condition, that thou give up thy Right to him, and Authorise all his Actions in like manner" (87).⁴ In transferring rights, the subjects obligate themselves not to interfere with the sovereign's governing; in authorizing the sovereign, they sign their names to each of the sovereign's acts. By giving the sovereign authority to act in their names, they make the sovereign's acts their own. The people don't simply consent to have someone legislate coercive legislation; they themselves legislate by means of their authorized representative.

Hobbes deploys this notion of authorization in his arguments against Parliamentary limits on the sovereign. Given authorization, subjects cannot be wronged

⁴ Lest we think that authorizing and giving up the right to govern oneself are one and the same, Hobbes gives separate accounts of these actions in separate chapters (rights transfers in Ch. 14 and authorization in Ch. 16 of *Leviathan*).

by their sovereign nor make reasonable complaints against the sovereign, for the sovereign's acts are one's own, and therefore one ought only to complain against oneself:

"Because every Subject is by this Institution Author of all the Actions, and Judgments of the Sovereigne Instituted; it followes, that whatsoever he doth, it can be no injury to any of his Subjects; nor ought he to be by any of them accused of Injustice. For he that doth any thing by authority from another, doth therein no injury to him by whose authority he acteth... He that complaineth of injury from his Sovereigne, complaineth of that whereof he himselfe is Author; and therefore ought not to accuse any man but himselfe; no nor himselfe of injury; because to do injury to ones selfe, is impossible. It is true that they that have Sovereigne power, may commit Iniquity; but not Injustice, or Injury in the proper signification" (*L* 90).

Moreover, a successful rebellion would wrong the authorized sovereign, for it would punish the sovereign for the rebels' own actions:

"No man that hath Sovereigne power can justly be put to death, or otherwise in any manner by his Subjects punished. For seeing every Subject is Author of the actions of his Sovereigne; he punisheth another, for the actions committed by himself" (*L* 90).

And again, rebellion against the sovereign is unjust because it is unjust to do anything for which one may be punished by one's own authority:

"if he that attempteth to depose his Sovereign, be killed, or punished by him for such attempt, he is author of his own punishment, as being by the Institution, Author of all his Sovereign shall do: And because it is injustice for a man to do any thing, for which he may be punished by his own authority, he is also upon that title, unjust" (*L* 89).

However this argument is supposed to work, Hobbes plainly uses that the fact that the sovereign's acts are attributable to subject as a premise in this further argument that it is wrong to attempt to depose the sovereign.

Although later thinkers tried to resist Hobbes's absolutist conclusions, Hobbes's basic notion of representation by authorization was adopted by much of the contractarian tradition. For example, Samuel Pufendorf invokes authorization as the natural solution to a difficulty in the state of nature, viz.,

“the great Variety of Inclinations and Judgments, about discerning what is most expedient for the common End: to which there is join'd, in many Men, a Dullness of apprehending which, of several Means propos'd, is more advantageous than the rest” (*JNG* 7.2.5).

We cannot hope for a “Natural Conjunction” of wills, as with bees, “or that one Person only should exert his Will, and all the rest suppress theirs” (*JNG* 7.2.5).

“The only Method then, by which many Wills may be conceiv'd as join'd together, is at least this; that each Member of Society submit his Will to the Will of one Person, or of one Council; so that whatever this Person or this Council shall resolve, in Matters which necessarily concern the common Safety, shall be deem'd the Will of all in general, and of each in particular. For when I have made over my Power to another, his Act and Choice is interpreted as mine.” (*JNG* 7.2.5)

John Locke also adopts the idea that the commonwealth's will counts as the will of the people in virtue of an act of authorization or giving of a right to represent:

“he [the citizen] has given a right to the Commonwealth to imploy his force, for the Execution of the Judgments of the Commonwealth, whenever he shall be called to it; which indeed are his own Judgments, they being made by himself, or his Representative” (*2TG* §88).

We will see presently that Rousseau marks a departure from the authorization theory.

For Rousseau, the citizen cannot simply transfer his will to another. Indeed, Rousseau rejects the very notion of representation that the sovereign represents the will of the people. For Rousseau, the sovereign *is* the will of the people.

But Pufendorf and Locke each anticipate aspects of Rousseau's view in their disagreements with Hobbes. Pufendorf anticipates Rousseau's thought that the will is not just a social construct, but an actual capacity for self-determination, and hence it cannot simply be transferred in an act. Locke anticipates Rousseau's thought that the will has a constitutive aim of the good of those whose will it is. I will consider these anticipations in turn.

The will as a capacity and its actualization

Pufendorf defuses some of the reactionary implications of Hobbes's account of authorization by insisting that a will is not the sort of thing that can be handed over to another. Contracting to let the sovereign represent me does not by itself make the sovereign's will my will. Only continual submission to the sovereign can do that:

“By the submission of their Wills, made on the part of the Subjects, their *Natural* Liberty of Choice is not extinguishe'd; by virtue of which they are still able, *de facto*, to resume what they once gave, and to deny and withdraw their Obedience which they promis'd: as likewise, that the Strength and Power of the Subjects are not, by any Natural Conveyance transferr'd really on the Sovereign, as if, for Instance, the Strength, which lay in the Shoulders of all the Subjects, should be removed to the Prince's Shoulders; and consequently that both the Wills and the Strength of those who are govern'd, are, as we may say, turn'd into the Scales, and render'd conformable to the Governour's Pleasure, only by these two *Moral Weights*, which we come now to examine.” (JNG 7.2.5)

The two moral weights are the covenant they have made (generating a moral obligation) and the fear of punishment. Whereas for Hobbes, the social contract authorizes the sovereign to bear my will, for Pufendorf, the social contract promises the sovereign my will. But whether the sovereign continues to represent my will depends on

whether I continue to submit to the sovereign. The significance of the social contract is the moral obligation it puts me under to continue to submit as well as any means it generates to coerce me into submitting.

Consequently, Hobbes's authorization-based arguments that subjects have no right against their sovereign and can never rightfully rebel don't work on Pufendorf's notion of authorization. To be sure, in rebelling I violate my promise to submit to the sovereign, just as in Hobbes, in rebelling I violate my contract to give the sovereign my right of self-governance. But authorization cannot generate any additional arguments against rebellion or resistance for Pufendorf because insofar as a rebel, the sovereign is by that very fact no longer the representative of my will.

Rousseau will press Pufendorf's thought that the will cannot be transferred to another against theories of authorization generally. For Rousseau, as for Pufendorf, the ground of the fact that the sovereign's will is my will cannot be some past act; it must be a fact about the present constitution of my will. Both insist on this point because to say otherwise is to disregard the nature of the will as a capacity of choice and treat it as a social (or moral) construct or a fiction. We can no more transfer our will once and for all than we can transfer our physical strength once and for all. A will is a natural thing. We can at most pledge to submit it to another.

But whereas for Pufendorf, my continuing to submit to the sovereign's will is sufficient for the sovereign's will to count as mine, for Rousseau, submission is not enough. I must participate in determining the sovereign's will. Hence, while Pufendorf allows that

my will may be represented by another will, Rousseau holds that it may not. For insofar as I have participated in determining the sovereign's will, it is not the will of another. I shall argue in this dissertation that in different ways, Kant and Hegel agree with Rousseau on this point: For the sovereign's will to be my will, I must actively participate in it in some way. They disagree, however, on what constitutes adequate participation.

The guise of the good

The makings of a second line of resistance to Hobbes's absolutism can already be found within Hobbes himself. In deriving an inalienable right to self-defense, Hobbes insists on a version of the guise of the good thesis:

“Whensoever a man Transferreth his Right, or Renounceth it; it is either in consideration of some Right reciprocally transferred to himself; or for some other good he hopeth for thereby. For it is a voluntary act: and of the voluntary acts of every man, the object is some *Good to himself*. And therefore there be some Rights, which no man can be understood by any words, or other signes, to have abandoned, or transferred. As first a man cannot lay down the right of resisting them, that assault him by force, to take away his life; because he cannot be understood to ayme thereby, at any Good to himself.” (L 65-66, emphasis Hobbes's)

Humans only act voluntarily in pursuit of some good for themselves. Of course, the good may fail to come about, but if it cannot even be understood how good might come from the act, it cannot be voluntary.

Hobbes never applied this notion of the guise of the good to his doctrine of authorization. If he had, he would not have been able to claim that when the sovereign administers the death penalty, it is the will of the subject who receives it:

“The Consent of a Subject to Sovereign Power, is contained in these words, *I Authorise, or take upon me, all his actions*; in which there is no restriction at all, of his own

former naturall Liberty: For by allowing him to *kill me*, I am not bound to kill my selfe when he commands me. 'Tis one thing to say, *Kill me, or my fellow if you please*; another thing to say, *I will kill my selfe, or my fellow.*” (L 112)

Here Hobbes is trying to show the compatibility of his claims that the sovereign has the right to kill subjects and that subjects cannot alienate their right to self-defense. The sovereign’s right to kill subjects cannot derive from the subjects’ having given the sovereign the right to their lives, for that right is inalienable. Nevertheless, the sovereign is authorized to kill the subjects: They authorize the sovereign to use lethal force against them, which is to say that they acknowledge that if the sovereign does so, it will only be their own will.

But if Hobbes were consistent about his guise of the good thesis, he could not give this argument. For by Hobbes’s own lights, I can see no good in being killed. Hence, it cannot be my will that I be killed. Hence, the sovereign’s will to kill me cannot be my will. Hence, it cannot be an act I have authorized.

Locke seems to have seen this:

“But though Men when they enter into Society, give up the Equality, Liberty, and Executive Power they had in the State of Nature, into the hands of the Society, to be so far disposed of by the Legislative, as the good of the Society shall require; yet it being only with an intention in every one the better to preserve himself his Liberty and Property; (For no rational Creature can be supposed to change his condition with an intention to be worse) the power of the Society, or *Legislative* constituted by them, *can never be suppos’d to extend farther than the common good.*” (2TG §131, emphasis Locke’s)

There are two ways of interpreting this argument. On the first, the guise of the good sets a limit on the power of the legislative through the act by which it is created: The people

must have intended to establish in the legislative charter that the legislative's power is limited to the common good. Hobbes would counter that any such limitation undermines the very good that the people hope to achieve through the social contract: peace.

On the second interpretation, the guise of the good directly sets a limit on the power of the legislative (regardless of how it was chartered) because the legislative is only legitimate if it is the will of the people, and it can only be the will of the people if it aims at their good. This version of the argument poses a much more difficult challenge for Hobbes.

However Locke intended the argument, Locke regarded the will of the state as so tightly connected with the common good that even though the legislative, and not the executive, has been authorized to bear the will of the people,⁵ the executive may occasionally bear the will of the people when he acts for their good:

“a Power in the hands of the Prince to provide for the publick good, in such Cases, which depending upon unforeseen and uncertain Occurrences, certain and unalterable Laws could not safely direct, whatsoever shall be done manifestly for the good of the People, and the establishing the Government upon its true Foundations, is, and always will be just *Prerogative*” (2TG §158).

Locke gives as an example a redistricting to yield a more representative Parliament, which “cannot be doubted to be the will and act of the Society, whoever permitted, or caused them so to do” (2TG §158). Here Locke attempts to unite two apparently conflicting thoughts about the public will. On the first, the public will is a product of a social contract whereby it is vested in a particular body. On the second, the public will is

⁵ The executive “has no Will, no Power, but that of the Law” (2TG §151)

determined by its connection with the public good: Whoever brings about the public good cannot be doubted to have embodied the public will.

One way of making these thoughts consistent would be to accept Pufendorf's version of the public will on which the public will consists in the submission of all to the sovereign will. On this account the redistricting of Parliament would count as the public will because everyone submits to it, and everyone submits to it because they see it as their own good. On this interpretation, the common good would not be the constitutive aim of the public will so much as the aim that best ensures that the public will submit.

But Locke's discussion of paternal power suggests a different interpretation on which the common good is constitutive of the public will. There Locke claims that guardians of those without understanding will for their wards:

“For God having given Man an Understanding to direct his Actions, has allowed him a freedom of Will, and liberty of Acting, as properly belonging thereunto, within the bounds of that Law he is under. But whilst he is in an Estate, wherein he has not *Understanding* of his own to direct his *Will*, he is not to have any Will of his own to follow: He that *understands* for him, must *will* for him too; he must prescribe to his Will, and regulate his Actions” (2TG §58).

Locke thinks that when people are unable to will for themselves, others may will for them.

Here the power of guardian seems parallel to the executive's prerogative, for the executive must also will for others when they cannot will for themselves. And as with prerogative,

willing for one's ward is predicated on seeking one's ward's good:

“this *power* so little belongs to the *Father* by any peculiar right of Nature, but only as he is Guardian of his Children, that when he quits his Care of them, he loses his power over them, which goes along with their Nourishment and Education, to which

it is inseparably annexed, and it belongs as much to the *Foster-Father* of an exposed Child, as to the Natural Father of another.” (2TG §65)

Strikingly, being governed by another who looks after one’s good is a kind of freedom for the unfree: “A *Child* is *Free* by his Father’s Title, by his Father’s Understanding, which is to govern him, till he hath it of his own” (2TG §61). Although this freedom is not the full freedom of adults with their own understanding, the child may be said to enjoy a sort of freedom in having his decisions determined by his own good. He is not subordinated to alien ends, for his father employs his understanding to discern the good of the child and pursue that.

When a child is unable to govern himself by his own understanding, a guardian may be said to understand and will for him, insofar as the guardian does so pursuant to the child’s interests, the child may be said thereby to enjoy a kind of freedom. In other words, the guardian’s acts may be said to be the child’s.

Locke’s account of prerogative can be understood to operate on the same principle as his account of guardianship. When, due to “unforeseen and uncertain Occurrences,” the public will as vested in the legislative is unable to reason for itself, the executive may act on its behalf. Insofar as the executive acts for the public good, the executive’s acts may be said to be the public will. If this is Locke’s doctrine, then the constitutive aim of the common good can be seen as the essence of the public will. For the requirement that the public will come from the authorized legislative is not absolute, but the requirement that the public will aim at the common good is.

Hence, Locke argues that both the legislative’s and executive’s expressions of the

people's will are conditioned by a constitutive aim of the common good, which he connects with the guise of the good thesis. For something to be the people's will at all, it must aim at their good. Rousseau, Kant, and Hegel will pick up on the thought of a constitutive aim, though they will disagree on whether to characterize it as the common good.

Summary of this dissertation

The seminal move for the philosophers whose views will be considered in this dissertation was Rousseau's rejection of the notion of authorization. For Rousseau, authorization is a kind of legal fiction whereby we give others rule over ourselves, and thereby lose our freedom. If we are to remain free, we must rule ourselves. For Rousseau then, the question of how we can rule ourselves becomes the central problem of political philosophy: "To find a form of association that will defend and protect the person and the goods of each associate with the full common force, and by means of which each, uniting with all, nevertheless obey only himself and remain as free as before." Rousseau's name for the solution to this problem is the general will, a will that legislates the laws and is the will of all the citizens. Under the general will, each citizen obeys only himself and is therefore free.

In my first chapter, I argue for a novel interpretation of Rousseau's general will as a genuine shared will, that is, a single, united capacity of the collective body of the people

to determine themselves to act⁶ based on their shared conception of the good.⁷ Hence, it is not merely a voting procedure (as some have held), since a voting procedure bears no essential relation to the common good. Nor is it merely a publicly shared conception of the common good issuing in some action (as many have held), for it is only our shared will if it is our shared action. For there to be a general will, there must be both a shared conception of the common good and a shared process (deliberation and voting) by which we determine ourselves to act. But such a general will is exceedingly hard to come by, for it requires that everyone participate directly in legislation, and the sustainability of the general will and freedom it affords depend upon a high degree of political motivation from the citizens.

In my second chapter I argue that Kant's version of the general will aims to solve the problems of direct participation and motivation. Kant's chief refinement of Rousseau's general will is to characterize it as an idea of practical reason, a necessary concept of practical reason that cannot be given in experience but which may be accepted from a practical interest, provided that it does not conflict with theoretical reason. The general will is part of the broader idea of the civil constitution, which must be accepted because of a command of our practical reason that Kant calls a fact or act of reason (*Factum der Vernunft*). I argue that a key part of the idea of the civil constitution is the idea of our participation in legislation. Since participation in legislation is an idea of practical reason,

⁶ We can see here echoes of Pufendorf's thought that a shared will requires not just a promise, but an actual continuous unity of wills. Rousseau has added to it the thought that active participation, and not just submission, is required for the shared will to count as ours.

⁷ Here we see echoes of Locke's thought that the common good is the constitutive aim of the general will.

we may regard ourselves as participating in legislation even if we do not do so empirically, provided that theoretical reason does not contradict it. Moreover, if our motivation to unite ourselves to a general will is an idea of practical reason, then it is guaranteed a priori. By reading Kant's *Doctrine of Right* as a refinement of Rousseau's arguments, we can also see a compelling derivation of Kant's Universal Principle of Right, a topic that has vexed Kant scholars.

If Rousseau's characterization of the general will makes it seem impossible to achieve (in the modern world anyway), Kant's solution to the problems of direct participation and motivation may seem too easy. According to Kant, the general will must be supposed to exist even when there is little empirical semblance of rule by the people. Hegel's diagnosis is that Kant has divided the human being into rational and sensible natures and attempted to honor only the former with freedom. Hence, freedom appears to us only in the form of a command to obey, not as a sensible awareness of our self-determination.

To flesh out this criticism, I undertake an investigation in the third and fourth chapters of Hegel's famous but often misunderstood criticisms of Kant's morality, which reveals three tensions in Kant's moral view, first about the relation between duties to self and to others, second about the need to reconcile our ends of virtue and happiness, and third about reason's ability to completely determine the law in abstraction from sensibility. I argue that in each case, Hegel points out a genuine tension arising from Kant's insistence that morality is a mere ought (practical necessity), rather than an is (actuality).

In my fifth chapter, I argue that Hegel's solution to these problems has its seeds in Kant's own view of an idea of reason, a concept which is the source of cognition. What Kant failed to appreciate, Hegel thinks, is that such ideas may in fact be given in experience. Indeed, the rationally constituted state is such an idea. Kant's ideas of reason are concepts of the unconditioned condition of some object of experience. I argue that Kant's concept of a natural end (a biological organism) crudely approximates an idea of reason, for it is a concept that partly explains its own existence (through its capacity for survival and reproduction). A concept that could completely explain its own existence in experience would be an unconditioned condition, an idea of reason. Hegel's idea of the state, I argue, is meant to be such an idea of reason, given empirically.

Because reason in the form of the state is (and can be known to be) an active force in this world, reason can answer the tensions Hegel finds in Kant between duties to self and to others, and between our needs for virtue and happiness. And because the rationally constituted state is reason incarnate, reason need not, as with Kant, legislate in abstraction from sensibility; reason is sensible. In this way, I argue, Hegel marries some of the virtues of the Kantian and Rousseauvian general wills. As with Rousseau, the general will has an empirical existence, and it is my will because I belong to it (the state) rather than because it belongs to me. As with Kant, my participation and motivation are guaranteed by the fact that the general will is an idea of reason. But unlike in Kant, my motivation and participation are guaranteed an empirical existence by the state because it forms me through institutions and education to be motivated to participate. But in

bringing together elements from Rousseau and Kant, Hegel sacrifices an appealing aspect of their views: individuals' direct participation in legislation (empirical in Rousseau and intelligible in Kant). Hegel argues that we share in the general will as participants in an organic whole, each with our own role.

Chapter 1

A Shared Capacity Account of Rousseau's General Will

“Y entonces fue instituido el divino sacramento
del matrimonio, con tales lazos, que sola la
muerte puede desatarlos. Y tiene tanta fuerza y
virtud este milagroso sacramento, que hace que
dos diferentes personas sean una misma carne, y
aun hace más en los buenos casados: que,
aunque tienen dos almas, no tienen más de una
voluntad.” (*El Ingenioso Hidalgo Don Quixote de la
Mancha*, I.XXXIII, p. 339)

[“And then was instituted the divine sacrament
of matrimony with such bonds that only death
can untie them. And this miraculous sacrament
has such force and virtue that it makes of two
different persons one single flesh, and it does
more in the good spouses, who, although they
have two souls, have but one will.”]

Introduction

Rousseau proposes the general will as an answer to a problem posed by humans' need to live together in mutual dependence. Insofar as one's meeting one's needs depends upon others' decisions, one is subject to others' wills and hence appears to be unfree. Rousseau suggests that in fact there is a way for each to enjoy the benefits of society and "nevertheless obey only himself and remain as free as before" (*SC* 1.6.4):⁸ to be ruled by the general will. If all are subject only to the general will, and if the general will is the will of each citizen, then each citizen is subject only to his own will and therefore free.

In spite of its centrality to Rousseau's political philosophy and massive influence on political philosophy to this day, precisely what the general will is has been the subject of a long debate. Opinions vary widely enough that it has been possible to question whether Rousseau is (in the title of one book) "Totalitarian or Liberal."⁹ Behind some of this debate are two strands in the *Social Contract* that suggest different accounts of the general will.¹⁰ On the one hand, Rousseau in some passages suggests that the content of the general will is determined procedurally: The general will is simply whatever the people vote for when the voting is carried out correctly. On the other hand, Rousseau also suggests that the general will necessarily wills the common good, the content of which does not seem to be determined simply by the voting procedure. It's hard to see how any

⁸ I follow an established convention for citing *The Social Contract*: book:chapter[:paragraph].

⁹ Chapman (1956).

¹⁰ Sreenivasan (2000), Bertram (2012), and Williams (2015) make a similar observations.

voting procedure can be guaranteed to will the common good.¹¹ Consequently, it's hard to see how these two strands of Rousseau's thinking are to be related.

Recent interpretations of Rousseau tend to take one of these strands as more fundamental and accommodate the other as best they can.¹² What I call **proceduralist accounts** identify the general will with the outcome of a certain voting procedure.¹³

What I call **common good accounts** identify the general will with the common good, usually as it is understood by the people.¹⁴ One reason for invoking the people's understanding in a common good account is that doing so provides for a kind of freedom: Unless the citizens understand that the law aims at the common good, they won't be able to recognize themselves as free under the law, and recognition of oneself as free seems an important part of being free.¹⁵ One family of common good accounts that has been

¹¹ Philosophers since at least Barry (1965) p. 292-3 have attempted to apply the Condorcet Jury Theorems to demonstrate that a majority vote (in the right conditions) is likely to arrive at the common good. See Cohen 2010 (78ff) for the limitations of this line of thinking.

¹² David Lay Williams (2015) notes, "As commentary has evolved in the twentieth and twenty-first centuries, there have been two broad camps of interpreters: (1) those who treat the general will as a procedure for generating the substantive content of the general will, and (2) those who treat the general will as an expression of a prior commitment to substantive values" (219). However, I divide the accounts differently than Williams. Where Williams puts formal and procedural accounts into the same category, I distinguish accounts on which the content of the general will is determined by a formal principle from accounts on which it is determined by a political procedure. Our different classifications reflect our different argumentative agendas.

¹³ For example, Giddin (1983) pp. 44ff and Sreenivasan (2000). Sreenivasan refers to his approach as prioritizing the practical as opposed to the pure strand in Rousseau's descriptions of the general will.

¹⁴ For example, Dent (2005, p. 138), Jones (1987) p. 115, Charvet (1995) p. 140, Levine (1976) p. 45-95, Masters (1968) p. 326ff, and Melzer (1990) ch. 9. Most common good accounts identify the general will not with the common good itself, but with our (rational) understanding of our common good.

¹⁵ See Neuhauser (1993) for an account of this as subjective freedom. Neuhauser connects it with Hegel's notion of subjective freedom. As I will point out later, another advantage of invoking the people's understanding in the definition of the common good is that it concords with Rousseau's understanding of the nature of a rational will as a faculty for self-determination in accordance with a conception of one's own good.

popular in recent years identifies the general will with what the publicly shared conception of the common good favors.¹⁶

The divergence of interpretations between procedural and common good accounts is sometimes motivated by the assumption that one of the two strands in Rousseau's account of the general will must be fundamental. Evidence against one strand's sufficiency for Rousseau's purposes is taken to be evidence against that strand's being fundamental. For example, consider Dent's (1989) claim that the fact that even a unanimous vote might dispossess a minority provides "further reason for supposing that participation in some form of procedure for arriving at rules applicable to all is not the crux of their legitimacy, is not the decisive mark of their 'coming from all' and being the declarations of a truly 'general will'" (183). Dent seems to presuppose that there must be a single crux of the legitimacy of laws, and since no procedure is sufficient to account for the general will's aiming at the common good, he concludes that the crux must be the voters' aiming at the common good itself:

"We can only take their contribution or agreement seriously if it is rationally made, on good grounds, clearly understood. But this means that the whole weight of the issue in determining what shall count as 'coming from all' must fall on establishing what is a good ground and what it dictates. The actual participation of people in an actual procedure is altogether secondary." (201)

Dent moves directly from the claim that a vote can only be legitimate if made on good grounds to the conclusion that good grounds, and not the vote, is fundamental. Like many

¹⁶ For example, Cohen (2010), Rawls (2008), Neuhaus (1993) p. 368.

other interpreters, he fails to consider that they might both be fundamental.¹⁷

My aim in this chapter will be to argue that neither of these strands is more fundamental than the other. I will argue for a **shared capacity account** of the general will that privileges neither the procedure nor the common good to which it should point over the other. On my account, the general will is a shared capacity for self-determination in accordance with a publicly shared conception of the common good, and to be governed by the general will is to be governed by a shared exercise of this shared capacity.

“Shared” is here to be taken not in the loose sense, as when two people who each have an ability to whistle may be said to share an ability, but in the strict sense, as when supreme court justices share the ability to decide cases. My ability to whistle can be exercised independently of yours, but a supreme court justice’s ability to decide cases can only be exercised with the rest of the justices. The shared capacity for self-determination is actualized in a (strictly) shared procedure whereby we all contribute to determining our will. For this shared procedure to be a capacity for self-determination, it must be guided by a publicly shared¹⁸ conception of the common good.

I shall defend my shared capacity account with two main arguments: First, it better accommodates both strands of Rousseau’s writings about the general will. Second, construing the general will as a shared capacity suggests a compelling interpretation of

¹⁷ See footnotes 6, 7, and 9 for examples. See footnote 12 for exceptions.

¹⁸ Here again, because the conception of the good is publicly shared, the sharing is strict. Part of having a publicly shared conception of the common good is knowing that others have this conception, and since knowing is a factive state, I cannot have a publicly shared conception of the common good unless others have it. A publicly shared conception of the common good is something that belongs to me only as member of a group.

the kind of freedom Rousseau claims we enjoy under the general will. In other words, the two strands of Rousseau's general will must both be accommodated as fundamental not only on interpretive grounds (i.e., to be faithful to Rousseau's insistence upon them), but also on philosophical grounds (i.e., to produce a satisfying account of a kind of freedom). Although there have been scattered previous attempts to weave together the shared procedure and constitutive aim of the common good,¹⁹ my account is unique in showing how these two strands fit together naturally in the idea of a shared capacity for self-determination, a will whose actualization is willing, and in showing that this way of interpreting the general will provides a better account of the freedom Rousseau thinks we enjoy under the general will.

The Two Strands

Before considering procedural and common good accounts of the general will, it will help to get clear on Rousseau's commitment to the two strands. First, consider the

¹⁹ Bertram (2012) likewise argues that two strands in Rousseau's description of the general will are not alternatives but complements (p. 411), but Bertram offers a very different picture than mine. On his account, a democratic vote is compatible with everyone's self-determination because it serves three purposes everyone wills: It ensures that we all do the same thing (which is our highest priority in some cases), satisfies a demand of fairness, and serves an epistemic purpose: Voting (in the right conditions) is a reliable means to discover the common good, which we all will (pp. 409-411). In other words, it is a means to various ends, rather than, as on my view, essential to the constitution of my freedom.

Although Melzer (1990) offers a common good account, he claims that a majority vote of all the citizens is a necessary means of expressing the general will (p. 170). Melzer neglects to tell us why a majority vote is necessary to express the general will.

Noone (1980) interprets the common good as our real will and the outcome of the voting procedure as our actual will. This distinction between real and actual will is familiar from accounts of particular wills: "If the unforeseen consequences of an act are or would be disastrous, it is claimed that the actual will was not the real will. It is on the basis of this distinction that forcible frustration of an actual will is sometimes justified" (p. 74). This characterization in some ways approaches my own, but instead of exhibiting the unity of the two strands in the concept of a will, it simply gestures towards a familiar duality.

necessary connection between the general will and the common good:²⁰ “the general will is always upright and always tends²¹ to the public utility. One always wants one’s good, but one does not always see it” (*SC* 2.3.1).

Rousseau does not define the good that “one always wants,” but various remarks sketch a rough idea. It seems to include one’s preservation and freedom:

“This common freedom is a consequence of man’s nature. Its first law is to attend to his own preservation, his first cares are those he owes himself, and since, as soon as he has reached the age of reason he is sole judge of the means proper to preserve himself, he becomes his own master.” (*SC* 1.2.2)

Although Rousseau does not invoke the good here, it seems reasonable to suppose that Rousseau thinks of the preservation that the first law of man’s nature tells him to seek is a central good for humans, and that since freedom is one of “the primary instruments of his self-preservation” (*SC* 1.6.2) freedom is likewise good.²² Rousseau goes on to argue that one reason that I cannot sell myself into slavery is that doing so could not possibly be good for me, for no good could compensate for the loss of my freedom (*SC* 1.4). Moreover, Rousseau seems to think that our use of freedom to preserve ourselves belongs to our nature and hence that to give up that freedom would be to fail to preserve ourselves as the kind of beings that we are, i.e. free beings: “To renounce one’s freedom is to renounce one’s quality as a man” (*SC* 1.4.6).

²⁰ I will not attempt to define the common good. Interpretations of the common good vary with interpretations of the general will. The interpretation of the common good corresponding to each interpretation of the general will should be clear enough to follow my argument.

²¹ Tends toward but does not necessarily accomplish. One may fail in many ways, e.g. misidentifying the means to one’s good.

²² Of course, freedom’s goodness for us need not be exhausted by its use in preserving our lives.

Hence, for Rousseau, the good is not simply whatever I might prefer. I might prefer for someone else to make decisions for me, but that would not make it my good to give someone power to do so, for in doing so I would give up my freedom and endanger my self-preservation. Whatever contributions my preferences make to defining my good, there are limits to what can count as my good set by my free human nature.

Rousseau initially describes the common good as based on an agreement of particular interests:

“While the opposition of particular interests made the establishment of societies necessary, it is the agreement of these same interests that made it possible. What these different interests have in common is what forms the social bond, and if there were not some point on which all interests agree, no society could exist.” (*SC* 2.1.1)

This description leaves many questions open,²³ but assuming that one’s interests and one’s good are tightly related (Rousseau nowhere distinguishes them), the common good involves (but may not be limited to) an intersection of particular interests.

Rousseau’s argument for the claim that the general will always tends to the public utility is based on his argument that sovereignty is indivisible:²⁴

“Either the will is general or it is not; it is either the will of the body of the people, or that of only a part. In the first case, the declaration of this will is an act of

²³ Attempting to answer them now would bias the discussion of various interpretations of the general will. My own interpretation of the common good will insist that an account of it cannot be given independently of an account of the general will, since what is good for a will depends upon how it is constituted. In other words, it is a mistake to think of the common good as completely determinate independently of the nature of the general will, and the general will as merely a means to pursue this prior end. Nevertheless, my interpretation does not insist upon the communitarian thesis that the good is determined by the particular features of a given community. I am merely committed to the thesis that the common good cannot be characterized independently of the necessary conditions for the existence of a general will.

²⁴ “From the preceding it follows that the general will is always upright and always tends to the public utility” (*SC* 2.3.1). The preceding at this point are the arguments that sovereignty is inalienable and indivisible, and in particular the argument I cite here for the latter claim.

sovereignty and constitutes law; in the second case it is merely a particular will, or an act of the magistracy; at most it is a decree.” (*SC* 2.2.1)

In other words, for something to be the general will, it must be everyone’s will in some sense. If *x* is not everyone’s will, then *x* is not the general will but the particular will of some. We can now see why the general will must will the common good: If, as Rousseau claims, one always wills²⁵ one’s good, then whatever everyone wills is everyone’s good, or the common good. And the general will is always the will of everyone. Hence, the general will always wills the common good. Rousseau begins from a version of the guise of the good thesis about individual wills and argues that this same thesis applies to the general will. If each wills her own particular good, then whatever the people all will together must aim at everyone’s good.

The claim that the general will aims at the common good is complicated by Rousseau insistence that the general will can only be expressed in a vote of all the citizens. If the standard for such a vote were unanimity, there would be no puzzle: If in voting, everyone aims at her own particular good, then when the voting is unanimous, the result of the vote will be guaranteed to aim²⁶ at the common good. But according to Rousseau, except for the social contract itself a majority is sufficient (*SC* 4.2.5); “Except for this primitive contract, the vote of the majority always obligates the rest” (*SC* 4.2.7).²⁷

²⁵ I have substituted “wills” for “wants” in the original quotation above. The context suggests that, if there is a want/will distinction to be made, Rousseau means “wills”: “the general will is always upright and always tends to the public utility. One always wants one’s good, but one does not always see it” (*SC* 2.3.1).

²⁶ Again, it could miss the common good if the citizens are not well-informed, but its aim would be the common good.

²⁷ Rousseau thinks that unanimity is desirable and even likely in the best sort of states, but it is not necessary for freedom. For an interpreter who thinks that unanimity is essential to Rousseau’s general will and finds an insidious tyrannical spirit lurking in it, consider Talmon (1960) pp. 46ff.

Moreover, Rousseau continues, in a proper vote, the citizens are not being asked “whether they approve the proposal or reject it, but whether it does or does not conform to the general will that is theirs;²⁸ everyone states his opinion about this by casting his ballot, and the tally of the votes yields the declaration of the general will” (*SC* 4.2.8). Thus, unanimous vote account of why the general will always aims at the common good is wrong on two counts: 1) It requires unanimity, and 2) it assumes that voting is a matter of declaring one’s own private good. We need an account of the general will that better accommodates Rousseau’s commitments.

A Procedural Interpretation

Rousseau’s claim that “the tally of the votes yields the declaration of the general will” (*SC* 4.2.8) has been taken to suggest a procedural interpretation of the general will,²⁹ according to which the general will just is the result of a certain voting procedure. The general will is everyone’s will in the sense that we all participate in determining it, and we all consent to abide by the results (*SC* 4.2.8). Rousseau adds, “when an opinion contrary to my own prevails, it proves nothing more than that I made a mistake and that what I took to be the general will was not” (*SC* 4.2.8). The procedural interpretation of the

²⁸ Sreenivasan (2000) disputes this claim on the grounds that Rousseau elsewhere seems to accept people’s voting entirely on considerations of private interest: “Why is the general will always right, and why do all constantly want the happiness of each of them, if not because everyone applies the word *each* to himself and thinks of himself as he votes for all? This proves that the quality of right and the notion of justice it produces are derived from the preference each person gives himself, and thus from the nature of man” (*SC* 2.4.5). But as Sreenivasan acknowledges, the context of this passage restricts the proposed laws to ones that would be universally binding on and hence equally burdensome/beneficial for all (p. 548). In the passage I’ve cited from Book IV, Rousseau clearly thinks that the proposed laws cannot be guaranteed to fulfill this universality condition, and hence that the voter is not exempted from the task of judging whether the universality condition has been fulfilled.

²⁹ For example, Giddin (1983) pp. 44ff and Sreenivasan (2000).

general will easily accounts for this claim: If the general will just is the result of a certain voting procedure, then it is clear how the result of the vote is proof of the content of the general will.

Joshua Cohen worries that if the general will just is the outcome of the procedure, then the question voters answer cannot be, as Rousseau insists, whether the proposal “does or does not conform to the general will that is theirs,” since that would amount to asking voters to express in their vote their prediction about the outcome of the vote. Cohen writes, “it is perverse to think of individual votes on a question as predictions of the outcome of the collective vote.”³⁰ For Cohen, this is reason to reject the procedural interpretation.³¹

Cohen is right that proper voting should not be a prediction of the outcome of the vote, but this response to the procedural interpretation is too quick. If I ask myself, as I begin to deliberate, “What is my will on this matter?” I am not expressing my intention to predict the outcome of my own deliberation. I am merely expressing the open-endedness of the process, the awareness that I myself will freely decide my will. A view on which the general will just is the outcome of the voting process can allow that voters ask themselves the question “What is the general will on this matter?” in the same spirit. The question

³⁰ Cohen (2010) p. 77. Williams (2015) cites Cohen’s argument on this point approvingly (p. 225).

³¹ On the other hand, Robert Paul Wolff (1970) thinks that Rousseau is making the very proceduralist mistake of which Cohen tries to exonerate him: “When the assembly is asked ‘whether (the proposition before them) is conformable or not to the general will,’ we may view them either as being asked for their opinion of the value of the proposition for the general good, or else as being asked to make a prediction of the outcome of the vote. I suggest that Rousseau himself confused these two senses, and was thereby led into the manifestly false assumption that the majority opinion of the assembly would successfully express what the minority were really striving for, and hence be binding on everyone who voted for or against.” p. 57.

does not instruct them to predict the outcome of their voting. It expresses the fact that they will freely determine the general will and that the perspective from which they are to determine the general will is not the perspective of the private individual but the perspective of citizen.³²

Similarly, from the proceduralist's perspective, when Rousseau says that the majority's going against me proves that I was mistaken about the general will, there is no need to interpret the mistake as a failure in the voter's deliberative activity. Suppose I must make decisions for others who have not specified their preferences on a matter. Suppose that my suspicions about what they would choose and my understanding of what would be good for them come apart. I take it that in the absence of instructions, I should hold to what I take to be good for them in making decisions for them. For it is far better to err on the side of using my best judgment about what is good for them than to err on the other side ("Sorry, I knew it wasn't good for you, but I thought it's what you would have chosen."). Now if you in fact would not have chosen what I take to be good for you, there is a sense in which I was mistaken in the decision I made for you: not that I should have decided otherwise, given my knowledge at the time, but simply in that I did not accurately represent your will. The voter in Rousseau's society is in roughly this sort of situation: He is asked to participate in a decision that will bind others. What the law should in fact be

³² Perhaps Rousseau should have worded the question, "What accords with the common good?" Then it would be clear that the general will (which on the procedural interpretation is indeterminate until the voting has taken place) is distinct from the consideration guiding deliberation, the common good (which is determinate before voting). Still, Rousseau's wording is hardly grounds for rejecting the procedural interpretation.

depends upon the will of the people, and given that he does not yet have information from the people about their will on the matter (nor could he, because the vote is the process by which their will is made up), he ought to vote based on his best understanding of what is good for the people.

This may seem an unexpected sense of being mistaken, but it matches the argumentative context fairly well. Rousseau's claim that the vote proves that "I made a mistake" comes in the context of a question about how people can be "free and subject to laws to which they have not consented" (4.2.7). The worry is that when I am part of the minority of voters on an issue, I am ruled by a will that is not my own. This worry presupposes that my vote expresses my will. Rousseau's response is that the relationship between my vote and my will is at best thought of as a kind of opinion, since the actual content of my will is determined by the vote itself: "The Citizen consents to all the laws, even to those passed in spite of him" because "the steady will of all the members of the State is the general will" (4.2.8). But because of the role that my vote plays in the determination of the general will, it's not an opinion in the sense of a "best guess," nor is it mistaken in the sense that I should have guessed otherwise. It's only mistaken insofar as it is thought of as an attempt to speak for the people or purports to express what my or our will is (as the worry presupposes).

A bigger worry for the proceduralist account is that Rousseau seems to deny the procedural account in the very next paragraph: "This [that the tally of votes declares the general will] presupposes, it is true, that all the characteristics of the general will are still

in the majority: once they no longer are, then regardless of which side one takes, there is no longer any freedom” (4.2.9). The idea that certain characteristics of the general will (about which more in the following paragraph) ought to but may not adhere in the majority seems hard to square with the idea that the general will is a procedure. It sounds like the procedure only yields the declaration of the general will under certain conditions.

Gopal Sreenivasan (2010) tries to tackle this objection by incorporating these conditions into a sophisticated account of the procedure. Rousseau does not specify what the “characteristics of the general will” are in this passage, but elsewhere in the text he suggests several conditions that must be met if the vote is to express the general will. For example, when a faction “is so large that it prevails over all the rest, ... there is no longer a general will, and the opinion that prevails is nothing but a private opinion” (*SC* 2.3.3). Sreenivasan glosses this as a requirement that the voters think for themselves. People who vote with a faction violate this requirement, and the procedure not having been followed, the result is not an expression of the general will. Sreenivasan identifies the following as Rousseau’s conditions for the expression of the General Will:

1. The subject matter of deliberation is perfectly general.
2. The conclusions of deliberation apply equally to all the citizens.
3. All citizens participate in deliberation.
4. All parties to deliberation think for themselves.

But Sreenivasan himself admits that these constraints do not seem to guarantee that the

resulting procedure aims at the common good.³³

Sreenivasan leaves out of his account two other conditions for the existence of a general will suggested by Rousseau: that socioeconomic inequality be limited (*SC* 1.9.8fn.) and that the people have the right sort of patriotic spirit (*GP* 4.1), which Rousseau thinks can be cultivated by a long process of civic education. Socioeconomic inequality must be limited so that the people's interests will coincide on the same laws, and so that no one can use economic means to control others' votes (*SC* 2.11.2). Patriotic spirit is important to motivate participation and prioritization of the common good.

Sreenivasan would need a different kind of procedural account to accommodate these other conditions, for unlike Sreenivasan's constraints, they aren't rules for how to vote, but rather the background conditions against which a voting procedure can yield the general will. If we fail to meet them, we cannot sit down to legislate properly until society has already undergone a massive overhaul.

This suggests a different sort of procedural account, one on which the general will is determined by the outcome of a procedure against certain background assumptions about the society. Giddin (1983) gives just such an interpretation of the general will.³⁴

³³ p. 574ff. Alternatively, Sreenivasan claims, if one additionally accepts a procedural account of the common good, on which the common good just is whatever people decide on through the procedure, then the general will is guaranteed to aim at the common good. But, Sreenivasan argues, a procedural account of the common good does not help the procedural account of the good, for if the common good is just whatever the people decide on through the procedure, then there is nothing to guarantee that the common good is my individual good, and this guarantee is crucial to Rousseau's argument for signing the social contract in the first place. Moreover, if Rousseau conceived of the content of the common good as determined procedurally, it would be mysterious why Rousseau gives an argument that the general will aims at the common good.

³⁴ Ch. 2.

But a pair of interpretive issues conspire to cast doubt on this sort of account. First, Rousseau doesn't give a very sharp definition of these background constraints, and hence it would be difficult to be sure whether one's society fulfilled them enough for the legislature to express a general will. Consider Rousseau's specification of the amount of material inequality that should be allowed in a society:

“as regards wealth, no citizen [should] be rich enough to be able to buy another, and none so poor that he is compelled to sell himself: Which assumes moderation in goods and influence on the part of the great, and moderation in avarice and covetousness on the part of the lowly” (*SC* 2.11.2).

How much money a person needs to insure him against having to sell himself depends not just upon how much he has relative to the rich, but also on the degree of avarice and covetousness among the lowly.³⁵ If the definition of material inequality is vague, the requirement that people have the right patriotic spirit is left even vaguer. A good first step would be for citizens to be patriotic enough to vote for the common interest over private interest. But unless they roughly share an understanding of the common good, this won't yield a result that each will be able to view as their own will.

Vagueness does not in itself speak against a procedural account of the general will, but these particular examples of vagueness suggest that what is essential to the general will is not a list of background conditions so much as that these background conditions conspire to produce a vote aimed at the common good. Rather than propose, as these procedural accounts do, that the general will is a set of procedural constraints against

³⁵ For a helpful discussion of this idea, see Ci (2013)

background conditions that succeed in aiming the vote at the common good, why not include in the definition of the general will a constitutive aim at the common good? Of course, it would be nice to have a complete set of necessary and sufficient conditions for a voting procedure to aim at the common good, but what reason do we have for thinking that Rousseau intended to offer such a complete set? Not only did he not argue for the completeness of any set of conditions, he didn't even compile a list of them in one place. The scattered, at times offhand gestures towards what is required suggest that Rousseau was pointing out common pitfalls rather than giving us a recipe.

Second, procedural accounts have a hard time explaining Rousseau's claim that the general will is never annihilated or corrupted, even when voters use the legislative procedure to advance their private interest:

"Does it follow that the general will is annihilated or corrupted? No, it is always constant, unalterable, and pure; but it is subordinated to others that prevail over it. Each person, in detaching his interest from the common interest, sees clearly enough that he cannot separate them entirely, but his share of the public evil seems to him as nothing compared to the exclusive good he seeks to make his own. Except for this particular good, he wills the public good in his own interest just as strongly as anyone else." (4.1.6)

In this passage Rousseau clearly associates the existence of the general will not with a procedure but with each person's recognition of a common interest that he wills even while making an exception to it for private interest. If the general will just is a procedure in the presence of certain background conditions, it's hard to see how an individual's willing of the common interest could count as the general will in any sense.

This is not merely a concern with squaring the various things that Rousseau claims

about the general will. There are two reasons for Rousseau to insist that the general will aims at the common good. First, unless I am assured that the general will aims at the common good, it's not clear why I have reason to agree to the social contract, through which Rousseau promises I will have my person and goods defended by the whole and remain as free as before I signed it (*SC* 1.6).

Second, if the general will is to be a will, it must have an aim in acting. As we have seen, Rousseau accepts a strong version of the guise of the good thesis: A will always aims at its own good. A voting procedure that is capable of behaving as erratically as a coin toss, now swinging one way, now the opposite is not, by Rousseau's lights, by itself a will. It could at most be a will under the right conditions (including e.g. socioeconomic equality, love of country, and a shared conception of the common good), which make it aim at the common good. To be a will, something must fairly consistently aim at its own good.

Rousseau appeals to his version of the guise of the good thesis in his arguments that man cannot sell himself into slavery: "To say a man gives himself gratuitously is to say something absurd and inconceivable; such an act is illegitimate and null, for the simple reason that whoever does so is not in his right mind" (*SC* 1.4.5). A man cannot sell himself into slavery because an act that departs so radically from his good cannot be viewed as his will, but only as madness. This is not to say that the madman no longer has a will in the sense of a capacity to determine himself to act, but only that the attempt to sell himself into slavery cannot be taken for an expression of it. Similarly, a voting

procedure's radical departure from the common good does not show that there is no general will, but only that the voting procedure has not expressed it: "Does it follow that the general will is annihilated or corrupted? No, it is always constant, unalterable, and pure; but it is subordinated to others that prevail over it" (*SC* 4.1.6). For the general will to be a will, it must aim at some good, and for it to be general, it must aim at the common good.

Of course, one might reasonably wonder whether it matters that the way we organize our society take the form of a will and whether we need to think of the general will as having the structure of a will. In a letter to Mirabeau, Rousseau suggests that "the great problem of politics" is "to find a form of government that might place the law above men" (*LM* 4). It's not immediately obvious why the source for such a law must be something that can properly be called a unified will. Many customs and traditions that govern societies do not seem to have arisen from any unified will, but rather from the contributions of many wills, often unconscious of their contributions. I will argue that for Rousseau's account of freedom under the general will to make sense, we need to think of the general will as a shared will, but I will only be in a position to make that argument once we've seen the limitations of common good accounts.

Common Good Accounts

Common good accounts define the general will in terms of the common good or our rational understanding thereof, independently of the results of any vote. One of the most thoroughly developed and representative recent examples of a common good account is

Joshua Cohen's.³⁶ On Cohen's interpretation, for a general will to exist citizens must share and know that they share a conception of their common good that they are motivated to prioritize over their particular interests in the relevant circumstances. The general will itself is the capacity to act for reasons based in this shared conception of their common good.³⁷ From Cohen's account, it is immediately apparent why the general will always aims at the common good, for the general will is defined in relation to it.

It is thus also apparent why we each have a general will. For not only do we each have the public conception of the common good, we also have the capacity to act on reasons based in it. And it is apparent why we might agree to be governed by the public conception of the common good: For something to be the common good is for it to be good for all, including for me and those I care about. Whereas it is not clear why I would agree to a voting procedure that provided no guarantee against a tyranny of the majority, it is clear why I would agree to be governed by a conception of what is by my own lights good for all of us.

Common good accounts are also able to give a convincing reading of Rousseau's claim, considered in the last section, that the general will persists even when the vote is corrupted such that the general will is no longer expressed in the vote. Rousseau indicates that the general will survives in such cases in the individuals' willing of the common interest:

³⁶ See footnotes 4 and 5 for more examples of common goods accounts. As I shall argue, it gets an important aspect of the general will right, but it fails for reasons inherent in all common good accounts. Hence, my criticism of Cohen's account will apply to common good accounts more generally.

³⁷ Cohen (2010) p. 61

“Each person, in detaching his interest from the common interest, sees clearly enough that he cannot separate them entirely, but his share of the public evil seems to him as nothing compared to the exclusive good he seeks to make his own. Except for this particular good, he wills the public good in his own interest just as strongly as anyone else.” (*SC* 4.1.6)

The interpretation of the general will as a publicly shared conception of the common good (or interest) accommodates this passage easily: Individuals retain their shared conception of the common good, but they subordinate it to their private interest.

But on common good accounts like Cohen’s, the voting procedure that Rousseau advocates becomes more puzzling. First, on such views, it’s not clear that there must be any authoritative body at all, let alone one in which all citizens participate. Cohen merely suggests that citizens must have “a reasonable confidence that the institutions conform to their shared conception of the common good, and those social institutions do in fact generally conform to it.”³⁸ As long as these institutions conform to the shared conception of the common good, it seems that they need not even form one government. If a general will is simply a capacity to act from the right kind of reasons, anyone, including non-governmental organizations, can express the general will simply by acting on those reasons. And provided that we have reasonable confidence that they are acting in conformity with those reasons, we can accept that such organizations are expressions of the general will. Yet Rousseau insists that only the sovereign assembly can express the general will (*SC* 2.1, 3.15.5-6).

Second, even assuming that there must be an authoritative declaration of the general

³⁸ P. 58

will, it is not clear from common good accounts why a vote of all the citizens is required to declare it. A suitably chosen representative body can make an authoritative decision based on a publicly shared conception of the good. Yet Rousseau insists that everyone be given an equal vote in the legislative assembly and that no representative legislature can be legitimate (2.1, 3.15.5-6).

Cohen responds to this difficulty by arguing that direct participation in lawmaking “is about preserving sovereignty, about ensuring its stability” and not “a defining condition in the conception of sovereignty itself.”³⁹ I find it difficult to square Cohen’s interpretation with passages like the following:

“Sovereignty cannot be represented for the same reason that it cannot be alienated; it consists in its very essence in the general will, and the will does not admit of being represented: either it is the same or it is different; there is no middle ground. The people’s deputies therefore are not and cannot be its representatives, they are merely its agents... Any law the People has not ratified in person is null; it is not a law. The English people thinks it is free; it is greatly mistaken, it is free only during the election of the members of Parliament; as soon as they are elected, it is enslaved, it is nothing.” (3.15.5).

This passage strongly suggests that what determines whether something is my will is not simply its content. Parliament’s will cannot be the English people’s will because they are simply not the same (identical) faculty of willing: “the will does not admit of being represented: either it is the same, or it is different; there is no middle ground.” Of course Parliament’s will is capable of voting for the same law (having the same content) as the

³⁹ p. 152. Neuhouser (1993) p. 390 and Dent (2005) similarly distance themselves from Rousseau’s insistence upon direct participation in legislation.

people's, but in doing so it remains Parliament's will (faculty of willing) and not the people's. If Rousseau's concern here is, as Cohen claims, with "preserving sovereignty" and "ensuring its stability," he should have said that the people's deputies cannot be relied upon to represent it, not that they "are not and cannot be its representatives."

Rousseau makes this same point when he says, "I say, then, that sovereignty, since it is nothing but the exercise of the general will, can never be alienated, and that the sovereign, which is nothing but a collective being, can be represented only by itself; power may well be transferred, but not will" (*SC* 2.1.1). Notice that the problem does not seem to be that no one can be relied upon to act in the sovereign's interests, but that the sovereign is a particular collective being, and whichever being purports to represent the sovereign is not that collective being and hence not one who can exercise the general will. Only after making this point does Rousseau add that no one can be relied upon to will the same things as the general will: "Indeed, while it is not impossible that a particular will agree with the general will on some point, it is in any event impossible for this agreement to be lasting and steady" (*SC* 2.1.3). Note here that the "indeed" (or "in fact" [*en effet*]) suggests that Rousseau is making an additional point beyond the one he just made: Not only is a representative's will not the people's will, it cannot even be relied upon to agree with it.

Some may wish to assert that Rousseau's allowing representatives in *Considerations on the Government of Poland* suggests that he was not absolutely committed to direct participation in legislation. But notice that what Rousseau allows (and it seems to be a compromise to adapt the ideal to Poland's circumstances) is not representation in the

sense of making decisions for another, but only in the sense of reporting another's decisions: Representatives are sent to vote for the people but bound to vote in accordance with the people's instructions. In other words, representatives do not think for the people like members of Parliament; they are simply the means by which the people who send them participate in legislation.

I have argued that a procedural account of the general will does not do justice to Rousseau's claim that the general will aims at the common good and that a common good account does not do justice to Rousseau's claim that a vote by all the citizens is necessary to declare the general will. On textual grounds, we should prefer an interpretation that combines these two seemingly opposed strands. In the next section I provide such an interpretation.

A Shared Capacity Account

When I argued that a procedural account of the general will does not account for its being a genuine will, I took "will" to refer to a kind of capacity we recognize in humans and perhaps other animals. Cohen's account avoids this objection by appealing to another sense of "will". What Cohen has described as the general will is clearly not a shared capacity, but simply a shared conception of the good that we each prioritize. It is not uncommon to use the word "will" in this sense: referring to a content, rather than to the capacity for willing that content. For example, we say, "Our will on this matter is the same." But the above passage, in which Rousseau speaks of sovereignty as "the exercise of the general will" and the sovereign as a "collective being" (*SC* 2.1.1), suggests that

Rousseau is thinking of the general will not merely as a content, but as a shared⁴⁰ or collective capacity.

We can get a sense for what it would be for people to share a capacity for willing by starting with the example of a request. Suppose I'm sick and want you to bring me soup. There's a straightforward sense in which your bringing me the soup is our will: Each of our wills (capacities for willing) was involved in the process of determining you to the same action. This example can be further specified to illustrate shallower and deeper senses in which our action may issue from a shared capacity. On the one hand, you may have very different reasons from me for assenting to my request. Perhaps you are bringing me soup in the hopes of extracting a future favor from me, rather than simply because you care about my well-being. On the other hand, you may bring me the soup for my sake. In the former case, the willing is shared in a less robust way: Our wills are united in content, but not in the reasoning that determines that content. What is united is not a full capacity of willing, but only the content that results from our independent capacities for willing. The process by which I made up my mind to bring you soup involved your request as an input, but it also involved considerations that from your perspective have no bearing on the goodness of my bringing you soup. In the latter case, by contrast, the reasoning is thoroughly shared: We arrived at a decision together based on reasons we shared. In the latter case, we are closer to sharing an act of willing: Shared content and shared process for determining the content.

⁴⁰ Again, shared in the strict sense: not a type of capacity of which we each have tokens, but a single token capacity that we have only together as a collectivity.

This reflection suggests that Cohen's insistence upon a *publicly* shared conception of the common good gets something very important right about the general will: For willing to be shared (in the deeper sense), the reasoning determining it must be shared, not just in the loose sense that we are each motivated by the same reasons to arrive at the same conclusion, but in the strong sense that we reason together, and that requires that our motivations on the matter be known to each to be the same, i.e., publicly shared. But a shared capacity for willing requires more than that we act on public reasons; we must participate together in the process of reasoning by which the reasons turn into action.

Even shared reasoning is not quite enough for the exercise of a shared will unless that shared process includes arriving at a decision together. Suppose that because of a miscommunication, one of us thinks we have arrived at a decision, but the other thinks we have not. The willing is in this case not shared. Willing together (actualizing a shared capacity of willing) involves the ability⁴¹ to recognize that we have arrived at a decision together. If I can't recognize your bringing me the soup as the carrying out of our shared will because I don't think we made that decision together, then it is not the exercise of our shared will in this robust sense.

⁴¹ I weaken this claim to mere possibility to cover cases like the following: I tell you that I'm in favor of buying a new TV, and if you decide you agree with me, you can go ahead and pick it up on the way home. I may not be aware of whether we have decided to buy the TV until after you've bought it, but once I see the TV I will recognize the purchase as based on our shared decision. Thus, from the moment I communicate my vote to you, I am able to recognize the purchasing of the TV as our action, should it occur. Rousseau seems to think that the voting in small states (which he prefers) is to be done simultaneously, but he speaks admiringly of the voting of Roman Centuries, one of which was chosen by lot to vote the first day, after which the rest voted another day: "In this way the authority of example was withdrawn from rank and given to lot in conformity with the principle of democracy. This practice led to yet another advantage; namely that the Citizens from the countryside had time between the two elections to inform themselves about the merit of the Candidate provisionally nominated, so that they might cast an informed vote" (*SC* 4.4.30-31).

I call this account of a shared will a **shared capacity account**. A shared will requires sharing not only reasons and reasoning, but a capacity to determine our shared will. This capacity involves each participant's ability to participate in shared reasoning and to recognize when that shared reasoning has determined itself to action. In Rousseau's general will, this is precisely what voting accomplishes. The citizens reason together about what to do⁴² and by voting bring their deliberation to a conclusion in such a way that they are all aware of their shared decision and can recognize it as the outcome of their deliberation.

The shared capacity account brings elements from procedural and common good accounts into unity. Procedural accounts are inadequate because they are incompatible with the guise of the good thesis. No voting procedure can be guaranteed to aim at the common good. Common good accounts solve this by making the will out to be a set of reasons based in the common good and the motivation to act on them. But common good accounts lack the self-conscious unity of a genuine capacity for willing. Without a recognized shared procedure for determining our will together, we don't share a will in this robust sense, for without such a procedure, citizens cannot recognize another citizen's actions as their own. Thus, what procedural accounts lack is a publicly shared conception of the common good, and what common good accounts lack is a shared procedure. The shared capacity account includes both.

⁴² Some may bristle at the suggestion that Rousseau's sovereign assembly deliberates, since Rousseau at times suggests that such deliberation should be avoided. For a good discussion of why Rousseau is in fact committed to popular deliberation, see Cohen (2010) p. 170-2.

Like common good accounts, my account provides an interpretation of Rousseau's insistence that the general will is never annihilated, that "it is always constant, unalterable, and pure" (*SC* 4.1.6). We have seen that common good accounts tend to interpret this claim by saying that the general will is our shared conception of the common good, which isn't lost when we subordinate it to other interests. On my interpretation, Rousseau is asserting that the general will as capacity is "constant, unalterable, and pure." We are not to understand the failures of the procedure as corruption of the capacity, but as the hijacking of the voting procedure for another purpose. In support of Rousseau's claim, he points out that each still has an interest in the public good. That shared interest is essential to our capacity for shared decision-making. To be governed in accordance with that shared interest, however, is not sufficient for being governed by ourselves.

A passage discussing another sort of group will, the will of government,⁴³ is helpful in this context. Rousseau describes the government quite clearly as sharing acts, and he claims that the disposition of the individual wills that form government is essential to the government's capacity for shared action:

"For the body of the Government to have existence, a real life that distinguishes it from the body of the State, for all of its members to be able to act in concert and to assume responsibility for the end for which it is instituted, it has to have a particular *self*, a sensibility common to its members, a force, a will of its own that tends to its preservation. Such a particular existence presupposes assemblies, councils, power to deliberate, to decide, rights, titles, privileges that belong exclusively to the Prince and make the magistrate's position more honorable in proportion as it is more demanding." (*SC* 3.1.20)

⁴³ For Rousseau, government is the executive, whereas the general will is the legislative sovereign.

While some of the elements in this list (esp. the privileges and titles) are peculiar to government, in its broadest outlines, Rousseau's account is applicable to the general will as well: The members of the state must be able to act in concert and to assume responsibility for the end for which the state is instituted. To do so, they need a self, which presupposes a set of institutions (and customs and traditions, as Rousseau suggests elsewhere regarding the general will⁴⁴) that on the one hand organize them for shared action, and on the other, motivate them to act for a common goal (in the case of government, in part by rewarding service with honor).

The capacity for shared action thus constituted may fail to act: Some member of the shared will may be tempted by private interest to subordinate the good of the shared will to her own good. But as long as there is still a shared self, including a shared procedure and shared motivation (however dormant), the capacity for shared action remains. The fact that someone still "wills the public good in his own interest" is evidence that the sensibility common to the members of the state has not vanished, that the capacity for shared willing is still present. But that does not mean that Rousseau is identifying the general will with a publicly shared conception of the common good and some motivation to pursue it. These latter are but elements in a capacity for shared action.

Rousseau sometimes speaks of the general will as existing within each individual: "Even in selling his vote for money [the citizen] does not extinguish the general will within himself, he evades it" (*SC* 4.1.6). It can be tempting to think that because the

⁴⁴ E.g. *GP* 4

general will is something within each of us, its content cannot depend upon the outcome of a vote. After all, the vote does not happen within me. But the above discussion of the government's will should disabuse us of this assumption. Rousseau does not mean that the general will is a type, tokens of which exist in each citizen. The general will (as capacity) is a whole of institutions and motivations that create a shared self that can "act in concert and to assume responsibility for the end for which it is instituted." I am a part of it as much as it is a part of me.

In this section I have introduced a shared capacity account of the general will which combines elements of procedural and common good accounts. In addition to these textual reasons I have given for preferring the shared capacity account, I have argued that the shared capacity account corresponds to something we naturally think of as a shared will. I have pressed this argument for two reasons: First, to emphasize how naturally the two strands of Rousseau's thinking on the general will fit together into something we would call a will, and thereby urge that the account is not just ad hoc. The second reason will emerge when I argue that construing the general will as a shared capacity of willing provides a compelling account of the kind of freedom Rousseau thinks we enjoy under the general will.

Must All Citizens Participate?

Before proceeding, it will help to get clearer on the motivation for one aspect of Rousseau's position: the requirement that everyone participate in the vote. Suppose an advocate of a common good account concedes that an authoritative voice must declare

our shared will so that we can all recognize it as such. Must this authoritative voice be the outcome of everyone's direct participation in the decision-making process for us to regard the declared will as ours? I have shown how shared participation yields a particularly robust sense of our sharing a will, but given the difficulty of involving everyone in legislation, is such a robust sharing necessary?

As Rousseau was well aware, Hobbes suggested a different way to unite the wills of all in one: authorization.⁴⁵ In Hobbes's social contract, subjects authorize the sovereign to act in their name and thereby make themselves the authors of all sovereign acts. If it is possible to authorize the sovereign to act in our name, then it may not be necessary for us all to deliberate together for us to recognize the law as our will. Would Cohen's common good account describe a genuine shared will if we add a requirement that there be an authorized sovereign (perhaps a representative body)?⁴⁶

As we've already seen, Rousseau rejects this sort of account on the grounds that one cannot give one's own will away: "power may well be transferred, but not will" (*SC* 2.1.2). But is Rousseau justified in holding that authorization is incompatible with freedom?

Consider an everyday case of authorization: If we all hand you money and tell you to buy dinner for us wherever you judge best, it seems reasonable to say that we bought dinner. And when the dinner turns out to be salad, it seems to follow that we bought

⁴⁵ Locke and Kant both endorse the authorization model, though with different conditions of authorization from Hobbes's.

⁴⁶ It might be thought that Cohen could derive the need for an authorized sovereign from his requirement that "Citizens have reasonable confidence that the institutions conform to their shared conception of the common good, and those social institutions do in fact generally conform to it" (58). But Cohen himself provides no such argument.

salad.

But is this latter claim correct *de dicto* or *de re*? It would be odd to say that we chose to buy salad. Authorizing you to determine what is left indeterminate in our act of authorization does not seem to have conferred upon you our will. We willed to buy something, and it was salad, but we willed to buy salad *de re* and not *de dicto*.

This *de dicto/de re* distinction matches the division Rousseau makes between legislative power (sovereign) and executive power (or government). Decisions that do not apply to all universally (like which punishment to give a particular criminal) cannot be made by the sovereign and must therefore be made by the government. The government may not make these decisions however it sees fit. It does so under the guidance of the law. Nevertheless, the fact that the government was merely carrying out the law does not mean that the executive action is at all levels of description the sovereign's will. Rousseau insists that the executive will is not the general will, even though its duty is precisely to carry out the general will. Just as your buying dinner may be attributed to us, but not your buying salad, so the government's punishing criminals may be attributed to the sovereign but not his punishing of this particular criminal.

Even when I can authorize someone to make decisions for me, those decisions cannot be attributed to my will *de dicto*. I cannot authorize someone to legislate for me and still think of the legislation as my own (except perhaps at a higher level of description: I willed that someone legislate). For this law to be my will, I must legislate it myself. For the general will to be a shared will, it cannot be an authorized will.

The General Will and the Individual Will

I have shown how the shared capacity account brings together aspects of Rousseau's discussion of the general will in a way that corresponds to a familiar and robust sense of the word "will." But to be fully satisfied with this account, we need to see why Rousseau thought we must be governed by such a shared capacity. The purpose of Rousseau's general will is to make law compatible with freedom by making subjection to law subjection to one's own will.⁴⁷ In this last section of my paper, I will argue that if we interpret the general will as a shared capacity, we get a compelling account of freedom under it.

Accounts of the general will that acknowledge a role for majority votes face a puzzle: How could our will be my will when I dissent from the majority that decides it? How is obeying the majority any better than obeying an authorized legislator? In some respects, the relationship between subject and sovereign in Rousseau is quite similar to the relationship between slave and master. One of Rousseau's arguments against slavery is that it is incoherent to claim that one person has absolute rights over another without any reciprocal obligation:

"Is it not clear that one is under no obligation toward a person from whom one has

⁴⁷ At *SC* 1.8, Rousseau contrasts three kinds of freedom: natural freedom, civil freedom, and moral freedom. Although many interpreters of Rousseau describe moral freedom as a kind of political freedom, Rousseau contrasts moral freedom not with being subject to laws legislated by others, but with being subject to our own appetites. The fact that he dismisses the topic of moral freedom by saying, "the philosophical meaning of the word *freedom* is not my subject here" (*SC* 1.8.3) further suggests that Rousseau did not take moral freedom to be a kind of political freedom. For political freedom manifestly *is* Rousseau's subject here. Nevertheless, interpreters are not wrong in thinking that essential to the freedom afforded by the general will is the fact that even when we are subject to the law, we are still subject only to our own wills. Although I do not refer to our self-governance in the general will as "moral freedom" so as not to confuse it with freedom from appetite, this self-governance is the freedom I have in mind throughout this discussion.

the right to demand everything, and does not this condition alone, without equivalent and without trade-off, nullify the act? For what right can my slave have against me, since everything he has belongs to me, and his right being mine, this right of mine against myself is an expression devoid of all meaning?” (*SC* 1.4.6)

And yet Rousseau denies reciprocity of obligation between citizen and sovereign:

“Just as nature gives each man absolute power over his members, the social pact gives the body politic absolute power over all of its members, and it is this same power, which, directed by the general will, bears, as I have said, the name of sovereignty.” (*SC* 2.4.1; see also *SC* 1.7.2)

Indeed, the only article in the social pact is “the total alienation of each associate with all of his rights to the whole community” (*SC* 1.6.6).

As Hilail Gildin (1983) writes, there seems to be an inconsistency between Rousseau’s claim that the general will cannot be delegated because no will can be delegated and his claim that in the social contract, we give the general will the right to govern us and declare our will: “Becoming a member of political society means submitting to the rule of a will that is other than my own. Why should the general will be held to be unable to do what every individual had to do in order to bring it into being?”⁴⁸

Gildin suggests that the difference between subjection to slavery and subjection to the general will lies in the fact that the general will by its nature necessarily acts in the subjects’ interests (*SC* 1.7.5). We have seen that this is guaranteed by a constitutive standard endorsed by my shared capacity account: No matter what the result of the vote is, if it departs too radically from our public conception of the common good, it is not an expression of the general will. Thus, the general will is never a tyranny of the majority.

⁴⁸ P. 51

Still, the fact that the general will is guaranteed to act in the interest of all does not make it my will. In fact, Rousseau differentiates such a relation of wills from the freedom he is after in his discussion of the family. Children may be ruled by their parents because they cannot rationally decide their own wills and because their parents naturally act in their interest from love. However, once they reach the age of reason, children become their own masters. Adult children who stay under their parents' rule "*alienate their freedom* only for the sake of their utility" (SC 1.2.3, emphasis mine). Contrast this with the social pact, under which "each, uniting with all, nevertheless obey[s] only himself and remain[s] as free as before" (SC 1.6.4). Paternalism can protect the interests of the governed, but it trades freedom for utility. The general will is not supposed to involve such a compromise.

To see how I could be free under a majority I disagree with, consider two attitudes I could have towards legislation. On the one hand, I could approach public deliberation as an opportunity to refine my views on the common good and to convince others of the correct views, and I could regard voting as expressing my individual will on the matter (aimed at the common good). My will is determined by my own judgments about the common good, and if the majority will opposes mine, then my will is overruled. If this is my attitude, and my vote counts, we can say that the general will is partly mine, just as we may say of the owner of a voting share in a company that the company and its decisions are partly hers: She has a small share of control over the company. In this case, the shared will is each individual's will in the very weak sense that each individual partly owns the shared will, and the sense becomes weaker the more shareholders there are.

On the other hand, suppose that I value our making decisions about the common good together even more than I value getting these decisions right. In this case I regard my participation in public deliberation not primarily as a means of making up my own mind about the public good and proselytizing correct views, but as the process by which we arrive at a collective decision. On this approach, I care more about our deciding about the common good together than I care that we decide upon the laws that I perceive to be best, and not (just) because I think that together we'll choose better than I could alone.⁴⁹ Rather, I care about our deciding together on the common good for its own sake. We see this kind of attitude in committed relationships like marriage. Spouses often make decisions together not (just) because the decisions will better serve the common good if they are made together or because neither can be trusted to make decisions in the others' interest, but because they value deciding together about the common good for its own sake.⁵⁰

In such cases, when my will that we decide together about the common good exceeds my will that we arrive at a certain outcome, it seems appropriate to say that my will is not determined until the shared will is determined; my mind is not made up until our mind is made up. This is not to say that I do not have my own views. But those views do not issue in a decision about what is to be done until we have decided together. In this case, we may say that the shared will is mine in a sense that exceeds its partly belonging to me. The

⁴⁹ Not, for example, for reasons associated with Condorcet's Jury Theorems.

⁵⁰ As we are about to see, they may value deciding together for its own sake because deciding together constitutes a kind of freedom.

shared will is my will in that I belong to it: My will is part of the whole, not just as a vote included in the whole, but because my will is completed and determined by the whole. My will is indeterminate with respect to legislative questions when taken independently of the whole, and yet it is still active with the whole in answering them. Rousseau describes such a will as having relative existence:

“Civil man is only a fractional unity dependent on the denominator; his value is determined by his relation to the whole, which is the social body. Good social institutions are those that best know how to denature man, to take his absolute existence from him in order to give him a relative one and transport the *I* into the common unity, with the result that each individual believes himself no longer one but a part of the unity and no longer feels except within the whole.” (*E* 39-40)

On my interpretation of this passage, to denature a will involves more than aiming the will at the common good rather than at private interest; it involves disposing the will to identify as a contributor to the collective judgment rather than as a determiner of individual judgment. The natural will is determined by its own perspective; the citizen's will is determined by the perspective of the whole. The citizen recognizes the collective decision as his own because it is the completion of the activity he is participating in.

Natural man can regard the general will as his will only in the shareholder sense because his particular will is complete by itself: Not valuing the shared decision-making process as much as the outcome, he cannot but regard the resistance of the majority opinion to his own will as a foreign imposition. But civil man's will with respect to legislative questions is not fully determinate independently of his fellow citizens', because civil man sees his own will as made up only together with everyone else's. He may differ

from the majority in his opinion on the common good, but he holds his own opinion to be of secondary importance: He has a partial, relative will with respect to the body politic. The general will is therefore in a robust sense each civil man's will. Hence, in the state only civil man enjoys the freedom of being subject only to his own will.

To be clear, civil man does not simply will that the citizens make decisions together. Civil man wills that the citizens make decisions about the common good together, and that requires that the citizens aim at a publicly shared conception of the common good. For if I will that we make decisions on the common good together and I cannot but see you as aiming at something other than the common good, then my deciding together with you will not satisfy my will, and hence I will not be free. Since making decisions together involves sharing the reasoning for the decisions, there are limits to the kind of compromise that such a shared will could require me to make. A couple might prefer to make decisions about their child's good together over getting all those decisions right without neglecting their child's good. For if the decision is to be shared, they must share in the conception of the child's good at which it aims, and they must each be satisfied enough with the reasoning that they can sign off on the decision, even if they don't see it as optimal.

Indeed, insofar as the couple's harmonious joint decision-making is itself good for the child's welfare, we may say that in an important sense, there is no conflict of priorities at all. There is merely a recognition that it is more important to the child's good that the decision be made together than that the decision be the best possible decision. The same

may be said of citizens deciding together. Suppose that the only means of passing what seems the best law is to exclude some citizens who oppose it from the vote. Presumably, disenfranchising citizens could itself cause great harm to the common good. Hence, the common good may best be served by settling for less than ideal laws.

It might be objected that on this account, it seems to follow that I would enjoy the relevant freedom if I ceased to care about our aiming at the common good together, if I simply willed that we decide together (regardless of the aim), or, for that matter, if I simply willed that you decide for the both of us.

Rousseau thinks we cannot will such things with full understanding because, according to his guise of the good thesis, we always will our own preservation and hence our own freedom, one of the chief means to our preservation. To will to let legislative decisions be made without regard for whether the decision aims at the common good would be a betrayal of this first law of our nature.⁵¹ Moreover, on such proposals, there would be no sense in which I govern myself either because I would not participate at all, or because what I participated in would not be a will but a mere procedure: It would not be the self-determination of a capacity for practical reasoning in accordance with a conception of the good. I could not say that I am ruled by my own will because what ruled me would not be a will at all.

In other words, it is crucial that the general will be a genuine shared will. If the

⁵¹ “To say a man gives himself gratuitously is to say something absurd and inconceivable; such an act is illegitimate and null, for the simple reason that whoever does so is not in his right mind. To say the same of a whole people is to assume a people of madmen: madness does not make right” (SC 1.4.4).

general will is not a genuine shared will, then it cannot be my will in the appropriate sense. For if it is a genuine will but not shared, then it is an alien will to those who are not part of it, and hence they do not enjoy freedom in being subjected to its laws. And if it is shared but not a genuine will (i.e. a mere procedure), then it lacks the publicly shared conception of the common good necessary to make it continuous with my will. Such a procedure could at most be the procedure I will to abide by, not my will; for it is not a will at all. Our will is my will in a genuine shared will because the activity that I participate in is the activity of willing, of practical reasoning. The outcome is not just a result, but a will, and it is my will because I am a part of it, and it holds itself to a conception of the good that I recognize as mine.

Of course, in a genuine shared will, our will is not my will in the sense that we usually think of when we think about our own wills. But the kind of freedom I am suggesting is the kind of freedom that people enjoy in a successful marriage where decisions are made together. Opinions about what to do may significantly differ between spouses, but insofar as each values deciding together more than having her own view win the day, they do not experience the compromise as a compromise of their freedom. The reason for this, I am suggesting, is that they identify as participants in a shared will. They experience their shared will's freedom as their own, even when their own opinions about the common good do not prevail on a particular occasion.

The shared capacity account I am advocating contrasts with a suggestion from Cohen about where to draw a line between Rousseau's "philosophical liberalism" and

“sociological communitarianism”. For Cohen, Rousseau is a philosophical liberal because his arguments for entering the social contract are addressed to “free agents considered apart from... national passions and attachments.”⁵² His communitarian streak, which requires that the actual citizens be patriotically attached to their state, is (according to Cohen) sociological rather than philosophical in that he thinks that the national passions and attachments are necessary not for the state to be legitimate, but for the state to be stable given human nature:

“the need for attention to public business and the importance of common happiness as an ingredient in individual happiness are offered... as devices to prevent the dissolution of the society of the general will: as conditions of the stability of a legitimate order, not of the philosophical conception of political legitimacy.”⁵³

I might agree with Cohen that my personal failure to attend to public business or lend importance to the common happiness as an ingredient in my individual happiness would not by itself undermine political legitimacy for Rousseau. But on my shared capacity account, my attention to and participation in public business is essential to the unique kind of freedom Rousseau aimed at, and the availability of this kind freedom is the mark of political legitimacy.

Moreover, the love of the fatherland that Rousseau insisted was an indispensable virtue of the citizen can be seen as rationally and not merely sociologically necessary for this freedom. For an attitude that could motivate me not only to value the common good above my particular interest, but in doing so to value joint decision-making with others

⁵² P. 95

⁵³ p. 57

above getting all of those decisions right is a kind of love. Respect for others could motivate me to seek a kind of legislative process that yields just results; it's harder to see how it could motivate me to insist on being a participant in such a process. Hence, it's harder to see how could not motivate me to value joint decision-making even at the expense of getting all the decisions right. It is no accident that a marital relationship is a natural example of shared willing: Willingness to will together seems to require (or perhaps constitute) a kind of love, whether of another individual or of country.

This interpretation explains the remarkable fact that in *Emile*, written concurrently with *The Social Contract* and published one month later, Emile, having been taught the entire contents of *The Social Contract* and having traveled for two years in search of a country where he can live freely concludes that political life isn't worth the trouble:

“The more I examine the work of men in their institutions, the more I see that they make themselves slaves by dint of wanting to be independent and that they use up their freedom in vain efforts to ensure it. In order not to yield to the torrent of things, they involve themselves in countless attachments. Then as soon as they want to take a step, they cannot and are surprised at depending on everything.” (E 471-2)

How could Emile say this in spite of his thorough knowledge of the contents of *The Social Contract* and his awareness (from his travels) of Geneva's institutions, which Rousseau took to exemplify much of what he advocated.⁵⁴ Emile is rational, just, and benevolent. He wants nothing that is incompatible with the like for all. He seems like precisely the citizen required for an account like Cohen's. But we must take seriously Rousseau's distinction, at the beginning of *Emile*, between a natural man brought up to live in society and a civil

⁵⁴ See the Preface to the 2nd *Discourse*.

man. Emile has not been denatured; he does not desire to spend his time making decisions about the public good with his fellow citizens. Without that desire, the life of a citizen is not autonomous.

Conclusion

I have argued that although there is only a very limited sense in which the general will belongs to me, corresponding to a very limited freedom, there is a more robust sense in which I belong to it, corresponding to a more robust freedom. This more robust sense requires that the general will be a genuine shared will, for only if it is shared can it be mine, and only if it is a will can it be my will. Thus, my defense of the shared capacity account of the general will consists of two main arguments. First, it accounts for both Rousseau's insistence on a democratic procedure and his insistence that the general will aims at the common good. Second, it makes possible a compelling account of a robust freedom under which I am ruled not just by a procedure I consent to or in accordance with principles I hold, but by something I can regard as my own will. This account of freedom under the general will also explains very naturally why Rousseau insists on a high level of patriotism and why a rational person like Emile might reasonably choose not to pursue the life of a citizen. We may doubt whether such a general will is achievable in today's states, but it is a unique and attractive vision of freedom in society with others.

Chapter 2

Kant's General Will

My broadest aim in this chapter is to show how Kant solved two difficulties in Rousseau's general will: the apparent unfeasibility of requiring everyone to participate in legislation and the challenge of getting everyone to share a conception of the common good. Along the way, we'll see how Rousseau solves two problems in Kant's political philosophy, or to speak more precisely, how reading Kant as a Rousseauvian solves two interpretive puzzles. First, what is the source of Kant's Universal Principle of Right (UPR), the fundamental principle of Kant's political philosophy, and what is its relationship to the Categorical Imperative (CI), the fundamental principle of Kant's moral philosophy as a whole? The answer will be that the UPR can be derived from arguments about freedom in the *Groundwork* and 2nd *Critique* by specifying the freedom as external. Hence, the UPR is a specification of the CI as it applies to external freedom. Second, is some kind of equality (of rights, outcome, opportunity, etc.) the core value of Kant's political philosophy? The answer will be no; freedom is rather the source of the demand for equality in Kant. It's worth elaborating on these two questions so that we can see how Kant's Rousseauvian inheritance can answer them.

The Source of the UPR

The UPR states that “any action is *right* if it can coexist with everyone's freedom in accordance with a universal law, or if on its maxim the freedom of choice of each can coexist with everyone's freedom in accordance with a universal law” (*MM* 6:231). The thought is that any action is right as long as it does not trespass the boundaries of another's freedom, and the boundaries of our freedom are set in accordance with universal law. I can do whatever I want unless my so doing might conflict with your doing whatever you want. Wherever there is a possibility of such conflict, a universal law must determine which of us must cede to the other. For example, my freedom to wave my fist around must cede to your freedom from having your nose smashed in.

One of the most vexed questions among interpreters of Kant is about the relationship between the UPR and the CI, “act only in accordance with that maxim through which you can at the same time will that it become a universal law” (*G* 4:421). Some interpreters insist that the UPR cannot be derived from the CI; others that it can or must be.⁵⁵ Behind all this debate is Kant's failure to provide a clear derivation of the UPR. In the absence of clear derivation, the most obvious source would seem to be the CI. After all, in a section in the Introduction to *The Metaphysics of Morals* that discusses “concepts common to both parts of *The Metaphysics of Morals*” (those pertaining to right

⁵⁵ For example, Pogge 2002, Wood 2002 and Willaschek 1997, 2009 and 2012 argue that the UPR cannot be derived from the CI. Pogge 2002, Ripstein 2009, and Nance 2012 argue that it can be derived from the CI, perhaps with some additional premises, though the justification of the UPR does not depend on the CI. Gerhard Seel 2009 argues that the UPR can only be derived from the CI and that the justification of the UPR depends upon the CI.

and to virtue) Kant asserts that the CI “as such only affirms what obligation is” (*MM* 6:225). Moreover, the UPR bears a clear resemblance to the CI.

Still, differences between laws of right and ethics have made a direct derivation seem difficult. Kant says that right is analytically accompanied by an “authorization to use coercion” (*MM* 6:231). But it has seemed to some interpreters impossible to derive an authorization to use coercion from the CI.⁵⁶ Moreover, Kant says that the UPR itself is analytic, while the CI is synthetic. Why should a synthetic principle be needed to derive an analytic one?

In this chapter I will argue that we can find the source of the UPR in the arguments for the CI in the *Groundwork* and 2nd *Critique*. On my view, the argument for the CI is not particular to ethics or right, to internal or external freedom. It is a general argument that applies equally to both. When applied to external freedom, the argument yields the UPR. The result is a view that bears some resemblance to a suggestion from Arthur Ripstein: The UPR is an application of the CI to beings who inhabit space.⁵⁷ A helpful clue to interpreting the argument for the CI as applying direction to external freedom and right is that the argument for the CI is very similar to Rousseau’s own argument for a condition of universal law in the state. This Rousseauvian inheritance sheds light on some obscure aspects of his argument.

⁵⁶ Willaschek 2009

⁵⁷ However, it’s not clear to me that space plays quite as important a role in Kant’s view as Ripstein gives it. More fundamental than our occupying space, it seems to me, is the fact of our being able to affect each other’s actions in some way. If we could affect each other’s actions in some way independent of our being in space, the UPR would still be the necessary principle of our freedom in relation to one another.

Equality and Liberty

As we've already seen, the UPR claims that what's right is whatever is compatible with everyone's enjoying freedom under universal laws. The phrase "in accordance with universal law" signals that any rightful freedom for one person must be compatible with like rightful freedom for everyone else. My freedom of speech is compatible with everyone else's freedom of speech and with the rest of our freedoms, and thus in accordance with universal law. My freedom to kill others, however, is not in accordance with universal law because my killing others deprives them of their own freedom to kill (as well as the rest of their freedoms). There could be no universal law that would protect everyone's freedom to kill, for one person's freedom to kill interferes with another's.

But where does the requirement that the law be universal come from? Or if you think that it is the nature of law to be universal, why must freedom be ordered into laws, rather than non-universal rules? The requirement that law be universal no doubt seems fair, but the aim of government is not fairness for Kant, but freedom itself. If the aim is simply to preserve freedom, where does the requirement that it be under universal law come from?⁵⁸

If the appeal to universal law is not grounded in freedom itself, then Kant's view is no exception to the thesis suggested by Ronald Dworkin and defended at length by Will Kymlicka that "every plausible political theory has the same ultimate value, which is

⁵⁸ One obvious possible source is the Categorical Imperative (CI). One of my goals in this chapter will be to show that the UPR is in fact a particular application of the CI.

equality.”⁵⁹ Theories differ on the relative importance of liberty and welfare and how to construe them, but they agree that whatever combination of values they endorse is to be protected for all equally, and so their ultimate value is equality, or so the thesis claims. The question for Kant’s view is whether he can insist on a condition of universal law, the equality of all under the law, while maintaining that freedom, and not equality, is the fundamental principle in the *Doctrine of Right*.

Again, Kant’s Rousseauvian inheritance points the way to an answer. Rousseau insisted that the purpose of equality was to make freedom possible, and not simply that freedom, as a good, must be distributed equally (thereby making equality the higher principle). Seeing how Kant adopts this argument points the way to a reading of the requirement of universality in the UPR as deriving from the necessary conditions of freedom and not as an independent value.

Kant’s Contributions to the Notion of the General Will

I will address the interpretive and philosophical questions I’ve just discussed by appealing to what Kant inherited from Rousseau. Kant was not, however, a mere inheritor of Rousseau’s ideas on the general will and freedom; he was also an innovator. Kant’s central innovation with respect to the general will is to elaborate an alternative to direct participation in legislation.

Like Rousseau, Kant held that for laws to be legitimate, they must be the will of all the citizens, the general will (*allgemeiner Wille*). He argues for this claim in the section of

⁵⁹ Kymlicka (2002), p. 3. This thesis is one of the central theses of Kymlicka’s book. See also Dworkin (1977), p. 179-83 and (1983), p. 24, 31-35.

the *Doctrine of Right* on public right:

“The legislative authority can belong only to the united will of the people. For since all right is to proceed from it, it *cannot* do anyone wrong by its law. Now when someone makes arrangements about *another*, it is always possible for him to do the other wrong; but he can never do wrong in what he decides upon with regard to himself (for *volenti non fit iniuria*). Therefore only the concurring and united will of all, insofar as each decides the same thing for all and all for each, and so only the general united will of the people, can be legislative.” (*MM* 6:314)

If a legislator is other than the citizens, then the legislator makes arrangements for another. Whenever someone makes arrangements for another, it is possible for the legislator to wrong the citizens by legislating something to which they do not consent. But this possibility conflicts with the idea of a legitimate legislator, one whose laws are the very source of all right and hence can never wrong the citizens. Hence, the legislator who makes arrangements for the people must be the people themselves.

It would be easy to read this as an argument for a particular kind of legislature, much like the argument in Rousseau, one that involves direct democracy and unanimity in decisions. And Kant does indeed go on to suggest that (in the ideal case) the citizens of the state vote on the laws. But Kant does not require that all citizens vote,⁶⁰ and he acknowledges the legitimacy of a variety of constitutions in which the people do not literally legislate, including monarchy:

“A people cannot offer any resistance to the legislative head of a state [*das*

⁶⁰ Any ‘passive citizen,’ “whose preservation in existence (his being fed and protected) depends not on his management of his own business but on arrangements made by another (except the state)” is excluded from voting, though his rights as a human being are to be protected (*MM* 6:314). It’s doubtful whether Kant’s status of “passive citizen” is justified by his other views, but the fact that he thinks that there can be passive citizens is evidence that he did not intend to argue that everyone must participate in the legislature.

gesetzgebende Oberhaupt des Staates] which would be consistent with right, since a rightful condition is possible only by submission to its general legislative will. There is, therefore, no right to *sedition* (*seditio*), still less to *rebellion* (*rebellion*), and least of all is there a right against the general legislative will [*ihn*] as an individual person (the monarch), *to attack his person* or even his life (*monarcho-machismus sub specie tyrannicidii*).” (*MM* 6:320, translation modified⁶¹)

Note that Kant does not simply say that we must obey the legislative head of state; he says that a monarch may be the voice of the general legislative will. Kant clarifies later that on his view a republican⁶² form of government is the only form that is “literally a state” (*MM* 6:341), but the concept of a general legislative will or general will is applicable even outside such ideal conditions.

Since Kant rejects Rousseau’s insistence that only a direct democracy is legitimate, he needs another account of how it is possible for the citizens to legislate the law by which they are bound. In the first part of the chapter, I argue that the possibility of the Kantian general will is provided by the distinction between the world of sense (the phenomenal world) and the world of understanding (the noumenal world).⁶³ Whether we participate in the legislative process empirically or not, we can participate noumenally. Indeed, as Kant argues, because free participation cannot be given empirically, a rightful constitution

⁶¹ Mary Gregor has translated this pronoun (*ihn*) as “the head of state,” interpreting it as referring to “the legislative head of state” [*das gesetzgebende Oberhaupt des Staates*]. This interpretation would not militate against the point I am making here, but I think it is mistaken, for *Oberhaupt* is a neuter noun, and *ihn* is a masculine pronoun. The only plausible masculine referent of this pronoun is the “general legislative will” [*allgemein-gesetzgebenden Willen*].

⁶² As we’ll soon see, in Kant’s use of the term, a republican government is one in which the executive and legislative powers are separate.

⁶³ For the purposes of this interpretation I think it is unimportant whether one takes a two-aspect or two-world view.

is only possible by supposing noumenal legislation. To see why, we must work through Kant's argument that freedom is the *ratio essendi* of the moral law (specifically the juridical law, the law of external freedom), and likewise of the general will as legislator of that law.

Even after showing that noumenal legislation is possible, we might ask whether it is actual, whether the citizens in fact do legislate the law by which they are bound. The answer is to be found in Kant's argument that the moral law is the *ratio cognoscendi* of our freedom. In the second part of this chapter, I will argue that the actuality of the Kantian general will is given to practical cognition in the same way that the actuality of our inner freedom is given in ethics, viz., as an idea of reason of which we are aware by a fact or act of reason (*Factum der Vernunft*).

My chief argument for these two points (regarding the possibility and actuality of the general will) will find the basis of right in the arguments for the Categorical Imperative in the *Groundwork* and the 2nd *Critique*. Those arguments, I will try to show, apply equally to virtue and to right, to internal freedom and to external freedom. Hence, the phenomenal/noumenal distinction is just as essential to external freedom as to internal, and it is essential precisely to make possible a general will by which we legislate laws of freedom for ourselves, by which we are autonomous.

The Need for a General Will

“Regardless of how a government is constituted, if there is a single person in it who is not subject to

the law, all the others are
necessarily at his discretion”

(Rousseau, *D2* 115)

The complete case for my claim that the possibility of Kant’s general will depends upon the distinction between the world of sense and the world of the understanding begins at the very foundations of the *Doctrine of Right*. This may seem odd because the general will does not seem to come up anywhere in the Introduction to the *Doctrine of Right*, in which Kant gives his basic characterization of right. Kant begins the *Doctrine of Right* with the Universal Principle of Right (UPR), which does not obviously invoke a general will: “Any action is *right* if it can coexist with everyone’s freedom in accordance with a universal law, or if on its maxim the freedom of choice of each can coexist with everyone’s freedom in accordance with a universal law” (*MM* 6:231). Kant’s general will does not seem to show up until a puzzle about property rights suggests the need for an omnilateral will.

Although Kant’s version of the general will [*allgemeine Wille*] is obviously a Rousseauvian inheritance, the order of and relationships among Kant’s arguments in the *Doctrine of Right* may seem to suggest that the genetic homology does not translate into a functional analogy. The need for the general will to legislate is on Kant’s view a consequence of a principle of right, rather than, as in Rousseau, the direct answer to the question, how can we live freely in society with one another? Kant’s answer to this question would seem to be, We can live freely in society with one another if we are subject

to universal laws of freedom, laws that are justified by their being necessary to our freedom. The question of who makes the laws arises only later. Where Rousseau begins with an account of the general will and derives principles of right from it, Kant begins with a principle of right and derives an account of the general will from it, or so it may seem.

This impression may be reinforced by the fact that Kant does not seem to appeal to sovereign representation to justify rightful coercion. Kant derives a right⁶⁴ to coerce others not to wrong us directly from the UPR. Whatever is not compatible with like freedom for all is, by the UPR, wrong, and hence not something one has a right to do. When one has no right to do something, it is no violation of one's right to be prevented from doing it. Hence, it is no violation of one's right to be prevented from violating others' rights. And therefore, my preventing another from violating my rights is itself right: "If a certain use of freedom is itself a hindrance to freedom in accordance with universal laws (i.e., wrong), coercion that is opposed to this (as a *hindering of a hindrance to freedom*) is consistent with freedom in accordance with universal laws, that is, it is right" (MM 6:231). Nothing in this argument seems to require an appeal to a general will. For Rousseau, on the other hand, the right to coerce me to obey the law derives in part from my having legislated the law.⁶⁵

⁶⁴ Or at least a provisional right

⁶⁵ See the previous chapter for Rousseau's argument. Rousseau also famously appeals to a tacit clause of the social contract "that whoever refuses to obey the general will be constrained to do so by the whole body: which means nothing other than that he will be forced to be free; for this is the condition that, by giving each Citizen to the Fatherland, guarantees him against all personal dependence; the artifice that... alone renders legitimate civil engagements" (SC 1.7.8). The tacit clause captures Kant's thought that freedom requires a right to coerce, but limits such coercion to rights legislated by the general will to people who have

I think this apparent dissimilarity of deployment conceals what is in fact a basic similarity of approach. The general will comes up late in the *Doctrine of Right* because only at a later stage of the argument is it clear that there must be publicly given law, and Kant only invokes the notion of a general will when discussing the public authority that gives the law. Nevertheless, I will argue that the thought that our wills must be united in giving laws of freedom is a background assumption motivating the requirement of universal law in UPR itself. Seeing that this is so and the relationship between this claim and the arguments in the *Groundwork* will help us to see why the general will presupposes the distinction between the world of sense and the world of understanding.

First, we need to map Rousseau's argument. Rousseau resolves the tension between freedom and the coercive force necessary for an orderly society with his notion of the general will. Rousseau's problem was to characterize a community in which, in spite of mutual dependence, each citizen is free, that is, obeys only her own will. His solution is that the citizens are free if they are themselves the authors of all the laws that govern their interactions with each other. I must be the author of the laws because if I am not the author of the laws that govern me, then someone else controls my will, and I am not free.

At the same time, you must be the author of the laws if I am to be free. For I am not free unless you are bound to abide by the laws that govern our interactions.⁶⁶ If, as

Rousseau and Kant agree, one can only be bound to do something if one binds oneself to

assented to the social contract. I will argue in this chapter that Kant likewise limits such coercion by the notion of a kind of proto-general will in which we participate a priori.

⁶⁶ See the previous footnote

do it, then you can only be bound by the laws that govern our interactions insofar as you bind yourself not to do so, or in other words, insofar as you give yourself the law that governs our interactions. Hence, I can only be free if you are the author of the law that governs our interactions. And since, as we have already seen, I can only be free if I am (also) the author of the law, I can only be free insofar as *we* are the authors of the law.

Hence, the freedom of any one individual depends upon the freedom of all the other individuals because the freedom of each depends upon others' being obligated not to infringe it, and others are only obligated not to infringe one's freedom insofar as they freely legislate the laws that obligate them. If you are not free under the law, then you are not bound not to interfere with me. If you are not bound not to interfere with me, then I am not free either, for my freedom is constituted by your obligation not to infringe it.⁶⁷ Only if everyone is free can everyone be bound by the law and thereby free from each others' interference.

For us to be authors of the laws that govern our interactions, we must agree on a set of laws. How can everyone agree on all the laws? A key part of the answer in Rousseau is the requirement that the law apply to all universally:

“The commitments which bind us to the social body are obligatory only because

⁶⁷ Contrast this view of freedom with the following alternatives: 1) I am free insofar as you are unable to interfere with my freedom, and 2) I am free insofar as you are unwilling to interfere with my freedom. The first alternative expresses a sufficient but not a necessary condition of freedom; if it were necessary, our external freedom would be limited to almost nothing. The second alternative does not even express a sufficient condition of freedom. If my choice depends upon the disposition of your will, then it is not free. This is as true of the slave whose choice happens not to be constrained by a benevolent master as it is of the tyrant whose choice happens not to be constrained by the intimidated public: “One believes himself the others' master, and yet is more a slave than they” (*SC* 1.1.1). Freedom must not be confused with having the upper hand: “rich, he needs their services; poor, he needs their help, and moderate means do not enable him to do without them” (*D2* 2.27).

they are mutual, and their nature is such that in fulfilling them one cannot work for others without also working for oneself. Why is the general will always upright, and why do all consistently will each one's happiness, if not because there is no one who does not appropriate the word *each* to himself, and think of himself as he votes for all? Which proves that the equality of right and the notion of justice which it produces follows from each one's preference for himself and hence from the nature of man; that the general will, to be truly such, must be so in its object as well as in its essence, that it must issue from all in order to apply to all, and that it loses its natural rectitude when it tends toward some individual and determinate object; for then, judging what is foreign to us, we have no true principle of equity to guide us." (*SC* 2.4.5)⁶⁸

A key thought in this passage is that once we assume that the law must apply equally to all, each person's self-interest starts to point toward the same laws.⁶⁹ When our self-interest points toward the same laws, we can all agree on them. In other words, the requirement that the law apply universally makes it possible for the law to issue from all. Being governed only by the law that issues from my own will is a kind of freedom for Rousseau. Thus, the universality of law makes the freedom of self-legislation possible.

Rousseau does not explicitly argue that insisting upon laws with universal form is the *only* way to get the people to agree on law. It's certainly plausible that small groups of people would agree to rules of interaction that do not have universal form. But laws for a

⁶⁸ "This is a marvelous paragraph," observes John Rawls (2007, p. 232) in his lectures on Rousseau. Indeed, I shall be adding evidence for the marvelousness of this paragraph by drawing from it a lesson beside the one that Rawls drew in his lectures.

⁶⁹ There are obvious ways to challenge this thesis. I may prefer that everyone be forced to mow their lawns weekly, and others may not. One of Kant's innovations to answer this challenge in the context of the CI is to urge that the motivation for the proposed law be considered as part of the law. Could I will that everyone be forced to do something just because it suits someone's (or even a majority's) aesthetic preferences? Whether this innovation can solve all the problems has been debated extensively in the literature on Kant's Categorical Imperative. In the context of Right, Kant has another solution in mind: It must be possible for the law to be chosen by all the citizens, which requires that it not conflict with their necessary concern to realize freedom. We'll see more on this below.

state need to be the subject of agreement not only of the citizens present at the moment of legislation, but also of all future citizens. Accidental agreement cannot be guaranteed to accommodate future citizens.⁷⁰ Hence, some principle is needed to ensure agreement.

The key insight from Rousseau is that we don't need to appeal to some further standard beyond freedom to derive the universality condition. Given that we live in community with one another, our choices may conflict. If you may interfere with my choices, I am not able to determine my own activity and hence not free. Thus, if I am to be free, there must be limits on your freedom, and vice versa. At this point, it might be tempting to say that the limits should take universal form because that's the only fair way to draw the limits. But that would betray Rousseau's project. To say that I must accept limits on my freedom in the interest of fairness or even in the interest of another's freedom is to deny that I am subject only to obligations I give myself, only to my own will. Fortunately, as Rousseau points out, we do not need an independent value (such as fairness) to determine how to draw the limits. My freedom can be preserved in its limitation if I am the author of the limits on my freedom. We can all be the authors of the limits on our freedom if we can all agree on limits. And we can all agree on the limits if they are drawn in accordance with universal law.

We can sum up Rousseau's argument as follows:

Rousseau's Argument for the Universality Condition

⁷⁰ Indeed, an accidental agreement cannot be guaranteed to hold over time even if we add no new citizens. I may agree today and change my mind later. If I am to remain free within a political order, I must continue to agree to it as long as I am part of it. No contingent agreement can guarantee this.

1. If I am to be free, there must be laws of freedom.
2. If the laws are to be compatible with my freedom, I must legislate them.
3. If I am to legislate the laws, then we must agree on them.
4. If we are to agree on the laws, they must apply universally.

Now if we allow that Kant is making the same move as Rousseau, the point of Kant's condition of universal law is to make it possible that everyone legislate the law by eliminating from the candidates for limits on our freedom any options upon which we could not agree. As Kant writes in the *Groundwork*, "this lawgiving must, however, be found in every rational being himself and be able to arise from his will" (*G* 4:434).⁷¹ We should note, however, that Kant's condition is more restrictive than Rousseau's. For Rousseau, the options to be eliminated were those that benefit one person over another. Kant argues that we must eliminate any options that would aim at welfare at all, because it cannot be guaranteed that everyone will in fact agree to any limitation of their freedom in exchange for the intended welfare. The point here is not just that we cannot be guaranteed to desire or need the same things; even if we did, the fact that I desire or need something does not mean that I will it. What we need to ensure that everyone legislates the law is not that the law benefit everyone, but that everyone *will* the law.

Kant puts this point in terms of a distinction between wish and choice: "The concept of right... does not signify the relation of one's choice to the mere wish (hence also to the mere need) of the other, as in actions of beneficence or callousness, but only a

⁷¹ This quotation refers to a possible kingdom of ends. I'll soon argue that it may be applied to earthly kingdoms even though they cannot be cognized as kingdoms of ends.

relation to the other's *choice*" (MM 6:230). The difference between wish and choice, Kant writes, is that only the latter "is joined with one's consciousness of the ability to bring about its object by one's action" (MM 6:213).⁷² Hence, in choice, one determines oneself "to do or refrain from doing as one pleases" (MM 6:213). Laws of right reconcile our choices and not (necessarily) our wishes.⁷³ Hence, while Rousseau frames his argument in terms of making sure the happiness and self-interest of all coincide, for Kant, considerations of happiness and self-interest are irrelevant because there can be no guarantee that everyone assents to laws that make the happiness and self-interest of all coincide.⁷⁴

Since Kant refers to our ends (including happiness) as the matter of maxims or choices, Kant says that in right,

"no account at all is taken of the *matter* of choice, that is, of the end each has in mind with the object he wants... All that is in question is the *form* in the relation of choice on the part of both, insofar as choice is regarded merely as *free*, and whether the action of one can be united with the freedom of the other in accordance with a universal law." (MM 6:230)

Only under the condition that the law abstracts from all matter of the will can we be

⁷² In *Anthropology from a Pragmatic Point of View*, Kant defines wish as follows: "*Desire (appetition)* is the self-determination of a subject's power through the representation of something in the future as an effect of this representation. Habitual sensible desire is called *inclination*. Desiring without exercising power to produce the object is *wish*." (A 7:251)

⁷³ Kant explains in the following paragraph how our choices (and *a fortiori* our wishes) can be incompatible: "That choice which can be determined by *pure reason* is called free choice. That which can be determined only be *inclination* (sensible impulse, *stimulus*) would be animal choice (*arbitrium brutum*). Human choice, however, is a choice that can indeed be *affected* but not *determined* by impulses, and is therefore of itself (apart from an acquired proficiency of reason) not pure but can still be determined to actions by pure will." (MM 6:213)

⁷⁴ Not to mention that it is not clear how any laws could be guaranteed to make the happiness and self-interest of all coincide.

guaranteed to agree to the law, since there is no matter of the will that I can be guaranteed to agree to pursue. The only thing that the free will can be guaranteed to pursue, according to Rousseau's argument, is agreement itself.⁷⁵

I have been suggesting that Kant is making an argument very similar to Rousseau's. There are some circumstantial reasons for finding my suggestion plausible: We don't have an argument for the UPR in Kant. He simply announces it at the beginning of the *Doctrine of Right*. So there is no explicit textual competitor to the Rousseauvian interpretation. And since Kant was an avid reader of Rousseau and was by his own confession greatly influenced by his moral and political philosophy, Rousseau is not a bad person to turn to for more insight into things Kant left unsaid.

But my main evidence that they are making the same move is that a parallel argument may be found in the *Groundwork* and *2nd Critique*. Kant does not provide the argument for the UPR in the *Doctrine of Right*, just as he does not provide the argument for the Principle of Virtue in the *Doctrine of Virtue*.⁷⁶ Kant gives his arguments for the foundations of these *Doctrines* in the *Groundwork* and the *2nd Critique*. The fact that some of Kant's arguments in the *Groundwork* and the *2nd Critique* may be applied to external lawgiving as well as internal lawgiving is easy to overlook because Kant's examples tend to

⁷⁵ Recall that the freedom of any one individual depends upon the freedom of all the other individuals because, as Rousseau and Kant both say, our obligations are only binding because they are mutual. Anyone who is not free under the law is not bound by it. Thus, if you are not free under the law, then you are not bound not to interfere with me. If you are not bound not to interfere with me, then I am not free either. Only if everyone is free can everyone be bound by the law and thereby free from each others' interference.

⁷⁶ *MM* 6:395. It's worth noting, in connection with my contention that the UPR's requirement of universality must be based in freedom alone (not, e.g. in considerations of fairness), that Kant describes the UPR as analytic because "I need not go beyond the concept of freedom" to see that it is true (*MM* 6:396).

focus on internal lawgiving, lawgiving that concerns the motivation for an action. But the *Groundwork* is the groundwork of *the Metaphysics of Morals* as a whole, and not just of the *Doctrine of Virtue*.

In both the *Groundwork* and the *2nd Critique*, Kant derives the need for universal law from the idea of a free will. In *Groundwork* III, Kant begins this argument with the claim that a will is a kind of causality and must therefore act under law. A free will is a kind of will whose law is given by itself, an autonomous will. In other words, if I am to be free, I must give myself laws.

At this point Kant refers his reader back to *Groundwork* II, where Kant had argued that the principle of autonomy was the CI, but since this argument is somewhat clearer in the *2nd Critique*, I will use the latter rendition. The argument begins by distinguishing material from formal principles. Material principles “presuppose an *object* (matter) of the faculty of desire as the determining ground of the will” (*KpV* 5:21). In other words, they are grounded in ends that the subject bound to them is supposed to have. Kant argues that such principles “are, without exception, empirical and can furnish no practical laws” because the object of our faculty of desire can be known only a posteriori, whereas practical laws have “objective necessity, which must be cognized a priori” (*KpV* 5:21-22).⁷⁷

⁷⁷ Note, however, that Kant also seems to think that the Formula of Humanity can provide an a priori matter of the law, rational nature as an end in itself. I take it that for Kant the crux of the issue is whether not whether any matter of the will can be assumed a priori, but whether matter or form is the determining ground of the will. Rational nature is a necessary object of the will only because the formal law requires it. Hence, when matter is the ultimate determining ground of the will, it is not the necessity of having such a matter of the will (as derived from the formal law) but a merely contingent fact (my brute desire) that determines the will. Hence, Kant defines the Formula of Humanity, though it has matter, is not a material principle, for it does not presuppose an object as the determining ground of the will (cf. *KpV* 5:21).

From this Kant concludes, “If a rational being is to think of his maxims as practical universal laws, he can think of them only as principles that contain the determining ground of the will not by their matter but only by their form” (*KpV* 5:27). The form, it turns out, is the form of universal law.

In broad terms, the argument then is, if something is to be a practical law, it must be binding a priori. No matter of the will can be known a priori, and therefore principles grounded in any matter of the will do not qualify as laws. But aside from its matter, a law only has the form of universal law. Therefore, if there is to be a practical law, it must be the law of willing in accordance with the form of law, namely, universality.

The complete argument looks somewhat like Rousseau’s argument:

Rousseau’s Argument for the Universality Condition

5. If I am to be free, there must be laws of freedom.
6. If the laws are to be compatible with my freedom, I must legislate them.
7. If I am to legislate the laws, then we must agree on them.
8. If we are to agree on the laws, they must apply universally.

Kant’s Groundwork/2nd Critique Argument

1. If I am to be a will, there must be a law of my willing.
2. If my will is to be free, I must legislate the law.
3. If I am to legislate the law (if the law is to be a law of practical reason), then the law must be a priori.
4. If the law is to be a priori, then it must apply universally (and hence, abstract

from any matter of the will).

There are of course a couple of differences between this argument and Rousseau's argument, one of which is only apparent: Where 3 & 4 of Rousseau's argument speak of agreement, 3 & 4 of Kant's argument speak of aprioricity. But of course aprioricity is the condition under which every possible rational being can (and as Kant will argue, must) agree.

Indeed, in a remark situated right after "Theorem III" in which Kant begins to argue for the consequent of 4, Kant stresses just this point. There he responds to the proposal that since everyone desires happiness, the desire for happiness might be a universal practical law. The problem with this proposal is not that we don't know a priori that everyone desires happiness, for we do: "To be happy is necessarily the demand of every rational but finite being and therefore an unavoidable determining ground of its faculty of desire" (*KpV* 5:25). The problem is rather that in desiring our own happiness, we are not desiring one and the same thing; we are not agreeing on an end:

"For then the will of all has not one and the same object but each has his own (his own welfare)... In this way there results a harmony like that which a certain satirical poem depicts in the unanimity between a married couple bent on going to ruin: '*O marvellous harmony, what he wants she wants too*' and so forth, or like what is said of the pledge of King Francis I to the Emperor Charles V: 'What my brother Charles would have (Milan), that I would also have.' Empirical determining grounds are not fit for any universal external legislation and are no more fit for internal lawgiving; for each puts at the basis of inclination his subject... To discover a law that under this condition would govern them all – that is to say, with omnilateral concord – is quite impossible." (*KpV* 5:28)

In this passage Kant makes explicit that concord, agreement, is necessary for a maxim to

qualify as a law. And in doing so, he draws our attention explicitly to the parallel between internal and external lawgiving.

There is also an important difference in how the requirement of universality appears in relation to the requirement of agreement. For Rousseau, a requirement of universal form is simply a clever means of ensuring agreement. For Kant, universality is the very form of law; it is all that is left once we abstract from any content of the law. We may say that in Rousseau, the need for agreement motivates the introduction of a requirement of universal form, whereas in Kant, the need for agreement motivates abstraction from matter, leaving only universal form.⁷⁸ But in Kant's case no less than in Rousseau's, the need for agreement drives the argument. The problem with laws with merely given matter is that "we can always be released from the precept if we give up the purpose" on which it is conditioned (*G* 4:420). In other words, the rational will does not bind itself to such laws because the law is conditioned on a purpose, and the rational will is not bound to adopt that purpose. The law does not come from all rational wills (we do not agree) because it does not apply to all rational wills (it isn't universal with respect to practical reasoners). As Rousseau argued, the law must apply to all to come from all.

We might put Kant's argument as follows:

Kant's Groundwork/2nd Critique Argument

1. If I am to be a will, there must be a law of my willing.
2. If my will is to be free, I must legislate the law.

⁷⁸ To avert confusion, we must observe that although for Kant all laws have universal form, not all laws bind all wills (universally). For Kant, only laws that abstract from matter bind all wills.

3. If I am to legislate the law (if the law is to be a law of practical reason), then the law must be the subject of an a priori agreement.
4. If the law is to be the subject of a priori agreement, then it must apply universally (and hence, abstract from any matter of the will).

Now this argument looks exceedingly similar to Rousseau's argument, and given this similarity, the fact that Kant himself drew attention to a point of similarity with regard to step 3, the fact that Kant seems to have thought that the *Groundwork* was an adequate foundation for the *Metaphysics of Morals* as a whole, and the fact that Kant was an avid reader of Rousseau, it's quite tempting to think that Kant's reasoning for the UPR is the reasoning given in Rousseau's argument.

A Necessary, but not a Sufficient Condition

Although the Rousseauvian argument shows that the requirement that the law apply universally is a necessary condition for freedom, the condition of universal law does not seem to be a sufficient condition for freedom. In Rousseau, the requirement that the law apply universally, by aligning private interests towards a common interest, made it possible for everyone to legislate the law together. The possibility of legislating the law together, however, is not the same as the fact of our actually legislating the law together. Only the latter constitutes us as free. I am free when I choose for myself and not when the person who chooses for me chooses as I would have chosen. Hence, Rousseau claims, law only makes everyone free if everyone participates in the legislative process, such that the

legislation is in fact an act of every citizen's will and not merely in accordance with what they might have willed.

Similarly, for Kant, the universal law condition (the condition which requires that the law have universal form and abstract from material conditions) in the UPR is justified only by the necessity of that for which it is necessary, namely, the necessity that everyone legislate the law. For everyone to legislate the law, we must be able to agree on the law. To ensure our agreement on the law, the scope of possible laws must be restricted to those which abstract from any matter of the will and hence apply universally. The idea of a kind of general will is logically prior to the UPR: We begin with the idea of external freedom, derive from it the necessity of an a priori agreement, and only then derive from the latter the necessity of universal application. The united (or general) will of every citizen is the presupposed aim of the UPR. It is only because the requirement that laws apply universally ensures that our wills are united a priori that this requirement is binding on us.

While it seemed that Kant was pursuing a different argumentative strategy from Rousseau's in beginning with a principle of right and deriving from it the need for a general will, in fact, Kant seems to follow Rousseau's argument: he begins with the demand for freedom, derives from it the need for a united or general will, and derives principles of right from the requirements of the general will.

Of course, it is only after Kant has shown the deficiencies of mere private right that we see that there must be an *institutionalized* general will, a sovereign with the authority to

legislate positive law for all. Hence, it is not surprising that Kant's use of the term *allgemeiner Wille* only emerges with the need for such legislation. But even before Kant has argued for the institutionalized general will of public right, the principles of innate and private right are binding only insofar as we may each be said necessarily to legislate them.

Acting Under the Idea of Freedom

Kant's argument for the general will begins with the assumption that we are to be free and argues from that assumption that we must agree on laws that apply universally. But we must agree on such laws only if we are to be free in the first place. As it stands, Kant's argument does not show that the UPR is categorically binding on us. For that, Kant needs an argument to show that we are free or must think of ourselves as free.

Before we examine his argument, it will help to see that he rejects Rousseau's picture of freedom not because it is hard to get everyone together to participate in voting, but because such empirical voting could not fully constitute a general will. Kant thinks that the general will is an idea of reason which cannot be empirically given. Hence, Kant will argue, Rousseau's voting procedure is neither necessary nor sufficient for a general will.

Kant argues that we must think of ourselves as free and that we are free in the *Groundwork*. In the 2nd *Critique*, he (apparently) gives up on the second argument and offers a replacement. Because I think that the first argument from the *Groundwork* (that we must think of ourselves as free) remains illuminating, I will examine it before proceeding to the argument from the 2nd *Critique* (that we are free). The argument in the *Groundwork* begins

by asserting that a rational will must always act under the idea of freedom:

“Now I assert that to every rational being having a will we must necessarily lend the idea of freedom also, under which alone he acts. For in such a being we think of a reason that is practical, that is, has causality with respect to its objects. Now, one cannot possibly think of a reason that would consciously receive direction from any other quarter with respect to its judgments, since the subject would then attribute the determination of his judgment not to his reason but to an impulse. Reason must regard itself as the author of its principles independently of alien influences; consequently, as practical reason or as the will of a rational being it must be regarded of itself as free, that is, the will of such a being cannot be a will of his own except under the idea of freedom, and such a will must in a practical respect thus be attributed to every rational being” (G 4:448).

Before I offer my interpretation of this passage, I must ask the reader to set aside for now the objection that this passage can only refer to internal freedom, since only internal freedom may be characterized as the freedom of *reason* from control by external impulses. External freedom is not concerned with motivations, and hence external freedom may seem to be best characterized not as the freedom of my *reason* from impulses in the form of other human beings, but the freedom of my *person* from such impulses. I believe that this objection is ultimately misguided, but I will not be in a position to answer it until I have given my reading of the passage.

Kant’s central claim in this passage is that a rational will cannot think of its judgment as being determined by an impulse from outside because then it would not attribute that judgment to itself. Freedom and attributability are essentially connected: “A *person* is a subject whose actions can be *imputed* to him. *Moral* personality is therefore nothing other than the freedom of a rational being under moral laws” (MM 6:223). An

action may be imputed to me only if I determined myself to it, and I can determine myself to act only if I am free.

It is essential here to distinguish between being influenced by an impulse and receiving direction from an impulse. The word translated as “direction” is *Lenkung*, which literally means “steering,” or in a more abstract sense, “control” or “governing.” Kant’s thought is not that my reason cannot regard an action as its own if it is in some way moved by an impulse, but only that it cannot regard an action as its own if it loses steering. What it means for reason to have steering of the action is for reason to give itself the law of its willing. Within the bounds the law sets, impulses may have an influence on what we do, but reason retains steering or governance of the will insofar as reason legislates the will’s law.

In his moral philosophy, which includes both right and ethics, Kant considers two kinds of impulse that a rational will cannot regard as determining itself from outside: an inclination and another human being.⁷⁹ Freedom from determination by the former is internal freedom; freedom from determination by the latter is external freedom.

⁷⁹ There are other things that might control the way a human body moves, e.g., the wind. One could think that such things do not violate our freedom because they do not determine our will, only its circumstances. But the same could be said of your breaking my leg, and that is clearly a violation of my external freedom. Kant does not discuss our freedom from the effects non-rational things as such, but his discussion of the Antinomy of Practical Reason may be seen as an attempt to grapple with this question: How can I regard myself as free when the world does not allow me to actualize my will? But even if we regard this as a third kind of freedom, it is clearly of a different sort, for there is no corresponding fact of reason by which we know it, even if only practically. Fichte thinks that things like the wind are no threat to our freedom because they are not rational and hence do not depart from their determinate course. If the wind gets the better of us, it is because we failed to take account of it, not because it acted upon us, for the wind is passive. But he also thinks that our freedom with respect to things like the wind depends upon our knowledge of it (*GVR* 114-116). As I argue in the following chapters, Hegel thinks that this sort of freedom in community with nature through knowledge is a collective achievement of a rational state.

Corresponding to these two kinds of freedom are two kinds of imputability or personality.⁸⁰ When someone demands money of me at gunpoint and I decide to give it to her out of a motive of duty to preserve my life, my external freedom is compromised, but not my internal freedom, for although my action was determined in violation of the law of my external freedom, it was determined in keeping with the law of my internal freedom.⁸¹ From the perspective of a court of law, the deed of forfeiting my money is not imputed to me. But from the perspective of my conscience, the action of saving my life from duty can be imputed to me.

The law of external freedom is what determines what is attributable to me from the perspective of right. When everyone obeys the law, we are only responsible for our own behavior. When you violate the law of my external freedom, you may become responsible for behavior that would otherwise only be attributable to me. For example, in his revised view on lying, Kant claims that when you tell a murderer the truth, you are not responsible for his actions, but if you lie to the murderer, even from a desire to prevent him from murdering, “you are legally accountable for all the consequences that might

⁸⁰ The “moral personality” Kant refers to in the above passage from the *Metaphysics of Morals* is contrasted with “psychological personality”: “psychological personality is merely the ability to be conscious of one’s identity in different conditions of one’s existence” (*MM* 6:223). There are two kinds of moral laws, external and internal, corresponding to two kinds of personality. Kant calls the first civil personality, the “attribute of not needing to be represented by another where rights are concerned” (314). If I cannot represent myself, then my own deeds cannot be imputed to me. (The fact that in some cases a representative’s deeds may be attributed to someone even when their own deeds cannot is one on which the present investigation will shed further light.) Kant seems to have no special name for civil personality’s ethical analogue (he usually simply refers to it as “personality,” but he describes it as “freedom and independence from the mechanism of the whole of nature, regarded nevertheless as also a capacity of being subject to special laws – namely pure practical laws given by his own reason” (*KpV* 5:87).

⁸¹ If I hand the money over merely out of fear without regard for the permissibility of so doing, my internal freedom has also been compromised.

arise from it” (*VR* 8:427). From a non-moral perspective, you may just as easily play a role in causing a murder by telling the truth as by telling a lie. We are all in causal connection with one another and regularly contribute to determining what they will in fact do. But from the moral perspective, we distinguish what may be imputed to whom on the basis of laws of freedom. When you transgress these laws, you may enter another’s sphere of freedom and become responsible for what would ordinarily only be imputable to that person.

Moreover, although Kant is not explicit about this, it seems reasonable to say that on Kant’s view, I can in some cases cease to be responsible for behavior that came about as a result of your encroaching upon my external freedom. If I bump into you and break your nose, I am usually legally responsible. But if someone else pushed me into you, my body’s motions are not my responsibility, but the pusher’s. If you simply ask me for money and I hand it over, the deed is attributable to me. The money is now yours. But if you seriously threatened me, the deed of giving you the money is no longer attributable to me, but to you, and hence from a juridical perspective, what happened was not a giving, but a taking or stealing. In the former case, as in the latter case, you prompted me to give you the money, but in the former case, your prompting did not trespass my rights, and thus my action is entirely attributable to me.

Of course, I don’t lose all legal responsibility whenever my rights are infringed. Your stealing my money does not make you responsible for anything I might do in response. It may make you responsible for my legal efforts to get the money back, and hence to

compensate me for them. The question of how exactly any particular infringement of my rights affects my responsibility for my deeds and property is complex and cannot be tackled here. The key point is that laws play a role in determining what is imputable to someone.

The legislation of laws of external freedom is also a deed that may be imputed to me. Insofar as, *per impossibile*, I will that there be no laws of freedom at all, I will that others be at liberty to interfere with my will that there be no laws at all (through coercion, threats, etc.). But if others are at liberty to interfere, then there is no longer a clear sense in which my will that there be no laws at all is my own (since it belongs as much to those at liberty to interfere as it does to me). Even supposing that no one else tries to influence my will regarding the laws, the very supposition that they are permitted to interfere implies that my will is determined not just by itself, but with their leave, that it does not belong to me alone.⁸² Indeed, the sense in which there is a civil person to determine itself begins to fall apart.

There is of course a kind of psychological personality and a kind of ethical personality that might be preserved on this hypothesis. I could still recognize my preference that there be no laws of freedom as arising from my own thoughts or inclinations. I could perhaps even (confusedly, according to Kant) think that my

⁸² Indeed, the very sense in which what is expressed could count as a will at may begin to break down as soon as we allow that others may interfere with its activities. For something to constitute a will, and especially a rational will, it seems to need to have a certain unity which would be lacking in something whose determination may be subject to the control of any number of other beings who wrestle for control of it.

anarchism is the right ethical position and so be moved to support it regardless of the consequences. But these internal senses of personality have no bearing on right. The inner motivations of my actions are not publicly available, and hence no sense of civil person based on them is available to adjudicate the question of which laws everyone might or might not necessarily agree to.

The only sense in which a person can be recognized as legislating (or rejecting) external law already presupposes the independence of the individual. If the individual did not accept the laws of freedom that constitute her as an independent person capable of having an attitude toward the law imputed to her, she would deny the very status in virtue of which she could challenge those laws. The challenge is therefore contradictory and cannot be made.

Kant makes a parallel point in addressing the possibility of selling oneself into slavery:

“A contract by which one party would completely renounce its freedom for the other’s advantage would be self-contradictory, that is, null and void, since by it one party would cease to be a person and so would have no duty to keep the contract but would recognize only force.” (*MM* 6:283)

If I contract to renounce my freedom, I thereby renounce my personality, the possibility of having any action imputed to me. If no action can be attributed to me, then keeping or failing to keep the contract cannot be imputable to me, and therefore I have no duty to keep the contract. But a contract by which I acquire no duty is not a contract. Hence, such a contract is impossible. Similarly, the rejection of laws of freedom would be the rejection of the very laws that constitute me as a civil person capable of rejecting such

laws. Hence, such rejection is impossible. All civil persons necessarily will laws of freedom, for they are only civil persons if they are subject to laws of freedom, and they are only subject to laws of freedom if they give themselves the laws.⁸³

We can think of Kant's argument on this point as a modification of the original Rousseauvian argument for the universality condition. The Rousseauvian argument begins with the claim that for me to be free, there must be laws of freedom, for without laws to govern our interactions, I am subject to the will of others. This is a very different premise from the one Kant's *Groundwork*/2nd *Critique* argument begins with: If I am to be a will, there must be a law that governs my willing. And at first blush, that might sound like

⁸³ Though the terminology is his own, Kant is taking advantage of a claim deployed by two of the greatest influences on his political philosophy, Rousseau and Hobbes: If willing X cannot be construed as rational, then even if someone seems to will X, he cannot rightly be interpreted to will X. We've already seen Rousseau use this claim in arguing that it is impossible to sell oneself into slavery: "To say that a man gives himself gratuitously is to say something absurd and inconceivable; such an act is illegitimate and null, for the simple reason that whoever does so is not in his right mind" (*SC* 1.4.4). Hobbes makes a similar argument against the possibility of laying down one's right to self-defense:

"Whosoever a man Transferreth his Right, or Renounceth it; it is either in consideration of some Right reciprocally transferred to himself; or for some other good he hopeth for thereby. For it is a voluntary act: and of the voluntary acts of every man, the object is some *Good to himself*. And therefore there be some Rights, which no man can be understood by any words, or other signes, to have abandoned, or transferred. As first a man cannot lay down the right of resisting them, that assault him by force, to take away his life... And therefore if a man by words, or other signes, seem to despoyle himself of the End, for which those signes were intended; he is not to be understood as if he meant it, or that it was his will; but that he was ignorant of how such words and actions were to be interpreted." (*L* 65-66)

As is clear from this passage, Hobbes's argument rests on a version of the guise of the good thesis, and Rousseau's argument seems to rest on his own version of the same. But Kant's conception of the good is radically different from Hobbes's and Rousseau's, such that his argument must take a different form. For Hobbes, the chief political good is the means to everything men naturally desire, the "security of a mans person" and "the means of so preserving life, as not to be weary of it" (*L* 65-66). Rousseau agrees that our first good is preservation, an end given to us by nature (*SC* 1.2.2), but he thinks that our force and freedom are the primary and indispensable instruments of our self-preservation and hence cannot be given up (*SC* 1.6.3). For Kant, however, the good is not first given to reason by nature, but is rather constructed by practical reason (citation). Hence, where Hobbes and Rousseau start their argument that I cannot be construed to will X from the claim that X is not good for me, for Kant the fact that X is not good for me is itself grounded in the fact that I cannot will X. Hence, when Kant argues against contracting oneself into slavery, he does not stop with the fact that such a contract could not be good for one but argues that such a contract is impossible.

an unlikely premise for an argument for external freedom. The law to which Kant refers in this premise may seem to be a motivational law, not a juridical law that abstracts from motivations. For surely I am a will whether or not I live under juridical laws.

But if the argument regarding civil personhood I have just given is a reasonable interpretation of Kant, then I am not a civil person unless there are civil laws, and if I am not a civil person, then no external act can properly be attributed to me, as opposed to those who make or let me perform it. Hence, the first two premises of Kant's *Groundwork*/2nd *Critique* argument can substitute for the first two premises of the Rousseauvian argument:

1. If I am to be a (civil) will, there must be a(n external) law that governs my willing.⁸⁴
2. If I am to be free (a civil person to whom external acts may be attributed), I must legislate the law of my (external) freedom.

We can now return to the objection that the *Groundwork* passage in which I find this argument cannot be applied to external freedom because it specifically discusses the

⁸⁴ The difference between having a civil will and being a civil person can be understood from our mugging example. When the mugger threatens my life, the mugger overrides the civil law governing my will with his own law. Since the mugger's command is not a law I give myself, actions that I perform according to the mugger's command are not attributable to me, and I therefore do not exhibit civil personality in those actions. Nevertheless, I still have a civil will in virtue of the fact that I am responsive to external law. Though my action was not my own in the sense that I was not free with respect to it, it came about through my responsiveness to an external law. Indeed, the distinction between being given an external law by a mugger and losing one's responsiveness to external law (however temporarily) is what allows us to distinguish this case from the case of temporary insanity or some other such condition which might perhaps deny my responsibility but would also cast doubt on my fitness for full membership in society. My handing the mugger my money is so unproblematically not to be attributed to me as a transfer of property precisely because my motivations for doing so are precisely the sort of motivations that make me susceptible to the rule of law to begin with.

freedom of reason from external impulses. Civil personality is constituted by the juridical law, legislation of which, as we have already seen, must be an a priori law of reason if it is to have validity at all. Hence, my person is constituted by a law of reason that I give myself. Kant's statement that in a rational being under the idea of freedom, "we think of a reason that is practical, that is, has causality with respect to its objects" applies to external freedom as well as internal, for my actions can only be attributed to me in the sense relevant to external freedom if the right to perform them is reserved to me by a juridical law that my reason gives me. The violation of this law my reason gives itself by another person is always a usurpation of the steering of my will from reason in this sense, regardless of whether my reason was steering my will in the sense relevant to ethics.

Unfortunately, the argument does not stop here. For although Kant has shown that a civil person cannot be thought not to will laws of freedom, his argument has not yet shown that a human being is a person, a free being. Kant addresses a parallel problem in this argument as addressed to inner freedom in the *Groundwork*:

"We have finally traced the determinate concept of morality back to the idea of freedom; but we could not even prove the latter as something real in ourselves and in human nature; we saw only that we must presuppose it if we want to think of a being as rational and endowed with consciousness of his causality with respect to actions, that is, with a will." (*G* 4:448-9)

In other words, at this stage in the argument, Kant has shown that if we are to think of ourselves as rational wills, we must think of ourselves as free and therefore as subject to the moral law. But he has not shown us that we are in fact free.⁸⁵ If freedom turns out to

⁸⁵ Nor does he intend to, for by his own admission, that is impossible (*G* 4:448 fn.).

be impossible, then all we will have shown is that I cannot make sense of myself as a practical reasoner.⁸⁶

To make this argument specific to external freedom, we've shown that we cannot think of someone as a civil person, as someone to whom deeds may be attributed, without thinking of that person as free (in an external sense) and as willing universal laws of freedom. But we have not yet shown that anyone is in fact a civil person, an externally free being to whom deeds may be imputed. If freedom turns out to be impossible, that is, if everyone's a priori willing laws of external freedom is impossible, then we will simply have shown that we cannot make sense of ourselves as civil persons and externally free.

The solution to this problem in the *Groundwork* involves the availability of a second

⁸⁶ In a move that has puzzled interpreters, Kant makes this point by asking,

“why, then, ought I to subject myself to this principle [the CI] and do so as a rational being in the first place [my translation of *überhaupt*]...? I am willing to admit that no interest *impels* me to do so, for that would not give a categorical imperative; but I must still necessarily *take* an interest in it and have insight into how this comes about.” (G 4:449)

Readers are often surprised that Kant regards the demand for an interest impelling me to subject myself to the CI as a real challenge to the argument he has just given us. After all, if a rational will can make sense of itself as acting only under the idea of freedom and a free being is bound by the moral law, then Kant seems to have proved that we must be able to take an interest in the moral law.

But Kant has not at this point shown us why we must regard our wills as rational. Indeed, Hume manifestly didn't think of his will as rational in the sense that Kant is suggesting. Kant's question is, Why must I attribute any actions to myself in that stronger sense? Why must I think of myself “as rational and endowed with consciousness of [my] causality with respect to actions” (G 4:449)?

The first answers to this question that Kant considers take the form of interests because Kant is looking for an answer from practical reason as opposed to theoretical reason. Theoretical reason cannot give us cognition of our freedom. Practical reason is reason applied to questions about what to do, and such questions are answered by considering the relevant interests, from the perspective of practical reason. Hence, the answer to the question of why we should apply the high standard of attributability would seem to be associated with an interest. The closest thing Kant finds to an interest motivating such a standard of attributability is an interest in worthiness to be happy. But this interest presupposes the moral law and thus cannot be used in an argument that the moral law is binding on us. The solution in the *Groundwork* (which Kant abandons in the 2nd *Critique*) is that we can distinguish two standpoints from which to cognize our powers, as belonging to the world of sense, and as belonging to the world of understanding (G 4:453). When we think of ourselves as free, we take up the second standpoint and from it cognize our autonomy and through it, the moral law.

standpoint on the causality of our will, the standpoint of the world of the understanding, from which reason may regard itself as giving itself the law. From this standpoint we can see how our freedom is possible. Moreover, since reason shows a spontaneity in its ideas that goes “far beyond anything that sensibility can ever afford it,” we must regard ourselves as intelligence as belonging to the world of understanding. Since freedom is such an idea of practical reason, we must regard ourselves as under the idea of freedom from the perspective of the world of understanding.

Kant seems to abandon this last claim on the 2nd *Critique*. Nevertheless, Kant consistently claims that the possibility of our freedom depends upon the distinction between a world of sense and world of understanding, for the a priori legislation of the moral law for myself can only be coherently thought to occur in the world of understanding.

My thesis is that these arguments apply equally to external and internal freedom. We have seen that there can be no right unless everyone legislates the rules of right. I have not yet argued that this cannot be achieved in the world of sense, but it is already clear that a phenomenal legislation would be exceedingly difficult to achieve. However, if such legislation could be understood from the perspective of the world of understanding, then we could be licensed to regard ourselves as externally free and subject to laws of right regardless of people’s empirically expressed attitudes toward the law.

I will argue that this is precisely what Kant had in mind by arguing the general will is an idea of reason and hence, a representation of something that cannot be given in

experience, but which must nevertheless be assumed if we are to consistently think of ourselves as externally free. The fact that the general will is an idea of reason will explain both how Kant thinks that universal agreement on the laws can be achieved and why it is not necessary that everyone participate empirically in a democratic legislative process.

Ideas of Practical Reason

An idea of reason is “a necessary concept of reason... to which no congruent object can be given in the senses” (A327/B383). Kant uses the same notion of an idea of practical reason in the *Doctrine of Right* when he refers to the civil constitution as an idea of reason:

“What can be represented only by pure reason and must be counted among *ideas*, to which no object given in experience can be adequate – and a perfectly *rightful constitution* among human beings is of this sort – is the thing in itself.” (MM 6:371)

The same notion of an idea is also invoked in “On the Common Saying” when Kant refers to an original contract as an idea:

“Now this is an *original contract*, on which alone a civil and hence thoroughly rightful constitution among human beings can be based and a commonwealth established. But it is by no means necessary that this contract..., as a coalition of every particular and private will within a people into a common and public will (for the sake of a merely rightful legislation), be presupposed as a *fact* (as a fact it is indeed not possible)... It is instead *only an idea* of reason, which, however, has its undoubted practical reality, namely to bind every legislator to give his laws in such a way that they *could* have arisen from the united will of a whole people and to regard each subject, insofar as he wants to be a citizen, as if he has joined in voting for such a will.” (TP 8:297)

In other words, the civil constitution as a whole as well as the contract from which it is derived and which generates the general will are ideas of reason and therefore cannot be

given empirically.

To understand why no object given in experience can be adequate to an idea of reason, we must trace these ideas to their source in reason. For Kant, reason is “the faculty of principles” (A299/B356). A principle, Kant says, is something that allows us to “cognize the particular in the universal through concepts” (A300/B357). Kant gives the example of a syllogism:

Major Premise: All humans are mortal.

Minor Premise: Caius is human.

Conclusion: Caius is mortal. (This example comes from A322/B378)

In this syllogism we cognize the particular (Caius) in the universal (mortal) through a mediating concept (human). Kant explains that each of these propositions comes from a separate faculty of the mind: The major premise from the understanding, the minor premise from the power of judgment, and the conclusion from reason:

“I think first the **rule** (the *major*) through the understanding. Second, I **subsume** a cognition [Caius] under the condition [human] of the rule (the *minor*) by means of the **power of judgment**. Finally, I determine my cognition [Socrates] through the predicate [mortal] of the rule (the *conclusion*), hence *a priori* through **reason**.”
(A304/B360-1)

The purpose of this syllogism could be to discover that Caius is mortal, to learn a fact about a particular. But there’s another purpose for constructing syllogisms even when we know the conclusion in advance, namely, to see “whether [the conclusion] does not issue from judgments already given” (A304/B361 P). Given the judgment “Caius is mortal,” reason searches in the understanding for a universal rule from which to derive

this judgment. The rule “All humans are mortal” serves the purpose, because Caius is human. Kant identifies this major premise as the principle of the syllogism (A300/B357).⁸⁷

By seeking to derive given judgments from other judgments, Kant claims, reason “seeks to reduce the great manifoldness of understanding’s cognition to the smallest number of principles (universal conditions) and thereby to bring about the highest unity of this cognition” (A305/B361 P).

This process of seeking for the conditions of conditioned principles is familiar from many branches of science. In physics, various formulae for describing a more particular set of phenomena can be derived from a more general formula, and physicists continue to seek more fundamental formulae to unite theories that have still not been subsumed under one principle. Reason aims at reducing the manifold of cognition to as few

⁸⁷ It might seem puzzling that reason would be the faculty of principles when Kant identifies the understanding as the source of the major premise, the principle of the syllogism. I take it that reason is the faculty of principles because rules from the understanding are principles only when they are situated by reason within a syllogism. Without being related by reason to a particular cognition, a rule of the understanding does not allow us to “cognize the particular in the universal through concepts.” Kant clarifies this when he distinguishes between what can be called a principle only comparatively and what can be called a principle absolutely. Absolute principles “yield synthetic cognitions from concepts” (A301/B358); that is, they allow us to cognize the particular in the universal from mere concepts. Comparative principles do not yield synthetic cognition from concepts; they allow us to cognize the particular in the universal only by the mediation of intuition of the particular. The proposition “All humans are mortal” can be called a principle only comparatively, for it does not yield cognition of the particular (Caius) from mere concepts; it allows us to cognize the particular (Caius) in the universal (mortal) only by the mediation of our intuition of Caius and judgment that Caius is human. An absolute principle would be a rule that allows us to derive cognition of the particular from mere concepts. Anselm’s definition of God is a purported absolute principle, for it is purported to yield cognition of a particular (God) from concepts alone. Kant holds that no rules of the understanding are absolute principles. Even the *a priori* rules of the understanding, such as the rule that space is infinite or that everything that happens has a cause are not cognitions from concepts, “for they would not even be possible *a priori* if we did not bring in pure intuition (in mathematics) or the conditions of a possible experience in general” (A301/B357). Hence, the rules of the understanding are only comparative principles, only principles relative to other propositions with which reason combines them to yield cognition of particulars.

principles as possible.

Likewise, in syllogistic reasoning, we can seek the principles behind principles. In our example, we can seek for a prosyllogism that concludes “All humans are mortal.” For every comparative principle under which we can subsume a cognition, we can repeat the process and seek for a syllogism that derives the comparative principle (that is, the major premise), from a still higher principle. (We can also repeat the process for the minor premise, in this case, seeking an episyllogism that concludes, “Caius is human.”) For this process to come to an end, we would need to arrive at rules such that there is no middle term, or in Kant’s terms, no condition left to place between the subject and predicate. The totality of such rules would give the unconditioned condition of the conclusion, a complete explanation Caius’s mortality. Hence, reason’s aim is “to find, for understanding’s conditioned cognition, the unconditioned whereby cognition’s unity is completed” (A307/B364 P).

Kant argues that reason’s search for the unconditioned leads us to several concepts of the unconditioned, called ideas of reason, concepts of “what alone can complete the series of conditions when we proceed to trace these conditions to their grounds. This is the course [on] which our human reason, by its very nature, leads all of us” (A584/B612).

Ideas or concepts of reason refer to objects that can never be found within experience: “For a concept of reason concerns a cognition whereof any empirical cognition... is only a part” (A310-1/B367 P). Anything we find in experience is conditioned. Ideas refer to the unconditioned totality of conditions of what we find in

experience. Hence, any empirical cognition can at most be a part of the object of an idea of reason. The cosmological argument for the existence of God fails to acknowledge this. The mistake in the cosmological argument is to construe reason's seeking the unconditioned as reason's cognizing the unconditioned. The fact that reason seeks the unconditioned totality of conditions an object of experience does not mean that reason has cognition of such an unconditioned totality of conditions.

Moreover, we do not cognize the objects of the ideas of reason a priori. The ideas of reason are not constitutive, like the categories of the understanding: The categories can give us a priori knowledge of objects of experience because they constitute experience. Ideas of reason do not constitute experience. The mistake in the ontological argument for the existence of God is to think that an idea of reason can be constitutive of experience of its object.

The General Will as Idea of Practical Reason

We are now in a position to understand how an original contract is an idea of reason. Kant writes that an original contract “must be the ultimate basis on which a civil constitution is established,” and that it must be “adopted as with universal agreement and so by a contract” (*TP* 8:296). As we have seen, Kant denies that it need be presupposed as a fact: “as a fact it is indeed not possible” (*TP* 8:297). This may seem an odd claim. Why should an agreement to a constitution be empirically impossible? It is of course scarcely likely that a whole people should unanimously agree on a constitution, but not

impossible.⁸⁸

However possible such an agreement might be, a merely contingent agreement on an original contract could not be the basis for my external freedom. For if my external freedom depended upon someone else's contingent agreement to respect it, then it would not be external freedom at all. External freedom is "independence from being constrained by another's choice... insofar as it can coexist with the freedom of every other in accordance with a universal law" (*MM* 6:237). Hence, my external freedom cannot be dependent upon someone's contingent choice to honor that freedom.

My external freedom can, according to Kant, depend upon the availability of an authority with the ability to enforce universal law because when there is no such enforcement, no one is obligated to honor others' external freedom without assurance that others will honor theirs. Such an obligation would be incompatible with my innate freedom, my being my own master. The fact that my external freedom depends upon the existence of an authority with the ability and disposition to enforce universal law might seem to doom Kant's account of external freedom, for my external freedom seems to be conditioned by someone's contingent choice.

As with other ideas of reason, the distinction between the world of sense and the world of understanding saves the idea of the original contract from this problem.

According to Kant, we must distinguish the world of sense or empirical world, which is constituted in part by the form of our intuition and the form of our understanding, from

⁸⁸ It is also impossible to guarantee that such a constitution will bind all future generations, but it is not impossible that all future generations would in fact happen agree to it.

the world of the understanding, which is independent of the conditions of our cognizing it.⁸⁹ Although an a priori necessary original contract cannot be given in the world of sense, we cannot rule out that such a contract belongs to the world of the understanding. For the original contract is not in itself contradictory, but only contradicts the law of the world of sense that every event has a cause.

If we are to think of ourselves as free, we must think of ourselves under the idea of the original contract, that is, as having agreed necessarily (and therefore unanimously) to regard the laws as having been unanimously legislated by all the citizens. From the standpoint of the world of sense, the constitution of course depends upon contingent wills. But from the standpoint of the world of understanding it is possible to conceive a constitution as having arisen from the a priori united will of all. Only from the latter perspective can we see ourselves as externally free.

The original contract is an idea of practical reason because it comprises a totality of conditions and therefore an unconditioned condition of an object of practical reason: the individual person as externally free. We have already seen that the general will that the original contract creates is a necessary condition for us to think ourselves as externally free. We can now see that this original contract must itself be unconditioned, for if our agreement were conditional, then our freedom would be conditioned by something external to it and would cease to be freedom.

⁸⁹ For my purposes, it is unimportant whether we take these to be two different worlds or one world considered from two different standpoints.

An Act of Reason

Still, the fact that the original contract must be unconditioned if we are to be free and the fact that it is not contradictory to suppose such a contract do not show that there is such a contract. The last step in the parallel argument from Kant's ethics is an appeal to an act or fact⁹⁰ of reason (*Factum der Vernunft*): We know that we're internally free because we are aware that we legislate the moral law for ourselves.⁹¹ Kant describes our consciousness of the moral law as a non-empirical awareness:

“Consciousness of this fundamental law may be called a fact of reason because one cannot reason it out from antecedent data of reason, for example, from consciousness of freedom (since this is not antecedently given to us) and because it instead forces itself upon us of itself as a synthetic a priori proposition that is not based on any intuition, either pure or empirical... However, in order to avoid misinterpretation in regarding this law as *given*, it must be noted carefully that it is not an empirical fact but the sole fact of pure reason which, by it, announces itself as originally lawgiving.” (*KpV* 5:31)

The moral law is not a given empirical fact because it is not, like empirical incentives, a sensible inclination towards some end. Our consciousness of it registers sensibly as respect for the law, but such respect is an awareness of an activity of our own reason. The purpose of the distinction between the standpoints of the world of sense and of the world of understanding is simply to show how we could make sense of the moral law:

⁹⁰ The word *Factum* can refer both to a fact and to an act. The fact that the moral law is something with respect to which our reason is active (by legislating it), suggests that “act” is the better translation, but the way that Kant talks about it as the basis of our practical cognition of our freedom suggests “fact” as the more natural translation. I will leave the passages I quote unmodified.

⁹¹ See *KpV* 5:47.

“The moral law... provides a fact absolutely inexplicable from any data of the sensible world and from the whole compass of our theoretical use of reason, a fact that points to a pure world of the understanding and indeed, even *determines* it *positively* and lets us cognize something of it, namely a law.” (*KpV* 5:43)

Hence, we derive our knowledge of our internal freedom from our awareness that the moral law is binding on us, and not vice versa.

Similarly, in the domain of right, we must derive our knowledge of our external freedom from our awareness that the law of right is binding on us, and not vice versa. Unfortunately, Kant is nowhere very explicit on this point, but two passages make it relatively clear. The first comes from Kant’s argument for private property, which depends upon what Kant calls a “postulate of practical reason with regard to rights,” which says that “It is a duty of right to act towards others so that what is external (usable) could also become someone’s” (*MM* 6:252). Since this is a principle of practical reason, it may be described as reason’s will: “Reason wills that this hold as a principle, and it does this as *practical* reason, which extends itself a priori by this postulate of reason” (*MM* 6:247).

As a consequence of this principle, Kant claims, there must be a kind of non-empirical (intelligible) possession, possession of an object even when I am not holding it. Because we know of its possibility only through the postulate of practical reason with regard to rights, theoretical reason cannot discern the possibility of such merely intelligible possession. In other words, we know about intelligible possession, our freedom to possess objects as property, not because we are directly conscious of our freedom, but because we are conscious of practical reason’s will that things be ownable:

“No one need be surprised that *theoretical* principles about external objects that are

mine or yours get lost in the intelligible and represent no extension of cognition, since no theoretical deduction can be given for the possibility of the concept of freedom on which they are based. It can only be inferred from the practical law of reason (the categorical imperative), as a fact of reason [*Factum der Vernunft*]” (MM 6:252).

The argumentative strategy is precisely the same as the strategy in the 2nd Critique: Kant begins with a categorical imperative, a fact or act of practical reason and derives from it the possibility of something that cannot be given empirically: my freedom with respect to external objects. The fact that Kant is deploying this argumentative strategy within the *Doctrine of Right*, I contend, gives us reason to think that he has the same argumentative strategy in mind for other noumenal features of the *Doctrine of Right*, including the general will.

The second passage gives an indication of which imperative provides us with practical cognition of the general will. As we have already seen, in the conclusion to the Appendix to the *Doctrine of Right*, Kant claims that the civil constitution as an idea of reason “to which no object given in experience can be adequate” (MM 6:371). He then claims that “The *idea* of a civil constitution as such... is also an absolute command that practical reason, judging according to concepts of right, gives to every people” (MM 6:372). The command, which Kant describes is “a categorical imperative” (MM 6:371), is, “Obey the authority that has power over you” (MM 6:371). This idea, which presents itself in practical reason in the form of a categorical imperative, is a fact or act of reason [*Factum der Vernunft*]. By our awareness that we bind ourselves to this categorical imperative, we have practical cognition that the authority who has power over us is our

general will.

Awareness of the imperative to “obey the authority that has power over you” is the fact or act of reason that gives us practical cognition of the general will as embodied in a publicly recognized sovereign. As I said earlier, however, Kant’s arguments seem to require that our wills be united in legislation of the UPR itself. Practical cognition of that united will would need to be provided by some further fact or act of reason. To my mind the best candidate is this division of duties in the Introduction to the *Doctrine of Right* (following Ulpian):

1. “Be an honorable human being... a duty expressed by the saying, ‘Do not make yourself a mere means for others but be at the same time an end for them.’”
2. “Do not wrong anyone... even if, to avoid doing so, you should have to stop associating with others and shun all society.”
3. “(If you cannot help associating with others), enter into a society with them in which each can keep what is his.” (*MM* 6:236-7)

The first, I suggest, is the source of our practical cognition of ourselves as externally free, that is, as not subject to anyone else’s will.

Kant says that these “three classical formulae serve also as principles for dividing the system of duties of right into *internal* duties, *external* duties, and duties that involve the derivation of the latter from the principle of the former by subsumption” (*MM* 6:327). This characterization is suggestive of the Rousseauvian interpretation of Kant’s argument with which I began. We begin with a conception of my freedom as autonomy

(1) and then consider that my freedom is incompatible with associating with others on terms that violate their freedom (2). There are two possible solutions: Spread out so as not to have to associate, or (3) arrive at principles limiting my freedom to its compatibility with others' freedom, the fundamental principles of which, as I argued above, must be the subject of an a priori agreement. Kant's claim that the third category of duty derives external duties from internal duties confirms my claim that my being subject to a principle of freedom under universal law derives ultimately from the requirements of my own freedom.

There is an important difference between internal and external freedom that can mislead one into thinking that external freedom does not require a fact or act of reason: Kant thinks that it can never be known (even to ourselves) whether we have acted virtuously, i.e., in accordance with the ethical law, whereas we can know whether we have acted rightfully. This can make it seem that right and the external freedom it constitutes are empirically available and belong to the phenomenal realm, and that can lead to doubt about whether any fact or act of reason is necessary to give us cognition of our external freedom.

The error in this reasoning is in supposing that our ability to judge the rightfulness of an act implies that right and external freedom are empirically available. Kant denies this in the 1st *Critique*:

“One cannot say that the common concept is sensible and contains a mere appearance, for right cannot appear at all; rather its concept lies in the understanding and represents a constitution (the moral constitution) of actions that pertains to them in themselves” (A43-4/B61).

Judgments about the rightfulness of actions are judgments about the actions considered as things in themselves, that is, about actions as they belong to the intelligible or noumenal world. It is true that we advert to their sensible characteristics to form the judgments, but we also appeal to the idea of a rightful constitution (including an original contract establishing a legitimate legislative power) that cannot be given phenomenally. Without ideas of reason, no judgment about right could be given, nor could there be a corresponding judgment about external freedom. Our external freedom, like our internal freedom, is known to us only through facts or acts of reason.

Kant and Rousseau on Direct Participation in Legislation

We've already seen that Rousseau does not share the claim that reason requires us to seek a civil constitution, and it might be tempting to think that Kant's disagreement with Rousseau about the need for direct participation in the empirical process hangs on this point. But conceding this point alone would not show us that the citizens necessarily ratify the law. For that we need Kant's phenomenal/noumenal distinction and his arguments, in particular in the analogies of experience, that our postulates about the noumenal world are compatible with the phenomenal world. Without such a distinction and such arguments, a rational requirement that we seek the civil constitution would not show that we are free under an imperfect constitution, but only that we necessarily will to be free.

Kant's distinction between the world of understanding and the world of sense puts a new spin on an argumentative strategy that he shares with Rousseau and Hobbes, one which we saw at work in his argument against a contract to sell oneself into slavery.

Whereas Hobbes and Rousseau began arguments with the claim that X cannot be thought to be good for someone and argued that therefore X could not be interpreted to be their will, for Kant, the fact that X is not good for me is itself grounded in the fact that I cannot will X. Hence, when Kant argues against contracting oneself into slavery, he does not start with the fact that such a contract could not be good for one but must find a contradiction in the very contract itself.

Aside from defining the good in terms of practical reason, Kant's second contribution to this argumentative strategy is to distinguish between the standpoints of the world of sense and the world of understanding. Hobbes and Rousseau were committed to the thesis that the will's content is subject to rational constraint. Kant connects this thesis with his own distinction between things as they appear to us (appearances) and things as they are in themselves.

Appearances are not subject to rational constraint in the way that the will is according to Hobbes and Rousseau. In science, we don't correct observational data to match the laws but the laws to match observational data. Reason does play a role in our understanding of appearances, for appearances are subject to physical law, and reason plays a role in our search for the physical laws governing appearances. However, its role is not constitutive, but regulative: It guides our discovery of the laws without suggesting revisions of the appearances themselves.

On the other hand, things in themselves are outside the domain of physical law. Hence, there is no reason to suppose that they are determined independently of reason.

Indeed, rationality itself, Kant argues, belongs to the world of the understanding. It is “pure self-activity” and hence belongs “not to the world of sense but to the world of understanding” (G 4:452). Hence, although we can have no theoretical cognition of things in themselves, theoretical reason can offer no objection to thinking of things in themselves as being subject to rational constraint.

As an appearance, then, the empirical will or will insofar as it is given in the world of appearances (which Kant calls *homo phaenomenon*) is not subject to rational constraint in the way that Hobbes and Rousseau insist. To be sure, reason plays a role in our interpretation of the empirical will, but that role is always subordinate to the data of experience. Reason corrects our understanding of the empirical will, but it does not correct the will itself.

However, as a thing in itself, the intelligible will (which Kant calls *homo noumenon*) is subject to rational constraint. For this we have the *Factum der Vernunft* as evidence. The intelligible will’s content is therefore constrained by rationality itself. The intelligible will is therefore the proper object of arguments like Rousseau’s and Hobbes’s.

The distinction between the world of sense and the world of understanding allows Kant to press this argumentative strategy further than Rousseau or Hobbes. Note that what Rousseau tries to show in his argument against slavery is that there is a certain act a will cannot perform, not that there is an act that a will necessarily performs. This difference is crucial to explaining why Rousseau insists upon direct participation in the legislative process, and why Kant does not. Because Kant thinks that there are acts that the will necessarily performs even when the empirical will denies the act, he can explain

how we all necessarily participate in the legislative process, albeit not empirically. Empirically, the law is chosen by whatever authority holds the sovereign power. But noumenally, the idea of a civil constitution licenses us to regard our wills as united in ratifying any legislation, that is, as the noumenal cause of “the obligation in accordance with the law.”⁹² Thus, Kant agrees with Rousseau that our will must participate in the legislative process, but the necessary participation need only be noumenal.

Rousseau shares with Kant (and Hobbes) the idea that the rightfully binding content of the will is a matter of rational interpretation, but his will is thoroughly phenomenally grounded. There are things that the phenomenal will cannot do because they defy rational interpretation. And there are situations in which one decision (like signing the social contract) takes precedence over an incompatible act of will (breaking the law). But there is no activity that the will necessarily performs regardless of its empirical activity. To suppose that the will is necessarily active regardless of its empirical activity requires a noumenal/phenomenal distinction and some sort of argument for why we are licensed to suppose that there is noumenal activity.

I take it that this is why Kant does not settle for regarding the sovereign as merely authorized by an original contract, where authorization would consist in the agreement to let another decide my will. Kant does speak of an original contract, as we have seen, but reason’s activity extends through the contract to the ratification of each law itself:

⁹² “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 law.” (*MM* 6:227)

“Therefore only the concurring and united will of all, insofar as each decides the same thing for all and all for each, and so only the general united will of the people, can be legislative” (*MM* 6:314).

Conclusion

In this chapter I hope to have shown that Kant is a more thoroughgoing inheritor of Rousseau’s political philosophy than is often thought and that an understanding of this inheritance can contribute to our understanding of key interpretive and philosophical questions. First, the similarities between Kant’s and Rousseau’s arguments point the way to interpreting the *Groundwork* and the 2nd *Critique* as the sources of the UPR, clarifying its relation to the CI. Second, understanding the source of the UPR shows that freedom and not equality is the highest principle of Kant’s political philosophy, which, assuming the plausibility of Kant’s or a Kantian view, makes it a counterexample to the claim some of have recently espoused that equality is the highest principle of any plausible political theory. Third, seeing the centrality of a shared will in Kant’s political philosophy enables us to see him as solving two problems for Rousseau’s general will. Because for Kant, the civil constitution is an idea of practical reason, the reason of all necessarily endorses the constitution of the legislature. Moreover, because noumenal participation is compatible with phenomenal non-participation, practical reason licenses us to regard ourselves as free under the law through our noumenal participation in lawgiving, rather than requiring the empirical participation Rousseau insisted on.

If my interpretation is correct, Kant's *Doctrine of Right* is an answer to roughly the same basic problem as Rousseau's *Social Contract*: How can I live in society with others and obey only myself? The starting point of Kant's answer is the same as Rousseau's: The laws governing society must issue from everyone's will. The differences between Kant and Rousseau lie in their metaphysics of the will.

Chapter 3

Hegel's Criticisms of Kant's Morality Part 1: Moral Dilemmas and Duties to Self

Not everyone will see Kant's version of the general will as a genuine improvement over Rousseau's. Kant's replacement of direct empirical democracy with noumenal willing will doubtless strike many as substituting the incredible for the impractical. One of Kant's ablest critics on the introduction of the noumenal was Hegel.

Hegel and Kant are in agreement that the essence of the human will is freedom, that freedom is a kind of self-determination, and that reason is the source of this self-determination. But Hegel rejects Kant's account of noumenal freedom. Perhaps the best way to understand their disagreement is to start from Kant's moral epistemology. As we just saw, on Kant's account, the moral law is the *ratio cognoscendi* of our freedom. The moral law is an imperative, an ought. Since a mere ought is not an empirical is, Kant ends up tracing the reality of this ought to the noumenal realm. As I will argue in this chapter and the next, Hegel tries to cut off this argument to the noumenal at its starting point: The moral law as mere ought. Taking the mere ought as the fundamental moral phenomenon leads to various contradictions, or so Hegel argues. Instead, we should take the actualization of these imperatives in what he calls the Ethical Life as the *ratio cognoscendi* of our freedom.

Because Hegel attacks the theory of noumenal freedom primarily by striking at its roots in moral epistemology, his arguments are not directed at the noumenal Kantian general will, nor are they specific to right; rather, they take on all of Kant's morality, right and ethics, insofar as it is known through the moral ought as fact of reason. Hence, although this chapter and the next are aimed at discovering the grounds of Hegel's rejection of Kant's general will, it deals with Kant's morality more broadly and rarely touches specifically on the general will. It will be helpful to bear in mind throughout that in arguing that a mere ought cannot be the *ratio cognoscendi* of a coherent moral system, Hegel is undermining Kant's noumenal theories of freedom, both internal and external, and setting the stage for his own account of the general will.

In addition to the attack on Kant's moral epistemology, there is a second line of criticism bearing directly on the general will, a line that runs through the arguments that will be considered in this chapter and the next. This second line argues against a conception of the general will as the unanimous wills of all individuals. We saw in the first chapter that Rousseau construes the general will as a shared capacity of willing so as to allow for disagreement. Even when my opinion does not prevail, I can still regard the general will as my own because I am a part of the shared capacity, and I prioritize choosing together. Kant has no need for such a shared capacity because he claims that our wills are a priori unanimous in ratifying the laws of the land. For Kant, the general will is just a group of individual wills that necessarily will the same thing. They don't share a capacity for willing; their individual capacities simply harmonize.

Hegel's second line of criticism is against the sufficiency of the individual will qua individual. According to Hegel, the individual will is not capable of resolving moral dilemmas, reconciling duties to self with the universality of law, or reconciling duty with the need for happiness. Each of these tasks can only be performed by the general will. Since each of these tasks is essential to constituting the individual will as rational and hence as free, the individual will must be thought of not as an independent constituent of the general will (as on Kant's model), but rather as essentially a part of the general will (as on Rousseau's).⁹³

Hegel's criticisms of Kant's notion of the will are scattered throughout many different texts, but I will focus on the criticisms offered in Hegel's mature exposition of his system, the *Encyclopedia*. I do so for several reasons. First, there may be some doubt about whether Hegel remained committed to some of his earlier criticisms of Kant's practical philosophy, whereas the *Encyclopedia* clearly represents Hegel's considered views on the subject. The *Encyclopedia* was published in three editions during Hegel's lifetime (1817, 1827, and 1830), and each was refined while serving as the outline for many years of lectures.

Second, in the *Encyclopedia*, Hegel's criticisms of Kant serve to introduce and partly justify Hegel's own views on the nature of individual and universal will. Since understanding Hegel's own position is our goal, the criticisms of Kant in the *Encyclopedia* are a natural starting point.

⁹³ However, as we shall see in the fifth chapter, Hegel has a different idea from Rousseau about what it takes to be a part of the general will.

Third, the *Encyclopedia* contains several criticisms of Kant's view not directly addressed in the *Philosophy of Right*, which is an expansion of the relevant section of the *Encyclopedia*. The most salient objection to Kant in the *Philosophy of Right* is his claim that particular duties cannot in fact be derived from the merely formal categorical imperative. This objection, known as the "formalism objection," has received much attention in the literature. But immediately after making this objection, Hegel refers us to the *Encyclopedia* and the *Phenomenology of Spirit* for arguments he does not cover in the *Philosophy of Right*:

"The further antinomies and shapes assumed by this perennial *obligation*, among which the merely moral point of view of *relationship* simply drifts to and fro without being able to resolve them [the antinomies] and get beyond obligation, are developed in my *Phenomenology of Spirit*, pp. 550ff.; cf. *Encyclopaedia of the Philosophical Sciences*, §§420ff" (*PR* 135R).

Interestingly, none of the *Encyclopedia* criticisms of Kant's Morality seem quite the same as the formalism objection in the *Philosophy of Right*.

The Formalism Objection

One tendency in the philosophical literature on Hegel's critique of Kant's moral philosophy has been to focus on the formalism objection. Kant claims that moral worth is not in our ends, the matter of our will, but in the principle of the will, which is formal (*G* 4:400). Moreover, if the moral law is to be categorical, it cannot presuppose any particular purposes as its condition (for then it would be hypothetical) (*G* 4:420-1).

Without the matter of purpose, it can issue only a formal constraint on willing, namely that the will's maxims have the form of a law, universality. Hegel's formalism objection is the claim that particular duties cannot in fact be derived from this formal constraint:

“However essential it may be to emphasize the pure and unconditional self-determination of the will as the root of duty... to cling on to a merely moral point of view without making the transition to the concept of ethics reduces this gain to an *empty formalism*, and moral science to an empty rhetoric of *duty for duty’s sake*. From this point of view, no immanent theory of duties is possible. One may indeed bring in material *from outside* and thereby arrive at *particular* duties, but it is impossible to make the transition to the determination of particular duties from the above determination of duty as *absence of contradiction*, as *formal correspondence with itself*, which is no different from the specification of *abstract indeterminacy*; and even if such a particular content for action is taken into consideration, there is no criterion within that principle for deciding whether or not this content is a duty.” (PR 135R)

Perhaps the literature has focused on this objection because its import for Kant’s theory is easier to understand than that of Hegel’s other objections; perhaps also because it is one of the few challenges to Kant’s view expressly addressed at Kant in the *Philosophy of Right*. Most of what might constitute Hegel’s criticism of Kant’s view is not explicitly addressed at Kant in particular, but at what Hegel refers to as “Morality,” a family of views to which Hegel thinks Kant’s view belongs. In Hegel’s view, Morality traces its origins as far back as Socrates, whose reasoned moral inquiry and faithfulness to his own conscience set him at odds with the ethical life of Athens. Hegel thought that Morality began to develop with great force and disruptiveness in the Reformation, and Kant brought Morality to a new level of depth and clarity. Thus Kant receives special attention in Hegel’s criticism of Morality, but not everything that Hegel attributes to the Moral standpoint is directed at Kant’s view. The formalism objection stands out because it is one of the few objections in which Hegel identifies Kant by name as his target.

There is no consensus among philosophers about how to understand the formalism

objection or whether it succeeds. Many have argued that Hegel's charge is based on a misconception of how the universalization is supposed to work.⁹⁴ As interesting as this debate has been, the formalism objection should not be considered the totality of Hegel's objections to Kant's moral theory, nor I think, the centerpiece.

Hegel suggests as much immediately after discussing the formalism objection when he refers us to the *Phenomenology of Spirit* and the *Encyclopedia of the Philosophical Sciences*.⁹⁵ The criticisms in the *Encyclopedia* are divided into four lettered points which will occupy us for the remainder of this chapter and the next. One striking feature of these four points is that none of them seems to be the formalism objection. Hegel had already clearly articulated the formalism objection by the time he published the *Encyclopedia*.⁹⁶ Thus, Hegel's decision not to make the formalism objection the centerpiece of his criticism of Morality in the *Encyclopedia* suggests that it is not the core of Hegel's argument. Given that the formalism objection occupies only one paragraph of the *Philosophy of Right*, we should be careful not to give it too much weight in our interpretation of the argument.

Unless, that is, the point of the formalism objection is quite different from what interpreters have sometimes thought. The unifying thread of each of the *Encyclopedia*

⁹⁴ The argument often runs as follows: The contradiction is not between the universalization of a maxim to steal and the independent existence of the institution of property; rather, the contradiction is internal to the universalized maxim itself: In willing to acquire something by stealing property I will that there be a system of property that makes it available to me for stealing, but in universalizing my maxim I will that there not be a system of property.

⁹⁵ See the passage from *PR* 135R quoted immediately before the beginning of this section.

⁹⁶ See, for example, *The Phenomenology of Spirit* ¶¶429-437, published in 1807, ten years before the first edition of the *Encyclopedia*, and *The Scientific Ways of Treating Natural Law, Its Place in Moral Philosophy, and Its Relation to the Positive Sciences of Law* p. 436-439/76-80, published 1802-3.

criticisms is the idea that a merely contingent good is in itself defective, that the good must realize itself. Hegel repeatedly criticizes Kant's idea of duty or the good as a mere "ought." If we interpret the formalism objection as a continuation of this thought, then it is not the claim that a merely formal principle cannot tell us what we ought to do, but that a merely formal principle can *only* tell us what we ought to do. It is, by its formality, always a mere ought, never a necessary actuality. Why Hegel thinks of that as a defect in Kant's view is the subject of this chapter and the next. Hegel's account of ethics as a necessary actuality is the subject of the fifth chapter.

The Encyclopedia Criticisms

I'll begin by setting the third of the *Encyclopedia* criticisms of morality aside because it is only indirectly helpful for getting us to Hegel's view:

"But the subject is not only in its existence⁹⁷ a particular in general; it is also a form of its existence to be *abstract* self-certainty, abstract reflection of freedom into itself. The subject is thus distinct from the reason of the will, and capable of making the universal itself into a particular and thus into a semblance. The good is thus posited as a contingency for the subject, who can therefore resolve on something opposed to the good, can be *evil*." (*EG* §508)

Hegel may have in mind the following: It seems incomprehensible on Kant's view that anyone could knowingly choose evil (cf. *R* 6:43), yet Kant himself acknowledges that people choose evil.

⁹⁷ Translation modified. Wallace and Miller have "reality" for the German is *Daseyn*. I have replaced it with "existence" to match the Brinkman and Dahlstrom translation of the term in the *Encyclopedia Logic* and the di Giovanni translation of the term in *The Science of Logic*. While it is not clear to me that Hegel intended to use the term in the technical sense developed in his logic, "existence" seems a better rendering of the German in any case.

But this does not seem to me to be Hegel's main point in the *Encyclopedia* criticisms as a whole, for the recurring theme in each of the other three criticisms is, as I suggested before, a concern with morality's being merely an "ought," that is, with the fact that on Kant's view, it is merely contingent whether duties are in fact performed.⁹⁸ Moreover, Hegel says that the fourth criticism is the upshot of the third criticism,⁹⁹ which suggests that this third criticism is intended as supporting background for the fourth. Hegel's primary attack is aimed at the idea of duty or the good as a mere "ought." The contingency of doing good is the explanation for why the good is contingently realized.

The overarching goal of this chapter and the next will be to illustrate why Hegel thinks that morality's being a mere "ought" is a problem for Kant. I will begin with the first criticism, which examines Kant's denial that there are moral dilemmas. I identify two aspects of Kant's moral philosophy to which this criticism may be responding: first, the characterization of duties of ethics as duties to have ends, and second, the question of how to resolve possible conflicts among internal and external duties. In either case, allowing the agent a role in resolving the conflict creates problems because the duties are

⁹⁸ Of course, for Kant, duties are practically necessary, that is, they are necessarily performed by practically rational beings insofar as rationality determines their actions. But because finite rational beings do not necessarily behave rationally, duties are not necessarily performed.

⁹⁹ The fourth criticism follows immediately after the third and immediately connects it with concerns about the contingency of the performance of duty:

"It is therefore contingent whether [external objectivity] harmonizes with the subjective aims, whether the *good* realizes itself in it, while *evil*, the aim that is in and for itself null, is null in it;—contingent too whether the subject finds in it his well-being, and more precisely whether in it the *good* subject becomes *happy* and the *evil* subject *unhappy*. But at the same time the world *ought* to let what is essential, the good action, be carried out in it, as it *ought* to grant the *good* subject the satisfaction of his particular interest, but refuse it to the *evil* subject, just as it *ought* to nullify evil itself." (*EG* §509)

supposed to have absolute validity. The solution, Hegel will argue, is that a harmony of duties is not something that merely ought to exist or ought to be generated by the subject's own will, but something that actually does exist wherever duties are well-constituted.

The second criticism takes aim at a tension in the idea of a self-legislated universal law. Insofar as it is self-legislated, Hegel claims, there is a special duty to the self, but insofar as it is universal, there can be no special duty to the self. This problem would be solved if morality were more than an ought, for then morality could provide for the self in such a way as to harmonize duties to the self with the demands of the universal.

Although the fourth and final criticism continues to argue against Morality as a mere ought, it differs from the previous two criticisms in that it discusses an issue of which Kant was deeply aware and about which Kant wrote many times throughout his critical period as an "Antinomy of Practical Reason." Since interpreters have claimed not to be able to see what problem Kant is trying to respond to, and since Kant seems to have modified his response to the problem over time, it will be worth dedicating a separate chapter to tracing the contours of the problem. In this chapter we will focus on two alleged problems internal to Kant's theory of duty; in the next we will discuss the problem of how to Kant's theory of duty relates to the highest good.

Hegel's First Criticism: Moral Dilemmas

"Because of the indeterminate determining of the good, there are in general *varieties* of good and *several duties*, the diversity of which involves dialectical opposition between them and brings them into *collision*. At the same time, because of the unity of the good, they *ought* to stand in harmony, and yet each of them, though it is a

particular duty, is, as duty and as good, absolute. The subject *ought* to be the dialectic which *decisively concludes* a combination of them by excluding the others and thus sublating this absolute validity.” (EG §508)

Hegel seems to offer two criticisms here. First, duties ought to be harmonized but in fact conflict. Second, the subject ought to be able to solve this by excluding some, but apparently cannot do this while preserving the absolute validity of duty. Hegel may have something like the central conflict in *Antigone* in mind, a work he admired and referenced often. Antigone must choose whether to obey the political ruler Creon, who has ordered that the corpse of her rebellious brother not be buried, or the gods, who, according to tradition, have ordered that families see to the burial of their deceased. Antigone thus finds herself torn between conflicting political and religious/familial duties.

What makes this conflict of duties more wrenching than a mere conflict of the agent’s interests is that in a conflict of the agent’s interests, the agent must simply set priorities, decide what she values the most. Duty, on the other hand, is not an enticement, but a command. I cannot dismiss a duty by simply deciding that I don’t value it as much as another duty. It lays claim to me whether I obey or not, just as Creon lays claim to Antigone’s obedience even after Antigone decides to disobey Creon to honor the gods. This is what Hegel means when he suggests that if the subject resolves the problem by a decision, the absolute validity of the law will be sublated. Antigone cannot resolve her problem simply by making a decision as to which duty to follow, since duty ought not to be a matter of decision. It ought to be absolutely valid, whether Antigone obeys or not.

Kant takes it as a logical truth that duties cannot conflict:

“But since duty and obligation are concepts that express the objective practical

necessity of certain actions and two rules opposed to each other cannot be necessary at the same time, if it is a duty to act in accordance with one rule, to act in accordance with the opposite rule is not a duty but even contrary to duty; so a *collision of duties* and obligations is inconceivable (*obligationes non colliduntur*).” (MM 6:224)

But Kant goes on to concede that though duties themselves may not be in conflict, there may be conflicts among the grounds of our duties:

“However, a subject may have, in a rule he prescribes to himself, two *grounds* of obligations (*rationes obligandi*), one or the other of which is not sufficient to put him under obligation (*rationes obligandi non obligantes*), so that one of them is not a duty. – When two such grounds conflict with each other, practical philosophy says, not that the stronger obligation takes precedence (*fortior obligatio vincit*) but that the stronger *ground of obligation* prevails (*fortior obligandi ratio vincit*).” (MM 6:224)

In other words, Antigone cannot have both a duty to bury and a duty not to bury her brother’s body, but there can be grounds for a duty to bury her brother’s body and grounds for a duty not to bury it. Only one duty can emerge from the conflicting grounds.

Just what are these grounds? Three possibilities must be treated separately. Kant may have in mind a conflict among wide duties, between a wide and a narrow duty, or among narrow duties.

Wide duties

A wide duty is one that “leaves a playroom (*latitude*) for free choice in following (complying with) the law, that is, ... the law cannot specify precisely in what way one is to act and how much duty one is to do by the action for an end that is also a duty” (MM 6:390). In other words, a wide duty is a duty to have an end without an accompanying obligation to perform any particular action to achieve that end. A narrow duty, by

contrast, is a duty to perform a particular action.

Kant hastens to disabuse his readers of a possible misunderstanding. One might think that a duty to have a particular end doesn't have a lot of teeth because one can have an end without assigning it a very high priority. A millionaire could lay claim to having the end of eliminating poverty simply by contributing one dollar to that end, reserving the rest of her riches for her own entertainment. She has the end of eliminating poverty; it's just pretty low on her list. Kant's wide duties are not satisfied so easily: "A wide duty is not to be taken as permission to make exceptions to the maxim of actions but only as permission to limit one maxim of duty by another (e.g., love of one's neighbor in general by love of one's parents)" (*MM* 6:390). This sounds a bit like what Kant meant when he claimed that in cases where the grounds of duties conflict, the stronger grounds give rise to duty.

There is, however, one small incongruity: Kant said that in cases of conflicting grounds of duty, the stronger grounds prevail, whereas he says that wide duties involve a "permission to limit one maxim of duty by another" that "leaves playroom (*latitude*) for free choice in following (complying with) the law" (*MM* 6:390). Apparently conflicts among wide duties are not generally to be resolved by comparing the strength of grounds but by a free choice. In at least some cases, a comparison of the strength of grounds seems preferable: Am I permitted to ignore the urgent need of a stranger I could help with little trouble to myself to go home and practice the piano, in conformity with my duty to cultivate my talents?

Hegel may have something like this in mind when he objects that the subject's resolution of conflicting duties by excluding some "sublates [their] absolute validity."

Duties are supposed to be necessary. If I get to choose which among a set of incompatible duties to perform, they lose this absolute validity. In an earlier version of Hegel's criticism, he wrote, "If the man of many virtues tries to make a hierarchy of his creditors, all of whom he cannot satisfy, he declares himself as less indebted to those he subordinates than to the others which he calls higher. Virtues therefore may cease to be absolutely obligatory and thus may become vices" (*GC* 294). Giving the individual playroom to choose between the virtues when they conflict is at tension with the idea that a duty is absolute. Granted, Kant has an answer to this worry from a technical standpoint: It is not the action, but the end, which is necessary, and I am not permitted to omit the end, only the particular action that serves it. But Hegel's point seems to be that that technical solution distorts the character of duty. When a friend fails to help me in my moment of need, it is one thing for him to say, "I'm sorry I couldn't be there because duty required that I be elsewhere." It is another for him to say, "I'm sorry I wasn't there. The ends of duty required me to be in two places, so I was permitted to choose the other end." It seems that duty itself should decide between helping me and whatever else my friend had to do. Leaving this decision in the hands of my friend seems to put him in an awful bind rather than freeing him.

That said, Kant's claim that the grounds of duty must decide between conflicting duties suggests another strategy for dealing with certain kinds of conflicts of duty, so we should consider cases involving narrow duties as well.

Narrow and wide duties

In the *Metaphysics of Morals* Kant states that duties of right are narrow, whereas duties of virtue are wide (*MM* 6:390), suggesting that our duties to obey the law of the land are all to be prioritized over other duties. Kant refers to the narrow duties as “unremitting” (*G* 4:424), suggesting that narrow duties take precedence over wider duties. This might seem to suggest an implausible view about the priority of duties: Surely there are occasions on which the ground of my duty of right ought to yield to the ground of another duty. It is a (narrow) duty of right not to trespass on others’ property and a (wide) duty of virtue to provide for someone’s safety, but surely it is not my duty to refrain from trespassing on someone’s property when that is the only way to save someone from a fire. There are of course traditional legal doctrines for dealing with this problem that Kant would no doubt like to accommodate. The question is whether he can.

The example suggests a problem with Kant’s claim that duties of virtue are wide, or alternatively, a problem with Kant’s division of morality into duties of right and duties of virtue. My duty to provide for someone’s safety seems narrow in that we can specify exactly what must be done, and yet it seems not to be a duty of right, since its end is not external freedom.¹⁰⁰ Still, labeling this duty a narrow duty of ethics does nothing to resolve its apparent contradiction with the narrow duty of right not to trespass.

Kant might have avoided the problem by acknowledging the possibility of a conflict

¹⁰⁰ A further wrinkle in this issue is the possibility that others could have saved my friend from the fire. Such cases fit uncomfortably in Kant’s framework. For an interesting discussion and alternative proposal, see Korsgaard (2018).

between right and ethics. Sometimes internal and external freedom are at odds, or so the thought goes. But Kant denies the possibility of a conflict of duties in the introduction to the *Metaphysics of Morals*, which addresses what is common to right and ethics, and his notion of duty as categorically imperative or practically necessary won't allow Kant such an easy answer.

One way to avoid the conclusion that I must not trespass even when doing so would save my friend is to deny that there is any narrow duty in such extreme cases. On this proposal, the narrow duty not to trespass is strictly speaking a duty not to trespass unless doing so could save lives.¹⁰¹ So there is in fact no conflict at all between this narrow duty and a duty to provide for one's friends' safety. There are two ways this limit on the narrow duty not to trespass could be thought to obtain. First, this limit might be written into the legislation or the judicial process by which it is enforced. This would make the limits on the duty uncontroversial, but there is no guarantee that the law will in fact make such allowances, so we cannot be guaranteed that the written or enforced law will offer a satisfactory list of narrow duties.

Alternatively, we might suppose that limits on the narrow duty not to trespass are determined by reason, regardless of how the publicly declared law determines them. And indeed, Kant seems to support some exceptions not written into the public law, as evidenced by an exception he gives to the imperative to obey the sovereign: "Obey the authority who has power over you (in whatever does not conflict with inner morality)"

¹⁰¹ We might build in other exceptions for avoiding severe injuries, etc. as well.

(*MM* 6:371). Evidently, inner morality can trump the law of the land.

Or perhaps more accurately, the law of the land can never be at odds with inner morality, for the standard of legislation is that it be possible that the people could agree to it: “What a people cannot decree for itself, a legislator also cannot decree for a people” (*TP* 8:304). Kant thinks that this is the principle by which the people has to appraise its rights negatively, “that is, [to] appraise merely what may be regarded as *not ordained* by the supreme legislation, as with its best will” (*TP* 8:304). When purported legislation seems to require something incompatible with inner morality, we are licensed to deny that the purported legislation is in fact the general will, in fact the law, and in fact binding on our will. The legislator decreeing such a law for a people cannot be interpreted as expressing the general will and therefore speaks without authority. So perhaps trespassing on another’s property to save the lives of my friends is permitted because the law cannot be interpreted as requiring me not to trespass in such circumstances.

But this solution seems to create difficulties for Kant’s view, for it is in tension with Kant’s insistence that even when legislation falls afoul of the principle of right, we may not resist it. As we have already seen, our being authorized to judge for ourselves according to the principle of right is connected with our inalienable rights, and in particular with freedom of the pen::

“A nonrecalcitrant subject must be able to assume that his ruler does not *want* to do him any wrong. Since every human being still has his inalienable rights, which he can never give up even if he wanted to and about which he is authorized to judge for himself, while, on that assumption, the wrong that in his opinion is done to him occurs only from the supreme power’s error or ignorance of certain consequences of

his laws, a citizen must have, with the approval of the ruler himself, the authorization to make known publicly his opinions about what it is in the ruler's arrangements that seems to him to be a wrong against the commonwealth. For, to assume that the head of state could never err or be ignorant of something would be to represent him as favored with divine inspiration and raised above humanity... To want to deny them this freedom [of the pen] is not only tantamount to taking from them any claim to a right with respect to the supreme commander (according to Hobbes), but is also to withhold from the latter – whose will gives order to the subjects as citizens only by representing the general will of the people – all knowledge of matters that he himself would change if he knew about them and to put him in contradiction with himself.” (TP 8:304)

The argument seems to be as follows: No ruler can be thought to be free from error and ignorance. Therefore, no (sane) ruler can think himself to be free from error and ignorance. If a ruler shuts down freedom of the pen, he shuts himself off from knowledge of ways in which he is unintentionally wronging me. But his shutting himself off from such knowledge is intentional and can therefore, in the context of his awareness of his finitude and imperfection, only be taken as deliberately wronging people. I have a right to judge whether or not I am being wronged in my inalienable rights. Suppose that I judge that my ruler is wronging me in my inalienable rights: Then for me to remain nonrecalcitrant, I must assume that he wrongs me unintentionally. If I am not to be recalcitrant, I must be able to assume that my ruler does not want to wrong me. Therefore, if I am not to be recalcitrant, the ruler must not shut down freedom of the pen.

Much hinges on what Kant means by “must” in the sentence, “A nonrecalcitrant subject must be able to assume that his ruler does not *want* to do him any wrong.” Kant

may simply mean that as a matter of psychological (or sociological) fact, citizens get recalcitrant when they think that their ruler intentionally wrongs them. But it seems more likely, given that this is the metaphysics and not the anthropology of morals,¹⁰² that Kant uses “must” in a moral register: For me to be obligated to the ruler, I must be able to assume that the ruler means me no wrong. Freedom of speech cannot be curtailed because without it, the conditions for my obligating myself to the ruler are not met.

Aside from the arguments’ appearance in the *Metaphysics of Morals* (rather than the *Anthropology*), three other considerations support my interpretation: First, taken as a psychological (or sociological) fact, the first premise is baldly false. A subject may without recalcitrance assume that his ruler *does* want to do him wrong, if for example, he is too intimidated to be recalcitrant. Second, the argument connects the premise under consideration to the premise that I may judge about my inalienable rights for myself. But it is unclear how a psychological (or sociological) fact about our tendency to recalcitrance and a normative claim about what I may judge combine to yield the intended conclusion of the argument.

Third, the normative interpretation of this sentence premise fits naturally with Kant’s moral theory: If I am to be obligated to the ruler, I must obligate myself. Granted that Kant argues that reason necessarily obligates us to whatever ruler establishes an approximately rightful condition, we still must judge for ourselves whether anyone has established an approximately rightful condition. Presumably judging that the ruler does

¹⁰² Hence, not concerned with psychological (or sociological) facts

not want to do wrong is part of such a judgment.

But there's a tension between this affirmation of the individual's right to judge about his inalienable rights and Kant's claim that we may never resist the sovereign, even when the sovereign does wrong us, "since each resistance would take place in conformity with a maxim that, made universal, would annihilate any civil constitution and eradicate the condition in which alone people can be in possession of rights generally" (*TP* 8:299).

Kant tries to avoid the problem by distinguishing between a right to publish our views on the wrongs the state commits from a right to resist the state. When a law violates the principle of right we may publish our judgments to that effect, but we may not resist the law:

"Hence, a law given about this is not to be regarded as the real will of the monarch, to whom counterrepresentations can accordingly be made. In all cases however, where something of this sort was nevertheless arranged by the supreme legislation, general and public judgments could be passed on it, but resistance to it in word or deed could never be summoned." (*TP* 8:305)

Not only does this position make granting exceptions to laws of right difficult, it leads to an obvious puzzle: What about cases where exercising freedom of the pen is itself an act of resistance? For example, what if the law in question is a law requiring that people not publish criticism of the laws?

If the people are not permitted to resist a law curtailing freedom of the pen, then what becomes of "the sole palladium of the people's rights" (8:304)? Kant's argument for freedom of the pen begins, as we have seen, from the inalienable rights belonging to every human being, "which he can never give up even if he wanted to and about which he is

authorized to judge for himself.” If these rights can issue in no action, not even an expression of one’s judgment that they have been violated where freedom of the pen has been proscribed, in what sense are they inalienable, or perhaps more to the point, in what sense are they rights? And perhaps more urgently, what becomes of autonomy, our legislation of the law to ourselves, if the legislation of the law is unmoored from our own judgments about how it applies?

On the other hand, if we suppose that Kant allows resistance because the conditions of one’s own legislating obedience to the sovereign are absent, then whether one is obligated to the state becomes a matter of individual judgment, which is in direct tension with the idea of the general will, under which I am free because I am subject only to my own individual will and to the general will which is mine. If my neighbor is obligated to respect my law-established rights only if he judges that he is in a rightful condition, then my rights depend on his judgment, and hence, I am unfree:

“A people cannot offer any resistance to the legislative head of a state which would be consistent with right, since a rightful condition is possible only by submission to its general legislative will... The reason a people has a duty to put up with even what is held to be an unbearable abuse of supreme authority is that its resistance to the highest legislation can never be regarded as other than contrary to law, and indeed as abolishing the entire legal constitution.” (*MM* 6:320)

This dilemma is a conflict between a purported duty and a purported right, rather than between two purported duties, but the problem is common to conflicts of duty. What is Antigone to do when she judges that the law proclaimed by the sovereign conflicts with her moral duty? If her duty to the state trumps moral duty, then internal freedom is

undermined. If her moral duty trumps her duty to the state, then external freedom is undermined.

This latter problem, the problem of subjecting others' rights to the individual's judgment about whether what the state demands is morally right, is an example of what Hegel is referring to when he suggests a tension between the absoluteness of duty and the subject's legislation of a consistent law,

“At the same time, because of the unity of the good, they [duties] *ought* to stand in harmony, and yet each of them, though it is a particular duty, is, as duty and as good, absolute. The subject *ought* to be the dialectic which *decisively concludes* a combination of them by excluding the others and thus sublating this absolute validity.”

In other words, what makes a duty both duty and good for Kant is that it is necessary. If it is not necessary, it isn't a duty at all, nor is it good. Duties must stand in harmony with each other, for two contradictory things cannot both be necessary. And since the subject's legislation is the source of duty, any apparent conflict of duty is to be resolved by the subject's own legislative power: What appeared to be a duty to the state, for example, is in fact not a duty because it conflicts with inner duty. Yet in “sublating this absolute validity,” the subject denies it the absolute status that makes it a duty at all, with devastating effects on the external freedom it constitutes. For if the law Creon declares is not valid absolutely, but only if Antigone judges it to conform to inner morality, then there is no public authority, but only the (sometimes) overlapping judgments of individuals, hence no general will and no rights established by it, and hence no freedom.

This argument for a dilemma between ethical duty and the sovereign's commands

can perhaps be generalized to other social demands. For family and work can make similar demands conflicting with similar duties. Hegel writes in *The Spirit of Christianity*, “In proportion as the multiplicity of human relationships grows, the mass of virtues also increases, and in consequence the mass of inevitable conflicts and the impossibility of fulfilment” (GC 294). Just as my external freedom in the state is predicated on obedience to the sovereign, so my freedom in more particular domains plausibly requires obedience to a kind of common will. Think here of the discussion of marriage from the chapter on Rousseau. In a marriage guided by a Rousseauvian shared will, each spouse gives up a degree of control over the outcome of the decision-making process because without equal willingness to give up such control, life together could not be governed by a shared will and would hence be incompatible with freedom.¹⁰³ Of course, the shared sense of the common good that guides their process might include Kant’s moral principle. But if they are in honest disagreement about what the moral principle asks of them,¹⁰⁴ they may need to compromise.¹⁰⁵ Or rather, they may need to choose between their duty to compromise and their apparent moral duty. Of course, they can decide to go their separate ways, severing the tie that creates the tension. But in some situations, it seems one would have a duty not to do so.

Hegel’s Response

¹⁰³ Assuming, that is, that there is no miraculously continuous unanimity

¹⁰⁴ Kant thought that morality’s demands were clear, but the denial of at least some confusion or disagreement over morality’s demands hardly seems plausible.

¹⁰⁵ This assumes, of course, that they are making a decision on which they must be united. Otherwise they can each act according to conscience.

We have seen two kinds of cases of alleged tensions moral dilemmas create for Kant's view. First, Kant's insistence that the individual has playroom for choosing between duties of virtue seems to conflict with the absoluteness of duty. Second, conflicts between duties of virtue and the demands of society (whether as the state or some other community or relationship) seem to put internal and external freedom at odds. If the former is given priority, then the absolute validity of duty is subordinated to the subject's judgment. If the latter is given priority, what remains of Kant's notion of freedom as autonomy? In each of these examples, the core problem is the same: First, our duties (or grounds of duty) may conflict with each other, and second, the absoluteness of duty conflicts with the subject's ability to resolve the first conflict. Since the subject's reason is, according to Kant, author of the law of duty, there is no other source to turn to for a resolution of the conflict.

Hegel's response to this problem in Kant's view consists of three parts. First, unlike Kant, Hegel accepts the existence of real moral dilemmas. Antigone may be right that she must bury her brother, but that does not detract from the fact that she must also not bury her brother. Antigone is bound both to bury her brother and not to bury him. And this means that Antigone and the state are caught in a contradiction; without a sublation of this contradiction, a development of the state that allows both Creon's authority and the law of the gods their due, Antigone and the state experience themselves as unfree. Antigone is not just unfree in the sense that her rightful external freedom is curtailed. She is also internally unfree: She cannot be subject to the internal moral law she has given

herself, for it demands two mutually incompatible things from her.

Second, Hegel affirms that it is the state and not the individual (qua individual)¹⁰⁶ that has the ability and the responsibility to resolve the contradiction, or in other words, it is the state and not the individual that has the power to shape duty. If the individual is to be free, whether internally or externally, the individual must belong to a state that harmonizes duties. This freedom may be a matter of degree: There are states with greater and lesser harmony of duties, and so with greater and lesser freedom. But there is not, in addition to this degree of freedom, an absolute internal freedom that the individual enjoys in spite of limited external freedom.

Third, although the individual caught in a moral dilemma to a degree unfree, the individual's unfreedom is in many cases not ultimately a tragedy. As we will see, Hegel thinks that rational state unfolds as a series of responses to contradictions within itself. The moral dilemmas that confront the individual may be expressions of a contradiction in the state that is a rationally necessary stage in the development of a fully rational state. The conflict between Antigone and Creon is emblematic of such a contradiction within the state: The rational state must honor the demands of both internal and external freedom. The result of this conflict will ultimately be a new kind of state capable of honoring these demands. Because the moral dilemma is rationally necessary, I can recognize it as what my own reason calls me to experience. Perhaps Hegel would call this a measure of freedom in unfreedom.

¹⁰⁶ This qualification is necessary because Hegel and Kant both think that the state's acts are acts of the individual in a sense that we have been and will continue to investigate throughout this dissertation.

Hegel traces Kant's problem to his insistence that duties ought to harmonize with each other and that the subject ought to harmonize them. The problem, in other words, is that morality is a mere ought. The sign of a genuine ethical order,¹⁰⁷ Hegel thinks, is that duties do harmonize with each other, such that the subject need not try to harmonize them, and therefore need not undermine their absolute validity. As we shall see in the fifth chapter, such a genuine moral order is not merely a utopia we hope for, but reason's destiny.

This issue of moral dilemmas also illustrates Hegel's thought that the individual will can only be rational in the context of a rational general will. If Hegel is right that the individual will cannot resolve moral dilemmas, then the individual will is consistent with itself only when it is in a context that resolves these dilemmas for it by making sure that duties do not conflict. We'll see later why Hegel thinks that a social context can resolve these dilemmas only if it is governed by a rational general will. If he's right, then an individual will is only rational (and hence free) when it is governed by a rational general will. Hence, the general will is not the united activity of independently existing free and rational individual wills. Free and rational individual wills presuppose a general will.

Hegel's Second Criticism: The Self-Legislated Universal Law

Hegel's second criticism examines what might be thought of as a particular kind of moral dilemma between duties to self and duties to others, but because Kant seems to think that the universality of duty will prevent this kind of dilemmas, Hegel frames it as a

¹⁰⁷ As we will see, Hegel uses the term "Ethical Life" [*Sittlichkeit*].

question of whether there is a special duty to the self, or whether all duties are universal. Hegel's claim will be that there is a special duty to the self, and that it can only be reconciled with the universality of duty if self and community are brought into the right relation.

I will divide Hegel's second criticism into four claims:

Claim 1: "To the subject, who in existence of his freedom essentially becomes a *particular*, his *interest* and *well-being* ought, for the sake of this existence¹⁰⁸ of his freedom, be an essential aim and therefore a duty."

Claim 2: "But at the same time in the aim of the *good*, which is not the particular but only the universal of the will, the particular interest *ought* not to be a moment."

Claim 3: "Because of this independence of the two determinations, it is likewise contingent whether they harmonize."

Claim 4: "But they *ought* to harmonize, because in general the subject, as individual and universal, is *in itself* one identity." (*Encyclopedia* §509)

Hegel has Kant in mind here, as we can see from his use of some of Kant's terminology.¹⁰⁹ Kant defines well-being [*Wohl*] in contradistinction from the good [*Gute*] in the *2nd Critique*:

"*Well-being* or *ill-being* always signifies only a reference to our state of *agreeableness* or *disagreeableness*, of gratification or pain, and if we desire or avoid an object on this account we do so only insofar as it is referred to our sensibility and to the feeling of

¹⁰⁸ Translation modified. See fn. 5.

¹⁰⁹ Of course, some of the terminology, such as "moment" and "particular," as well as Hegel's association of the "universal" with the good, is not Kant's. I will try to clarify these terms as I work through the dialectic.

pleasure or displeasure it causes. But *good* or *evil* always signifies a reference to the *will* insofar as it is determined by the *law of reason* to make something its object.” (*KpV* 5:60)

Well-being encompasses those ends that we are moved to adopt by a feeling of pleasure or displeasure, hence by sensibility, our receptive faculty. The good encompasses those ends that we are moved to adopt by the law of reason, our spontaneous faculty, independently of any sensible inclination. Since duty is that to which we are bound by the law of reason, the good encompasses those ends that we have a duty to pursue.

The point of Kant’s distinction is that well-being ought to be subordinate to the good:

“The human being... needs reason in order to take into consideration at all times his well-being and woe; but besides this he has it for a higher purpose: namely, not only to reflect upon what is good or evil in itself as well – about which only pure reason, not sensibly interested at all, can judge – but also to distinguish the latter appraisal altogether from the former and to make it the supreme condition of the former.” (*KpV* 5:62)

Already from this distinction and subordination it seems that Kant would want to reject Hegel’s first claim in the criticism, that well-being ought to be an essential aim and therefore a duty. The whole point of Kant’s distinction is that the ends of duty comprise the good, that the good is independent of well-being, and that well-being ought to be subordinated to the good. Hegel seems to be expressing something like this in Claim 2:

“But at the same time in the aim of the *good*, which is not the particular but only the universal of the will, the particular interest *ought* not to be a moment.”¹¹⁰ If pursuing

¹¹⁰ What exactly Hegel means by “moment” in this passage is not entirely clear, but Hegel’s apparent claim to find tension between Claim 1 and Claim 2 suggests that by “not a moment,” Hegel means “not an

one's own well-being were a duty, then it would likewise be a part of the good, and it would make no sense to talk of subordinating the one to the other.

But it's important to bear in mind that although the well-being/good distinction is a distinction between two kinds of ends, they are not distinguished by their content, but by their motivation. An end belongs to well-being insofar as it is one to which we are inclined and to the good insofar as it is one which we are obligated to pursue. A single end may therefore be part of both well-being and the good insofar as it is both agreeable and required by the law of reason. To subordinate well-being to the good simply means that we must subordinate sensible motivation to rational (moral) motivation, and consequently that we must subordinate ends that are not demanded by the law to those that are. Thus, there is no contradiction in saying that well-being is also good,¹¹¹ insofar as we may pursue it from a motive of duty.

In fact, Kant asserts that the pursuit of one's own well-being is a duty. Kant thinks that we have an indirect duty to secure our own happiness (*G* 4:399).¹¹² Happiness is

essential part of." I take it that Hegel's phrase "not the particular but only the universal of the will" is a shorthand for Kant's claim that what defines the good is not whatever interest the particular will has in an object, but the necessity of the object from a principle of reason. Since something is a necessary object of reason if and only if it is a universal object of reason, Hegel refers to the good as the universal of the will.

¹¹¹ There is a looseness of expression here, for Kant at one point claims that "good or evil is, strictly speaking, referred to actions, not to the person's state of feeling" (*KpV* 5:60). In this strict sense, well-being is not itself good, but an action the end of which is well-being may be good.

¹¹² In the *Metaphysics of Morals* Kant seems to take another approach: Pursuing our own happiness is not a duty because we naturally seek it (*MM* 6:386, *MM* 6:387). This suggests that it would otherwise be a duty. However, as Kant himself argues in the *Groundwork*, there are times when we fail to pursue our own happiness because we succumb to a particular inclination (*G* 4:399). Kant's argument that we can be tempted not to pursue our own happiness is very convincing, which leaves us a choice between concluding that Kant should accept that pursuing our happiness is in fact a duty or that the reason it is not a duty is other than he has stated. If we choose the former, then we are asserting that Kant should agree with Hegel's Claim 1.

But this interpretation does some violence to the text. In the very same passage in the *Metaphysics of*

complete well-being throughout our whole existence (*KpV* 5:61, *MM* 6:480), and therefore our indirect duty to secure our own happiness is an indirect duty to assure our own well-being. The duty is indirect because we must assure our happiness simply in order to perform direct duties: Happiness includes “means for the fulfillment of one’s duty” and lack of happiness includes “temptations to transgress one’s duty” (*KpV* 5:93). Securing the necessary means to a direct duty is an indirect duty. “However,” Kant asserts, “it can never be a direct duty to promote one’s happiness” (*KpV* 5:93).

Hence, if we read Claim 1 as asserting that securing one’s well-being ought to be an indirect duty, it doesn’t cause much trouble for Kant. For treating well-being as an object of indirect duty is perfectly compatible with a distinction that defines the good independently of well-being. However, Claim 1 seems to be asserting that one’s well-being ought to be a direct and not an indirect duty: “To the subject, who in existence of his freedom essentially becomes a *particular*, his *interest* and *well-being* ought, for the sake of this existence of his freedom, be an essential aim and therefore a duty.” This claim seems to be asserting a stronger, essential connection between the subject’s well-being and his freedom. The subject is essentially a free being, and the being of his freedom is therefore no more or less than his own being. His well-being is therefore the well-being of his freedom. Well-being is not merely a means to freedom or the removal of an obstacle to

Morals, Kant considers the suggestion (mentioned already) that we have a duty to pursue our own happiness to keep ourselves out of temptation and concludes that “then the end is not the subject’s happiness but his morality, and happiness is merely a means for removing obstacles to his morality” (*MM* 6:388). In other words, Kant continues to acknowledge that there may be circumstances in which we do not naturally pursue our own happiness and to insist that in those circumstances, happiness (and prosperity) is “not directly a duty, but indirectly it can well be a duty” (*MM* 6:388).

freedom. It is the subject's freedom actualized, and hence, Hegel suggests, a direct duty.

Although Kant asserts that securing one's own happiness is not a direct duty, we can perhaps see this strain in his thinking in his discussion of the duty not to kill oneself. In *The Metaphysics of Morals*, Kant bases this duty in the fact that one's existence is the existence of morality in the world:

“A human being cannot renounce his personality as long as he is a subject of duty, hence as long as he lives; and it is a contradiction that he should be authorized to withdraw from all obligation, that is, freely to act as if no authorization were needed for this action. To annihilate the subject of morality in one's own person is to root out the existence of morality itself from the world, as far as one can, even though morality is an end in itself. Consequently, disposing of oneself as a mere means to some discretionary end is debasing humanity in one's person (*homo noumenon*), to which the human being (*homo phaenomenon*) was nevertheless entrusted for preservation.” (MM 6:422-3)

There is a tempting misreading of Kant's assertion that to commit suicide is to “withdraw from all obligation” such that it is not really an argument for a duty to oneself. According to this interpretation, the wrongs committed in suicide are none other than violation of the various outstanding duties the suicide leaves unfulfilled. Obviously, I cannot perform my duties if I die before performing them. If suicide were permitted, then these duties would not be obligatory. They would become hypothetical imperatives: If you do not want to commit suicide, you must perform these duties. Hence, for duties to be categorical, suicide must be impermissible.¹¹³

¹¹³ There is a parallel here to the argument for the indirect duty to pursue one's own happiness: The argument enjoins or forbids one pursuit on the basis of its contribution to a further obligation. But there is also an important difference. Whereas unhappiness is merely a temptation to neglect obligations, suicide constitutes a neglect of obligations. And consequently, whereas the (indirect) duty to pursue our happiness is clearly subordinate to our other obligations, the duty not to commit suicide, as a necessary condition for

Still, this argument would not show the act of suicide to be wrong qua suicide, but only qua neglect of other duties. Kant clearly intends for his argument to show more, for by the time Kant gives the argument, he has already acknowledged and set aside the wrong that suicide constitutes qua neglect of duties in order to consider whether there is a duty to oneself not to commit suicide:

“But since what is in question here is only a violation of duty to oneself, the question is whether, if I set aside all those relations [to God, superior, and fellow citizens], a human being is still bound to preserve his life simply by virtue of his quality as a person and whether he must acknowledge in this a duty (and indeed a strict duty) to himself.” (*MM* 6:422)

Here the concern is not with the neglect of further obligations (perhaps to others), but with one’s duty to oneself, independent of any obligations to others. Kant’s commitment to such a duty is evident when he applies the argument for a duty not to commit suicide to the selling of one’s own teeth and to castration “in order to get an easier livelihood as a singer” (*MM* 6:423). Kant does not apply this duty to bodily organs by way of an argument that the particular organs are necessary to perform any further moral duty. Rather, Kant claims that such acts “are ways of partially murdering oneself.” Regardless of whether we think his claim plausible, the way that Kant derives these duties strongly implies that Kant does not conceive of our duties towards our body as indirect duties, duties to secure the necessary means to one’s performance of direct duty. Rather, he thinks that we owe our body a kind of respect because it has been entrusted to us as the

fulfilling any obligations we would leave at the moment of death, cannot be subordinated to those obligations.

existence of morality in the world.

Our lives and bodies are, according to this argument, not merely a means to the performance of duty, but the very subject of morality. In the *Groundwork* Kant argues for a duty not to commit suicide on the grounds that “A human being... is not a thing and hence not something that can be used *merely* as a means” (G 4:429). A human may use himself as a means only if he also treats himself as an end. Hence, our lives and our bodies do not get their value from their instrumental relation to a set of ends external to them, but as the essential homes of morality as an end in itself: “To annihilate the subject of morality in one’s own person is to root out the existence of morality itself from the world, as far as one can, even though morality is an end in itself.”

The phrase “as far as one can” [*so viel an ihm ist*], which could also be translated, “as far as it is up to one,” suggests that however suicide eradicates the existence of morality from the world, it is the most I can do toward that end. One might well wonder how this could be; surely I could take out a few dozen other people with me if I really stretch myself, and that would eliminate the subject of morality in many people and hence far more of the existence of morality than mere suicide would eliminate.

Kant’s thought, I take it, is that my own body is the subject of morality as it exists for me. In a discussion of the nature of duty from the section before, Kant argues that all of my obligations to others depend upon obligations to myself:

“For I can recognize that I am under obligation to others only insofar as I at the same time put myself under obligation, since the law by virtue of which I regard myself as being under obligation proceeds in every case from my own practical reason; and in being constrained by my own reason, I am also the one constraining

myself.” (MM 6:417-8)

In other words, morality exists for me only because I give it to myself. I am the essential author of the moral law insofar as it binds me, and I am the unique subject of the moral legislation of my own will. Morality is, for me, the law that I legislate, not the law that you legislate. I am the unique existence of morality insofar as I legislate it. In committing suicide, I root out that existence of morality from the world. It would be wrong to kill others, but the wrong would be very different, for in their case I would not be assaulting the existence of morality itself as it binds me.

Morality is the law we give ourselves and hence our freedom. So rooting out the existence of morality itself from the world is tantamount to rooting out the existence of our freedom. This language recalls that of Hegel’s Claim 1: “To the subject, who in existence of his freedom essentially becomes a *particular*, his *interest* and *well-being* ought, for the sake of this existence of his freedom, be an essential aim and therefore a duty.” Still, the well-being that Kant says we have a duty to secure seems to be our health and wholeness, whereas Hegel refers to the subject’s “interest and well-being,” which seems more in line with Kant’s use of “well-being” [*Wohl*] to refer to pleasure.

A further argument against suicide will show that Kant is committed to a connection between welfare (health and wholeness) and well-being. Kant claims that one way to see the wrongfulness of suicide from feeling “sick of life” is to see that the destination of our sensible interests is “to impel toward the furtherance of life,” and therefore a law by which they would lead us to suicide would contradict itself (G 4:422). Setting aside how the argument is supposed to work, the claim that our sensible interests are for impelling us

toward the furtherance of life connects welfare with well-being: The purpose of our inclination towards the pleasant is to impel us towards our welfare. In other words, the object of our well-being just is our welfare. Of course, well-being does not always align with welfare: We take pleasure in things that are not good for us. But very often these are in fact cases in which our immediate well-being is not aligned with our long-term well-being. Hence, the need for maxims of happiness that unify our well-being into a whole. Note in this connection that in speaking of the subject's interest alongside his well-being, Hegel suggests a rational ordering of inclinations like that which Kant refers to as happiness.

In sum, Kant seems committed to precisely the kind of duty Hegel suggests in Claim 1, for he is committed to a duty to preserve one's own life not merely as a means to performing duty, but as the existence of one's own freedom; the object of our well-being, properly shaped by reason, just is this self-preservation. Of course, Kant does not think that we have a duty to secure our well-being because we find it pleasant. We have a duty to secure our well-being regardless of the feeling it provides us. And this is just what Hegel seems to be suggesting in Claim 1. When Kant says that there is no direct duty to secure one's happiness, he must have in mind that there is no direct duty to do so because of feeling itself, for there is a direct duty to secure the kind of life that makes us happy.

Universality and the Special Duties to Self It Can Accommodate

Hegel is suggesting that this duty to secure one's own well-being (the kind of life that makes us happy) is in tension with the universality of the law, in which one's well-being

does not get special consideration: “But at the same time in the aim of the *good*, which is not the particular but only the universal of the will, the particular interest *ought* not to be a moment.” In other words, the moral law subjects maxims of the will to a universality condition so as to align all wills towards something that all can think of as good. As we saw in the previous chapter, the point of universalization, for Kant as for Rousseau, is to make it possible for us to agree on the good. Kant makes this clear in a response to the proposal that the desire for happiness is the universal practical law:

“Whereas elsewhere a universal law of nature makes everything harmonious, here, if one wanted to give the maxim the universality of a law, the most extreme opposite of harmony would follow, the worst conflict, and the complete annihilation of the maxim itself and its purpose. For then the will of all has not one and the same object but each has his own (his own welfare [*Wohlbefinden*])... In this way there results a harmony like that which a certain satirical poem depicts in the unanimity between a married couple bent on going to ruin: ‘*O marvellous harmony, what he wants she wants too*’ and so forth, or like what is said of the pledge of King Francis I to the Emperor Charles V: ‘What my brother Charles would have (Milan), that I would also have.’”
(*KpV* 5:28)

Under properly universalized maxims, the objects of people’s wills harmonize in the sense that they come to have the same ends. The problem with Francis I and Charles V is that they don’t actually want the same end. The moral law is supposed to ensure that we can agree on the good without equivocation so that there is no conflict among our actions. We achieve this agreement by constructing the good in accordance with universal law. If my happiness is to be acknowledged good, then others’ must be too. When Hegel says that on Kant’s view my particular interests ought not to be a moment in the good, he means that my particular interests do not count as good simply because they matter to me, but only

because they can be made part of a system of universal law, that is, only insofar as they lose their particularity and become universal. The good and duty come into focus when we set aside the particularity of our interests and think of them in universal terms.

The same thesis about universality and the harmony of ends is illustrated in Kant's notion of a kingdom of ends and in the *Doctrine of Virtue*, in which Kant begins his categorization of duties of virtue with something like the elements of the highest good, perfection (only somewhat broader than virtue¹¹⁴) and happiness, and then argues for the exclusion of my own happiness (because I pursue it naturally) and others' perfection (because I can't pursue it). In other words, Kant seems to think that when considering one's moral ends, the appropriate place to start is from a universal perspective: the complete good. The ends of the complete good are all ends that everyone must endorse, but they are not all obligatory for me because some are impossible for me and some are naturally necessary for me independently of any obligation.

Hegel's accusation was that Kant's theory cannot accommodate a special duty to secure one's own well-being. To see what he has in mind, we must first set aside some kinds of special duties that Kant's commitment to universal law can accommodate. I may have a special duty to myself when the object of my duty is something that others cannot achieve for me, namely, my virtue. But this is not the ground of the special duty to secure one's own well-being that Hegel has in mind, for in many circumstances, others can secure my well-being.

¹¹⁴ Perfection is the ability "to set [one's] end in accordance with [one's] own concepts of duty" (*MM* 6:386).

Even when others can secure my well-being, there may be grounds for denying any obligation on their part to do so, for example, when to do so would require that they infringe upon my freedom. Kant sometimes talks about our duty with regard to others' happiness as a duty not to interfere with their pursuit of happiness rather than a duty to pursue others' happiness. And there may be people whose happiness we are obligated to pursue in ways that we are not obligated to pursue others', such as our friends and family.

We do not need to consider all such cases of special duties, however, to see a feature of them built into the idea of harmonious universal law. Even when I am not required to pursue another's well-being, I must still be able to see another's well-being as good, and hence, as something that must not be infringed against. The good must be the subject of universal agreement: We must be able to agree that it is good that you pursue your well-being and correlatively that it is bad for others to interfere in that pursuit. Hence, wherever it is possible for me to interfere with your duty to yourself, I have an obligation not to do so. Hegel's thought, I take it, is that Kant's morality is at once committed to this feature of special duties and in contradiction with it.

When Well-Being Conflicts with Universality

The *Metaphysics of Morals* argument regarding suicide yields a special obligation towards my own life, as opposed to life in general. Of course, I do have a duty not to kill others or let them die when I can help it, but that duty cannot have the same justification, for to kill another would not be for me to "withdraw from all obligation" with "no authorization" or "to root out the existence of morality itself from the world, as far as [I]

can, even though morality is an end in itself.”

As we just saw, insofar as they are governed by universal law, special duties to oneself are accompanied by others’ duties not to interfere.¹¹⁵ Hence, insofar as this special duty to myself is governed by universal law, there is a correlative duty on the part of others not to interfere with my preservation of my own life.

Of course, these duties of self-preservation and non-interference with another’s self-preservation can come into conflict with each other. Suppose we’ve been shipwrecked and there’s only one lifejacket. Let us set aside the question of our taking the lifejacket from each other once one of us has it and consider only the question of who is to grab it in the first place. How should I proceed? My duty of self-preservation might suggest that I ought to grab it, but my duty not to interfere with your self-preservation suggests that I ought not. In spite of my special duty to myself, the universality of law denies me license to give myself special treatment. Whatever the law requires, its universality requires that we treat each other as equals in standing. Your self-preservation is as good as mine.

Hegel’s claim is that this insistence on universality is in tension with the argument Kant gives for the duty of self-preservation. Notice that Kant’s argument for the duty of self-preservation does not seem to invoke universality the way that many of his arguments for other moral duties does. A maxim lacks universal form if I couldn’t will that it be a universal law, binding on all wills. But what’s wrong with maxims of suicide is not that they could not bind *all* wills. Rather, maxims of suicide seem to yield contradictions

¹¹⁵ At least if it is possible to interfere

independently of whether they are considered as universal laws. In the argument in the *Metaphysics of Morals*, the problem is a failure to treat morality as an end-in-itself. In the argument in the *Groundwork*, the problem seems to be a contradiction between the natural end of my self-love and the end to which my maxim puts it. Although Kant mentions universal law, the universalization of the maxim is not required to generate the contradiction; the maxim is already in contradiction with the natural purpose before the maxim has been universalized.¹¹⁶

Hence, Kant gives an argument for a duty of self-preservation that is independent of any considerations of universality, and precisely because it is independent of considerations of universality, there is no obvious way to reconcile it with the demands of universality. When a typical maxim fails a test of universality, that simply shows us that the maxim is impermissible; universality is the very standard for determining permissibility. But in the case of maxims of self-preservation, Kant seems to have introduced an independent standard of obligatoriness (which implies permissibility) rooted not in the universality of law, but in the existence of the subject of lawgiving. When a maxim of universal self-preservation is impossible, we cannot simply say that in this case, maxims of self-preservation are not permissible, for we have independent

¹¹⁶ Interestingly, Kant doesn't stop by pointing out the contradiction, but suggests that the contradiction is a threat to the subsistence of the subject: "A nature whose law it would be to destroy life itself by means of the same feeling whose destination is to impel toward the furtherance of life would contradict itself and would therefore not subsist as nature" (*G* 4:422). Interpreters have puzzled over why we should be committed to the natural purpose or destination of self-love. I suggest that Kant's real concern is with our nature's subsistence, which, as we've seen in the *Doctrine of Virtue*, is the subject of morality in the world. Suicide is in conflict with universal law not because it is not universalizable but because it is incompatible with the existence of the legislator and subject of universal law.

grounds for asserting that it is not only permissible, but a duty.

There are a number of ways for Kant to respond to this charge. Kant might assert that in this case “the stronger *ground of obligation* prevails” (*MM* 6:224). But how could the duty treating everyone’s life as equally deserving of preservation have stronger grounds than duty of self-preservation if the latter is the ground of the former? That is, the reason that I have a duty not to interfere with your self-preservation is that it is a correlate of your duty to preserve your own life. If we deny the latter, the ground of the former vanishes with it. Moreover, why should the requirement that the maxim have universal form take precedence over the ground of that requirement: the existence of morality in my person? A morality that issues an imperative without universal form ceases to be morality. But a morality that issues an imperative that requires the death of the lawgiver ceases to be.

Kant might point out that making it my end to preserve my life and making it my end to preserve your life are both wide duties, and hence “the law cannot specify precisely in what way one is to act and how much duty one is to do by the action for an end that is also a duty” (*MM* 6:390). Or he might suggest that we resolve the issue by formulating more precise maxims. After all, the duty of self-preservation Kant seems to have in mind is opposed to “disposing of oneself as a mere means to some discretionary end” (*MM* 6:423). The shipwreck case is not about discretionary ends.

Neither the attempt to make the duty less precise (by characterizing it as wide) nor to make it more precise (by considering a more specific maxim) seems to accord with the

derivation of the duty itself. The argument is that to fail to preserve my life is to root out the existence of morality itself from the world, as far as I can (or to allow it to be rooted out). A morality that allowed such a maxim would be in contradiction with its very existence. This problem is not solved by my merely having an end of preserving my life with playroom for when I may subordinate it to other ends. In permitting me to subordinate my existence to other purportedly moral ends, the moral law is permitting itself to come to an end. It is not undermining its essence, universality, but it is undermining its existence, and is thus in contradiction with itself. Nor, for the same reason, is the problem solved by specifying other ends to which my existence may be subordinated.

It is no use to point out that morality will continue to exist in the will of the person I save. Although morality in me requires of me the same exact things that morality in you requires of you, my morality and your morality are separate acts of legislation. I legislate the moral law to myself, and you legislate the moral law to yourself. Sacrificing morality's existence in me to save morality's existence in you would make morality in me a mere means to the end of preserving morality in you.

The condition of universality is the condition that is supposed to enable us to agree on laws. But as Rousseau pointed out, the universality condition is only necessary and not sufficient for us to agree; there must also be an overlap of interests:

“For while the opposition of particular interests made the establishment of societies necessary, it is the agreement of these same interests that made it possible. What these different interests have in common is what forms the social bond, and if there were not some point on which all interests agree, no society could exist.” (*SC* 2.1.1)

When there is no common interest between us, there can be no agreement on a law. I cannot regard your neglecting my life to save your own as good, not only because I desire my own happiness, but because I have a duty to preserve my own life.

This special duty to the self is just what Kant meant to avoid by appealing to universal law: The CI ought to harmonize duties such that we can agree on the good. Kant tries to steer away from a special duty to the self that could conflict with universal duties every time he introduces a duty to preserve one's life. He specifies the wrong as "disposing of oneself as a mere means to some discretionary end" in the *Metaphysics of Morals* and "mak[ing] use of a person *merely as a means* to maintain a tolerable condition up to the end of life" in the *Groundwork*. But the argument he gives in the *Metaphysics of Morals* seems to generate a much more robust duty than the one Kant acknowledges.

Hegel's thought in Claim 1, I take it, is that autonomy requires both the universal form of the maxim adopted and the existence of a will to adopt it. That is, autonomy depends not just on my giving myself a *law*, but on *my giving myself* a law. A contradiction may be generated either by the failure of the maxim to have universal form or by the failure of the maxim to accord with the existence of the act of moral legislation. A law that puts my life at risk, thereby conflicting with all of my other ends, moral and non-moral, and with the very process of legislation we are imagining, conflicts with my autonomy even before we consider its universalization.

A tension arises, Hegel asserts, because the existence of the self is independent of the universality of law: "Because of this independence of the two determinations, it is

likewise contingent whether they harmonize” (Claim 3). Because there is no guarantee that what is good from a universal perspective is also good for my existence as a particular and vice versa, there can be a conflict between the two.

It will not do to subordinate the claim of universality to the claim of existence nor the claim of existence to the claim of universality. The end-in-itself is precisely a(n existing) will under universal law: “But they [particular well-being and universal good] *ought* to harmonize, because in general the subject, as individual and universal, is *in itself* one identity” (Claim 4).

Only if this conflict between the universality of law and the existence of the free individual can be avoided can the individual’s reason avoid contradiction. And only if reason avoids contradiction is it what it most is, namely, reason. This conflict can only¹¹⁷ be avoided if, as Rousseau points out, there is an overlap of interests, a common good that keeps the demands of universality and the demands of my existence consonant. Hence, my reason is most what it is when it exists in the context of a common good. The common good is constitutive of my reason. Again we see that the free and rational individual will exist only in the context of a general will. Hence, the individual will is essentially a part of the general will whose activity cannot be reduced to the unanimous activity of each of its members.

Although an overlap of interests may come about by chance, mere chance cannot be the constitutive ground of a lasting freedom. Sustained overlap of interests comes about

¹¹⁷ Technically living in complete isolation from others is another way of avoiding the conflict, but for reasons we’ll see, Hegel thinks that one cannot achieve freedom in complete isolation from others.

as a result of a well-designed society. Kant's Kingdom of Ends was just such a society, but for Kant, the Kingdom of Ends was an end to strive for, a mere ought, and not an actuality. Hence, it could not serve as the solution to the tension of which Hegel accuses Kant's morality. Hegel's proposed solution is to show that a different sort of Kingdom of Ends is not merely an ought, but actual.

Conclusion

We have seen two arguments against Kant's morality. First, Kant denies that there are moral dilemmas, for at most one action can be necessary, and the individual's reason is sufficient to determine this. Against this claim, Hegel asserts that allowing the individual's reason to resolve seeming moral dilemmas undermines the absolute validity of duty. The answer is to deny morality the status of a mere ought: Whatever morality ought to be like, when harmonious conditions are lacking, morality is not in fact consistent with itself.

Second, Kant insists on a standard of universal law that is supposed to allow us a common perspective, but as legislators and subjects of that law, we are individuals, and hence a concern with our own well-being is fundamental to the activity of binding ourselves to the law. This generates obligations that may come in tension with the obligations generated by the universality of law. Such a tension would be avoided if the moral law were not merely a law of how people ought to behave, but of how they do in fact behave, for then the universality of law would ensure the universal welfare of its subjects.

Kant's general will is a noumenal reality we are justified in assuming because of a fact of reason, a mere ought. In attacking the coherence of a moral system built around morality as a mere ought, Hegel is attacking the justification for assuming Kant's general will. In the next chapter, we will see a third attack on this justification rooted in a tension Kant himself identified as the Antinomy of Practical Reason.

Chapter 4

Hegel's Criticisms of Kant's Morality Part 2: The Antinomy of Practical Reason

O Duty,

Why hast thou not the visage of a sweetie or a cutie?

Why glitter thy spectacles so ominously?

Why art thou clad so abominously?

Why art thou so different from Venus

And why do thou and I have so few interests mutually in

common between us?

Ogden Nash, "Kind of an Ode to Duty"

In the second part of last chapter we considered duties regarding our well-being in Kant's moral philosophy. On my interpretation, Hegel alleges that the duty to preserve freedom in its existence as an individual can conflict with the form of law, universality. This chapter we will consider the significance of well-being apart from being an object of duty; we will consider well-being as happiness. Unlike the alleged problems of last chapter, this problem was the subject of considerable commentary by Kant himself. He referred to it as the Antinomy of Practical Reason. Since many Kantians have found Kant's preoccupation with this problem puzzling, it will help to examine Kant's arguments for the Antinomy in detail before considering Hegel's response.

The Antinomy of Practical Reason presents a problem about how to reconcile the moral (and hence rational) aim of virtue with our sensible interest in our own happiness. Since many Kant interpreters have not been moved by the problem the Antinomy raises, I will argue that the Antinomy, if unresolved, constitutes a serious (though perhaps not decisive) problem for Kant's view. The Antinomy of Practical Reason is again the expression of a worry about morality's merely being an ought: People ought to be virtuous, and virtue ought to be rewarded with happiness,¹¹⁸ but on Kant's account, there is no guarantee of this. Kant himself viewed this as a serious problem. In the wake of this objection, I offer Hegel's view as an attractive alternative to Kant's that avoids the Antinomy: The universal form of law is not the only condition for our affirming a purported imperative as a duty. Its ability to provide for our well-being is another condition.

The Antinomy of Practical Reason

Hegel's fourth criticism takes up Kant's Antinomy of Practical Reason:

"External objectivity, in accordance likewise with the distinction that has emerged between it and the subject will (§503), constitutes the other extreme, independent of the internal determinations of the will, a peculiar world for itself. It is therefore contingent whether it harmonizes with the subjective aims, whether the *good* realizes itself in it, while *evil*, the aim that is in and for itself null, is null in it;—contingent too whether the subject finds in it his well-being, and more precisely whether in it the *good* subject becomes *happy* and the *evil* subject *unhappy*. But at the same time the

¹¹⁸ In speaking of what ought to happen, Hegel may be borrowing from Kant's 1st *Critique* treatment of the antinomy, in which he characterizes hope as "the inference that something **is** (which determines the ultimate final end) **because something ought to happen**" (A806/B834). Kant goes on to explain that virtue would be proportioned to happiness if everyone obeyed the moral law because their own behavior would bring about the corresponding happiness.

world *ought* to let what is essential, the good action, be carried out in it, as it *ought* to grant the *good* subject the satisfaction of his particular interest, but refuse it to the *evil* subject, just as it *ought* to nullify evil itself.” (EG §509)

Hegel’s claim that the good subject ought to become happy alludes to Kant’s discussion of the highest good.¹¹⁹ Kant himself thought that the fact that the virtuous are not happy in this world leads to an Antinomy of Practical Reason that can only be solved if reason postulates the immortality of the soul and a God to reward the virtuous with happiness.

Many Kant interpreters have not been moved by the problem the Antinomy raises: Commentators have cast doubt on the derivation of the highest good, on the alleged duty to adopt it as our end, and on the need for the postulates for us to be able to do so.

Although the primary purpose of this chapter is not to sort through all of these problems, we will best be able to appreciate the value of Hegel’s solution to the Antinomy if we are able to appreciate the urgency of the problem. Hence, I offer an interpretation of Kant’s Antinomy that seeks to answer some interpretive puzzles and argue that the Antinomy, if unresolved, constitutes a serious (though perhaps not decisive) problem for Kant’s view.

Kant proposes his Postulates of Practical Reason as a solution to the Antinomy. Although Hegel offers no argument in the *Encyclopedia* against Kant’s Postulates, the context of Hegel’s argument, leading into the section on Ethical Life, his own moral and political view, suggests that Ethical Life is his own solution to the problem raised in the Antinomy.

To conclude this chapter, I sketch out some features of Ethical Life that constitute it as a

¹¹⁹ Kant uses this sort of language in describing the highest good: “[Hope] comes down to the inference that something **is** (which determines the ultimate final end) **because something ought to happen**” (A806/B834). See also A808/B836.

solution to the Antinomy.

The highest good is a demand of practical reason. Reason is the faculty of principles, the faculty that seeks the unconditioned. Practical reason seeks the unconditioned condition of all willing. Willing always takes place in accordance with a principle that has both matter (its end) and form. A free will could never be determined by the matter of a principle, for a free will is independent of empirical conditions (*KpV* 5:29). Hence, if there is to be an unconditioned condition of all willing for such a being, it would have to be in the form of the principle that is binding on all free wills. Hence, the unconditioned condition of all willing (the Categorical Imperative) abstracts from all ends.

Nevertheless, practical reason does not dispense with ends entirely, for as finite practical reasoners, humans never will without an end:

“In the absence of all reference to an end no determination of the will can take place in human beings at all, since no such determination can occur without an effect, and its representation, though not as the determining ground of the power of choice nor as an end that comes first in intention, must nonetheless be admissible as the consequence of that power’s determination to an end through the law.” (*R* 6:4)

Although the moral law, and not the end, should determine the will, every determination of the will has an effect, and this effect must be represented by the will for the action to take place: I cannot act without representing to myself what my act should look like. This representation does not itself determine the will (I am determined to this act by the moral law and not by my inclination to perform the act). Nevertheless, the act represented is itself an end: It is that which I take myself to be doing.¹²⁰

¹²⁰ In a footnote Kant explains that the proposition that human beings must “consider in every action, besides the law, also an end” is itself possible “only because it contains the *a priori* principle of the cognition

The end of moral action is action in accordance with duty from a motive of duty. As Kant explains in the section on the “Typic of Pure Practical Judgment,” reason constructs the shape of action in accordance with duty by borrowing the concept of natural law from the understanding.¹²¹ The shape of action in accordance with duty varies according to circumstances, and so the ends of moral action appear to be diverse.

The need for a single unconditional good that unites the diverse ends of the moral law is a rational need for a certain kind of unity in the will’s activity. Reason, as the faculty of principles, is not content with a plurality but seeks always a unity.¹²² In discussing reason’s logical use, Kant writes: “Reason, in inferring, seeks to bring the greatest manifold of cognition of the understanding to the smallest number of principles (universal conditions), and thereby to effect the highest unity of the manifold” (A305/B361).¹²³ In its practical use, reason seeks to bring our cognition of the good to unity.

of the determining grounds of a power of free choice in experience in general, so far as experience, by exhibiting the effects of morality in its ends, gives an objective, although only practical, reality to the concept of morality as having causality in the world” (*R* 6:7 fn.). Unless Kant has reversed his position on the possibility of knowing our motives to be virtuous, Kant must mean by this not that we have practical cognition of morality’s actually causing action, but rather that we have practical cognition of morality’s possibly causing action (i.e., the fact of reason (*KpV* 5:31-2)). The thought is that to think of morality as a possible cause of action is to think of it as having an object (action), i.e., an end.

¹²¹ *KpV* 5:67-70

¹²² Note that this point is separate from the earlier claim that we must have an end in every action. It is compatible with that earlier claim that all of our actions have a different end. The need for our various ends to unite into one end is a demand of reason akin to its demands for unity in scientific principles. So far, no unity of all moral ends in virtue nor unity of moral ends with natural ends is supposed. These two kinds of unity are not necessary for moral action, but only for reason to satisfy its demand that our ends be united under one material principle.

¹²³ It is this search for unity that can lead us into illusion, for the unity sought is not a possible object of experience. Hence, in reason’s theoretical use, our cognition of the sought-for unity is limited by what is (and can be) given in experience. In reason’s practical use, however, reason need not limit itself to possible experience.

What reason does in constructing the highest good is not unlike what it does in uniting various needs into an end of happiness:

“Happiness...is appraised, as reason especially requires, not in terms of transitory feeling but of the influence this contingency has on our whole existence and our satisfaction with it... The human being is a being with needs, insofar as he belongs to the sensible world, and to this extent his reason certainly has a commission from the side of his sensibility which it cannot refuse, to attend to its interest and to form practical maxims with a view to happiness.” (*KpV* 5:61)

Reason’s aim in deriving maxims with a view to happiness is to maximize the satisfaction of our needs. The aim of maximizing the satisfaction of our needs, of providing us with satisfaction with our whole existence, is an aim unifying the needs, without which the various needs could not be said to constitute one end of happiness. Kant thinks that non-rational animals function without the unifying end of happiness, following one inclination after another without thought for satisfaction with the whole of their existence (*KpV* 5:61-2).

Just as reason unites our sensible needs into the end of happiness, reason unites the ends required of us by the moral law into the end of virtue. Kant describes virtue as a “capacity... to overcome all opposing sensible impulses” to the moral law (*MS* 6:397). It seems more accurate to say that the activity of this capacity, moral action, is the unconditional end, and indeed, Kant uses the word “virtue” to describe the activity, as when he suggests that we develop the capacity of virtue “by *practicing* virtue” (*MS* 6:397). Activities, not capacities, are practiced. I take it that part of what makes virtue one unified unconditional end is that it is the activity of one capacity to obey the moral law

from the right motive. The fact that this capacity can be strengthened by practice lends further unity to virtue. The actualization of our capacity for virtue strengthens the capacity, making virtue a sort of continual progress: “The utmost that finite practical reason can effect is to make sure of this unending progress of one’s maxims toward this model [holiness of will] and of their constancy in continual progress, that is, virtue” (*KpV* 5:32-3). This progress is our end not in the sense that it is our motivation for acting, for the moral motivation is always respect for the law. But it is our unconditional end in the sense that it is the effect of all our willing insofar as it is in accordance with the moral law.

Hence, if the categorical imperative is the unconditioned formal principle of willing, the unconditional good is the unconditioned condition of the matter of willing. For although the matter of willing is not what determines the free will to act, reason nevertheless seeks the matter’s unconditioned condition: “an end proceeds from morality just the same; for it cannot possibly be a matter of indifference to reason how to answer the question, *What is then the result of this right conduct of ours?*” (*R* 6:4). In other words, although we can be motivated to obey the moral law simply out of respect for it, reason is interested in what sort of thing we are to accomplish by this obedience. Indeed, without being able to unite the demands of morality into an object of some sort, Kant suggests, we would experience a certain rational dissatisfaction: “without this end, a power of choice ... can itself obtain no satisfaction” (*R* 6:4). We would appear to ourselves in one respect like the non-rational animals, following one command after another without any thought for the whole. Of course, in pursuing virtue, we are not to be motivated by virtue

as an end (the matter of the will's principle), but only by respect for the law.

Even after reason has united the necessary ends of the moral law under the heading of virtue, there is still the question of uniting virtue and happiness, an end we necessarily have as finite rational beings (*KpV* 5:25). Our complete end, the complete or highest good,¹²⁴ includes both virtue and happiness conditioned on virtue. Virtue does not contain happiness within it: The happiness of the virtuous is necessarily good, but the virtuous are not necessarily happy. Hence, our complete end is not contained in virtue, the unconditioned condition.

Only when all our ends are unified under a single concept of the highest good encompassing both our moral and natural ends can we see in all of our activities a rational unity. The highest good provides “a special point of reference for the unification of all ends...; only in this way can an objective practical reality be given to the combination, which we simply cannot do without, of the purposiveness [deriving] from freedom and the purposiveness deriving from nature” (*R* 6:5). Kant speaks of the need to make a unified highest good our end as “a natural need... to think for all our doings and nondoings taken as a whole some sort of ultimate end which reason can justify” (*R* 6:5). The need is natural because it belongs to our finite nature always to think an end of our will and rational because reason, as faculty of principles, seeks unity in this end. But the unity does not come naturally, for “the maxims of virtue and those of one's own happiness are quite heterogeneous with respect to their supreme practical principle...

¹²⁴ Kant allows that the term “highest good” may refer to the unconditional good or the complete good that includes both virtue and happiness, but he almost always uses the term to refer to the complete good.

They are so far from coinciding that they greatly restrict and infringe upon each other in the same subject” (*KpV* 5:112). A relationship we have worked hard to foster is destroyed when we obey morality’s demand to tell the truth. A fortune we would gladly have won is rejected because it could not be won in accordance with the moral law. We seem to be without a unified end, an object that would justify all of our activities. Our reason cannot be satisfied with these diverse goods unless it can unite them under some final end.

According to Kant, uniting virtue and happiness conditioned on virtue under the highest good is not simply a matter of adding the two together with the word “and.”¹²⁵ Here there is a disanalogy with reason’s role in constructing maxims of happiness. Reason constructs maxims of happiness as a maximum of satisfaction with one’s life. None of the elements of happiness need to bear a necessary relationship to the others because reason is simply performing the work of unification on a posteriori sensible materials.¹²⁶ But virtue and happiness are combined *a priori* in the idea of the highest good (*KpV* 5:113),¹²⁷

¹²⁵ What Reath refers to as “the secular conception of the highest good” combines virtue and happiness conditioned on virtue with something like a mere “and” (cf. 613). The secular conception ensures that happiness is consistent with virtue by limiting the former by the latter, but ensuring consistency between two ends is not the same as unifying them, i.e., making them one end. Thus, the secular conception fails to discover the necessity of their being combined in one concept (*KpV* 5:111). In fairness to Reath, he may have something like Stephen Engstrom’s claim that the highest good is “a hylomorphically constituted whole of virtue and happiness” in mind (Engstrom 2015, p. 153, cf. p. 138). On this conception of the highest good, virtue is the form and happiness the matter of the highest good, and hence they form one synthetic unity. But this hylomorphic unity does not seem to be the unity that Kant has in mind in this context, for he does not consider it among the four possible kinds of unity I will shortly discuss. And rightly so, for hylomorphic unity would unite form with matter, but what Kant is trying to unite in this context are two kinds of matter: virtue (the end necessitated by the formal law) and happiness (the end necessitated by sensible nature). Engstrom gives a more careful analysis on which he distinguishes between virtue as end and the moral principle as form in Engstrom 2016.

¹²⁶ It strikes me that even here, Kant might want more than a set of ends connected by the word “and,” for our satisfaction with life is not just a function the satisfaction of various ends, but of our satisfaction with the way those ends fit together. This shape of a life is not just one end among the others, for it has the other ends as its objects.

¹²⁷ This may seem odd, since it says that the content of our happiness is only known to us a posteriori. We must distinguish between our need for happiness and the content of our happiness. Our need for happiness

and so they must be related necessarily.

What is related necessarily is related either analytically or synthetically as ground and consequent:

“Two determinations *necessarily* combined in one concept must be connected as ground and consequent, and so connected that this *unity* is considered either as *analytic* (logical connection) or as synthetic (real *connection*), the former in accordance with the law of identity, the latter in accordance with the law of causality” (*KpV* 5:111).¹²⁸

The choice of virtue as either ground or consequent in either an analytic or synthetic unity yields four possible relations:

Analytic:

1. Virtue is consciousness that one’s maxim leads to happiness. (Epicureanism)
2. Happiness is consciousness of one’s virtue. (Stoicism)

Synthetic:

3. The desire for happiness is the motive to (cause of) maxims of virtue.
4. The maxim of virtue is the efficient cause of happiness. (*KpV* 5:113)¹²⁹

belongs to our nature as finite beings and is therefore a priori (*KpV* 5:25). But we only know a posteriori what will satisfy this need.

¹²⁸ Sussman complains about the move from the general synthetic relation of ground and consequent to the specific relation of efficient cause and effect (p. 219). Kant is quick here, but the thought is that the relation between the two must be sufficient to unite them into one end that I may count myself as pursuing in all of my action. Sussman’s proposal, that intelligible virtue serves as the ground of happiness in a purely normative sense, will not work, for that would only unite virtue and the goodness of happiness in one end, when what we wanted to unite were virtue and happiness itself.

¹²⁹ Sussman complains that (3) and (4) are not parallel (pp. 218-9). Why “the desire for happiness” in (3) and simply “happiness” in (4)? The reason is that these are each proposals about our ultimate end. Insofar as our ultimate end is happiness and virtue caused by it, happiness is the first aim of our will, and virtue its necessary consequence. If happiness is our ultimate end, then our motive will be the pursuit of happiness. If virtue is our ultimate end, then our motive will be a maxim of virtue. Sussman’s alternative proposal that

These all seem to be non-starters. The relations of analytic unity are incompatible with the crucial distinction between determination of the will by its matter and by the form of law. Kant's argument for the categorical imperative is diametrically opposed to the third, which make happiness the ground and virtue the consequent. And the fourth is not true because the happiness one in fact achieves in life depends not on the moral disposition of one's will but on "knowledge of the laws of nature and the physical ability to use them to one's purposes" (*KpV* 5:113).¹³⁰

This is the Antinomy of Practical Reason: Practical reason (and hence the moral law) demands that we make the highest good our end, but the highest good is apparently impossible. We cannot make something that we know to be impossible our end. Since ought implies can, the impossibility of making the highest good our end means that we are not in fact obligated to make the highest good our end. But since the obligation to do so was derived from pure practical reason, and hence, Kant asserts, from the moral law,¹³¹ the moral law itself appears to be false:

"Now, since the promotion of the highest good, which contains this connection

happiness itself may be the cause of maxims of virtue fails to unify the two ends in the desired way. For what we desired was a single end uniting the two such that in all of our actions we can think of ourselves as pursuing one thing. If happiness is that end, then the desire for happiness is always our primary motivation, whether we are pursuing happiness or virtue (or both). And if happiness is always our primary motivation, then virtue is impossible.

¹³⁰ Pauline Kleingeld and Stephen Engstrom each claim that virtue and happiness are related in the common good as follows: Since the virtue of all is the unconditional good and "universal happiness follows as the collective effect of universal virtue," the complete good includes both (Engstrom 2016, p. 106, cf. Kleingeld p. 40). Something like this line of thinking is suggested by Kant's 1st *Critique* discussion of the highest good, but it seems to me to miss the point of the 2nd *Critique* arguments. In the 2nd *Critique*, Kant arrives at the inclusion of happiness in the highest good prior to considering any causal relationship between virtue and happiness and then explicitly considers three alternatives to virtue's being the cause of happiness as possible ways of connecting the two.

¹³¹ This is one of the central points on which many Kantians disagree with Kant (cf. Sussman pp. 216-7, Murphy pp. 104-5, Beck p. 242-5). We will see evidence that Kant himself had doubts about this.

[between happiness and virtue] in its concept, is an a priori necessary object of our will and inseparably bound up with the moral law, the impossibility of the first must also prove the falsity of the second. If, therefore, the highest good is impossible in accordance with practical rules, then the moral law, which commands us to promote it, must be fantastic and directed to empty imaginary ends and must therefore in itself be false.” (*KpV* 5:114)

Unless a necessary connection between virtue and happiness is possible, the highest good cannot be willed, in which case there cannot be a duty to will the highest good. And since in that case the moral law commands what is not in fact a duty, the moral law is false.

Fortunately (whether for the moral law or for finite practical reasoners), the fourth relation, in which the maxim of virtue is the efficient cause of our happiness, is not in fact strictly impossible. Because theoretical reason cannot deny the possibility of the immortality of the soul and of the existence of an omniscient, omnipotent, omnibenevolent God, practical reason can postulate an afterlife in which the soul may continue to progress towards virtue and be rewarded with happiness accordingly and a God with the knowledge (of our virtue), the power, and the will to perform the rewarding. By God’s agency, happiness could be conditioned on our virtue, and since virtue could thereby be the cause of happiness, we can make the highest good our ultimate end.¹³²

¹³² Jeffrie Murphy and Andrews Reath argue that it is impossible for us to promote the highest good because we cannot know people’s virtue, and hence it is impossible for us to apportion happiness to virtue (Murphy pp. 107-8, Reath pp. 609-610). But our contribution to the highest good is not the apportionment of happiness to virtue, but virtue itself. Kant is explicit about this in “On the Common Saying,” where he describes the highest good as “universal happiness combined with and in conformity with the purest morality throughout the world” and then points out that this highest good “is within our control from one quarter [i.e. morality] but not from both taken together [i.e. morality and happiness combined with and in conformity with morality]” (8:279). This may be surprising because we expect that an obligatory end will add something to our duties, and morality is a duty independently of the highest good. But Kant is clear by the *Religion* (cf. *R* 6:4-7) that the point of the highest good is not to tell us what to do (for which the moral law is sufficient) but only to provide a satisfying answer to the question “*What is then the result of this right conduct of ours?*” ... Harmonizing with this end does not increase the number of morality’s virtues but rather provides these with a special point of reference for the unification of all ends” (*R* 6:5). For an excellent

Indeed, it is important to Kant's moral picture that our rational faith in God and immortality cannot rise to the status of knowledge. For if we knew that God would reward our virtue, Kant thinks, our moral motivation would be corrupted by awareness of this reward (*KpV* 5:147). The postulates are therefore not, for Kant, a second-rate solution to the Antinomy, for knowledge of the actuality of God and immortality would be no less a blow to Kant's moral system than knowledge of their impossibility. Only knowledge of their mere possibility is compatible with the right motivation.

Softening the Antinomy

As I mentioned earlier, Kant later took the threat to morality posed by the Antinomy to be considerably less dire than what he suggests in the 2nd *Critique*. Whereas in the 2nd *Critique* Kant claims that a failure to resolve the Antinomy would undermine the moral law itself, in the 3rd *Critique* Kant says that the impossibility of the highest good would do “damage to the moral disposition” (*KU* 5:452), and in the *Religion*, Kant merely claims that the impossibility of the highest good would be “a hindrance to moral resolve” (*R* 6:5). These claims are important for an evaluation of Hegel's criticism for two reasons. First, if the problem is less dire, then a failure to find a solution is less devastating for Kant's view. Second, Kant's later versions of the Antinomy explore possibilities that the 2nd *Critique* passed over, and in so doing illuminate the nature of the problem.

discussion of this point, see Kleingeld (2016, pp. 42-9). Murphy is right to say that the highest good is an idea of reason, but wrong to see it as an “aesthetic ideal” (p. 109), for although we are not capable of achieving the highest good ourselves, it is our end, the object of our moral willing, constructed out of our moral and natural ends. But he is not too far off insofar as by “aesthetic ideal” he is trying to indicate that the highest good is not the determining ground of any action.

If Kant is to relax his claim that the moral law itself would be undermined by the Antinomy, he owes us an explanation: He must reject one of three premises: 1) making the highest good our end is a duty, 2) we cannot make something that is impossible our end, or 3) ought implies can. This last option would be drastic for Kant. In the 3rd *Critique*, Kant seems to suggest the second option: The moral law “determines for us, and indeed does so *a priori*, a final end, to strive after which it makes obligatory for us” (*KU* 5:450). Here there is an obligation to pursue an end, but it need not be a possible end. Since the possibility of the happiness conditioned on virtue is what is in question, we are simply obligated to promote this happiness as much as we can consistently with the moral law:

“Now for us to promote this [final end] as much (insofar as happiness is concerned) as lies in our power to do so is commanded by the moral law, let the outcome of this effort be whatever it will. The fulfillment of duty consists in the form of the earnest will, not in the intermediate causes of success” (*KU* 5:451).¹³³

The problem with this second option is that it doesn’t really answer to the demands of reason: What reason sought was an end that could unite all of our activity, something we could see ourselves as aiming at in everything we do. For the end to have the proper unity, it is necessary that our virtue be the cause of our happiness: What we aim at in the highest good is virtue and happiness caused by virtue. But in this world, our virtue is not the cause of our happiness. Hence, if there is no God and immortality in which virtue can become the cause of our happiness, we cannot strive after the end of happiness caused by virtue at all. We can only strive after two separate ends, virtue and happiness,

¹³³ I’ve modified the translation to make it clearer (though perhaps more clunky).

with no unity between them.

Perhaps for this reason, in the later *Religion*, Kant instead seems to pursue the first option: He denies that making the highest good our end is a requirement of the moral law. Kant describes the need to make the highest good our end as deriving from but going beyond the moral law: “The proposition, ‘make the highest possible good in this world your own ultimate end,’ is a synthetic proposition *a priori* which is introduced by the moral law itself, and yet through it practical reason reaches beyond the law” (*R* 6:5). Here Kant claims that the need to make the highest good our end is a need arising from but not included in the moral law:

“But that every human being ought to make the highest possible *good* in the world his own *ultimate end* is a synthetic practical proposition *a priori*, that is, an objective-practical proposition given through pure reason, since it is a proposition that exceeds the concept of the duties in this world, and adds a consequence (an effect) of these duties that is not contained in the moral laws and cannot, therefore, be evolved out of them analytically. For these laws command absolutely, whatever their consequences; indeed, they even require that we abstract from such consequences entirely whenever a particular action is concerned, and thereby they make of duty an object of the highest respect, without proposing to us, or assigning, an end (and an ultimate end) such as would constitute some sort of inducement for it and an incentive to the fulfillment of our duty.” (*R* 6:7)

Here making the highest good our end is a duty (“ought”) that is not contained in the moral law itself, but derives from it.¹³⁴ This account makes sense of the fact that although

¹³⁴ Reath denies that Kant could accept a principle of moral desert that is not part of the moral law, for “to allow that there are moral principles that are independent of the law... would violate the autonomy of pure practical reason... by making it subject to principles that it does not generate out of itself” (p. 612). But if, as I am suggesting, it is pure practical reason itself that gives rise to a principle other than the moral law, then the principle does not violate reason’s autonomy. Reath tellingly slips from Kant’s “the unconditioned totality of the object of pure reason” to his abbreviation, “the *unconditioned object* of the Moral Law” (p. 597). By the *Religion*, if not before, it seems clear that the unconditional object of the Moral Law is virtue, and the

Kant derives the highest good from pure practical reason's seeking the unconditioned with respect to our ends, he never derives an obligation to make the highest good our end from the categorical imperative. Indeed, it's hard to see how such an obligation could be derived. Perhaps in the 2nd *Critique* Kant hastily assumed that a demand of pure practical reason is a demand of the moral law.¹³⁵ In fact there are two separate demands of pure practical reason: the highest good, as matter of the will, and the moral law, as form of the will. The failure of the former would not directly undermine the moral law; it would merely be a hindrance to our moral motivation.

Worthiness to Be Happy

Although in his later work, Kant does not think that the Antinomy threatens to undermine the moral law,¹³⁶ he still thinks that it constitutes a serious danger to the moral motivation, and that the Postulates of Practical Reason are a necessary answer to it. Since many Kantians are unpersuaded that there is a genuine Antinomy in need of a solution, it may be helpful to flesh out the problem Kant is raising.¹³⁷ One key concern regarding the Antinomy is how Kant gets from the claim that virtue is the unconditioned condition for willing one's happiness to the claim that virtue should be rewarded with happiness. As Reath writes, "This way of linking moral and natural ends evidently follows

unconditioned totality of the object of pure reason is its proper superset, virtue united with happiness conditioned on virtue.

¹³⁵ In the *Religion*, Kant suggests a slightly different demand of the moral law: "the moral law wills that the highest good possible through us be actualized" (*R* 6:5). Here what is necessary is not that we will the highest good, but only the highest good possible through us. Hence, this demand is compatible the impossibility of happiness's being proportioned to virtue.

¹³⁶ It is perhaps noteworthy that he does not refer to it as an antinomy in the *Religion* or the 3rd *Critique*.

¹³⁷ For examples, see Beck (1960, pp. 242-45), Murphy (1965), Auxter (1979), and Reath (1988).

from a principle of moral desert. But no reasons were ever given for thinking that the Moral Law generates such a principle, or provides any basis for relating virtue and happiness in this way.”¹³⁸

I have already given Kant’s argument that for virtue and happiness to be united in one end, virtue must be the cause of happiness, and I have shown that Kant backs away from the claim that the Moral Law contains a duty to adopt this highest good. Rather, practical reason seeks the highest good as a way of uniting our material ends. In this section, I will argue for the urgency (if not necessity) of the highest good as a claim of moral desert for Kant’s system. I will argue that this desert claim is connected with Kant’s discussion of worth in the *Groundwork*.

Kant describes the need for a final end as the need for something we can love in morality: “Now in this end human beings seek something they can *love*, even though it is being proposed to them through reason alone” (*R* 6:7 fn.). In the *Metaphysics of Morals* Kant clarifies that “love is a matter of *feeling*, not of willing, and I cannot love because I *will* to, still less because I *ought* to (I cannot be constrained to love); so a *duty to love* is an absurdity” (*MS* 6:401). In claiming that in the final end, human beings seek something they can love “even though it is being proposed to them through reason alone,” Kant seems to be asserting that in addition to satisfying reason’s demand for unity in our ends, the highest good as final end would satisfy a demand of our sensible nature.

The claim that in the final end, human beings seek something they can love parallels

¹³⁸ Pp. 611-612

a comment in the 2nd *Critique* in which Kant contrasts the grounds of our love for God with the grounds of our adoration:

“For nothing glorifies God more than... respect for his command, observance of the holy duty that his law lays upon us, when there is added to this his magnificent plan of crowning such a beautiful order with corresponding happiness. If the latter (to speak humanly) makes him worthy of love, by the former he is an object of worship (adoration).” (*KpV* 5:131)

The moral law is a proper object of respect regardless of the results of obeying it, but we can find something to love in the moral law only if it is properly connected with our happiness.

It’s important to note that Kant does not seem to think of this as merely a problem that each individual’s reason raises about the individual’s own highest good: A highest good including happiness is required, “not merely in the partial eyes of a person who makes himself an end but even in the judgment of an impartial reason, which regards a person in the world generally as an end in itself” (*KpV* 5:110). If the highest good of the individual is virtue and happiness conditioned on virtue, then the highest good of a possible world is the virtue of all and happiness conditioned on virtue, or as Kant glosses it, virtue and happiness in proportion¹³⁹ to virtue. In his most vivid illustration of the

¹³⁹ It is not clear that anything in Kant’s argument implies that it is bad that the vicious be happy, only that the virtuous be unhappy. However, Kant seems to think that it is good that the criminals (and perhaps the vicious?) be punished. He even suggests that the need to punish criminals is a ground for belief in an afterlife: “Belief in a future life does not, properly speaking, come first, so as to let the effect of criminal justice upon it be seen; on the contrary, it is from the necessity of punishment that the inference to a future life is drawn” (6:490 fn.). Kant’s discussion of this point in connection with criminality rather than vice may be motivated by the thought that enforcement of laws of right is necessary not just for the freedom of the criminal, but of his victim. Omission of punishment is a denial of external freedom to all. Whatever Kant thought about the goodness of the unhappiness of the vicious, in the two arguments he gives regarding the final end, it is evident only that the virtuous must be rewarded with happiness, not that the vicious must be punished. This point will become important in Hegel’s response to the Antinomy. Thanks to Christine

problem posed by the Antinomy of Practical Reason, Kant asks us to suppose a virtuous man who does not believe in a God or an afterlife in which the highest good is achieved:

“Deceit, violence, and envy will always surround him, even though he is himself honest, peaceable, and benevolent; and the righteous ones besides himself that he will still encounter will, in spite of all their worthiness to be happy, nevertheless be subject by nature, which pays no attention to that, to all the evils of poverty, illnesses, and untimely death, just like all the other animals on earth, and will always remain thus until one wide grave engulfs them all together (whether honest or dishonest, it makes no difference here) and flings them, who were capable of having believed themselves to be the final end of creation, back into the abyss of the purposeless chaos of matter from which they were drawn.” (*KU* 5:452).

This virtuous man who has no hope for an afterlife, Kant asserts, cannot help but have his moral motivation damaged by seeing the unhappiness of the virtuous around him. The unhappiness that damages the virtuous man’s moral motivation is not his own. The thought is that there is something disappointing about the unhappiness of virtuous people from an impartial standpoint: “a rational impartial spectator can nevermore take any delight in the sight of the uninterrupted prosperity of a being adorned with no feature of a pure and good will, and that a good will thus appears to constitute the indispensable condition even of the worthiness to be happy.”¹⁴⁰ This disappointment seems to be captured by the idea that the virtuous are worthy to be happy (*KpV* 5:123-4).

To see virtue as worthiness to be happy in this way is to see it as an essentially incomplete good. Kant repeatedly emphasizes that virtue is an unconditional good, that

Korsgaard and Lucas Stanczyk for helpful thoughts and references on this point.

¹⁴⁰ Note that Kant does not here assert that the vicious be unhappy, only that their happiness is not good.

its goodness does not depend upon anything that accompanies it. But when he says that it is not the complete highest good, he does not, I think, merely mean that other things are also good; he means that we cannot help but see virtue without happiness as disappointing. Happiness is not just one more dish in the feast of the good. It is the salt¹⁴¹ without which we still recognize that we must eat, but do so joylessly.

Here I think it's helpful to return to the distinction Kant makes in the *Religion* between our ability to respect the moral law and our ability to find in it something we can love. Hearing of great moral fortitude accompanied with great unhappiness may inspire respect for the moral law, but not love. We recognize the necessity of morality and admire the virtuous agent, but we are not pleased with the story.¹⁴²

This assessment is supported by Kant's discussion of the kingdom of ends in the *Groundwork*, where he arrives at a conception of virtue as worthiness to be happy through a discussion of two kinds of worth in the kingdom of ends, price and dignity. Price is relational worth. Objects have a price by virtue of their relation to human inclinations and needs. According to Kant's view, the worth we should ascribe to things (their worth in the kingdom of ends) is limited by the moral law. If the object of my inclination is contrary to the moral law, then it has no objective worth. But if I will the object of my inclination or need in accordance with virtue, then it has an objective price in the kingdom of ends.

¹⁴¹ A salt which of course does nothing to improve the repulsiveness of vice.

¹⁴² The displeasure I have in mind is a displeasure that things are/were so, not a displeasure in the telling of the story or its aesthetic qualities. It is best to think of non-fiction here.

If lawgiving determines the worth of things in the kingdom of ends, Kant thinks, then lawgiving itself must have unconditional worth: “The lawgiving itself, which determines all worth, must for that very reason have a dignity, that is, an unconditional, incomparable worth” (*G* 4:436). Note that Kant infers here from the worth of the objects of inclination and need to the worth of lawgiving. The worth of objects that satisfy our inclinations and needs is immediately evident to us because it is in our nature to regard them as valuable. But the worth of persons as moral legislators and of the act of moral legislation itself we discover only by a rational inference from the worth of objects that satisfy our inclinations and needs. It’s because regarding objects that satisfy our inclinations and needs as having worth rationally requires us to recognize lawgiving as having an inner worth that we recognize the dignity of the latter. Moral legislators in the kingdom of ends and their act of legislation are the *ratio essendi* of the worth of objects of inclination and need, but objects of inclination and need are the *ratio cognoscendi* of the worth of moral legislators and the moral law.¹⁴³

This might explain why we are so disappointed by the unhappiness of a virtuous person. We conceive of her virtuous will as valuable because it is the condition of the worth of her happiness. But when she is unhappy, there is no happiness for her virtue to condition. Because our awareness of the worth of virtue is mediated by our awareness of

¹⁴³ Kant himself uses the *ratio essendi*/*ratio cognoscendi* distinction to describe the relationship between freedom and the moral law. Freedom is the *ratio essendi* of the moral law, the reason for its being, since the free will legislates the moral law to itself. But although freedom is first in the order of being, it is second in the order of knowing: Reason is the *ratio cognoscendi* of freedom, the reason for our knowing that we are free. According to Kant, we cannot have experience of ourselves as free, but know ourselves to be free (for practical purposes) because we know ourselves to be under the moral law (*KpV* 5:4 fn.). Kant refers to our awareness of the moral law as binding our will as a fact of reason (*KpV* 5:31-2).

the worth of happiness, we inevitably see the worth of the unhappy person's virtue as containing a reference to her missing happiness. The worth of virtue, it seems, is worthiness.

This is not to say, of course, that the goodness of virtue is conditioned on happiness for Kant; Kant repeatedly denies this. It is rather to say that as finite rational beings we can only understand the worth of virtue in terms of its rational relationship to a possible happiness. Kant even describes the moral motivation in these terms:

“We do indeed find that we can take an interest in a personal characteristic that brings with it no interest at all in a condition, if only the former makes us fit to participate in the latter in case reason were to effect the distribution, that is, that mere worthiness to be happy, even without the motive of participating in this happiness, can interest us of itself” (*G* 4:450).

The prospect of actual happiness is not essential to our moral motivation, but we nevertheless conceive of the interest we take in morality in terms of a possible happiness. We will be virtuous not because it will make us happy, but because it will make a possible happiness objectively good. And because that is how we understand the worth of virtue, virtue unrewarded with happiness always appears incomplete to us.

So when Kant says that virtue alone is “not yet ... the whole and complete good as the object of the faculty of desire of rational finite beings” (*KpV* 5:110), he is not just making the obvious point that people want more from life than just virtue. He is saying that people cannot see virtue as the final end of the faculty of desire without relating virtue to the happiness it makes them worthy to enjoy:

“*happiness* is also required, and that not merely in the partial eyes of a person who makes himself an end but even in the judgment of an impartial reason, which

regards a person in the world generally as an end in itself. For, to need happiness, and to be also worthy of it, and yet not to participate in it cannot be consistent with the perfect volition of a rational being that would at the same time have all power, even if we think of such a being only for the sake of experiment.”¹⁴⁴ (*KpV* 5:110)

If the worth of virtue can only be cognized as worthiness to be happy, then we can see why Kant thinks the postulates of practical reason are required. For to aim at virtue and regard the worth of virtue as worthiness to be happy requires that one regard the reward for virtue as part of the highest good. The highest good would then require that virtue be rewarded. But who will see to it that virtue is rewarded? Virtue does not include happiness in itself. I cannot always see to it that my virtue is rewarded, for my own happiness might only be bought at the expense of my virtue. And the laws of nature certainly do not look after the happiness of the virtuous. Thus, the only way we can formulate for ourselves a highest good to aim at is by supposing that there is some time after this life in which virtue is rewarded and some power to reward it.

Kant is clear that such a reward is not part of the motivation for virtue. He is also clear (in the *Religion*) that such a reward need not be guaranteed, since respect for the moral law alone is sufficient to motivate us to action. We regard the law as necessary and the virtuous will as good regardless of whether it is rewarded. But we cannot love the law

¹⁴⁴ This passage is odd because Kant suggests that an impartial reason would require happiness, but the example he gives seems to be of a rational being's relationship to his own happiness, about which, we might think, he is not in a position to be impartial. One could suppose that this is simply a grammatical infelicity: Kant says that a rational being cannot consistently will not to participate in happiness it needs and is worthy of. Nevertheless, Kant may have intended for us to understand there to be two beings in question, one who is worthy of and lacks happiness, and another who has all power. But we could also read the thought experiment as referring to just one rational being who has two grounds (partial and impartial) for demanding happiness: 1) his need (inclination) for happiness and 2) his worthiness to be happy. His merely felt need for happiness is one to which he is naturally partial, but his worthiness to be happy is as such an objective ground for a demand for happiness

that deprives the good will of happiness. We do not wish that a good will deprived of happiness by its own virtue had acted otherwise, but under such circumstances we cannot but view the moral law as a harsh taskmaster, “which cannot occur without damage to the moral disposition” (*KU* 5:452-3).

The core problem here is a conflict between two applications of practical reason. On the one hand, reason seeks the unconditioned in the form of our maxim and derives the categorical imperative, and on the other hand, reason seeks the unconditioned in the matter of our maxim and derives the highest good.¹⁴⁵ But there’s a tension between the categorical imperative and the highest good. The categorical imperative is indifferent to whether our happiness is caused by our virtue, while the highest good requires it. The impossibility of the highest good is the impossibility of a unified end of practical reason, which would be unsatisfactory both to practical reason and to sensibility: to practical reason because it seeks the unconditioned totality of our ends; to sensibility because it seeks our happiness. Moreover, because our sensible need for happiness provides us with the *ratio cognoscendi* of the worth of virtue in the first place, our understanding of the worth of virtue is always mediated by its reference to a possible happiness. Virtue without happiness must therefore always seem to us to be an incomplete good, even from an impartial standpoint. This does not diminish our respect for law, Kant thinks, but it does make us hate it, which does damage to the moral motivation.

¹⁴⁵ “As pure practical reason it likewise seeks the unconditioned for the practically conditioned (which rests on inclinations and natural needs), not indeed as the determining ground of the will, but even when this is given (in the moral law), it seeks the unconditioned totality of the object of pure practical reason, under the name of the *highest good*.” (*KpV* 5:108)

Hegel's response

Hegel's diagnosis is that Kant's system endorses a desideratum of a theory of freedom that his system cannot fully meet: Kant acknowledges (as he should) that we cannot be satisfied with a moral law that doesn't necessitate a reward for obedience, but his theory of morality, divorced as it is from worldly conditions, cannot provide such a reward. The most that Kant can offer is an argument that theoretical reason does not strictly preclude the possibility of an otherworldly reward.¹⁴⁶ Hence, Kant's position alienates us from our own happiness: The desire for happiness becomes a temptation from without. But reason itself (in its application to the matter of practical reason) resists thinking of happiness as alien and seeks to establish a unity in the highest good.

Hegel claims that the problem with Kant's view is that it takes "the standpoint of the ought" (*EG* §512).¹⁴⁷ The subject ought to be virtuous, and virtue ought to be rewarded with happiness, but it is contingent whether either of these occurs in reality. As long as morality exhibits this contingency, the highest good must be exiled to another world, and we are condemned to pursue but never achieve it in this life.

The alternative would be for the moral law to be more than a mere ought, for virtue

¹⁴⁶ As will become even clearer next paragraph, Hegel denies that awareness that there is a reward for our virtue necessarily corrupts the moral motivation. But as we will see, the nature of the connection between virtue and its reward in Hegel allows for cases in which virtue is not rewarded. And indeed, Hegel gives special importance to these occasions because they demonstrate that the demands of duty go beyond our need for happiness (cf. Hegel's discussion of war in §323-324), but these demands can only go beyond our need for happiness because they are on the whole aligned with it.

¹⁴⁷ In the *Phenomenology*, Hegel gives a series of arguments against Kant's characterization of the highest good and the postulates. Even a cursory examination of these arguments would take us too far from the guiding thread of the general will, so I will adhere to the lines of thought suggested in the *Encyclopedia*.

in fact to be the cause of happiness in this world.¹⁴⁸ The most obvious objection to this proposal is that it is simply not the case, for virtue is in fact often the cause of unhappiness. Hegel's response will be to offer a different conception of the moral law on which duty must provide for our happiness as a condition of its being duty (*PR* §130). This proposal may make it seem that Hegel is simply abandoning the central features of Kant's view. After all, Kant's central aim in the *Groundwork* is to show that we can be motivated by reason independently of any consequence for our happiness. Duty is shown in its purest form when it conflicts with happiness.

Hegel does not want to return to the Humean thesis that reason is the slave of the passions. Nor does he want an intuitionistic hybrid position in which the ends of reason and sensibility must somehow be weighed against each other. Hegel recognized Kant's thesis that we have our freedom in duty as one of Kant's greatest contributions to philosophy: "In doing my duty, I am with myself and free. The merit and exalted viewpoint of Kant's moral philosophy are that it has emphasized the significance of duty" (*PR* 133A). Hegel agrees with Kant that we find our freedom in doing our duty and that our duty is legislated by reason. But he disagrees with Kant on what reason legislates.

¹⁴⁸ After an interesting discussion of the problem Kant's highest good presents and the failure of the postulates to answer to that problem, Birgit Recki comments,

"We should be fair here, however. It may be helpful to consider the troublesome question of what an alternative to Kant's view would look like, i.e. what a defensible alternative would consist in (assuming one wishes neither to be unrealistic about the claim to happiness and lose sight of the real condition of the world, nor to lose sight of the demands of morality)." (p. 86)

My aim in this section is to present the outlines of such an alternative as Hegel envisioned it. Whether Hegel's view loses sight of the real condition of the world or of the demands of morality will receive further examination in the next chapter.

Kant insisted that duty is legislated in reason's formal principle. Reason's material idea simply tells us what the world must become for our reason and sensibility to be satisfied with the moral law.¹⁴⁹ For Hegel, reason's principle is its legislative principle.¹⁵⁰

The material principle is legislative in that legislation occurs only when reason unites the form of law (universality) with the needs of particular sensibility. Hegel's view, then, is that the ability of duty to reward obedience is a condition on its being duty. But doesn't this still return us to pre-critical moral philosophy in making happiness the motivation for performing one's duty? Although Hegel's moral psychology differs from Kant's, this charge misses Hegel's point. We need to distinguish between the conditions that make something a duty and the motivation for doing one's duty. That a duty is connected with happiness may be a condition on its being a duty without being the proper motivation for performance. Modifying Kant's position, we can make a simple argument for why happiness might be thought to be a condition on something's being a duty: Duty is something the agent legislates for herself. To legislate it, she must see it as good. To see duty as good, she must see that duty treats those who perform it (the virtuous) well.¹⁵¹ Thus, for something to be a duty, it must be compatible with the happiness of those who obey it. Nothing about this argument suggests that one's motivation for obeying such a

¹⁴⁹ Here I give Kant's later view from the 3rd *Critique* and the *Religion*. In the 2nd *Critique*, the material principle tells us what the world must be capable of becoming for us to have duties.

¹⁵⁰ Although the claim that reason has a material principle may seem like rational intuitionism, the two views share little in common. Hegel does not think, as do the intuitionists, that basic moral propositions are self-evident, nor that moral properties are non-natural.

¹⁵¹ This need derives not from her own desire for a reward but from her inability to affirm as good a law that makes the virtuous unhappy, as Kant's example of the virtuous atheist emphasizes.

duty is the desired happiness.¹⁵²

In other words, Hegel eliminates the asymmetry between the goodness of the moral law and the goodness of happiness expressed in the thought that the moral law is a condition on the goodness of happiness but not vice versa or in the thought that the moral law is the *ratio essendi* of the goodness of happiness, whereas happiness is merely the *ratio cognoscendi* of the goodness of the moral law. Hegel accepts that the moral law is a condition on the goodness of happiness, but the opposite is also true. Happiness is a condition on the goodness of the moral law:

“Welfare has no validity for itself as the existence [*Dasein*] of the individual and particular will, but only as *universal* welfare and essentially as *universal in itself*, i.e. in accordance with freedom; welfare is not a good without *right*. Similarly, right is not the good without welfare (*fiat iustitia* should not have *pereat mundus* as its consequence).” (§130)

We may succinctly sum up the differences between Kant’s and Hegel’s notions of the subject and its freedom by recalling that Kant’s evidence that we are free is that the moral law infringes upon our happiness: A person threatened with immediate execution unless he gives false testimony against an honorable man “must admit without hesitation that it would be possible for him. He judges, therefore, that he can do something because he is aware that he ought to do it and cognizes freedom within him, which, without the moral law, would have remained unknown to him” (*KpV* 5:30). The irony, Hegel would say, is that Kant presents as proof of our freedom a scenario in which one is clearly unfree in an

¹⁵² It is another question just what that motivation is. Since Hegel rejects the Kant’s merely formal law, the proper motivation cannot merely be respect for lawful form.

important sense: one's duty and one's happiness are radically opposed.¹⁵³ Kant wants to show that one is free in such a situation because one's needs as a finite being do not determine one's will; one is free from one's needs, from one's finitude. But this freedom from finitude comes at the cost of denying that reason takes an interest in one's happiness.

The early Hegel put the point this way:

“Between the Shaman of the Tungus, the European prelate who rules church and state, the Voguls, and the Puritans, on the one hand, and the man who listens to his own command of duty, on the other, the difference is not that the former make themselves slaves, while the latter is free, but that the former have their lord outside themselves, while the latter carries his lord in himself, yet at the same time is his own slave. For the particular—impulses, inclinations, pathological love, sensuous experience, or whatever else it is called—the universal is necessarily and always something alien and objective.”¹⁵⁴

Although our ability to act from duty against self-interest illustrates an important aspect of our self-determination, complete self-determination requires that our duties be (in general) properly aligned with our sensible inclination, for only then is the whole self (rational and sensible) free. Self-love is an ingredient in our self-determination, and therefore someone whose duty has become radically opposed to his happiness is to that extent unfree.

But what kind of moral principle could secure the happiness of those who obey it?

The answer was suggested by Kant himself in idea of the kingdom of ends, or as Hegel

¹⁵³ Kant could say that I am unfree in that I lack external freedom, the freedom with which right is concerned, since I am being threatened with execution for doing what I must do and therefore cannot have willed. I am therefore being subjected to another's will.

¹⁵⁴ GC 211. Hegel's mention of Shaman, European prelate, Voguls, and Puritans is a reference to Kant's own mention of them in the *Religion*, but Kant's conclusion is that the man who listens to his own command of duty, unlike the others, is free (R 6:176).

calls it, Ethical Life. The tendency of duty to promote the happiness of those who obey it depends upon political and social context. Respect for others' property claims tends to promote the happiness of the respecer in middle-class, law-governed communities. But it does not always promote the happiness of the respecer in communities where the respect is not reciprocated or where the distribution of resources is such as to leave respecers of the law impoverished. In Rousseau's words,

“Considering things in human terms, the laws of justice are vain among men for want of natural sanction; they only benefit the wicked and harm the just when he observes them toward everyone while no one observes them toward him.

Conventions and laws are therefore needed to unite rights with duties and to bring justice back to its object.”¹⁵⁵

In the kingdom of ends (or to use Hegel's term, Ethical Life), however, everyone's obedience to the law is the cause of everyone's happiness. According to Hegel, to qualify as a duty, an imperative must exist in a social context such that in performing my duty, I contribute to maintaining the social context that makes the virtuous (of which I am one) happy:¹⁵⁶ Whereas “in morality, it is merely an *obligation* that the right of my own knowledge and volition, and my welfare, should be united with my duties and exist objectively,” “In the ethical realm, a human being has rights in so far as he has duties, and duties in so far as he has rights” (*PR* §155).¹⁵⁷

¹⁵⁵ *Social Contract* 2.6.2

¹⁵⁶ Note however that Hegel's solution does not make the reward depend upon virtue in quite the way Kant hoped. Both those who act in accordance with duty from a motive of duty and those who so act from another motive receive the benefits. But Kant's argument that the happiness of the virtuous was good did not show that the happiness of the vicious was bad. Still, Hegel's virtuous will enjoy a special kind of happiness not available to the vicious, for they will experience themselves as free in the performance of their duties, while the vicious will not.

¹⁵⁷ It is striking how well what Reath's description of the secular conception of the Highest Good applies to

It would be a mistake to assume that Hegel's argument requires that each individual duty carry its own reward for the virtuous. The need for a unified end of our rational and sensible nature would be satisfied by a principle that ensured the happiness of the virtuous. Some particular duty derivable from the principle, considered in isolation, may be a net cost to some of the virtuous as long as the whole system of duties is tends to take care of their happiness. Indeed, Hegel may even allow that there are occasions on which duty requires great sacrifice to our happiness, so long as this sacrifice is appropriately connected with happiness. Citizens may have a duty to risk their lives defending their country against an invader because such a defense is a necessary condition for the state, which provides for their happiness, to exist.¹⁵⁸ Our ability to find something we can love in the law depends upon its general character rather than upon its having the desired effect in any one particular case.

Hegel's Ethical Life plays the roles of both the kingdom of ends and God in Kant's resolution of the Antinomy. At various places, Kant had already proposed that the kingdom of ends was part of the means for rewarding virtue with happiness.¹⁵⁹ But of course, the kingdom of ends cannot peer into people's souls to discover their degree of virtue and reward it appropriately. God was needed for that. Hegel seems to be content that the virtuous generally be rewarded without worrying that those who obey the law

Hegel's Ethical Life (cf. pp. 614-619). In summing up his discussion, he says, "The aim of this system of institutions would be to create conditions which would be conducive to moral conduct, in part by making it fully rational to act from moral principles" (p. 619). I take this agreement as a testament to Kant's influence on Hegel and Reath's interpretive insight.

¹⁵⁸ *PR* §323-324

¹⁵⁹ Cf. (*KrVB* 837-38)

from the wrong motive might receive unearned happiness. Kant's arguments for the highest good do not show that the unhappiness of the vicious is good anyway. All that was necessary was that our virtue cause our happiness, not that our vice cause unhappiness.¹⁶⁰

God's other role in Kant's solution to the Antinomy is to set up the kingdom of ends, to provide the conditions for people to associate according to the moral law. Here Ethical Life takes over God's role in that it provides the conditions of its own existence. Ethical Life is not merely an association of persons who each abide by the moral law. In Ethical Life, as in any kingdom, the people act together to provide for the preservation of Ethical Life. Moreover, as we will see in the next chapter, Ethical Life brings itself into existence.

Although Ethical Life brings itself into existence through its own rational necessity,¹⁶¹ not every given political and social context approximates the kingdom of ends to the same degree. For Kant, the failure of a political and social context to provide laws of freedom limits my external freedom, but not my internal freedom: I always legislate the law of my internal freedom, the categorical imperative. But for Hegel who insists that legislation in Ethical Life is the necessary condition for duty, a deficiency in my political and social context does not merely constitute a deficiency in my external freedom. It is also a deficiency in my internal freedom: Reason gives me genuine duties only in the proper

¹⁶⁰ Moreover, Hegel's moral psychology is importantly different from Kant's. Hegel resists Kant's strict rational/sensible duality, so that for Hegel, the right sort of feeling can register a rational motivation in irrational form (*PR* §11). For example, in the family, we first experience duty in love. To perform familial duty from a motive of love is a primitive form of virtue for Hegel (*PR* §158, 158A). Love is not simply an arational feeling: We can articulate the reason why love is an appropriate feeling in a family context. And since part of the job of Ethical Life is to instill such motives into people, people will, broadly speaking, be virtuous.

¹⁶¹ See the fifth chapter.

political and social context.

Hence, Hegel's response to the Antinomy is not to point to some guarantee that Kantian virtue will be rewarded with happiness. In choosing reason's material principle as the proper source of law, Hegel is accommodating the demands of virtue to needs of happiness. In doing so, he is giving up on an aspect of the attractive Kantian thesis that the demands of ethics are universal. Hegel thinks about both ethics and right what Kant thinks about right: It must be instituted by a power capable of enforcing it for it to be binding. Where no power has instituted the relevant law, it is not binding on the individual. Of course, Hegel is not a complete relativist, just as Kant is not a relativist about right: There is something rationally deficient in a form of Ethical Life that does not include the proper laws of freedom, and reason can call upon members of such a form of Ethical Life to institute the proper laws. But in the absence of such laws, there is no duty to live as if they existed.

The claim that ethical duties, and hence internal freedom, depends upon one's political and social context gives rise to a further worry: How is Hegel's picture compatible with autonomy? If my political and social context gives the law, the law seems to be heteronomous. Hence, for Hegel to stay true to the Kantian conviction that duty can only be legislated autonomously, Hegel must show that the will of the state, which gives the laws of my social and political context, is in fact my will.

We saw in the chapter on Kant's general will that Kant has his own version of this claim: The state's legislative will, or general will, is my will. But for Kant, the general will

is my will in the sense that each individual's faculty of reason necessarily joins itself to a unanimous willing of the laws of right. The necessity of reason's so doing takes the form of an *a priori* command of reason, "Obey the authority who has power over you" (*MS* 6:371). As *a priori*, this command addresses itself to us independently of our political circumstances, though it can move us to act only when there is an authority over us.

But this command cannot form part of Hegel's own view, for as we have just seen, Hegel denies that reason gives imperatives independently of any context. For Hegel, the laws of duty are not prescribed to us *a priori* but are given to us by our political and social context. Hence, the fact that the general will is my will is not a fact that every individual's faculty of reason can unanimously constitute *a priori*. Instead, Hegel claims, the identity of the general will and my will is constituted in the reverse direction: The general (or universal) will's faculty of reason constitutes individual reason.

In the next chapter, I will flesh out this claim, which once again takes advantage of Kant's own thoughts about reason. I will argue that for Hegel, reason is not in its most fundamental sense a faculty of the individual's mind. Rather, reason is most fundamentally an idea that can give rise to itself (e.g. as Ethical Life) in the world. The individual's faculty of reason enables the individual to recognize the necessity of his or her own activity as part of reason's self-actualization.

Before concluding, we should also note that in the Antinomy of Practical Reason, as in the criticisms of last chapter, it is apparent that individual practical reason is insufficient to accomplish the task it sets itself. Indeed, unlike in the criticisms of last

chapter, in the Antinomy Kant himself acknowledges this. Individual practical reason cannot reconcile the ends it sets itself; instead, it must appeal to another reason. For Hegel, this other reason is none other than the general will. And since the individual will presupposes this general will, it is to be understood not as a set of individual wills with the same determination, but rather as the substance within which these individual wills exist.

Conclusion

In the previous chapter we saw two arguments against Kant's morality: First, individual reason cannot satisfyingly resolve moral dilemmas, and second, individual reason's duties to self are in tension with the universality of law. In this chapter we've seen a third problem for Kant's view: Reason's formal law is in tension with reason's own demand that all of our ends (virtue and happiness) be united into one whole. And this tension makes it hard for us to love the law and hence to see the value of virtue.

In each case, Hegel insists that the problem is Kant's insistence that duty is a mere ought: Duties ought to harmonize with each other and with happiness, and the subject ought to harmonize them. This thought that duty is a mere ought is the key presupposition of Kant's moral philosophy. It is the fact or act of reason, the *ratio cognoscendi* of our freedom. Hegel thinks that if we eliminate this presupposition, we can eliminate all of these problems at once. Instead of taking a mere ought as the fundamental moral phenomenon, we should look to a different kind of act of reason which Hegel calls the Ethical Life. The Ethical Life is not a mere ought, but an actuality, and it contains our duties within it. We turn now to consider what kind of act of reason

the Ethical Life could be.

Chapter 5

Hegel's General Will

“It is the nature of love to make one consider oneself and the object loved as a single whole of which one is but a part; and to transfer the care one previously took of oneself to the preservation of this whole. One keeps for oneself only a part of one's care.” (Descartes 2018, 311)

In the last chapter we saw Hegel's arguments that a morality of mere *oughts* is in contradiction with itself. Hegel thinks that only an actualized morality, a morality in which duty is an *is*, and not just an *ought*, can resolve these contradictions. But this suggestion might seem to ruin Kant's project of grounding morality in reason alone. It was plain how pure reason could legislate a mere ought a priori, but how can it legislate an actuality? Actualities are known to us not a priori from reason, but a posteriori from sensibility. Moreover, since grounding morality in reason alone is what entitles Kant to claim that each of us legislates the moral law and therefore that we are free under it, despair of the project of grounding morality in reason might lead one to despair of the project of establishing our freedom under the law.

But Hegel does not reject Kant's thought that morality (or Ethical Life, Hegel's replacement for morality) is grounded in reason alone. Rather, Hegel rejects Kant's conception of reason. On Kant's view, our reason provides only the form of cognition, whether the cognition be theoretical or practical. This means that our cognition, theoretical and practical, is discursive, requiring both concepts from reason (form) and intuitions and inclinations from sensibility (matter). Hegel argues against this merely formal conception of reason. On Hegel's conception, reason provides not only the form of cognition, but also the matter.

Hence, Hegel rejects Kant's claim that the moral principle is merely formal. Whereas Kant thinks that practical reason legislates a merely formal principle, the Categorical Imperative, Hegel thinks that practical reason gives rise to the properly constituted social order itself, and that the duties that this social order prescribes to its members are themselves therefore rational. Hegel calls a properly constituted social order "Ethical Life" [*Sittlichkeit*]. Ethical Life can take different forms: The family is, Hegel thinks, the most basic form of Ethical Life, but the state is the most developed form.

For Kant, the law of the land must be the general will, and hence my will, because my will joins the wills of all in a noumenal act of ratifying it as law, and that act of noumenal legislation is known to us through the act or fact of reason [*Factum der Vernunft*], the command we give ourselves to obey the authority who has power over us. For Hegel, by contrast, reason's legislative activity is not carried out in the merely intelligible world, but in the world of sense. Hence, Hegel must reject Kant's alternative to Rousseau's

empirical participation in legislation. For Hegel, there is no noumenal realm to which we can relegate the general will.

But Hegel does not return to Rousseau's version of direct participation in legislation. Instead, Hegel articulates a new notion of the relationship between individual will and general will, one that takes advantage of an aspect of Rousseau's theory that Kant had abandoned: the idea that the general will is mine not because it belongs to me, but because I belong to it. Kant had no need for this thought because his notion of an idea of reason allowed him to insist that the general will was unanimous. Hegel develops a version of Rousseau's thought that allows him to claim that the legislative will is mine even when I did not directly participate in the legislation.

My first aim in this chapter is to explain why Hegel thinks that the state is rational in the sense that it has the features of what Kant referred to as an idea of reason. I'll begin with an overview of what Kant thinks an idea of reason is. I'll then argue that what Kant describes as the concept of a natural purpose, of an organism, has some of the features of an idea of reason. Then I will explain how Hegel's concept of the state has these features of a natural purpose and supplements them with further features necessary to qualify the state as an idea of reason.

Once we can see how the state is rational in itself, we will be in a position to see how the state is rational for its citizens, that is, how the state's will is the will of its citizens, an expression of their practical reason, not because their wills unite noumenally in ratifying the law, as in Kant, but because their wills stand in an organic relation to the will of the

state of which they are part. We will see two arguments that the state's will is the will of its citizens. First, the state and its laws are the solution to the three criticisms of Kant's morality investigated in the last two chapters. Hence, only in the context of the state's rationality does the individual will gain rational determinacy (as opposed to being in conflict with itself). Because the state's rationality is constitutive of the individual's practical reason, the individual's practical reason is not separable from it; hence, the state's rationality (manifest in its will) is the individual's own will.

Second, the account of the rational state as an idea of reason shows that the state is responsible not only for constituting the individual as rational, but also for bringing the individual into existence in the first place. The state brings the individual into existence as both end in itself and means to the state's activities. Thus, the state and the individual stand in an organic relationship to one another; the individual is properly speaking an organ of the state. Hence, the state's will is the individual's in the sense that the individual's will belongs to it as a rational organ: What the individual is, properly speaking, is a part of the state's will.

This last thought will help us understand why Hegel disagrees with Rousseau on the need for direct participation in legislation. The arguments mentioned in the previous paragraph suffice to show that direct participation is not necessary. But Hegel also thinks that direct participation is not ideal, for it is not rational. For Hegel, the most rational way to structure the state is organically, i.e., by dividing labor among different social roles. Direct participation in legislation lacks this organic structure and is hence irrational.

Ideas of Reason in Kant

As we saw in chapter 3,¹⁶² for Kant, reason is “the faculty of principles” (A299/B356), that is, the faculty that allows us to “cognize the particular in the universal through concepts” (A300/B357) by inferring judgments about particulars from universal principles. In so doing, Kant claims, reason “seeks to reduce the great manifoldness of understanding’s cognition to the smallest number of principles (universal conditions) and thereby to bring about the highest unity of this cognition” (A305/B361 P). The ultimate end of this search is what Kant calls the unconditioned, cognition from principles themselves.

Kant argues that reason’s search for the unconditioned leads us to several concepts of the unconditioned, called ideas of reason, concepts of “what alone can complete the series of conditions when we proceed to trace these conditions to their grounds. This is the course [on] which our human reason, by its very nature, leads all of us” (A584/B612). Among these ideas, Kant includes the soul (“the subject that is itself no longer a predicate”), the world (“the presupposition that presupposes nothing further”), and God (“an aggregate of those members of a division which require nothing further in order to complete the division of a concept”) (A323/B379-80 P).

Ideas or concepts of reason refer to objects that can never be found within experience, for anything we find in experience is conditioned. Ideas refer to the unconditioned totality of conditions of what we find in experience. Hence, any empirical

¹⁶² See the section “Ideas of Practical Reason.”

cognition can at most be a part of the object of an idea of reason.

Moreover, we do not cognize the objects of the ideas of reason a priori. The ideas of reason are not constitutive, like the categories of the understanding: The categories can give us a priori knowledge of objects of experience because they are necessary for constituting experience. Ideas of reason do not constitute experience.

Since we cannot have cognition of the object of an idea of reason a priori or a posteriori, we cannot have cognition of the object of an idea of reason at all. Still, ideas of reason do have a regulative use for our investigation of nature: “to direct the understanding to a certain goal by reference to which the directional lines of all the understanding’s concepts converge in one point” (A644/B672 P). Kant thinks that we need to think of our psychology as united in the soul, nature as united in the world, and everything as grounded in God to unify our cognition scientifically. Although this claim goes well beyond the scope of this dissertation, the thought is roughly that reason underwrites the scientific but ultimately unprovable supposition that all of our mental powers and states belong to one being, that series of conditions in the world go on *in indefinitum*, and that they can be united according to principles of a systematic unity, “as if they had all arisen from one single all-encompassing being” (A682-6/B710-4). But the fact that we must think the unity of our cognition in these terms does not mean that we have cognition of entities corresponding to these ideas.

If an idea is a “concept of the *totality of conditions* for a given conditioned,” then the object of an idea is itself unconditioned (A322/B379). Hence, an idea is the concept of

something necessary, for if it were not necessary, then it would be conditioned. This thought need not lead to the worry that Kant must endorse the ontological proof for the existence of God: Kant's distinction between constitutive and regulative principles of reason is meant to diagnose the temptation that leads to the ontological proof. Both constitutive and regulative principles involve a kind of necessity, but in the constitutive principle, the necessity is understood to hold of the object, and in regulative principles, the necessity is understood to hold of the inquiring subject. Hence, where Anselm interprets the necessity of the idea of God as meaning that God must exist, Kant interprets it as meaning that we must search for God by seeking to unite scientific knowledge into a systematic whole. And where Mendelssohn interpreted the soul's necessity as meaning that the soul must exist, Kant interprets it as meaning that we must search for the soul by seeking an account of the mental that unites all of our powers and states in one. Both pursuits can never be completed, and hence we will never achieve cognition of God or the soul, but science advances by the search for them.

Another way Kant puts this point is by distinguishing between a concept in which its object is given and a concept in which its object is given problematically (cf. A498-9/B526-7). In the idea of God, God is given to us not as an object of cognition, but only problematically, as something that we must seek in the unity of nature as a system. Hence, although Kant does not put it this way, we may say that the idea of God contains God's existence not as cognition, but only problematically, that is, as something we must necessarily search for.

It may help to articulate Kant's thought that reason seeks the unconditioned in other words. Why does reason seek cognition from pure concepts? When we know something merely through sensibility, we simply know that it is. Sensibility does not tell us why it is. My senses can tell me that the sun is to the south, but not why. Reason seeks to know why. Reason develops a theory of the solar system that accounts for the sun's location. Still, reason's job is not done, for we can still ask why there is a sun in the first place. If the answer appeals to other observed facts about the cosmos, then reason will want to know why those facts are so. Reason will not be completely satisfied with an account of the sun until it has been completely explained without appeal to facts that are merely given by experience. But that just means, that reason will not be satisfied until the necessity of empirical objects has been accounted for from pure concepts. Only then is there no more question to raise. An idea of reason is the concept of an object the existence and characteristics of which would be accounted for from mere concepts.

To sum up this discussion of reason in Kant: Reason's function is to cognize the particular in the universal, thereby unifying our various cognitions. Reason's task would come to an end only with the unconditioned condition of objects of experience. Hence, reason seeks the unconditioned. An idea of reason is the concept of an unconditioned totality of conditions, a concept which would provide us cognition from pure concepts if its object were given to us by reason and not merely the aim of reason.

Ideas of Practical Reason

While Kant insists that reason in its theoretical use cannot provide us with synthetic

cognitions from concepts, he suggests that a principle of practical reason might be the source of such cognitions:

“It is an ancient wish – who knows how long it will take until perhaps it is fulfilled – that in place of the endless manifold of civil laws, their principles may be sought out... But here the laws are only limitations of our freedom to conditions under which it agrees thoroughly with itself; hence they apply to something that is wholly our own work, and of which we can be the cause through that concept” (A301/B358).

Kant refers here to what he will call the Universal Principle of Right (UPR), the principle that requires us to limit freedom to the condition of universal law, the condition under which it thoroughly agrees with itself. Because such a principle derives from pure practical reason, and because we as pure practical reasoners are the cause of the civil laws required by this principle, Kant suggests, we can have practical cognition of the civil laws as necessitated by a principle.¹⁶³ The same is not true of “objects in themselves, as well as the nature of things” (A301/B358), of which we are not the cause through mere concepts (or at least, of which we do not know ourselves to be the cause).

Whereas ideas of theoretical reason are of merely regulative use, ideas of practical reason, because they regulate our own causality, can also give us practical cognition, cognition of the necessity of legislating certain civil laws, in this example. More generally, the all-encompassing idea of practical reason, determined according to the principle of practical reason, is the idea of the good. The idea of the good is the idea of that which

¹⁶³ Although Kant does not use this terminology here, we cannot have theoretical cognition of the civil laws as actually caused by ourselves according to the UPR. We can only have practical cognition of the necessity of establishing the civil laws from the UPR.

must be realized from the perspective of practical reason.

Note, however, that the necessity of the activity of practical reason itself is not itself explained. We can see, given that we are practical reasoners, why we must bring about the good, but we cannot see that there must be practical reasoners to bring the good about. Even setting aside the assumption that there are practical reasoners, we cannot see that the finite practical reasoners there are must be capable of overcoming temptation and realizing the good. And of course, Kant never intends to offer us reasons why there must be practical reasoners who do indeed realize the good. He does not think such cognition is available to us. Practical reason is not concerned with whether the good is actually realized because practical reason always adopts a standpoint presupposing its own freedom. But while the practical cognition afforded in ideas of practical reason exhausts practical reason's demand for explanation, it does not exhaust theoretical reason's demand for explanation. Moreover, practical cognition does not extend to awareness that anything empirical has been brought about by a free will. A free will, itself an idea of reason, cannot be given empirically and hence cannot be cognized as cause of empirical events.

Two Kinds of Empirical Concepts

Hegel thinks he can improve on Kant's picture by showing how an idea of reason can be given empirically, how an empirical object could be adequate to reason's demand for the unconditioned. To see how, consider two kinds of empirical concepts. The first is a mere abstract universal that applies to anything with the property or set of properties that

it picks out. For example, the concept block picks out a solid mass with at least one approximately flat surface. Individual blocks vary in their dimensions, number of flat surfaces, orientation, and location in space, but the concept abstracts away from these particularities.

The concept block does not by itself explain how any individual block comes to exist in the world. Admittedly, the concept block could be said to cause the existence of a block by the mediation of our faculty of desire: When we decide to make a block, our concept of a block becomes the cause of the existence of the block. But nothing internal to the concept of a block explains why there should be a being that desires to make one. At least with respect to the concept block, it is entirely contingent whether there is a being that desires to make one, and thus entirely contingent whether any block is produced by the concept of a block.

Like the concept block, the concept rabbit abstracts away from many of the particularities of the individual rabbit. But unlike the concept block, the concept rabbit is not entirely indifferent to the rabbit's existence. For the parts of the rabbit are not merely arranged in the proper form; they are arranged so as to *maintain* the rabbit's form: Their claws dig a hole for shelter from the elements and predators; their ears sense danger from afar; their long hind legs carry them to safety; their teeth bite through tough vegetation. And these parts are in turn maintained by the rabbit's other parts: Blood circulating through the body provides the nutrients from the lungs and the digestive system for maintaining each of these parts, and signals from the rabbit's brain direct them in their

characteristic activity. Moreover, the rabbit's activity and properties as a whole maintain each of its parts: Rabbits' foraging keeps each of the body parts alive and functioning; their crepuscularity helps them elude diurnal and nocturnal predators.

Following Kant, I will call such a concept the concept of a natural purpose. Kant describes a natural purpose as meeting two requirements:

1. "The possibility of its parts (as concerns both their existence and their form) must depend on their relation to the whole."
2. "The parts of the thing combine into the unity of a whole because they are reciprocally cause and effect of their form." (*KU* 373 P)

Hence, "a thing exists as a natural purpose if it is *both cause and effect of itself*" (*KU* 370). A rabbit is such a natural purpose. It does not follow from the concept rabbit that the rabbit exists, but given that a rabbit does exist, its concept partly explains why it persists, that is, continues to exist.

A similar relationship holds between the individual rabbit and the rabbit species. The individual rabbit maintains the existence of the rabbit species by reproducing. Each individual rabbit is brought into existence by other individual rabbits, so that the rabbit species and its characteristic properties cause the individual rabbits to come into existence. Thus, the concept rabbit explains the causal relations by means of which the species brings new particulars into existence and by which these maintain themselves long enough to reproduce themselves, ensuring the continuation of the species. The concept of the rabbit species is thus itself the concept of a natural purpose, and indeed a more

robust natural purpose, for the rabbit species is able to maintain itself over a far longer period of time and against far greater challenges than the individual rabbit.

The concept of a mere abstract universal (e.g. block) is a rule for the mind for sorting things that fall under the given concept from things that do not. The concept of a natural purpose is more than just a rule for the mind: It expresses natural laws that partly explain the self-preservation and reproduction of rabbits. It is, in addition to being a rule for the mind for sorting rabbits from non-rabbits, a rule by which rabbits distinguish themselves from the rest of nature.

Notice how the step from the concept of an abstract universal to the concept of a natural purpose is a step towards satisfying the demands of reason for the unconditioned. The cause or condition for the existence of the block is usually not a block, and even when it is, it is by an unusual circumstance that is itself a condition for the block's coming into existence. These conditions were in turn determined by their own conditions, and so on to infinity. The conditions for the existence of the block are an infinite sequence of causes not given to me in the block itself.

But the cause of a rabbit is in the first instance itself, for the rabbit has been maintaining its own existence. Moreover, the original causes of this rabbit's existence are other rabbits which themselves fall under the concept "rabbit." In a sense then, when I am given a rabbit, I am given not only the conditioned (rabbit), but also the condition (the rabbit species), and the condition of that condition (the rabbit species again), and so on. From the perspective of the species, this chain of conditions for the existence of the

rabbit, unlike for the block, circles back on itself. It hints in this self-contained infinity at an answer to reason's demand for the unconditioned. Kant says that the logical maxim of reason to seek the unconditioned would yield cognition of the unconditioned only if "when the conditioned is given, then *so* is also given (i.e., contained in the object and its connection) the whole series of conditions subordinated one to the other, which is itself [the series] unconditioned" (A307-8/B364). The concept rabbit contains the thought of a series (lineage) of rabbits, each the condition of the next condition, but each of which belongs to the same continuously existing species. Although there are external conditions for the rabbits' existence, we may say (following Kant) that when the conditioned is given, then so is also given (i.e., (partly) contained in the object and its connection) *a* whole series of conditions subordinated one to the other, which is itself [the series] (*partly*) *self-conditioned*. The concept of a rabbit is not an idea of reason yet, but it exhibits some of the structure that an idea of reason would be expected to have.

The thought that natural purposes exhibit some of the structure of an idea of reason would not have surprised Kant. After all, Kant describes the kind of unity that reason seeks as purposive:

"The supreme formal unity, which rests solely on concepts of reason, is the *purposive* unity of things, and reason's *speculative* interest necessitates our regarding all arrangement in the world *as if* it had sprung from the intention of a most supreme reason. For such a principle opens up to our reason, as applied to the realm of experiences, entirely new prospects for connecting the things of the world according to *teleological* laws, and for arriving thereby at their greatest systematic unity." (A687-8/B715-6 P)

If the systematic unity that reason seeks is purposive unity, then it is unsurprising

that reason would find a small measure of what it seeks in the purposive unity of a natural purpose. And indeed, that is just what Kant seems to say:

“A plant, an animal, the regular arrangement of the world edifice (hence presumably also the whole natural order) show distinctly that they are possible only according to ideas. They show that although no individual creature under the individual conditions of its existence is congruent with the idea of the most perfect creature of its kind... yet in the highest understanding these ideas are individual, unchangeable, thoroughly determined, and are the original causes of things.” (A317-8/B374 P)

Kant is speaking of Plato’s metaphysics here, but he is plainly in sympathy with him. He prefaces this quotation by saying: “in regard to nature itself he [Plato] rightly sees distinct proofs of its origin from ideas” (A317/B374). So Kant agrees that natural purposes have at least some of the structure of ideas of reason. Hegel is going to try to show us something that has the complete structure.

The concept of a natural purpose falls short of an idea of reason, or in other words, it fails to be an unconditioned condition of itself in two respects: First, the natural purpose is conditioned by its external environment. The rabbit maintains its own existence only if it can find the food and shelter it needs and is not killed by predators. Although it plays an active role in securing these conditions, it cannot secure them without the cooperation of the environment. But an idea of reason would contain the *complete* conditions for its object’s persistence. Second, even if the concept of a natural purpose completely explained the natural purpose’s own persistence, it does not explain *why* it should exist in the first place. Why are there rabbits but no unicorns? The answer

does not seem to be found in the concept rabbit.¹⁶⁴ The concept of a natural purpose lacks some of the conditions for its object's persistence as well as the conditions for its coming to exist in the first place. Hegel's aim is to argue that the concept of the rational state contains all of the conditions of its persistence and of its coming to exist in the first place. We will approach the concept of the state through a succession of increasingly rational concepts, each of which incorporates into its concept conditions for its existence that remained external to the previous concept.

Human Persistence

The concept of a natural purpose only explains the persistence of its object given appropriate external conditions. For an idea of reason, by contrast, none of the conditions for its persistence would be external. Hence, a natural purpose would need to become independent of external conditions to become an idea of reason. One way we could imagine this happening is for it to cease to have needs. But another more promising way we could imagine this happening is for the things that it needs to become a part of itself. For example, an animal could develop a shell and consequently cease to depend upon its environment for shelter.

But Hegel thinks that self-conscious beings need not *make* the object that meets their needs for that object to be part of themselves. We catch a first glimpse of this in the

¹⁶⁴ For Kant, the idea of God demands that we regard the world as if it had been brought about by divine intention, but only as a regulative principle. I take it that that means that in investigating why rabbits exist, we are to think about the purpose they might serve in their ecosystem. Such an investigation will lead us to observations about their function as consumers of vegetation and as food for predators, thereby furthering science.

section on “Self-Consciousness” in the *Phenomenology*, but Hegel gives us a much clearer presentation in the corresponding section of the *Encyclopedia*. There Hegel articulates his conception of freedom and reason:

“Freedom and reason consist in this: I raise myself to the form of $I = I$, I recognize everything as *mine*, as *I*, I grasp every object as a member in the system of what I myself am,—in short, they consist in this: I have in *one and the same* consciousness *I* and the *world*, in the world I find myself again and, conversely, in my consciousness I have what *is*, what has *objectivity*.” (EG §424Z)

Many philosophers have tried to defend the freedom of the will by showing that external objects do not determine the will to action, that the will is in a certain way independent of its environment. Hegel is proposing to defend the freedom of the will not by showing that the will is independent of its environment, but by showing that its environment is part of the will, that what seem to be external objects are in fact not external. Hence, their determination of the will is not a violation of the will’s freedom.

The free will Hegel is describing is an idea of reason: It is determined or conditioned only by itself. And it is completely self-determined because everything in its world is a part of itself. To make good on this picture of freedom, Hegel needs to show us how seemingly external worldly objects can in fact be part of the will itself. In the “Self-Consciousness” section of the *Encyclopedia*, Hegel outlines a first example of how the will can recognize seemingly external objects as *itself* through the satisfaction of desire, specifically, the desire to consume¹⁶⁵ something.

¹⁶⁵ For the present discussion, I will limit myself to consumption in the narrow (digestive) sense, which I take to be the paradigm case Hegel has in mind. Hegel may have a wider range of kinds of consumption of desired objects in mind, but consumption of food is a good case because it doesn’t depend on things that will only be developed later in the dialectic. For example, since property arises later on the basis of

Hegel describes such a desire as an awareness of a lack within the self-conscious subject that the object of desire would fulfill: “In the object, the subject beholds its *own lack*, its own one-sidedness, sees in the object something belonging to its own essence and yet missing from it” (*EG* §427Z). If the subject is lacking something, then the subject is not its complete self until it has consumed the object of its desire; the object of a desire is a missing part of the subject. Hegel describes the process of consumption as follows: The subject “tak[es] possession of the object whose independence is, so to speak, only a pretence, satisfies itself by consuming it and, since it is an end in itself, maintains itself in this process” (*EG* §427Z).

First, note that Hegel refers to the subject as an end in itself that maintains itself by consuming certain objects. In other words, the self-conscious subject has the structure of a natural purpose, a self-maintaining being, and the unity between subject and object being demonstrated is like the unity of a natural purpose. The outcome of consumption will be that the food is transformed into the subject’s own body and energy, a part of the subject.

Second, notice that Hegel claims that the object’s independence is only a pretence. In completing my lack by consuming the object, I demonstrate, Hegel thinks, that the object was not in fact independent of my desire to begin with, for it was powerless to resist my desire. It only seemed to be independent. Hegel goes on to put this point in other words: “By the satisfaction of desire, the implicit identity of the subject and the object is

developments in this dialectic, to invoke property at this point could create a vicious circularity in Hegel’s argument. This is not to say, however, that aspects of the significance of acquisition of property seem to be illustrated by this discussion of “consumption.” The features of previous stages in the dialectic tend to be preserved and revisited with new complexity.

posited, the one-sidedness of subjectivity and the seeming independence of the object are sublated” (*EG* §427Z).

We tend to think that an apple is an independent object (not a part of myself) right up until the moment that I digest it. But Hegel says that digestion is just proof that the apple was a part of me before it passed my lips. To understand his motivation, I think it is helpful to consider the differences between a vegetative and sensitive soul.

The vegetative soul of the plant does not experience desires for nutrients around it and, on the basis of those desires, act to capture them. The plant simply absorbs what happens to be available to it. The ammonium that will soon be absorbed by a plant is not already a part of the plant because it has not yet entered into the plant’s activity. Of course, plants tend to reach for sunlight and water, but this is not based on any kind of consciousness of the sunlight or water and desire to obtain them. It is a kind of irritability rather than sensitivity, a natural response to a stimulus without consciousness. Hence, the plant cannot be said to form a desire for the sunlight and water it approaches. It is not capable of including anything extra-corporeal in its activity.

By contrast, the sensitive soul of a being with desires is aware of the broader world outside its body and can form intentions regarding objects in accordance with its physical capacities. The sensitive soul need not wait for chance to put nutrients next to its mouth. It spots food from afar and chases it down. For the vegetative soul, nutrition begins in the root. For the sensitive soul, nutrition begins in the eye. The process by which the sensitive soul sustains itself does not begin with chewing, but with hunting or foraging.

Hegel's thought, as I understand it, is that the apple I conceive a desire to eat is already, by the very fact that I have determined myself to eat it, a part of my activity of nutrition, hence already a part of my life, hence, already a part of me. At least insofar as I succeed in following through on that activity; that is, eating it.

This is not to say that we can't draw a boundary between my body and the world on which the apple is not yet a part of my body. The thought is rather that who I am is not just a body, but a living activity that extends out into the world.

The point of this discussion is to argue that seemingly external objects can be internal to a being who makes them a part of its life. And that's a promising step towards finding an idea of reason or an unconditioned object in the world because an unconditioned object must contain all of its conditions, everything it depends on, within itself. The plant depends upon ammonium from outside itself to survive. It is conditioned by something outside itself. I depend upon the apple to survive, but if Hegel is right, from the moment I resolve to eat the apple, it's not external to me.¹⁶⁶ It's a part of me and hence not an external condition of my survival. In this respect, I am more like an idea of reason than the plant, because what is an external condition of the plant's survival is not an external condition of my survival.¹⁶⁷

¹⁶⁶ This claim is complicated by the possibility that someone else has resolved to eat the same apple. Hegel is about to get to that complication in the master/slave dialectic. Who the apple is in fact a part of depends upon who can come out of it as master. I'll discuss this presently.

¹⁶⁷ Hegel supports his claim that the object of my desire is already a part of me by this description of the concept of the object:

"The immediate object must, by its *own* nature, by its *concept*, sublate itself, since in its *individuality* it does not correspond to the *universality* of its concept. Self-consciousness is the *appearing* consciousness of the object itself. In the annihilation of the object by self-consciousness, the object perishes, therefore, by the power of its own concept which is *only internal* to it and, just because of that, seems to come to it *only from*

The problem with this way of recognizing other objects as implicitly oneself through the satisfaction of desire is that it requires their consumption and hence their destruction:

“The satisfaction of self-consciousness caught up in desire, since this self-consciousness does not yet possess the power to endure the Other as an independent entity, destroys the independence of the object, so that the form of the subjective does not attain any subsistence in the object.” (EG §428Z)

As long as self-consciousness must consume objects to incorporate them into to itself, it cannot recognize itself in persisting objects, and so none of the objects that are persisting conditions of the self-conscious subject are part of the self. The apple may be part of me, but the tree that produces the apple is not. Yet I depend upon the tree, for without the tree, there can be no apples.

This leads Hegel to propose another kind of object in which the subject can recognize herself in something external and a corresponding way for that object to be part of the subject. The object is another self-conscious being. And the way in which the self-conscious being becomes a part of the subject is, at first, slavery. In pledging to do the

outside.” (EG §427Z)

When Hegel says in this passage that the object’s individuality does not correspond to the universality of its concept, he means that the object does not contain the structure of the idea, the structure according to which the concept of the object contains the existence of the individual, or in other words, the structure according to which the concept of the object contains the unconditioned condition of the individual and therefore explains the individual’s existence. If the concept of the object were sufficient to explain its existence, then there would be some necessity to its existence. And if there is necessity to its existence, then I shouldn’t be able to deprive it of its existence. If I can make this object cease to exist without any resistance on its part, then the mere concept of this object cannot contain its existence as an individual.

Hegel says that the fact that the object’s individuality does not correspond to the universality of the concept implies that the object must sublate itself and that it perishes by the power of its own concept. The thought here is that an object whose concept does not contain its existence is, by its very concept conditioned by something external. Its very nature is to give itself up to external powers, and when it encounters an end-in-itself, an unconditioned condition like self-consciousness that desires it, it gives itself up to an external power that is not one more link in a chain of conditions but the terminus of such a chain. Hence, Hegel thinks, the object is by its very nature a part of the self-conscious subject.

master's bidding, Hegel thinks, the slave becomes a part of the life of the master.¹⁶⁸

We just saw that the sensitive soul's activity of nutrition includes the activity of foraging or hunting, by which the sensitive soul obtains food. The master's activity of nutrition likewise includes the activity of foraging or hunting, but the foraging or hunting are performed by the slave. The master obtains food by commanding the slave to get it for him. Thus, the slave is part of the master's activity of nutrition, and hence, Hegel thinks, a part of the master.

The idea isn't just that whatever plays a role in my activity is part of me. The ground that I walk on plays a role in my hunting and gathering too, but the ground is not a part of me. What sets the slave apart from the ground is that I give the slave my telos. The ground does not aim to help me hunt. It's just there. But the slave *does* aim to help me hunt. Moreover, the slave aims to help me because I make him do so. A passing hunter who decides to feed me some of his game is not thereby a part of my self-determined activity because I did not determine him to do so. What makes it possible to think of the slave as part of me is that the slave stands to me in the relation that the organs stand to the organism of which they are a part: They are caused by the organism and in turn cause the organism. I cause the slave to do what he does, and the slave causes me to continue to exist by taking care of my needs. Of course, I did not bring the slave into

¹⁶⁸ I think that the basic kind of slavery Hegel has in mind is not yet chattel slavery because property will only develop later in the dialectic. Indeed, property presupposes a kind of mutual recognition that makes reducing another to slavery unnecessary. More on this soon. The kind of slavery that Hegel has in mind here is a dyadic relationship based on fear rather than an institution generated by an entire community. Elements of Hegel's master/slave dialectic may well be exhibited in chattel slavery, since the dialectic preserves and revisits its own previous stages, but in illustrating what Hegel has in mind, it is helpful to choose the simplest cases.

existence, and so I am dependent upon other conditions for the slave. But once the slave is under my control, the slave is a part of my vital activity, my life.

I cannot here do justice to the rest of the master/slave dialectic. Nor can I do justice to arguments common to Rousseau, Kant, and Fichte that one cannot be free as master of slaves, but only in society with other free persons. The important thought for *our* purposes is that there is a more rational alternative to the master/slave relationship by which, rather than making another a subservient part of me, I recognize and am recognized by other free persons as an independent being, and this opens up a new possibility for our relationship to objects: property.

Once humans are brought together in associations of mutual recognition of freedom, a new way of recognizing oneself in nonrational beings becomes available: property. The problem with trying to make an inanimate object that I don't consume into a part of me is that inanimate objects have no loyalty. My spear won't resist someone else coming along and grabbing it. I know the apple is a part of me when I've eaten it. It's too late for anyone else to wrest it from me. Not so with tools. But if I am recognized by others as a free being, a new possibility emerges: They may recognize an object as my property, as subject only to my will. With this recognition, I no longer need to consume and destroy an object to know that it is subject to my will. I have the assurance that the object will obey me in others' recognition of it as my property.

Hegel uses the word "person" to refer to a human being who is recognized as a property owner. A person's property is part of the person. Hegel describes the person's

unity with external objects in terms of the person's placing his will in the object: "A person has the right to place his will in any thing. The thing thereby becomes *mine* and acquires my will as its substantial end [(since it has no such end within itself)], its determination, and its soul" (*PR* §44).

Normally we would think of the soul as the life of the body, but in the person the soul inhabits property as well. The land and the tree become one with me as I give them my telos by plowing and cultivating. They become organs in my extended body. Hegel's thought is that the development of a system of private property changes our form of being in an important way. We're not just rational organisms, but also legal entities called persons, and our property is a part of our person. The advantage of property over mere consumption as a way of discovering ourselves in external objects is that with property we can assimilate not just the apple to ourselves (by destroying it), but also the tree and the land that it grows on (without destroying them).

Here the objection might be raised that the fact that others will not interfere with a tree or some land they recognize me as owning does not mean that the tree or land is subject to my will. The tree may rot. The land may flood. The fact that others don't interfere with my property does not mean that it will obey me.

There are of course limits to our control over our property. But Hegel need not demonstrate that our property completely obeys our will to show that it is a part of us. I need not have total control over something for it to have my telos. The same is true of our own bodies. My arm may get numb and be unable to do some of the things that I would

like it to do, but it nevertheless remains a part of me. Similarly, my land might be subject to the occasional natural disaster and lose my crops, but it nevertheless remains a part of me.

The Ethical Life's Persistence

Still, although the person, or property owner, has more of the structure of the idea, of an unconditioned totality of conditions, than a human being who is not recognized as owning property, the person is still only imperfectly an idea. The person depends upon other persons in forms of sociality Hegel calls Ethical Life. The family is one form of Ethical Life. Only the rationally organized state itself, the highest form of Ethical Life, is fully unconditioned. Such a state, like the person, has the organic structure of a natural purpose in which parts depend upon the whole and in turn determine each other and the whole. The citizens and institutions of the state depend upon the state (not just the formal institutions, but also the economy, culture, and informal practices of the people) for protection, education, goods and services, etc. And the state as a whole is in turn constituted by its citizens and their activities. Moreover, the state has the structure of a person: It has property, including its entire territory, and a governing rational will to control it (in the form of governing bodies). As Hegel puts it, “Ethical Life is... the *concept of freedom which has become the existing world and the nature of self-consciousness*” (PR §142).

Moreover, the state has a power to master its environment that the individual human lacks. Through its sophisticated economy and technology, the state erects dikes to contain floods, develops disease-resistant crops, and in general, makes nature bow down before

it.¹⁶⁹ In giving nature its telos, it makes nature a part of it.

It will be helpful to consider a pair of objections to this claim. First, it might be objected that because the state has to work within natural law, it has to conform to nature, to something else. Hegel would concede the first point and deny the second: We do indeed work within natural law, but natural law is not to be thought of as other than our own will. The reasons for this are complicated and in part go beyond the scope of this dissertation. But we can get some sense for Hegel's position by starting with Kant's thesis that laws of nature are in large part contributions of the form of our understanding (and intuition). Since Hegel rejects things in themselves as contributing matter to these laws, it will turn out on Hegel's view that the laws of nature are wholly generated by reason. Moreover, the distinction between theoretical and practical reason is, Hegel argues, not absolute. So not only are laws of nature products of reason, they may in some sense be objects of practical reason or the will.

Hegel's thought is not that the state is not limited in what it can do, but that whatever limits the state is not other than itself. Even without invoking Hegel's arguments about the rationality of the laws of nature, there is a clear sense in which the laws of nature are involved in the constitution of the state's own existence. Just as you cannot give a good biological characterization of a rabbit without assuming the gravity that keeps it on the earth, you cannot characterize the state's own existence without assuming the laws

¹⁶⁹ Of course, today's critic might point out that whole of society is still dependent upon nature in an important way evidenced by worries about global warming, but we can see how the cooperative structure of society and the development of technology and the means of production make a nation in significant ways unconditioned by many of the vicissitudes of nature.

of thermodynamics that govern its power plants. These laws must enter into an account of what the rabbit or state is. Hence, it is a mistake to think of them as other than the rabbit or state.

The second objection is that there are all sorts of disasters that a state cannot control, for example, hurricanes or meteors. Hegel might want to handle these two examples differently. Hurricanes are regular occurrences for the state and as such belong to a description of its life. Although states cannot stop hurricanes from inflicting damage, they have ways of predicting hurricanes, managing the damage, and recovering from it. Meteorology, dykes, and FEMA are ways that we control such acts of nature.

But what about large meteors that could wipe out states or even humanity? Scientific solutions to such problems may await us, and they might be thought to make future states more rational. Still, they certainly weren't available to the states of Hegel's day. Here I think it's important that these sorts of rare events do not belong to a description of the life of the state. It's true that the state is not immune to destruction, but a state need not achieve that to be free. Recall that in seeking an idea of reason, we were looking for a concept that explained the existence and hence persistence of its object. We could appeal to two kinds of explanation of the persistence of an object: 1) an explanation of why x persists in any possible circumstances and 2) an explanation of why x persists in the circumstances specified by its own concept. The former sort of explanation, if it could be given, would characterize the state as invincible. But the latter sort of explanation would characterize the state as persistent in its normal life conditions. I suggest that Hegel had

the latter sort of explanation in mind.

There are of course other ways in which the state is affected by things outside it. It trades with other states, for example. But Hegel thinks that the rationally organized state performs these activities not out of necessity, but freely. So external factors condition only contingent aspects of the state. The state's existence and its fundamental features do not depend upon external conditions. Hence, the concept of the state contains a complete explanation of the state's persistence.

Ethical Life's Coming Into Existence

Not only does the concept of Ethical Life explain its own persistence, Hegel thinks, but the concept of Ethical Life also explains why it comes to exist in the first place. The question why something exists can be asked with emphasis on its nature or on its existence. We might want to know why *rabbits* exist, as opposed to something else, or why rabbits *exist*, as opposed to not existing. The former question might be answered by an explanation of the sources (perhaps evolutionary pressures) that shaped rabbit nature. The latter calls for an explanation of why there was something whose nature could be so shaped. One such explanation would trace rabbits back through the various species they evolved from to the original source(s) of life.

In the case of free beings, the answer to the former question is contained within their concept, for they give themselves their own ends. A free being chooses its way of life. In choosing the central aims of its life, the free being is, as it were, giving itself a nature. It is choosing the whole that will condition the parts and in turn be conditioned by them.

Hence, while the rabbit's existence is conditioned by whatever brought about the rabbit form of life, the free being itself is what brings about the free being's form of life.

Of course, the free being's choice of life is not arbitrary for Hegel, nor does it consist in a kind of independence from what we tend to think of as limitations: laws of nature, conditions of scarcity, etc. The freedom of the free being is rather an ability to own the conditions of its existence, to incorporate them into itself such that it no longer finds them alien. Not only does such freedom explain the free being's persistence, it also explains its nature.

Hegel argues that Ethical Life is a free being and hence that Ethical Life's end is not merely given to it. The Ethical Life's freedom is realized through the rationality of its members when they are politically organized. In this, Hegel expands on Rousseau's idea that individuals can join their wills together to form a general will. Although the details of Hegel's general will are different, Hegel thinks of his general will as a genuine rational will and hence as genuinely free. Since that will determines its own end, its form of life, it is the author of its own nature. The answer to the question of why the *Ethical Life* exists, as opposed to some other form of political organization, is contained in the concept of the Ethical Life: The Ethical Life exists because it is a free social being, and the form that free social beings give themselves is, ultimately,¹⁷⁰ the state.

But to have a complete explanation of the existence of Ethical Life, we need to know why free social beings come into existence in the first place. Hegel has such a story.

¹⁷⁰ Of course, Hegel thinks that there is a long process by which Ethical Life gives itself the form of the state. We begin in more primitive forms of social organization.

Unfortunately, the story is far too long to examine in detail in a dissertation, let alone a dissertation chapter. As I understand it, the story occupies the vast majority of Hegel's three-volume *Encyclopedia*. I cannot in this chapter get into the details of Hegel's argument that the Ethical Life must exist; I can only clarify what the argument is supposed to show.

The form of Hegel's explanation of the existence of Ethical Life is teleological. Whereas a scientist might explain why *Homo sapiens* exists by tracing its development from past species back to the origins of life, Hegel explains why Ethical Life exists by arguing that Ethical Life is the telos of free beings, and he explains the existence of free beings by arguing that freedom is the telos of self-conscious beings. This teleological chain is traced back ultimately to being and not-being themselves. Hence, the Ethical Life is a telos of being itself.

Hegel's teleological explanation from being to Ethical Life is driven by inner contradiction. Ethical Life is a necessary development of the concept of freedom because without Ethical Life, freedom is in contradiction with itself; hence outside Ethical Life, one cannot be genuinely free.¹⁷¹ Now Hegel thinks that things can exist in a state of contradiction, but only insofar as they are headed for a resolution of sorts. His thought is similar to the thought that an oak seedling is an oak-seedling only insofar as it is developing into an oak tree. Part of what freedom outside of Ethical Life consists in is its development towards Ethical Life. If self-contradictory freedom were not headed towards Ethical Life, then it would not be freedom. So insofar as there are free beings, Ethical Life

¹⁷¹ Hegel's argument that ethical life is a necessary development of the concept of freedom is given over the course of the *Philosophy of Right*.

must develop.

Moreover, there are necessary stages of the development of freedom prior to the rational state, such as the development of a rudimentary system of property. Since these stages are necessary to the development of a rational state, they are in themselves rationally necessary. Hence, in spite of their own inner contradictions, they exhibit a kind of rationality in virtue of their own rational necessity for the development of complete rationality. This thought contributes to Hegel's account of the rationality of the state. Even when states fall short of complete rationality, they can be considered rational insofar as they are necessary steps on the way to complete rationality. This may speak to earlier worries about challenges that contemporary states are unable to deal with. States that are still developing the ability to handle challenges to their existence are the incomplete unfolding of an idea of reason. Like the oak seedling, they are to be understood in terms of that into which they are growing.¹⁷²

Since being considered in itself has a kind of contradiction that must eventually lead to rational free beings, and hence the state, for there to be anything at all, or for anything not to be, Hegel argues, freedom, and hence the state, must ultimately develop. Thus, ethical life must come into existence in a way that rabbits need not. The state comes about as the result of a long series of necessary developments from internal contradictions in previous, imperfect realizations of being itself.

¹⁷² Kant similarly holds that the idea of a civil constitution is a kind of ideal towards which the sovereign should develop the state, but whereas Kant holds that this ideal can in principle never be given empirically, Hegel holds that it can be given empirically.

I take it that Hegel's argument is not supposed to deny the possibility of evolutionary accident. A meteor strike could have wiped out the ancestors of *Homo sapiens* before they had the chance to give rise to *Homo sapiens*. What could not happen, according to Hegel, is that a species of rational beings never comes to exist. For being by its nature develops towards rationality. If rational beings never come to exist, then there was never any being in the first place, for it is the very nature of being to set off a dialectic that results in freedom.

The Reason of Ethical Life Is the Individual's Reason

So far we have seen how Ethical Life is an idea of reason, but we have not yet seen what relationship this idea might have to an individual's practical reasoning. Why, after all, should I recognize the fact that this social order can explain its own existence as having any bearing on what I ought to do? Moreover, any answer to this question should be able to account for the difference between the import of this social order's rationality for a member of the social order and that for a non-member.

I will divide Hegel's argument that the Ethical Life's reason is my reason into two parts. The first part begins with roughly the Kantian understanding of the individual's reason and argues that the Ethical Life's reason must be part of the individual's reason. Because in this part, the individual reasoner is taken as given, some of the Ethical Life's role in creating and constituting the individual reasoner is passed over, and the Ethical Life's reason figures merely as a solution to some problems that face finite reason. In the second part of Hegel's argument we will consider the remainder of Ethical Life's role in

the life of the individual reasoner.

The Ethical Life's Reason Completes the Individual's Reason

We saw in the last chapter that Hegel is not satisfied with Kant's answer to the Antinomy of Practical Reason. Ethical Life is an alternative answer to this antinomy, a god that is capable of providing the individual with happiness in this life.¹⁷³ Kant had been careful to maintain that the moral law itself does not require a solution to the antinomies. That need was a consequence of our finitude, which brings with it a need to have an end. Reason itself can legislate the moral law even without the possibility of the highest good as an end. But to insist that the moral law applies to us regardless of the limitations inherent in our nature as finite beings is, Hegel argues, to give to human beings a law fit only for the nature of God, fit only for the kind of being we are not. Moreover, the fact that Kant refers to the antinomy as an antinomy of practical reason suggests that Kant recognizes that our finitude and our rationality cannot properly be separated.

The very fact that Ethical Life is necessary to solve the Antinomy of Practical Reason suggests that Ethical Life is internal to my practical reason. For without Ethical Life, my practical reason cannot be made self-consistent and hence cannot be rational. In other words, without Ethical Life, my practical reason is not what it is. Ethical Life is part of the complete account of what my will is. We must see how this is.

¹⁷³ Kant's other argument for faith in immortality is that we need immortality to achieve moral perfection. Hegel rejects the moral psychology on which Kant bases this argument. Since this topic goes considerably beyond my main topic, the general will, I will leave it aside.

For Kant, the key to solving the Antinomy of Practical Reason was to find a way to make my virtue the cause of my happiness. Only then could reason's demand for a unified end be answered. Hegel's Ethical Life is intended to unify the practical reasoner's ends in just this way.

Just as in a natural purpose, each part is generated and maintained by the whole and the other parts, so in Ethical Life, each individual is generated and maintained by the whole of Ethical Life and the other individuals that constitute it. Maintaining an individual's existence involves satisfying the individual's needs. For Hegel, happiness is the satisfaction of one's "natural subjective existence – [one's] needs, inclinations, passions, opinions, fancies, etc." (*PR* §123). Hence, to maintain an individual's existence is to contribute to the individual's happiness.

But the social whole's contribution to the individual's happiness goes beyond merely keeping her alive. For the social order to be fully rational, it must be free, and for the social order to be free, its members must be free. In the chapter on Kant we observed that unless others are free, I am not free, so for Hegel, unless the individuals that make up the state are free, the state itself is not free. The argument works slightly differently in the case of individuals and the state because the individuals are part of the state, which forms the individuals, but the gist is the same. For the state to be free, other wills (including the particular wills that make it up) must not be at liberty to interfere in the state's activity. For anyone not to be at liberty to do something, it must either be beyond their power or contrary to obligation. Since interfering in state affairs is plainly not beyond the citizen's

power, it must be contrary to obligation. But for citizens to be obligated to the state, the obligation must be one that they themselves legislate (a principle endorsed by Rousseau, Kant, and Hegel), and hence autonomous. There is one end that humans can be presupposed a priori to have in legislating such an obligation: freedom. Hence, the state can only be free if it has the freedom of the individual as its end.

For its members to be free, the state must provide for their happiness. We've seen Hegel's reasoning on this point in the last two chapters. First, Hegel endorses Kant's thought that freedom must act compatibly with the preservation of its own activity. Since happiness is (roughly) awareness of successful carrying out of one's activity, freedom must seek the object of its own happiness. Second, only when happiness and duty are united into one end can reason truly be satisfied. Hence, only when the law is properly connected with happiness can the will fully recognize the law as its own. As Kant himself puts it, "to need happiness, to be also worthy of it, and yet not to participate in it cannot be consistent with the perfect volition of a rational being that would at the same time have all power" (*KpV* 5:110). If the whole systematically denies a possible happiness to its members, then the will of the whole cannot be consistent with the will of the individual, in which case the individual is not free. In that case, the social whole is itself not free and hence not a genuine Ethical Life.

Just as in a natural purpose, the parts cause and maintain the whole, so in ethical life, the individuals cause and maintain the whole social order by, among other things, performing their duties. Duty is not, in ethical life, the only way that individuals maintain

the whole. The difference between duty and other actions which are necessary to maintain the whole lies in the demand duty makes on us that we do something regardless of whether it promotes our private interest. Our labor in the market economy is necessary to the maintenance of the social whole, but in the market economy we are permitted to pursue our private interest. And our mere habits are necessary to the maintenance of the social whole, but in habit we may not consciously aim at any good. A duty may accidentally coincide with private interest, but its necessity is not contingent on whether it happens to serve the private interest. In duty the pursuit of the good of the whole (or the good of another compatible with the good of whole, or both) is required of me regardless of whether it directly promotes my private interest.

Nevertheless, it is by performing our duties that we collectively enable the state to provide for our happiness. Hence, Hegel maintains Kant's separation between the motive of duty and the motive of happiness while establishing an empirical causal connection between the two. For Kant, an empirical causal connection would undermine the moral motivation (e.g. *KpV* 5:147). Although Hegel differs from Kant with respect to the moral motivation, Hegel's view is sensitive to the worry that too direct a system of rewards would compromise our ability to act from a motive of duty. Fortunately, that is not what the Ethical Life offers. In the Ethical Life, the virtue of all (or most) is the cause of the happiness of all (or most). Because the causal connection is achieved at the group level, it is to some degree indifferent to variations in the individual's virtuousness. Although this contradicts Kant's insistence upon a perfect proportion of virtue to happiness, it does

justice to the arguments behind Kant's insistence. There must be something I can love in the law, namely, the happiness it provides the virtuous. And relatedly, reason must give unity to my ends of happiness and virtue. That unity is to be found in the universal end of the state, which includes the virtue and happiness of all.

An individual who finds that an apparent duty will deprive him of all happiness is registering a flaw in Ethical Life, a remaining contradiction that must still be worked out for freedom to fully flower. But in the truly rational Ethical Life, the sacrifices we are asked to make are not wholly unrewarded. Even the soldier who lays down his life for his country is rewarded with the country's preservation from destruction, with the country's honoring his memory, etc. He may recognize his life as satisfactory in the most important respects.¹⁷⁴

This first argument based on puzzles about the relationship between duty and happiness is confirmed by a second Hegelian claim of chapter 3, that the CI is by itself inadequate to resolve some conflicts among duties. These conflicts seem to stand in need of resolution from the institutions that generate them: The state and the family need to become aligned for practical reason to speak with one voice about my duties towards them. Again, the need for external conditions to allow my practical reason to speak univocally suggests that these external conditions are part of my practical reason.

We have seen two arguments that the reason of Ethical Life completes my reason: 1) It is necessary to solve puzzles about the relationship between duty and happiness, and 2)

¹⁷⁴ However, the international difficulties that make wars necessary are themselves indicative of an incomplete development of the idea in Ethical Life.

it is necessary to resolve conflicts among duties, and hence for the individual's autonomy.

Of course, if Ethical Life were not itself rational, it would not be possible to view it as part of my practical reason. If I depended upon a mere accident for my practical reason to work, my dependence would not make the accident part of my practical reason.

Moreover, not just any kind of dependence upon some rational power makes that power a part of the dependent's reason. I may need your help to accomplish my goals, but that does not make you a part of my practical reason. What makes the reason of Ethical Life part of my reason is that its rational activity is a condition for my own reason's resolving what is to be done in the first place. Hegel's Ethical Life performs in the empirical world what Kant's reason performs only as an ought: building the kingdom of ends, and thereby establishing what I am to do within the kingdom.

It would be natural for Kant to respond at this point that the world we live in obviously falls short of the ideal of the kingdom of ends. Hegel has two responses to this worry. On the one hand, Hegel himself is critical of some aspects of the states of his day. Hegel's thought is not that the states of his day had no imperfections, but that whatever imperfections there were were mere accidents of history that did not obstruct from the philosopher's view the story of the rational unfolding of freedom in the world which had worked through the obvious contradictions within the concept. Just as the small defects we see in an individual human being do not prevent us from recognizing human nature as ordered around its own good, so the defects in the state do not prevent us from recognizing the state's nature:

“Any state, even if we pronounce it bad in the light of our own principles, and even

if we discover this or that defect in it, invariably has the essential moments of its existence within itself (provided it is one of the more advanced states of our time). But since it is easier to discover deficiencies than to comprehend the affirmative, one may easily fall into the mistake of overlooking the inner organism of the state in favour of individual aspects. The state is not a work of art; it exists in the world, and hence in the sphere of arbitrariness, contingency, and error, and bad behaviour may disfigure it in many respects. But the ugliest man, the criminal, the invalid, or the cripple is still a living human being; the affirmative aspect – life – survives in spite of such deficiencies, and it is with this affirmative aspect that we are here concerned.”
(*PR* 258R)

Here Hegel affirms that the key rational feature of a state is that it contains the essential moments of its existence within itself, which I take to mean that its concept contains its existence. The state may have various defects, but as long as it has this essence of the state, it is fundamentally rational.

Kantians may wonder how different this view is from Kant's. Kant also thought of the state as having a sort of ideal form, and he thought that we could treat empirical political organizations as states even though they fell short of this ideal. But whereas for Kant the state's nature is an ideal discovered by pure reason and never to be found empirically, for Hegel, the state's nature is to be discovered within the state by looking closely at the state's workings and the history of its rational development. Hence, for Kant, the idea of the state has no scientific explanatory value. We do not discover the rationality of the state, and the rationality of the state does not explain its existence. The significance of the idea of the state is merely practical: We strive to realize the ideal state. For Hegel, on the other hand, a proper understanding of the state is an understanding of why the state needed to come about. It may be an accident that Saxony was absorbed

into Prussia and not vice-versa, but it is not an accident that the modern state came to exist. Moreover, for Hegel, as opposed to Kant, we learn things about how the state must be ordered from the philosophical study of the empirical state. What we learn is not captured by Kant's formal rules because the rationality of the state must not only provide for the consistency of the state, but also for its existence.

This brings us to Hegel's second response to the Kantian worry: Most apparent imperfections in the state are really imperfections in our understanding of the state. Hegel felt that recognizing the rationality of the actual was a difficult philosophical task, and most people are not able to do so: "Uneducated people delight in argument and fault-finding, for it is easy to find fault, but difficult to recognize the good and its inner necessity" (*PR* 268A). The problem with fault-finding, even by brilliant philosophers like Kant, is that whatever they hold up as more rational than the actual is missing an important mark of rationality: the necessity of its actuality.

"Since philosophy is the *exploration of the rational*, it is for that very reason the *comprehension of the present and the actual*, not the setting up of a *world beyond* which exists God knows where – or rather, of which we can very well say that we know where it exists, namely in the errors of a one-sided and empty ratiocination." (*PR* Preface 20)

Since ideals which are not empirically realized are not actual, they are not true ideas of reason. They do not contain all of the conditions of their existence within themselves. Hence, they are not rational. Of course, one might predict that the state of the future will take on certain characteristics. But that is not typically what such critics of the present state are doing. And Hegel is dubious about whether philosophers are capable of foreseeing such developments.

Ethical Life Creates the Individual's Reason

Ethical Life's role in the individual's life goes beyond harmonizing duties among themselves and duty with happiness to allow the individual's reason to speak univocally. Ethical Life also creates the individual and forms the individual's capacity for reason. Individuals are brought into existence by members of Ethical Life according to reproductive customs of Ethical Life. They are cared for by the Ethical Life's medical resources and protected by the Ethical Life's police services. They are educated by various members of Ethical Life in various educational institutions of Ethical Life and in accordance with the traditions of Ethical Life. Their very existence and ability to reason are a product and end of Ethical Life. They are also in turn a means for the continuation of Ethical Life, for the individual's existence and reason are the means by which Ethical Life carries out its various activities. Hence, the individual's nature is to be member of a rational organic whole. The Ethical Life's reason is the individual's reason not just in the sense that it completes the individual's reason, but also in the sense that it begins the individual's reason.

This leads us to a final oft-repeated Hegelian complaint about Kant: The CI stops at a mere ought, whereas reason ought to be able to explain the existence of something. If the demand of reason is the cognition from concepts, then the CI does not, Hegel thinks, answer to it. Kant's CI is supposed to determine the necessity of an action from pure concepts, but the necessity is only a necessity "if reason completely determined the will" (*KpV* 5:20). But the actuality of the CI can never be exhibited in an empirical action

because there can be no theoretical cognition of a will completely determined by reason, a free will. So the practical cognition yielded by the CI turns out to be cognition not of the sensible world, but of an intelligible world. To give us the unconditioned, Kant has had to remove us from the empirical world it is supposed to condition.¹⁷⁵

By contrast, Hegel's Ethical Life yields cognition of this world from mere concepts, and among the things that Ethical Life determines, as we have just seen, is the existence of the individual as a free being. Hence, in addition to playing a role in legislating my duties, Ethical Life provides the condition that makes the performance of them truly necessary: The existence of the individual who is to perform them. Now if what makes the performance of duty truly rational is its necessity from concepts, then the rationality of a particular performance of duty derives not just from the necessity of a particular performance, given an individual to perform it, but also from the necessity of the existence of an individual to perform it. Hence, for me to recognize my performance of duty as fully rational I must not only recognize the law as rationally necessary, but also my own existence. Ethical Life is what rationally necessitates the existence of its members. Now if what makes my action fully rational is what makes it mine, then this is a further way in which I can recognize the reason of Ethical Life as my own.

A Rational Feeling

For Hegel, as for Kant, individuals who perform their duty recognize the rational as their will. For Kant, this involves recognizing the form of law as the form of their will. For

¹⁷⁵ Or, if you favor a two-aspect reading, Kant has had to remove us from an aspect of the world it is supposed to condition.

Hegel, the matter of the law must additionally be recognized as the matter of my will. Thus, rational motivation involves receptivity (recognition) and therefore sensibility, but sensibility is here compatible with my freedom because what sensibility discovers is not ultimately other than me: Without the whole of which I am essentially a part, I would not be who I am, and thus the whole is not other than me.

Consciousness of one's relationship to the whole need not take the form of an explicit and fully articulated understanding of the concept of the whole to constitute duty.

In the family, this consciousness takes the form of love, a feeling:

“Love means in general the consciousness of my unity with another, so that I am not isolated on my own [*für mich*], but gain my self-consciousness only through the renunciation of my independent existence [*meines Fürsichseins*] and through knowing myself as the unity of myself with another and of the other with me. But love is a feeling [*Empfindung*], that is, ethical life in its natural form.” (PR 158A)

In familial love, I am implicitly aware of myself as essentially part of a whole. I need not articulate this thought to myself for it to be true of me. My awareness is implicit in the way I feel that good or evil to my beloved is good or evil to me. Of course, there are many ways in which another's good or evil might be good or evil to me. There might be a causal connection with me, as when my job is affected by the well-being of my employer. But there is much more than a causal connection at work in familial love. There is also an emotional connection, caring for the other. This caring is, by itself, not distinctive of familial love, for I may care about a stranger or one who does not requite my love. What's special about familial love is that the spouses fuse their lives. They mutually become the means and ends of each others' activity in such a thoroughgoing way that they form one

organic whole. The individuals become the cause of the family as a whole, and the family becomes the cause of its individual parts. As an organic whole made up of rational individuals, the family is a rudimentary form of Ethical Life.

This shared life registers emotionally for the individual not only as a kind of concern for the well-being of the other, but also as feeling of safety. Within the family's unity, one feels safe caring more about another than oneself because one knows that the other requites one's caring. Without this safety, one's beloved's good would not be one's good in an unmitigated sense. My caring about my beloved could easily turn out to be bad for me: It could make me vulnerable to emotional hurt or to physical harm, should I overextend in providing for my beloved. To be aware that my beloved's good is my good and that mine is my beloved's is thus an awareness of our unity, of our being parts of one whole.

As Rousseau pointed out, the ability to concern oneself with duty, rather than one's own appetites brings with it a kind of freedom, "for the impulsion of mere appetite is slavery, and obedience to the law one has prescribed to oneself is freedom" (*SC* 1.8.3). Although the full development of this freedom requires the development of the state and the recognition of law as deriving from one's own will, the love which accompanies the fulfillment of familial duty is a preliminary recognition that familial duty is self-prescribed. In the family, I am set free from slavery to my own appetites by the committed love that allows me to attend to another instead. The feelings of caring and safety in caring constitute an awareness of myself as part of a whole. I view myself as essentially part of this whole and thus as connected to the other part by a shared essence.

Of course, Hegel thinks that the family and familial love are not the most developed, internally coherent form of Ethical Life, but love already exhibits its basic structure. In its most developed form, the modern state,¹⁷⁶ Ethical Life will be recognizable as a more fully articulated whole systematically providing for each of its parts. Some individuals will continue to recognize their unity with the whole only in feelings like patriotism, but others will, through education, learn to recognize the rational unity of the state with its citizens.

Earlier we saw that on Hegel's view we are capable of greater self-determination than mere Kantian autonomy. This greater self-determination is accompanied by an awareness of one's self-determination that is impossible on Kant's account. For Kant, my knowledge of my own freedom is limited to practical cognition of the law and our awareness that theoretical reason does not rule out my freedom. I know what I must do if I am free and that I am licensed to think myself free from a practical point of view, but I never know my empirical action to be free; theoretical cognition of my own freedom is impossible. For Hegel, by contrast, because the causality of reason is to be found in the

¹⁷⁶ Readers may wonder why a cosmopolitan order is not a more complete realization of Ethical Life. I don't find Hegel's remarks on this in the *Philosophy of Right* very persuasive. On the one hand, Hegel argues that it is essential to the state that it be individual, that this requires recognition by other states, and that combining states into a world state would undermine this (*PR* §322 and 322A). The individuality of the state is essential for the state to be independent and free. The thought seems to be that there is a kind of freedom that consists in one's consciousness of oneself as not subject to others' wills, and this self-consciousness requires that there be others. It's not clear to me why the citizens of the state qua private individuals (as opposed to qua members of the state) cannot in a way serve as these others. After all, they sometimes seem to have wills opposed to the will of the state. Hegel also thinks that interstate conflicts play an important role in making citizens conscious of their organic relation to the state. It's not clear to me that natural disaster cannot serve a similar role. Moreover, he observes that relations between states ought to be governed by right, "but since no power is present to decide what is right in itself in relation to the state and to actualize such decisions, this relation must always remain one of obligation. The relationship between states is a relationship of independent units which make mutual stipulations but at the same time stand above these stipulations" (*PR* 330A). This sounds almost like an admission that the world of his day is not fully rational. It seems like Hegel could have revised his view to uncertainty: Philosophy must wait to see whether the state grows old before it can paint its grey in grey.

world, I can know actions in the world to be products of my free will, to be self-determined.

To see why this greater possible awareness of our self-determination might matter to our freedom, consider the case of public law. In the last chapter we saw that the condition of the possibility of a universal will (and hence autonomy) is that it be possible for a law to be willed by the people. However, because the actuality of the universal will cannot be theoretically cognized, fulfillment of that condition is no guarantee that the present law of the land does derive from the universal will of the people. For that to be the case, the will of all the people would need to be the cause of the law. Because theoretical reason does not rule this possibility out, practical reason is able to license a rational faith that public law is legislated by a universal will. Note that this rational faith goes beyond the mere rational faith that, qua practical reasoner, I give myself the moral law. For it is possible that I give myself the moral law without giving myself the public law. Indeed, Kant suggests that is precisely the case when the public law is not one that the people could give themselves: All are bound to obey it by the moral law they give themselves, but it cannot possibly be a law legislated by their wills. The actuality of my will's joining with others in legislating the public law is an additional object of rational faith.

For Hegel, by contrast, the law of the land can be known to be the will of the people, and not merely rationally believed to be the will of the people. The key to such knowledge lies in the recognition that the state is a self-caused being and the recognition of ourselves as members of this being. Because we can know the public law to be our own will, we

experience even the Kantian freedom of autonomy as an object of knowledge. It seems reasonable to say that this knowledge is itself a subjective kind (or component) of freedom: Part of what it is to be free is to know oneself to be free.

Of course, Hegel does not hold that every law of a modern state is rationally necessary. As we have seen, Hegel holds that modern states have their contingent defects; the sovereign can make mistakes. Hegel's point is not that the people can know all of the laws to be their will, but rather that they can know the laws that are in fact rational to be their will.

Hegel's Disagreement with Rousseau

We have seen that for Kant, even though citizens already participate noumenally in the legislative process regardless of whether they participate phenomenally, in the ideal state, citizens have a right to vote (phenomenally) because there is no reason why they would deny themselves this right. For Hegel, on the other hand, the right to vote is far more restricted, and the representatives that those with the right to vote elect play a more limited role in legislation. Like Kant, Hegel has an explanation for how the state's will can be the individual's will that does not require that the individual participate in the way Rousseau envisioned, but for Hegel, direct political participation is not an ideal. Indeed, Hegel thinks that a system of direct participation by its very nature falls short of the idea of reason.

Before getting to Hegel's reasoning for why direct participation is rationally inferior to Hegel's own preferred legislative scheme, it will help to recall that the arguments by

which Hegel establishes that the reason of Ethical Life is the individual's reason did not depend upon the individual's participation in any legislative process. If direct participation is rationally necessary for the general will, then it is not built directly into the conception of what it is for something to be my will.

In a way, Kant opened the door for this sort of argument when he developed his conception of the will as practical reason. Recall that because practical reason is constitutive of the will for Kant, the individual's empirical will is only half of the story of what the individual wills. The individual claims not to will the universal law, but noumenally she wills it nonetheless. What reason does behind the noumenal curtain is every bit as attributable to the individual's will as the individual's actions.

In Hegel's version, noumenal willing is replaced by the Ethical Life. The Ethical Life simply is (part of) the individual's reason. But whereas Kant's noumenal reason could still be thought of as the unanimous willing of essentially individual wills, Hegel's empirical Ethical Life cannot. Whereas Kant introduced the thought of a kind of willing that goes on beyond the empirical individual's activity, Hegel suggested that that willing need not be thought of on the model of unanimity.

The alternative model, that of the organism, was suggested by Rousseau himself. Recall that for Rousseau, a general will in which I participate is mine not so much in the sense that it belongs to me, but in the sense that I belong to it. Hegel adopts Rousseau's view that the individual is a member of the general will and adds to it Kant's thought that the reasoning constitutive of my will is my will even if I don't perform it as an empirical

act. The general will is my will not because I voted on the law, but because I am in a rational organic unity with the legislators who did.

That said, one could always suppose that since Ethical Life is supposed to provide for the freedom of its members, it would be more rational for there to be direct democracy than not. Hegel disagrees, and his statement of disagreement will help us better to characterize the organic unity that constitutes the general will:

“The idea [*Vorstellung*] that *all* individuals ought to participate in deliberations and decisions on the universal concerns of the state – on the grounds that they are all members of the state and that the concerns of the state are the concerns of *everyone*, so that everyone has a *right* to share in them with his own knowledge and volition – seeks to implant [*setzen*] in the organism of the state a *democratic* element *devoid of rational form*, although it is only by virtue of its rational form that the state is an organism [*der nur durch solche Form es ist*]” (PR 308R).

Hegel plainly thinks there is a conflict between direct democracy and the organic structure of an idea of reason. Hegel acknowledges that man is “a *thinking* being with consciousness and volition of the *universal*. But this consciousness and volition remain empty and lack *fulfilment* and actual *life* until they are filled with particularity, and this is [to be found in] a particular estate and determination” (PR 308R). Hegel’s thought is this: the human has a universality that, say, the bee lacks in that the human is able to reason generally about what to do. So far, he agrees with Kant. But, he thinks, an individual’s ability to reason in universal terms does not yet tell an individual what to do, for a rational society is articulated into various roles, and our duties are to be found in those roles.

For Kant, professional duties are merely applications of general principles to particular cases: I owe my boss a certain amount of work because I promised that

amount, and I must always fulfill my promises. And I should file my tax returns because that is the law, and I must obey the law. In each of these cases, we begin from a merely given situation (I have made a promise; the law has been given) and think about what reason demands, given that situation. But as we have seen, Kant does not think that we can recognize the situation itself as caused by reason, for we cannot have cognition of ideas of reason in nature.

If we cannot recognize the situation as determined by reason, Hegel thinks, then we will not be able to see ourselves as fully determined by reason and hence as fully free. Did reason demand that I promise a certain amount of work? Perhaps not. Perhaps the labor market is exploitative, and I had no choice but to promise more than would be rational from the perspective of a universal legislator. Did reason demand that I file my tax returns? Perhaps not.¹⁷⁷ Perhaps a system of taxation based on individuals filing returns is part and parcel of a system of taxation that inevitably allows those with access to good accountants to pay at a lower rate than those without such access. Perhaps some other system is more rational.¹⁷⁸ If our duties are to be fully rational, then they cannot be

¹⁷⁷ Does this mean we are not obligated to file our tax returns on Hegel's account? Here I think we return to the sort of dilemma we investigated last chapter in connection with Hegel's assertion that a merely formal principle cannot harmonize all of our duties. Hegel's motivation for the criticism was not that a moral system should be able to tell people what to do in all circumstances. His thought was rather that Kant's moral system fell short of its own aim, for Kant himself wanted to assert that individuals are in a sense always free, always able to act from duty, where duty is what reason prescribes. Hegel's own system does not assert that people are always free in this sense. There can be genuine moral dilemmas, and in such cases, reason speaks equivocally. If the above case of an unjust tax system is such a moral dilemma, the consequence, I suspect, is not that I am at liberty not to file my taxes, but rather that I am not at liberty no matter what I do. My duties are mutually inconsistent.

¹⁷⁸ In order to avoid misleading the reader with these examples, it's worth noting here that Hegel thinks that there are limits to the level of detail that is subject to justification from the concept of the state. Hegel thinks Fichte oversteps these limits when he perfects "his *passport regulations* to the point of 'constructing', as the expression ran, the requirement that the passports of suspect persons should carry not only their personal description but also their painted likeness. In deliberations of this kind, no trace of philosophy

merely the application of general principles to arbitrary circumstances. The circumstances themselves must also be rational.

Still, you might wonder why making circumstances rational requires a division of labor into various estates. A division of labor certainly improves efficiency, but that is not the only value. Couldn't we make our particularity rational simply by introducing thoroughgoing equality? Hegel thinks not:

“But this consciousness and volition [of the universal] remain empty and lack *fulfilment* and actual *life* until they are filled with particularity, and this is [to be found in] a particular estate and determination. Otherwise, the individual remains a *generic category*, but only within the *next* generic category does he attain his *immanent* universal *actuality*.” (PR 308R)

Hegel's thought is roughly that until society is differentiated into various roles, it isn't rationally organized and hence cannot give rise to a robust set of duties for each. The term for “next generic category” here is “*nächste Gattung*,” or “proximate genus.” Hegel discusses proximate genera in the section of the *Science of Logic* on disjunctive judgments.

There he distinguishes between necessary and merely empirical disjunctive judgments:

“An empirical disjunctive judgment is without necessity; A is either B or C or D, etc., because the species B, C, D, etc., are *found beforehand*; strictly speaking, therefore, there is no question here of an ‘either or,’ for the completeness of these species is only a subjective one” (WL 12.81).

If an exobiologist finds that there are three kinds of animals on a planet, B, C, or D, then

remains” (PR Preface 21). Hegel's thought is not that there is no better or worse way to regulate passports, but that the issues that such details of the law are too affected by accidents of life “in the sphere of arbitrariness, contingency, and error” to be regarded as deriving from the concept of the state itself. That is, the explanation for such details of the law, even when they are as rational as they can be, will always involve mere historical contingencies. Hence such laws themselves will not be fully rational (fully explained by an idea of reason).

it is true that any animal on that planet is B, C, or D, but this statement bears no necessity.

Hegel is looking for necessary disjunctive statements.

However, in searching for necessary disjunctive statements, Hegel is not solely interested in statements in which B, C, and D exhaust all logical possibilities. For example, humans either have either straight, non-straight, or no hair. Presumably, that exhausts all of logical possibility and is thus necessarily true. But that is not the kind of necessity Hegel is interested in, for the fact that humans are divisible into those with straight hair, those with not-straight hair, and those without hair is an accidental property with respect to the human genus. They might as easily have all belonged to one of the three categories: “The empirical species... have their differences in some accidentality or other which is a principle external to them and is not, therefore, *their* principle, and consequently also not the immanent determinateness of the genus” (*WL* 12.81).

In other words, hair type, along with all other merely empirically discovered differentia, is a subjective way of dividing humans up because it doesn’t derive from the concept of the human in the right way. A necessary disjunctive judgment would find the grounds of disjunction within the concept of the human: “This totality has its *necessity* in the negative unity of the objective universal which ... possesses, immanent in it, the simple *principle* of differentiation by which the species are *determined* and connected” (*WL* 12.81). Here Hegel describes the necessity he is seeking as a kind of negative unity of the objective universal: Something within the genus itself necessitates the division into different species. Hegel calls a genus with such a negative universal a proximate genus, a

genus the concept of which immediately divides itself into disjuncts: “Hence the genus is the *proximate* genus of a species, for the latter possesses its specific difference in the essential determinateness of the genus” (*WL* 12.82).

A necessary disjunction, Hegel thinks, would show the species to be contrary and contradictory: “Species are contrary inasmuch as they are merely *diverse*... They are *contradictory*, inasmuch as they exclude one another” (*WL* 12.81). In other words, species are contrary insofar as they aren’t equivalent to each other, but they are contradictory insofar as they are mutually exclusive. A necessary disjunction divides a genus into species that are mutually exclusive.

At the same time, a necessary disjunction shows that the contrariness and contradictoriness of the species is by itself “one-sided and void of truth” (*WL* 12.81). One example Hegel gives of the necessary division of a genus into species is the difference of the sexes (*EL* §220). Readers may balk at the suggestion that biological sexes are contradictory (mutually exclusive). Clearly not all species divide themselves into two sexes. But I suspect that Hegel’s thought is not that all species divide themselves into two sexes, but rather that for at least some of the species that divide themselves into sexes, there is a reason why they do so. There is a biological need that the division into sexes fulfills, for example, the need for genetic diversity in a small population.

If the species deer necessarily divides itself into buck and doe, then to be buck is to have a counterpart in the doe. If there were no doe, there would be no deer, for the doe is essential to the genus of deer. Thus, if there were no doe, there would be no buck. The

buck is what it is by relation to the doe. Of course, if hunters killed all of the does, the bucks wouldn't immediately disappear, but they would immediately lose their essential function and therefore nature as bucks, just as an arm severed from a human body does not immediately disappear (or decompose), but immediately loses its essential function and therefore nature.

Objectors might point out that certain sexually dimorphic animals like the California condor are able to reproduce by parthenogenesis. Hence, the objection goes, the female may be essential to the species, but the male may not. Here a Hegelian might be tempted by the thought that the California condor is simply not the kind of species that exhibits a necessary disjunction in its sexes. But I think that reaction is premature. Although female California condors can reproduce by parthenogenesis, I suspect that male California condors are essential to maintaining the form of the species over significant periods of time. If the males were all killed off, the females would only be able to produce clones of themselves, which would greatly diminish the ability of the species to produce genetic diversity, jeopardizing the species. California condors are not able to reproduce and sustain themselves in such great numbers that their rate of chance mutations can supply adequate diversity, as with much smaller species.

Setting aside whatever debate sexual dimorphism might arouse, I suspect that Hegel would say that necessary disjunctions are evident in the roles of bee and ant colonies and even within the organs of the body, where something general (the hive, the body) must divide itself into constituent parts to have a certain kind of life. Hegel's thought is that

such division is rationally necessary; not that all ways of life must exhibit the particular divisions evident in the beehive, but that such divisions are essential to the existence and persistence of the particular genera that exhibit them.

Of course, there need not be bees or ants. But as we have seen, Hegel thinks that certain ways of life are more rational than others (e.g., the life of a self-conscious, free being is more rational than that of a plant). Because the life of a free being is rational, any disjunctions that are necessary to the existence of such a being are rationally necessary in a more complete sense. And in the case of the state, which is an idea of reason, its necessary disjunctions are absolutely rational.

When the disjunction is in this way essential to the genus, there is a kind of unity in contradiction between the species. The heart is not the lung, but the heart is only a heart because it is a part of a body that contains lungs. Hence, the lung is essential to the heart. In Hegel's paradoxical language, the lung is both contradictory to and one with the heart. Each becomes necessary to the other in a way that the cells of a sea sponge are not. There is no rational necessity to there being one more or one fewer cell in the sea sponge, but there is rational necessity to there being one more or one fewer heart in the body.

Let us return to the passage that referred to proximate genera:

“But this consciousness and volition [of the universal] remain empty and lack *fulfilment* and actual *life* until they are filled with particularity, and this is [to be found in] a particular estate and determination. Otherwise, the individual remains a *generic category*, but only within the *next* generic category does he attain his *immanent* universal *actuality*.” (PR 308R)

In applying the idea of a proximate genus to the estates, Hegel is saying that the

divisions of the ethical life into the various estates are rationally necessary from the concept of ethical life itself. Hegel thinks that the principle division of society is into three estates, the substantial estate (the agricultural estate), the estate of trade and industry, and the universal estate (the governing estate). Since ethical life necessarily divides itself into these estates, they have a deeper unity that underlies their division. Each estate is what it is only by relation to the other estates, and hence a member of any of these estates is what he or she is only by relation to the other estates.

This kind of unity is not enjoyed by a mass with no essential differentia. The farmer's farming is necessary to ethical life in a way that the citizen's voting is not in Rousseau. Of course, the citizen's voting is in one sense necessary to the Rousseauvian general will, because without everyone's participation, the general will cannot be expressed. But in another sense, the citizen's voting is unnecessary, for if the citizen were to die, there would be no need for a replacement in the legislature. It is necessary that all citizens vote, but it is not necessary that there be a particular number of legislators. The Rousseauvian general will continues to be what it is even if it loses many of its constituents. Thus, the citizens can view their duty to participate in voting as rational in the Kantian sense (necessary by reason's universal law, given the circumstances), but not in the Hegelian sense (the circumstances themselves are also rationally necessary).

By contrast, if the farmer in Ethical Life were to die, a replacement would be needed. The amount of farming required by a society is determined by the number of mouths to feed. Since one farmer feeds many mouths, the loss of a farmer reduces the

supply of food far more than the demand. Moreover, in a fully rational society, the number of mouths to feed is not arbitrary but determined by the needs and desires of society as a whole and its members. Thus, the farmer's farming is necessary to the state in a way that the citizen's voting is not. The Hegelian state will not continue to be what it is if it loses its farmers, for it will then not be able to subsist. In other words, Ethical Life in Hegel's state has a kind of unity that the Rousseauvian sovereign lacks: Its parts are necessary to it.¹⁷⁹ From the perspective of members of an estate, this unity means that their profession and professional duties are what they are only by relationship to the other estates.

When Hegel says that we achieve immanent universal actuality only within a proximate genus, he is referring to this unity. The problem with a state with no division of labor is that the members of the state do not experience their labor as necessary. The necessity of their labor is contingent on their existence, which has no necessity. They cannot view such labor as fully rational. By contrast, in a state that is rationally divided, everyone's labor is fully rationally necessary: It is not just what is rationally required of them given their circumstances. The very circumstances are rationally necessary.

Hegel's complaint about Rousseauvian direct participation then is that it fails to generate a general will. If the point of the vote is to unite individual wills, it cannot

¹⁷⁹ Of course, farmers can be replaced with other farmers. And in that sense, the members of the Hegelian state remain, with respect to the state in some way inessential. That is one reason why Hegel thinks that the family is an essential moment of ethical life. In the family, we experience a kind of unity with others in which we are irreplaceable. Still, the sort of unity that we experience with others in the state is not made less significant because we are replaceable.

succeed, because real unity is achieved only when the whole necessarily disjoins itself into the parts. Only when my vote is a necessary contribution to the decision is there a true unity. Of course, this doesn't mean that voting could not achieve some other purpose. It could be a source of information about what the people want. But Hegel is skeptical about voting's value for such purposes:

“If the term ‘the people’ denotes a particular category of members of the state, it refers to that category of citizens *who do not know their own will*. To know what one wills, and even more, to know what the will which has being in and for itself – i.e. reason – wills, is the fruit of profound cognition and insight, and this is the very thing which ‘the people’ lack.” (*PR* §301R)

Hegel did, however, think that a kind of representation of the estates in government was necessary both to supply information to the sovereign and to let the people see their perspective represented in government. While Hegel is opposed to direct democracy, certain kinds of robust representative democracy don't seem to conflict with the essential lines of Hegel's argument.

Conclusion

We began this chapter by considering how Hegel purports to discover ideas of reason in the empirical world. Ideas of reason are essential to Kant's characterization of the general will, and by locating them in the empirical world, Hegel intends to show us that the general will is empirical. Having shown the state as Ethical Life to be an idea of reason, we then considered Hegel's arguments that the Ethical Life's reason is my practical reason, and hence my will. We considered two kinds of arguments.

The first kind assumed Kant's morality as a starting point and showed (by means of

the criticisms from last chapter) that the Ethical Life's reason is essential to provide the conditions for the individual's practical reason to come to univocal conclusions. Because individual practical reason can only take place in the context of the Ethical Life's reason, Hegel concludes that Ethical Life is a part of the individual's complete practical reason.

The second kind of argument considered the source of what was assumed in the first kind of argument: the individual practical reasoner. It was shown that Ethical Life is that source. Only in Ethical Life is the individual practical reasoner's activity shown to be necessary, since only in Ethical Life is the individual practical reasoner's existence shown to be necessary. The individual's practical reason is therefore shown to be the end of and means to the Ethical Life, a part of a unified organic whole.

We concluded by considering Hegel's departure from Rousseau's and Kant's emphasis on the individual's direct participation in legislation. Hegel's position is prefigured in Rousseau's characterization of the individual as belonging to the general will. It is also prefigured in Kant's assertion that the individual's will as practical reason has a kind of activity beyond its empirical activity. For Hegel, the individual's will as practical reason has a kind of activity beyond its activity qua individual; it is the activity of the organic general will to which it belongs. This organic whole is characterized by necessary disjunction, the division of itself according to the rational needs of its own concept. Such rational division makes every part necessary and explains the unity of the parts, since each is a necessary correlate of each other part.

Conclusion

We have seen how Rousseau, Kant, and Hegel each developed a notion of the general will in reaction to their predecessors. Rousseau's general will builds on Locke's thought that the common good is the constitutive aim of the general will and Pufendorf's thought that the general will cannot exist only in virtue of a contract, but must involve a continuous unity of wills. Rousseau's innovation was to insist that the public will is a general will, a genuine shared capacity of willing and not just a shared object of willing. According to Rousseau, only if all the citizens participate in determining it on the basis of a publicly shared conception of the common good can it be a genuine shared capacity.

Kant's version of the general will solves two problems in implementing Rousseau's general will: First, people cannot be guaranteed to share a conception of the common good. Second, people may not all be able to participate in legislation. The first problem is solved by replacing the constitutive aim of the common good with the constitutive aim of freedom under universal law. The second is solved by the idea of noumenal participation in legislation. Both solutions have their roots in the notion of an idea of practical reason. The idea of our external freedom contains both the necessity of our willing in accordance with laws of freedom (universal laws) and our noumenal participation in legislation. We are licensed to accept this idea by facts or acts of reason, in particular the imperative to obey the authority with power over us.

Hegel's version of the general will is intended to solve problems in Kant's moral

theory more generally, problems about the relationship between duty and happiness and about reason's determination of our duties. These problems turn up in Kant's notion of the general will in his insistence that the general will's aim is freedom and not happiness. For Hegel, freedom and happiness are inextricably linked. Hegel argues that we can make out in the structure of the state an empirical idea of reason, that we are essentially parts of it, and that because of how the state creates us, legislates our duties, and harmonizes them with each other and our happiness, we can recognize the state's will as our own.

In offering my interpretations of Rousseau's, Kant's, and Hegel's views, I have made their arguments as compelling as I can. I have no more to offer as justification for the ways that each modified his predecessor's position on the general will. But because history has given Hegel the last word, it seems appropriate to offer a word of response on Kant's and Rousseau's behalf. For it strikes me that just as Hegel found inspiration in aspects of Rousseau that Kant had discarded, so today's philosophers should recognize the strengths of aspects Rousseau and Kant that Hegel discarded.

One strength of Kant's view that Hegel's view seems to give up is its insistence that the categorical imperative is binding on us even when the Ethical Life has failed to harmonize our duty with our happiness. We have seen Hegel's grounds for asserting that such situations put the individual's practical reason at odds with itself. We may accept this argument and accept as a consequence that in harmonizing duty and happiness the Ethical Life offers us a kind of freedom. But there is still, I think, great appeal in Kant's conviction that even when duty and happiness are not properly harmonized, reason

prescribes a form of willing, and hence, that a kind of freedom is available to us in making that form the ground of our actions. Assuming that the form of universal law includes the command to obey the authority who has power over you, this means that there is always a Kantian general will, even if it is at odds with itself.

A strength of Rousseau's view as compared with Kant's and Hegel's is its metaphysical parsimony. Whereas Kant and Hegel must each appeal to their very distinctive metaphysical doctrines to account for the general will, Rousseau gives us a general will that is much less dependent on controversial philosophical assumptions. Although it is hard for a contemporary state to meet Rousseau's requirements, they can be approximated, as he himself suggests in his suggested constitution for Poland. This metaphysical parsimony is particularly helpful in the realm of political philosophy, as Rawls suggested, because it allows a political theory to be the subject of consensus. Moreover, as we saw briefly in discussing Hegel's criticisms of Kant, Rousseau's theory of the general will may plausibly account for obligations in smaller social institutions like marriages or certain working environments.

To these virtues that set Kant's and Rousseau's views apart from their successors' may be added other contributions that were not so much rejected as forgotten: for example, the brilliance of Kant's argument that persons must act under the idea of freedom and hence under the moral law, or in other words that attributability implies laws of freedom. When applied to external freedom, this shows that the laws of external freedom must be presupposed in the very act of opposing such laws, and hence that

opposing such laws cannot be an externally free act. Although Hegel does not reject this argument, he nowhere explicitly endorses it.

Another argument that seems to get lost in Hegel's view is Rousseau's (and Kant's) argument that for me to be free, all must be free. Hegel of course has his own quite complex argument for this thesis, but Rousseau's original thought does not depend upon the rationality of the Ethical Life.

Many other contributions of Kant and Rousseau could no doubt be adduced, but I hope that these examples are sufficient to suggest an alternative to the reading on which Hegel marks the culmination of thought on the general will, as opposed to a rich addition to an already rich tradition.

Before concluding, it's worth returning to the contrast between these views and those of contemporary liberalism with which I began. As I remarked before, contemporary liberalism has largely abandoned its interest in the thought that the state's will must be the will of the people. This shift in interest is closely related to a turn away from theories of legitimacy towards theories of justice. Rousseau begins with the notion of freedom and asks how state power could be reconciled with it, that is, what form of state power could be legitimate. The principles regulating the use of state power arise directly from an account of legitimate state power. I have argued in this dissertation that Kant and Hegel share Rousseau's basic project. In setting aside the question of legitimacy, then, contemporary liberals are approaching questions of justice from a radically different angle.

Nevertheless, vestiges of a theory of a general will are still to be found in the work of its foremost proponent, John Rawls. Rawls counts among the influences on his thinking a Kantian thought about freedom, namely, “that a person is acting autonomously when the principles of his action are chosen by him as the most adequate possible expression of his nature as a free and equal rational being.”¹⁸⁰ Principles that are not fully expressive of our nature as free and equal rational beings are heteronomous; they are not self-rule, but rule by a principle that is not essentially our own.

Rawls expands on this Kantian thought by suggesting that if the principles that govern a well-ordered society are the most adequate possible expression of our nature as free and equal rational beings, then we may achieve a kind of freedom by rationally coming to accept them:

“We may over the course of life come freely to accept, as the outcome of reflective thought and reasoned judgment, the ideals, principles, and standards that specify our basic rights and liberties and effectively guide and moderate the political power to which we are subject. This is the outer limit of our freedom.”¹⁸¹

Crucial to the thought that this is the idea that the principles of justice “guide and moderate the political power to which we are subject.” In a well-ordered society, this happens because

“citizens have a normally effective sense of justice, that is, one that enables them to understand and apply the publicly recognized principles of justice, and for the most part to act accordingly as their position in society, with its duties and obligations,

¹⁸⁰ *Theory of Justice* p. 222

¹⁸¹ *Justice as Fairness* p. 94

requires.”¹⁸²

If understanding and applying publicly recognized principles of justice is being ruled by them, and if accepting those principles as a result of reflective thought and reasoned judgment is willing them, then Rawls can say that the will that rules me in a just society is my will. But he cannot say that I rule myself; only that I agree to the way that I am ruled. For the will that rules me is my will only in the sense that the ruling will and my will have the same content. It is not the same capacity of willing. Unless, that is, Rawls were to offer us a metaphysics of willing to explain how the power by which the conception of justice that rules society is in some way an activity of my will.

There is perhaps a Hegelian reading the third part of *A Theory of Justice* that draws out this possibility. There Rawls discusses the stability of the society governed by his conception of justice, which is to say its persistence. Rawls argues that a society governed by his conception of justice would be stable because people raised in such a society would come to adopt the conception of justice and act on it. Treatment of individuals in accordance with the conception of justice would be part of the reason why the individuals would adopt the conception of justice. Moreover, such a conception of justice and the society that implemented it would harmonize duties with happiness. These resemblances to Hegel’s version of the general will suggest the possibility of a Hegelian argument that Rawlsian principles of justice are the will of the people not only qua content, but qua capacity. On such an account, the members of a well-ordered Rawlsian

¹⁸² *Ibid.* p. 9

society could regard their own wills as organs of a united general will.

But Rawls never makes such an argument. He takes the Hegelian point about persistence as a reasonable demand on a political theory, but not as a step towards an account of the state as an empirical idea of reason. He deflates the Rousseauvian thought that we can be free if we govern ourselves to the contemporary thought that we can be free if we rationally accept the principles by which we are governed. And he derives his own claim that we can rationally accept only principles that express our nature as free and equal beings from Kant's thought that we know ourselves to will laws of freedom a priori.

Rawls's *Theory of Justice* would no doubt have lost some of its audience had Rawls hewn closer to Rousseau, Kant, and Hegel. Rawls's careful avoidance of metaphysical claims is surely part of the reason it has such widespread appeal. But the libertarian reaction to Rawls is perhaps a side-effect of his transforming a theory of freedom into a theory of fairness. Rawls has the materials for a particularly robust answer to the question of how his view respects the individual's freedom, but he sets some of them aside, which cannot but encourage his opponents to complain that freedom is being subordinated to fairness.

The promise of the idea of the general will is to unite considerations of justice and legitimacy under a single guiding thought: That humans are free and must therefore rule themselves. It is to the great credit of Rousseau, Kant, and Hegel that they derived rich political theories from this thought.

Bibliography

Works by Rousseau

- [D2] *Discourse on the Origins of Inequality* in *The Discourses and Other Early Political Writings*. Victor Gourevitch, ed. and trans. 2nd edition. *New York, NY, USA: Cambridge University Press*.
- [E] *Emile: Or, on Education*. Bloom, A, ed. and trans. New York: Basic Books, 1979.
- [GP] *Considerations on the Government of Poland* in *The Social Contract and Other Later Political Writings*. Victor Gourevitch, ed. and trans. 2nd edition. New York, NY, USA: Cambridge University Press, 2019.
- [SC] *The Social Contract* in *The Social Contract and Other Later Political Writings*. Victor Gourevitch, ed. and trans. 2nd edition. New York, NY, USA: Cambridge University Press, 2019.

Works by Kant

Citations (except to the first Critique) are to the Akademieausgabe ('Ak.'). Translations are from the Cambridge Edition of the Works of Immanuel Kant (ed. Guyer & Wood) (volume name in brackets if needed) unless otherwise indicated.

Critique of Pure Reason (cited with A/B page numbers) (when marked "P," the citation is to the Pluhar translation: Werner S. Pluhar, trans. Indianapolis, Ind.: Hackett Pub. Co., 1996.)

[A] *Anthropology from a Pragmatic Point of View*

[G] *Groundwork of the Metaphysics of Morals* [Practical Philosophy]

[KpV] *Critique of Practical Reason* [Practical Philosophy]

[KU] *Critique of Judgment* [Practical Philosophy] (when marked "P," the citation is to the Pluhar translation: Werner S. Pluhar, trans. Indianapolis, Ind.: Hackett Pub. Co., 1987.)

[MS] *Metaphysics of Morals* [Practical Philosophy]

[R] *Religion Within the Boundaries of Mere Reason*

[TP] “On the common saying: That may be correct in theory, but it is of no use in practice” [Practical Philosophy]

[VR] “On a supposed right to lie from philanthropy” [Practical Philosophy]

Works by Hegel

[EG] *Philosophy of Mind*. Michael Inwood, ed. W. Wallace and A.V. Miller, trans. Oxford, United Kingdom: Oxford University Press, 2007.

[GC] “The Spirit of Christianity and Its Fate” in *Early Theological Writings*. T.M. Knox, trans. Philadelphia: University of Pennsylvania Press, 1975.

[NG] *The Scientific Ways of Treating Natural Law, Its Place in Moral Philosophy, and Its Relation to the Positive Sciences of Law*. T. M. Knox, trans. Philadelphia: University of Pennsylvania Press, 1975.

[PR] *Elements of the Philosophy of Right*. Allen Wood, ed. H.B. Nisbet, trans. Cambridge, United Kingdom: Cambridge University Press, 1991.

[PS] *The Phenomenology of Spirit*. Terry Pinkard, ed. and trans. Cambridge, United Kingdom: Cambridge University Press, 2018.

[WL] *The Science of Logic*. George Di Giovanni, ed. and trans. New York: Cambridge University Press, 2010.

Works by Hobbes

[L] *Leviathan*. Richard Tuck, ed. Cambridge, United Kingdom: Cambridge University Press, 1991. (Page numbers refer to the original edition, given in brackets in this edition)

Works by Locke

[2TG] *Second Treatise of Government* in *Two Treatises of Government*

Works by Pufendorf

[JNG] *De Jure Naturae et Gentium Libri Octo*. William Abbott Oldfather, trans. London: Clarendon Press, 1934.

Works by Others

Auxter, T. (1979). "The Unimportance of Kant's Highest Good." *Journal of the History of Philosophy*, 17(2), 121-134.

Barry, B. (1965). *Political argument* (International library of philosophy and scientific method). New York: Humanities Press.

Beck, L. (1960). *A commentary on Kant's Critique of practical reason*. Chicago: University of Chicago Press.

Bertram, C. (2012). "Rousseau's Legacy in Two Conceptions of the General Will: Democratic and Transcendent." *The Review of Politics*, 74(3), 403-419.

Chapman, J. W. (1956). *Rousseau—Totalitarian or Liberal*. New York: Columbia University Press.

Charvet, J. (1995). "Rousseau, the Problem of overignty and the Limits of Political Obligation." In Wokler, R. *Rousseau and liberty* (pp. 139-151). New York: Manchester University Press ; Distributed in the USA and Canada by St. Martin's Press.

Ci, Jiwei. (2013). Agency and Other Stakes of Poverty. *The Journal of Political Philosophy*, 21(2), 125-150.

- Cohen, J. (2010). *Rousseau: A Free Community of Equals* (Founders of modern political and social thought). Oxford: Oxford University Press.
- Dent, N. (1989). *Rousseau: An introduction to his psychological, social, and political theory*. Oxford, UK; New York, NY, USA: B. Blackwell. Dent, N. (2005). *Rousseau* (Routledge Philosophers). London; New York: Routledge.
- Dworkin, Ronald. (1977). *Taking Rights Seriously*. Cambridge: Harvard University Press.
- Dworkin, Ronald. (1983). "In Defense of Equality." *Social Philosophy & Policy* 1 (1): 24.
- Engstrom, Stephen. (2015). "The Complete Object of Practical Knowledge." In *The Highest Good in Aristotle and Kant* (pp. 129-157). Oxford: Oxford University Press.
- Engstrom, Stephen. (2016). "The Determination of the Concept of the Highest Good." In Höwing, Thomas. *The highest good in Kant's philosophy* (pp. 33-49). Berlin: De Gruyter.
- Fichte, Johann Gottlieb. (2000). *Foundations of Natural Right: According to the Principles of the Wissenschaftslehre*. Frederick Neuhouser, ed., Michael Baur, trans. Cambridge, UK ; New York: Cambridge University Press. (Citations are to academy pagination)
- Giddin, H. (1983). *Rousseau's Social contract: The design of the argument*. Chicago: University of Chicago Press.
- Jones, W. (1987). "Rousseau's General Will and the Problem of Consent." *Journal of the History of Philosophy*, 25(1), 105-130.
- Kleingeld, Pauline. (2016). "Kant on 'Good', the Good, and the Duty to Promote the Highest Good." In Höwing, Thomas. *The highest good in Kant's philosophy* (pp. 33-49). Berlin: De Gruyter.
- Korsgaard, Christine. 2018. "The Claims of Animals and the Needs of Strangers: Two Cases of Imperfect Right."
- Kymlicka, Will. 2002. *Contemporary Political Philosophy : an Introduction*. 2nd ed. Oxford ; New York: Oxford University Press.

- Levine, A. (1976). *The politics of autonomy: A Kantian reading of Rousseau's Social contract*. Amherst: University of Massachusetts Press.
- Masters, R. (1968). *The political philosophy of Rousseau*. Princeton, N.J.: Princeton University Press.
- Melzer, A. (1990). *The natural goodness of man: On the system of Rousseau's thought*. Chicago: University of Chicago Press. Neuhouser, F. (1993). "Freedom, Dependence, and the General Will." *The Philosophical Review*, 102(3), 363-395.
- Murphy, J. G. (1965). "The Highest Good as Content for Kant's Ethical Formalism (Beck "versus" Silber)." *Kant-Studien*, 56(1), 102.
- Nance, Michael. 2012. "Kantian Right and the Categorical Imperative: Response to Willaschek." *International Journal of Philosophical Studies: IJPS* 20 (4): 541–56.
- Nelson, Eric. 2019. *The Theology of Liberalism: Political Philosophy and the Justice of God*. Cambridge, Massachusetts: The Belknap Press of Harvard University Press.
- Neuhouser, Frederick. (1993). "Freedom, Dependence, and the General Will." *The Philosophical Review*, 102(3), 363-395.
- Noone, J. (1980). *Rousseau's Social contract: A conceptual analysis*. Athens: University of Georgia Press.
- Pogge, Thomas W. (2002) 'Is Kant's Rechtslehre a "Comprehensive Liberalism"?' in Mark Timmons (ed.) *Kant's Metaphysics of Morals: Interpretative Essays*, Oxford: Oxford University Press, pp. 133–58.
- Rawls, J., & Freeman, S. (2007). *Lectures on the history of political philosophy*. Cambridge, Mass.: Belknap Press of Harvard University Press.
- Rawls, John. 1999. *A Theory of Justice*. Rev. ed. Cambridge, Mass.: Belknap Press of Harvard University Press.
- Rawls, John, and Erin Kelly. 2001. *Justice as Fairness: A Restatement*. Cambridge, Mass.: Belknap Press of Harvard University Press.

- Reath, A. (1988). "Two Conceptions of the Highest Good in Kant." *Journal of the History of Philosophy*, 26(4), 593-619.
- Recki, Birgit. (2016). "'Mixtum Compositum': On the Persistence of Kant's Dualism in the Doctrine of the Highest Good." In Höwing, Thomas. *The highest good in Kant's philosophy* (pp. 71-88). Berlin: De Gruyter.
- Ripstein, Arthur. 2009. *Force and Freedom: Kant's Legal and Political Philosophy*. Cambridge, Mass.: Harvard University Press.
- Seel, Gerhard. 2009. "How Does Kant Justify the Universal Objective Validity of the Law of Right?" *International Journal of Philosophical Studies: IJPS* 17 (1): 71–94.
- Sreenivasan, Gopal. (2000). "What Is the General Will?" *Philosophical Review*, 109(4), 545-581.
- Sussman, David. (2015). "Why Some Things Must Remain Unknown: Kant on Faith, Moral Motivation, and the Highest Good." In Aufderheide, J., & Bader, R. *The highest good in Aristotle and Kant*. Oxford, United Kingdom: Oxford University Press.
- Talmon, J. (1952). *The rise of totalitarian democracy* (Beacon studies in freedom and power). Boston: Beacon Press.
- Willaschek, Marcus (1997) 'Why the Doctrine of Right Does Not Belong in the Metaphysics of Morals: On some Basic Distinctions in Kant's Moral Philosophy', *Jahrbuch für Recht und Ethik* 5: 205–27.
- Willaschek, Marcus. 2009. "Right and Coercion: Can Kant's Conception of Right Be Derived from His Moral Theory?" *International Journal of Philosophical Studies: IJPS* 17 (1): 49–70.
- Willaschek, Marcus. 2012. "The Non-Derivability of Kantian Right from the Categorical Imperative: A Response to Nance." *International Journal of Philosophical Studies: IJPS* 20 (4): 557–64.

- Williams, David Lay. (2015). "The Substantive Elements of Rousseau's General Will." In Williams, D., & Farr, J. *The general will: The evolution of a concept* (pp. 219-246). New York, NY: Cambridge University Press.
- Wolff, R. (1970). *In defense of anarchism* (Harper Torchbooks; TB 1541). New York: Harper & Row.
- Wood, Allen (2002) 'The Final Form of Kant's Practical Philosophy', in Mark Timmons (ed.) *Kant's Metaphysics of Morals: Interpretative Essays*, Oxford: Oxford University Press, pp. 1–21.