KANT'S THEORY OF JUSTICE

I. INTRODUCTION

This chapter explores the unity of Kant’s practical philosophy by examining the relationship between politics and morality in his writings. Kant's attempt to ground principles of justice (Recht) in a notion of pure practical reason, it will be argued, constitutes a distinctive feature of his liberalism that can be contrasted with attempts to ground such principles in a notion of instrumental reason and self-interest (Hobbes) or in a theory of natural rights (Locke). According to Kant, “A system of politics cannot take a single step without first paying tribute to morality.” This tribute consists of nothing less than the concept of autonomy (or positive freedom)—the capacity to act from laws that the agent gives to himself—serving as the cornerstone for his notion of a just political order:

In fact, my external and rightful freedom should be defined as a warrant to obey no external laws except those to which I have been able to give my own consent. Similarly, external and rightful equality within a state is that relationship among citizens whereby no one can put anyone else under a legal obligation without submitting simultaneously to a law which requires that he can himself be put under the same kind of obligation by the other person.  

Of course, to maintain that the concept of autonomy provides the cornerstone for Kant's theory of justice is not to claim that citizens will in fact always act from laws that they could (rationally) give to themselves, either individually or collectively. Kant describes this ideal as the "kingdom of ends" or the "kingdom of God on earth" precisely to contrast it with the notion of a just political order. Rather, the claim is that his theory of justice
must assume that individuals are beings capable of acting in this way, if the project of grounding principles of justice in a notion of pure practical reason (in contrast to empirical practical reason) is to succeed.

This claim concerning the unity of Kant’s practical philosophy may seem suspect to those already familiar with his political theory. After all, Kant not only drew a sharp distinction between the realm of legality and the realm of morality, he also claimed that progress in the former does not insure any improvement in the latter. Moral improvement consists in greater conformity of the agent’s maxims to the categorical imperative, but no amount of coercive legislation can create a good will. Moreover, Kant believed that the task of creating a just political order could be solved by a “race of devils” as long as they possessed understanding—that is, Verstand in contrast to Vernunft. And, as commentators have been quick to note, if the task can be solved by devils it surely can also be settled by “Benthamites trying to maximize their utility function.” Such views would not seem to permit a great deal of unity in Kant’s practical philosophy and, therefore, also seem to threaten the distinctiveness of his political liberalism.

The first reply that might be made to this objection would be to insist on the familiar Kantian distinction between validity and genesis. That is, while the validity of a just political order lies in its conformity to “laws of freedom” and, ultimately, the constraints of practical reason, its historical genesis need not (and generally does not) conform to these requirements. However, although this sort of reply is open to Kant, it does not get to the heart of the objection. The question is whether the principles agreed to by interest-maximizing Benthamites would be the same as those specified in relation to a normative concept of practical reason. If so (as the “race of devils” passage suggests), the distinctiveness of Kant’s theory would be lost since practical reason would not provide a unique criterion for determining the legitimacy of the rules which define the political order.

A second reply might proceed as follows: As many commentators have noted, Kant is able to insist on the harmony between principles agreed to by a race of devils and principles grounded in practical reason only by assuming a teleological conception of history. The providential hand of nature insures that what individuals ought to create will be brought about despite their unwillingness. However, Kant’s predictions about the course of historical events have not fared any better than Marx’s. Nature has produced neither just political orders nor a condition of international perpetual peace. If Kant’s teleological conception of history is unjustified (as I believe it is), what consequences does this have for his assumption about the unity of practical philosophy? Although Kant claimed that his assumptions about what nature would inevitably produce were secondary or subsidiary to the moral principles grounded in practical reason, my own view is that his teleological con-
ception of history has influenced his substantive views on social justice. I develop this objection in the conclusions to this chapter where I indicate a number of problems raised by Kant’s two-world doctrine and the peculiar status he ascribes to property rights in his theory. However, rather than developing these criticisms here, let me indicate what I take to be some of the virtues of Kant’s political theory.

The distinctiveness of Kant’s liberal political theory, I will argue, lies in his attempt to ground principles of justice in the notion of pure practical reason or, as I prefer to speak of it, a conception of the person as a free and equal moral being. This project provides an alternative to the tradition of “possessive individualism,” be it that of Hobbesian rational self-interest or that of Locke’s natural law theory. Kant does not begin with a natural right to private property (Locke) nor does his theory of the social contract rest upon considerations of rational self-interest (Hobbes). Rather, Kant’s aim is to specify a set of basic rights and a criterion of political legitimacy (the social contract) with reference to a notion of practical reason that cannot be reduced to instrumental reason or self-interested rationality. At the same time he also rejects a more metaphysical conception of reason, such as can be found in the classical tradition of natural law.

The attempt to ground principles of justice in a notion of practical reason also provides the basis for Kant’s distinction between the right and the good, or questions of justice and questions of the good life. Whereas principles of justice are based on what all can agree to as free and equal moral beings, there exists a plurality of irreconcilable conceptions of the good. Kantian liberalism affirms a plurality of conceptions of the good within the limits of justice or right. Although this distinction is certainly controversial, it will be important to see how Kant defends it and where the strengths and weaknesses of his arguments lie.

Finally, a third distinct feature of Kant’s liberalism concerns the status of property rights in his theory. They are based not upon a natural right to the objects of one’s labor, nor upon considerations of social utility or efficiency. Rather, however problematic in the end, Kant develops a theory of property rights with reference to a conception of the person or moral autonomy. Further, property rights remain provisional and subject to the united agreement of all in the social contract and the realization of international peace among nations.

Several commentators have noted a tension between the theory of property rights (and justification of coercive legislation) developed in the Metaphysics of Morals and the idea of the social contract found primarily in Kant’s “occasional” writings on politics and history. I believe that this interpretation is mistaken, due in part to the failure to perceive the unique function of the “Universal Principle of Justice” (UPJ) in the Metaphysics of Morals, and
the idea of the social contract in Kant’s political theory. The UPJ is primarily introduced to provide a justification for the use of coercion, while the idea of the social contract provides a test for just legislation. However, both notions have their roots in the concept of autonomy or practical reason.

I should also add that the following exposition of Kant’s practical philosophy differs from several recent studies with regard to its method of interpretation. Rather than relying primarily upon Kant’s “occasional” writings, I take the *Metaphysics of Morals* to be his most important and systematic discussion of topics in political philosophy. The general neglect of this work is apparently due, at least in part, to the widespread conviction, voiced quite early by Schopenhauer, that Kant’s philosophical abilities were well beyond their prime at the time of its publication (1797) and that, consequently, it should not be included among his “critical” works. This conviction is often accompanied by the hope that such essays as “Theory and Practice” and “Perpetual Peace,” or even the third Critique, will offer a better point of departure for reconstructing Kant’s practical philosophy. Despite the many difficulties contained in the *Metaphysics of Morals*, I believe its neglect is unjustified and that we must look to it for Kant’s systematic views. This is not to suggest, however, that the occasional writings are unimportant. On the contrary, they often fill out and illuminate views that Kant formulates all too briefly in the *Metaphysics of Morals*.

II. JUSTICE AND MORALITY IN KANT

A. Autonomy and the Moral Law

In his introduction to the *Metaphysics of Morals*, Kant lists “moral personality” as one of the concepts common to both the Rechtslehre and the Tugendlehre. It is defined as “the freedom of a rational being under moral laws,” and contains two aspects: the capacity to have a conception of the good, that is, to frame and pursue ends of our own choosing rather than simply to adopt ends given to us by nature (*Mds*, 392; *KU*, 427); and the capacity to act in accordance with, as well as out of respect for, the Moral Law or, derivatively, laws that express our freedom and autonomy (*Mds*, 223). Moral personality can thus be defined as the capacity to act autonomously. Although the Rechtslehre does not assume individuals will in fact act autonomously, it does presuppose that individuals equally share this capacity. Similarly, the Rechtslehre presupposes the objective validity of the Moral Law (in Kant’s sense). Kant refers twice to the conclusion reached in the second Critique, namely, that we become aware of our freedom or autonomy.
only through the Moral Law (MdS, 226, 239). According to Kant, apart from the objective validity of the Moral Law, there would be no moral authority to compel others to form a civil society under public laws, or, more broadly, no moral basis for (coercively) enforcing claim-rights against others. However, although the Metaphysics of Morals presupposes both the concept of autonomy and the objective validity of the Moral Law, it does not offer any arguments for either. For these we must turn, however briefly, to Kant’s earlier writings on morality and, in particular, to the Critique of Practical Reason.

In a well-known footnote in the preface to the second Critique, Kant states that while freedom is the ratio essendi of the Moral Law, the Moral Law is the ratio cognoscendi of freedom, that is, the Moral Law requires freedom for its objective validity, yet we are able to assume our freedom only because we can first become conscious of the binding character of the Moral Law upon us (KpV, 4n). Kant develops this position (together with the related doctrine of the “Fact of Reason”) as an alternative to the “ vainly sought deduction” of freedom promised earlier in the Groundwork (KpV, 47). And, even if Kant did express some dissatisfaction with this alternative at the end of his career, it must be considered his final position on the matter. I cannot take up here the considerations that may have led Kant to abandon his efforts to provide a theoretical proof of freedom and replace it with the doctrine of the Fact of Reason. However, three remarks are relevant for our own “constructivist” interpretation of Kant’s moral philosophy, and for a clarification of the dependence of the Metaphysics of Morals on the second Critique, particularly for the notion of autonomy and the objective validity of the Moral Law.

First, the Fact of Reason, or consciousness of the Moral Law, is not a rational insight into a moral order that can be grasped apart from or prior to the conception of ourselves as moral agents. This is the interpretation argued for by Patrick Riley in opposition to constructivist views. Kant, however, consistently opposed such intuitionism as dogmatic and a violation of human autonomy. For Kant, the notion of moral personality entails that “a person is subject to no laws other than those that he (either alone or at least jointly with others) gives to himself” (MdS, 223), and Kant would have considered this a criticism of rational intuitionism, “divine command” theory, and other versions of moral realism. Rather, the objective validity of the Moral Law is based on the fact that it is a law that the rational agent gives to himself in accordance with the structure of practical reasoning. According to Kant, it is also a law which the moral agent becomes conscious of in the process of constructing maxims of the will (MdS, 225; KpV, 28). I will discuss the role of maxims for understanding the categorical imperative as a test procedure in section C; here I only wish to note its importance for understanding Kant’s
notion of consciousness of the Moral Law as a Fact of Reason: As moral agents we construct or adopt maxims or general rules of conduct. We are not simply beings who reason calculatively in light of pregiven ends. Rather, we are able to adopt various ends and pursue different courses of action. It is in the process of constructing maxims (or in reflection upon this process) that we become conscious of the binding character of the Moral Law. We experience guilt or shame when our conduct violates this capacity in others, and we experience admiration for others and respect for ourselves when we act upon maxims that acknowledge this capacity in others and in ourselves (see KpV, pt. 3).

Second, Kant’s position in the second Critique is that the objective validity of the Moral Law is not known through something outside of, or independent of, reflection upon our moral experience. This may be Kant’s most significant departure from the Groundwork. The doctrine of the Fact of Reason is not an argument from an independent or more narrow conception of rationality (e.g., theoretical reason or rational self-interest) to the validity of the Moral Law. The claim is rather that the Moral Law is a principle that represents to us what is already present and effective in common moral experience. (This explains, incidently, why Kant did not consider it an objection when a critical review noted that the second Critique did not provide a new principle of morality, but only a new formula [KpV, 8 n. 5].) The Fact of Reason is, to use a juridical metaphor, a fact confirmed by the testimony of witnesses who attest to the binding character of the Moral Law upon them. The achievement of the first Critique was to show that freedom (and hence the Moral Law) are at least theoretically possible, however, the objective validity of the Moral Law cannot be established by means of non-moral considerations. This interpretation also coincides with Kant’s most developed position in the Religion where he argues that to act against the Moral Law does not mean that we act irrationally (in the narrower sense), although in so acting we deny our autonomy.

Third, Kant’s argument from the objective validity of the Moral Law to the notions of freedom and autonomy should only be viewed as a deduction in a “weak sense.” That is, it does not demonstrate whether we are free and autonomous, but how it is possible to regard ourselves as free and autonomous agents. The Fact of Reason, once it has been introduced into court as evidence, can be used as a “credential” for establishing our freedom and autonomy. It adds to the negative conclusion of the first Critique—that freedom is at least possible—the positive conclusion of how there can be “a reason which determines the will directly through the condition of a universal lawful form of the maxims of the will” (KpV, 48). The “positive concept of freedom” (the capacity of reason to be of itself practical) is revealed “through the subjection of the maxim of every action to the condition of its fitness to be a universal law” (MdS, 214). Thus, by virtue of our conscious-
ness of the Moral Law as a Fact of Reason, we are entitled to view ourselves not only as part of a phenomenal world, but also as members of a noumenal realm. The notions of autonomy, the noumenal self, and the objective validity of the Moral Law are thus all presupposed in the Metaphysics of Morals:

In the theory of duties, man can and should be represented from the point of view of the property of his capacity for freedom, which is completely supersensible, and so simply from the point of view of his humanity considered as a personality, independently of physical determinations \((\text{homo noumenon})\). \((\text{MdS}, 239)\)

These preliminary (and controversial) remarks are no doubt insufficient given the complexity of Kant’s views, but they must suffice for our own purposes and we can now turn to a closer examination of the Metaphysics of Morals. I will return to some further questions in Kant’s moral philosophy (for example, the interpretation of the categorical imperative as a test procedure) when I consider the relation between the Moral Law and the Universal Principle of Justice.

B. The Moral Law and the Universal Principle of Justice

Contrary to many interpretations, and contrary even to some of Kant’s own remarks, the Metaphysics of Morals does not divide neatly into a discussion of “duties of justice” in the Rechtslehre and “duties of virtue” in the Tugendlehre.\(^{28}\) The central task of the Rechtslehre is not to define duties of justice, but rather to clarify the related, though different, distinction between “laws of justice” (Rechtsgesetze) and “ethical laws” (Sittengesetze) and the parallel distinction between “juridical legislation” and “ethical legislation.”

According to Kant, all duties (both juridical and ethical) can be determined through an application of the categorical imperative as a test procedure to our maxims. To be sure, in the Tugendlehre Kant amends this procedure for determining duties of virtue by introducing a material doctrine of ends \((\text{MdS}, 380–1)\), but for duties of justice he apparently accepts without alteration the formulations he has already provided in the Groundwork and the second Critique.\(^{29}\) Rather than introducing a further criterion for determining duties of justice, the Rechtslehre takes up a topic not treated in these earlier writings, namely, the moral possibility of using coercion against others. This was a topic of discussion within the natural law tradition which Kant took over and attempted to solve within the context of his own moral philosophy.\(^{30}\) The contrast between these two sets of problems (i.e., between determining duties and providing a moral justification for the use of force) is
evident in the fact that someone’s failure to act according to the requirements of justice (as determined by an application of the categorical imperative) does not by itself give someone else a moral permission to compel him or her to do (or to refrain from doing) those acts. Likewise, the moral obligation to perform an act (i.e., the fact that the contrary of its maxim is prohibited by an application of the categorical imperative) does not mean that an individual is morally justified in coercing someone else to observe that right. The task Kant sets for himself in the Rechtslehre is to specify the conditions under which it is morally permissible to use coercion against another.

Since, as we have seen, the principle of autonomy stands at the center of Kant’s moral philosophy—we are morally subject only to those laws that we could give to ourselves—this task must appear highly suspect, if not entirely self-defeating. How can the principle of autonomy provide a moral justification for the use of force? Moreover, in contrast to some of his contemporaries, notably Fichte, Kant does not attempt to deduce or derive claim-rights (and, with these, the legitimate use of force) directly from the Moral Law. Rather, and this is crucial for a proper understanding of the structure of the Rechtslehre, Kant regards the use of force as a long-standing and unsolved political problem and he attempts to provide a solution for it by means of the categories of his moral theory.

The basis of Kant’s solution to this problem can already be found in the Introduction to the Metaphysics of Morals in his distinction between the two forms of legislation (Gesetzgebung).

All legislation (whether it prescribes internal or external actions, and whether these either a priori through mere reason or through another person’s will) consists of two elements: firstly, a law that objectively represents the action to be done as necessary, that is, makes the action a duty; second, an incentive that subjectively links the ground determining the will to this action with the representation of the law. (MdS, 218)

As is even more clear from Kant’s notes to the Metaphysics, he is here invoking a distinction between a principle of judgment (Beurteilung; principium diiudicationis) and a principle of execution (Ausübung; principium executionis). The principle of judgment is the Moral Law which, through an application of the categorical imperative as a test procedure for maxims of action, specifies our duties. Insofar as it abstracts from consideration of the incentives to follow the law, “it is mere theoretical knowledge of the possible determinations of the will, that is, a knowledge of practical rules” (MdS, 218). The principle of execution, on the other hand, considers the mode of compliance or incentive through which an action is performed in accordance with the law.
Therefore (even though one legislation may agree with another with regard to actions that are required as duties; for example, the actions might in all cases be external ones) all legislation can nevertheless be differentiated with regard to the incentives. (MdS, 218-9)

The distinction between ethical and juridical legislation is a distinction regarding the principle of execution or mode of compliance, not the principle of judgment: Ethical legislation (whether of duties of justice, such as fulfilling contractual obligations, or duties of virtue, such as promoting the happiness of others) makes duty itself the incentive, whereas juridical legislation (which, according to Kant, only applies to certain duties of justice) makes use of incentives other than the idea of duty itself. These other incentives to compliance are “pathological” in that they consist of sanctions that appeal to our inclinations and disinclinations. “Therefore, the Rechtslehre and Tugendlehre are distinguished not so much by their differing duties as by the difference in the legislation that combines one or the other incentive with the law” (MdS, 220). Thus, to restate Kant’s problem more specifically, the task of the Rechtslehre is to specify the conditions under which it is morally permissible to employ external sanctions in order to bring about compliance with the specified duties of justice.

Corresponding to this distinction between ethical and juridical legislation is the distinction between ethical laws and juridical laws (MdS, 214). Ethical laws (and, again, these may be of virtue or of justice) are those in which duty or respect for the Moral Law becomes the determining ground of the will (or the motive for compliance), whereas juridical laws are directed only to external compliance (in abstraction from the agent’s motivation). Both types of law are “laws of freedom,” in contrast to laws of nature, and thus their validity (or principle of judgment) is found in their agreement with the Moral Law (as determined by the categorical imperative as a test procedure). Kant then adds:

whether we consider freedom in the external or in the internal exercise of the will, its laws, being pure practical laws of reason governing the free will in general, must at the same time be internal grounds of determination of the will, although these laws must not always be considered from this point of view. (MdS, 214, my emphasis)

That is, although any morally legitimate law must be one that could be an ethical law (because all laws are determined by the same principle of judgment), juridical laws can also be viewed from another perspective, namely, their capacity to become laws whose compliance is insured through external sanctions.
Although Kant’s distinction between ethical and juridical legislation (and ethical and juridical laws) is fairly clear, what still remains unclear is the way in which Kant provides a justification for the use of coercion. Kant’s solution to this problem is found in a “practical law” that has a unique (if problematic) status in Kant’s practical philosophy, the Universal Principle of Justice (or Right).\(^{32}\)

Every action is just (right) that in itself or in its maxim is such that the freedom of the will of each can coexist with the freedom of everyone in accordance with universal law. \((Mds, 230)\)

A second formulation, which he calls “the law of justice,” reads,

act externally in such a way that the free use of your will is compatible with everyone according to a universal law. \((Mds, 231)\)

The UPJ introduces Kant’s notion of “strict justice,” or right in a strong sense\(^{33}\)—that is, it does not merely define a moral permission, obligation, or prohibition (as the categorical imperative does), but identifies the conditions under which an individual has “the moral capacity to bind others” \((Mds, 237)\). That is, it states when coercion is morally justified, or when freedom can legitimately be restricted. Kant cites three conditions that must obtain if the UPJ is to apply, that is, if the use of coercion is to be justified: First, it refers only to external action insofar as these affect the freedom of others; second, it refers only to actions (or their maxims) that individuals can universally will—that is, it does not refer to the particular needs, wishes, or desires that an individual may have; and third, it refers to the formal character of an agent’s maxim (not to its material end). \((Mds, 230)\)

Although the UPJ seems quite straightforward, it is difficult to give it a clear interpretation and to locate its proper status in Kant’s practical philosophy.\(^{34}\) One reason for this confusion is to be found in the fact that the UPJ is frequently viewed as a principle for defining duties of justice, whereas it is primarily introduced by Kant in response to the question of the legitimate use of coercion. The UPJ states that a specific class of free actions may justifiably be restrained by force.

If a certain use of freedom is itself a hindrance to freedom according to universal laws (that is, unjust), then the use of coercion to counteract it, inasmuch as it is the prevention of a hindrance to freedom, is consistent with freedom according to universal laws; in other words, this use of coercion is just. \((Mds, 231)\)

However, whether an action falls under this category or not must itself be determined by the CI-procedure (see next section). What Kant lacks is an
account of the just use of force, and it is this, I suggest, that the UPJ is intended to provide. After introducing the UPJ, Kant can go on to say that the concept of justice (that is, “strict justice”) is not composed of two parts (i.e., a juridical obligation and the authorization to use coercion); rather, it “can be held to consist immediately of the possibility of the conjunction of universal reciprocal coercion with the freedom of others” (*MdS*, 232).

Strictly speaking, then, the UPJ is not another formulation of the categorical imperative. That is, it is not introduced by Kant as a test procedure for maxims in order to determine duties of justice. Rather, it has a quite different status and function in Kant’s moral philosophy: It is introduced as a solution to the problem of the legitimate use of coercion. Further, unlike the various formulations of the categorical imperative, the UPJ is not directed to the incentives to action, but only to external actions themselves.

Admittedly, this law imposes an obligation on me, but I am not at all expected, much less required, to restrict my freedom to these conditions for the sake of this obligation itself. Rather, reason says only that, in its very Idea, freedom is restricted in this way and may be so restricted by others in practice.... We may not and ought not to represent this law of justice as being itself an incentive. (*MdS*, 231)

On the other hand, as Kersting notes, the UPJ is also not a hypothetical imperative. It imposes upon us an unconditional obligation whose source can only be derived from the Moral Law.

Strict justice is admittedly founded on the consciousness of each person’s obligation under the law; but, if it is to remain pure, this consciousness may not and cannot be invoked as an incentive in order to determine the will to act in accordance with it. (*MdS*, 232)

In light of these considerations it is perhaps best to assign the UPJ a unique status in Kant’s practical philosophy. It is a practical law in a scientific doctrine (*doctrina scientiae*) of morals whose purpose is to specify the conditions under which the use of coercion is morally justified (see *MdS*, 375, 218). I will attempt to clarify the unique status of this principle in Kant’s practical philosophy, and the relationship between justice and morality more generally, by contrasting my own interpretation with three others in the recent secondary literature.

C. Three Interpretations of Justice and Morality in Kant

The first interpretation of the relation between justice and morality I will call, following Kersting, the “independence thesis.” It asserts that
Kant’s concept of justice, and the task of establishing a just political order in
general, can be clarified without reference to the basic categories of Kant’s
moral philosophy, and in particular without reference to the notion of auton-
omy (or “positive freedom”). The thesis is expressed in the following quota-
tion from Yovel’s *Kant and the Philosophy of History*:

Good citizenship is possible even in a kingdom of devils. It requires no
ethical community (kingdom of ends) and presupposes none. It is
something that can be imposed by coercion, while morality can be
rooted only in the free or spontaneous will of individuals. Therefore,
even to the best of states cannot be attributed a moral value per se, and
it is not in any political organization that the end of history is to be
placed.\textsuperscript{37}

This passage contains several assertions that need to be sorted out. Yovel is
certainly correct in claiming that, according to Kant, a just society does not
entail that its citizens act on the basis of moral incentives and thus does not
require the existence of an ethical community. Similarly, it is true that for
Kant the highest good and “end of history” is not a just political order. Nei-
ther of these observations, however, warrant the conclusion that no “moral
value per se” can be attributed to a political order founded on principles of
justice. Further, in contrast to Yovel’s interpretation of the “kingdom of dev-
ils” passage, I do not think that it can be understood to mean that the problem
of creating a just political order can be solved on the basis of considerations
of rational self-interest alone.\textsuperscript{38} Since this passage from “Perpetual Peace” is
central to Yovel’s interpretation, I will cite it at length.

As hard as it may sound, the problem of setting up a state can be
solved even by a nation of devils (so long as they possess understand-
ing). It may be stated as follows: “In order to organise a group of ration-
al beings who together require universal laws for their survival, but of
whom each separate individual is secretly inclined to exempt himself
from them, the constitution must be so designed that, although the citi-
zens are opposed to one another in their private attitudes, these oppos-
ing views may inhibit one another in such a way that the public con-
duct of the citizens will be the same as if they did not have such evil
attitudes.” (*KPW*, 112–13)

As I noted in my introduction, this passage is admittedly difficult to reconcile
with Kant’s position in the *Rechtslehre* and requires an interpretation of the
relation between Kant’s philosophy of history and his systematic theory of
justice (for which Yovel’s study would be indispensable). At this point I will
only offer three considerations which weigh against Yovel’s interpretation.

First, the claim that the problem of creating a just political order is merely a problem of “political technology” and can be solved on the basis of rational self-interest (that is, on the basis of Verstand rather than praktische Vernunft) conflicts with other claims made by Kant not only in other writings, but even within “Perpetual Peace”: “A true system of politics cannot therefore take a single step without first paying tribute to morality.... For all politics must bend the knee before right, although politics may hope in return to arrive, however slowly, at a stage of lasting brilliance.”

Second, as Yovel is also aware, the “kingdom of devils” passage is closely connected to Kant’s notion of a providentially guided nature or what Yovel calls “the cunning of nature.” Thus it is not rational self-interest alone, but self-interest together with the providential guidance of nature, that is able to solve the problem of a just political order. If Kant’s assumptions about the “mechanisms of nature” are wrong, or at least if these mechanisms do not yield the consequences he predicts, then we cannot look to rational self-interest alone to produce a just political order. In other words, the “critical turn” that Yovel finds in Kant’s decision to restrict the “cunning of nature” to the production of a just state (and not the production of an ethical community) would have to be extended further.

Third, two different readings of the “kingdom of devils” passage are possible, depending on what knowledge one wishes to attribute to the devils. If they know the Moral Law but choose not to act on it, in their rational calculations they may still devise a constitution that would be different from one in which their calculations were made without any knowledge of the Moral Law at all (that is, one that was based on rational self-interest alone).

In light of these considerations, we must give priority to Kant’s systematic presentation of the concept of justice as it is contained in the Rechtstheorie. As we have noted, the concept of moral personality or autonomy is a basic category for this work. Consequently, the only acceptable interpretation of the independence thesis is the formulation offered by Mary Gregor in her Laws of Freedom:

Law is independent of ethics in the sense that it has no need of ethical obligation in determining its duties. But it cannot be independent of the supreme moral principle; for if its laws were not derived from the categorical imperative, then the constraint exercised in juridical legislation would not be legal obligation but mere arbitrary violence.

A second interpretation of the relationship between justice and morality, more prevalent in the secondary literature than the independence thesis, is the teleological interpretation. Riley’s Kant’s Political Philosophy offers one of
the most sustained defenses of this position.\textsuperscript{42} I refer to this interpretation as teleological for two reasons. First, it is teleological in the sense that it is opposed to constructivist interpretations of Kant’s moral philosophy. Rather than viewing Kant’s moral theory as a construction based on a conception of the person as an autonomous moral agent, it maintains that Kant’s moral theory depends upon a rational intuition of “objective ends” given to us.\textsuperscript{43} Second, it is teleological in that it regards justice, and politics more generally, as something that exists “for the sake of” these objective ends. Thus, Riley writes,

public legal justice is instrumentally (purposively) related to morality in two ways: in a \textit{weaker sense}, it creates legal conditions for the exercise of a good will—it limits occasions for sin and creates occasions for morality; in a \textit{stronger sense}, it legally enforces part of what ought to be, even where a good will is absent.\textsuperscript{44}

We have already discussed the first aspect of a teleological reading in section A above: Kant rejects such a teleological conception because it violates the notion of human autonomy. Further, it runs counter to the basic principle of his critical philosophy, namely, that we cannot have knowledge beyond the bounds of sense experience. The second aspect of the teleological interpretation is more difficult to criticize since, like the independence thesis, it finds some support in Kant’s writings, especially “Perpetual Peace” and “The Conflict of the Faculties.” As before, the rule of interpretation must be to read these essays in light of Kant’s more systematic presentations, and that means the \textit{Rechtslehre} must be given priority. One reason for our earlier discussion of Kant’s distinction between two forms of legislation and between the principle of judgment and the principle of execution was to show that the validity of juridical and ethical laws is the same, namely, their conformity to the Moral Law. To describe justice as existing “for the sake of” morality is thus to diminish the intrinsic value of political rights and duties, just as the independence thesis denies them moral worth altogether. Political rights and obligations, according to Kant, are unconditional commands of pure practical reason, whether they produce a greater degree of \textit{internal} morality or not (although Kant was, I think, ultimately of the opinion that they would). In this regard it is also important to note that while Kant considered legality inferior to morality when it refers to the incentives upon which people act, this does not mean that a just political order is somehow inferior to and merely instrumental for the moral kingdom of ends.

A third, more promising interpretation of the relationship between morality and justice, and specifically the UPJ, is found in Onora Nell’s \textit{Acting on Principle}.\textsuperscript{45} This is partly because Nell rejects a teleological interpretation of Kant’s practical philosophy and argues for a constructivist interpre-
tation. Further, her study offers a comprehensive interpretation of the categorical imperative as a test procedure that, in contrast to numerous other interpretations, attempts to preserve both the "formality" and the "fertility" of Kant's moral theory. Since I consider this interpretation the most plausible reconstruction of Kant's theory, and since it provides the background for her rendering of the UPJ as a "restricted version" of the categorical imperative, I will offer a brief summary of her interpretation.

A moral theory that, like Kant's, proposes a universality test of principles for determining the moral status of particular acts confronts what Nell calls "the problem of relevant act descriptions."46 Since an act can be viewed under a number of different principles, and since different acts can be performed under the same principle, it is difficult to see how the moral status of an act can be determined by a universality test of principles. Kant's own solution to this problem, Nell argues, is that the agent's maxim provides the relevant act description or appropriate principle for assessing the moral status of an act.

Maxims, according to Kant, are "subjective principles," or general rules of conduct that underlie an agent's voluntary actions (MDs, 225). We need not assume that an agent is always conscious of the maxim upon which he or she acts, but, according to Kant, it must be possible for the agent to identify or supply a maxim upon reflection. Some examples of maxims that Kant offers in the Groundwork and second Critique are: to neglect developing my natural gifts if I find it convenient (CMS, 423), not to tolerate an unavenged offense (KPV, 19), and to increase my property by every safe means (KPV, 27). Accordingly, Nell proposes the following as a formal schema of a maxim: "To___if..." or "I will___if....," where "___" and "..." can be filled in with some act description and some agent description respectively (p. 35).

It is clear from these examples that maxims already involve a degree of generality and abstraction from a concrete context of action. They are not descriptions of intentions to carry out specific actions such as, "I will eat dinner with friends every Friday evening," or "I will drink beer at five o'clock," but more general descriptions of a course of life conduct under which the intentions to carry out more specific activities would fall.47 Thus, Kant states that maxims are "general determinations of the will which have under them several practical rules" (KPV, 19). Kant offers no mechanical method for identifying maxims (or practical rules of sufficient generality for applying the CI-procedure) and seems to hold that this requires an element of judgment that comes only with age and experience.48

Further, Kant assumes that the maxims to which the CI-procedure can be applied (or better, the agents acting on those maxims) are rational in two senses: the agent is able, upon reflection, to order his or her maxims and more specific practical rules into a coherent and consistent scheme, and the agent is able fully to intend both the conditions required for the completion
of the act and the consequences that can reasonably be predicted to follow from
the act.49 Both of these assumptions can be, more or less, inferred from
Kant’s Principle of Hypothetical Imperatives: “If I fully will the effect, I also
will the action required for it” (GMS, 417). With this solution to the problem
of relevant act descriptions in hand, we can turn to Nell’s reconstruction of
the CI-procedure as a test applied to an agent’s maxims.

In its most basic form, the categorical imperative or “supreme principle
of morality” reads, “Act only on that maxim which you can at the same time
will that it should become a universal law” (GMS, 321). The first step in the
CI-procedure is to formulate the intended maxim as a universal law. Thus, “I
will break promises if it is advantageous” can be formulated as “Everyone is
to break promises if it is advantageous,” and the schema “To___if…” has the
universalized counterpart, “Everyone to___if…”

Since nature, “in its most general sense,” means a system of effects
governed by universal law, Kant says that the categorical imperative may
also be formulated, “Act as if the maxim of your action were to become
through your will a universal law of nature” (GMS, 421). From Kant’s more
detailed description of this formula in the second Critique, it is clear that it is
indispensable for understanding how to apply the CI-procedure. He states:

The rule of judgment under laws of pure practical reason is: ask your-
self whether, if the action you propose should take place by a law of
nature of which you yourself were a part, you could regard it as possi-
ble through your own will. (KpV, 72)

The second step in the CI-procedure requires not only that we formulate the
intended maxim as a universal law, but also that we regard it as a law of
nature. We thus arrive at “Everyone will break promises if it is advan-
tageous,” and the corresponding schematization, “Everyone will___if…”
(which Nell calls the “universalized typified counterpart”). To be sure, max-
ims, as rules that underlie voluntary actions, belong to the realm of freedom
not nature, but Kant’s point is that unless we can regard the universalized
maxim as if it were a law of nature—that is, as a law that governs a normal
and predictable set of events—it will not be possible to decide whether the
intended maxim is in contradiction with it.

Kant elaborates upon this second step in the section of the second Cri-
tique entitled, “Of the Typic of Pure Practical Judgment,” a section which
is important for avoiding some common misunderstandings of the “law of
nature” formula. In this passage Kant is quite specific about what we may and
may not assume in viewing universalized maxims as laws of nature. On the
one hand, in treating the universalized maxim as a law of nature, we are not
making empirical assumptions about what might in fact happen if I were to act
on the intended maxim. Because I choose to break promises if I consider it advantageous does not mean that others will also act in the same way. Rather, we make a fictive or counterfactual assumption about a system or order of things in which everyone acts in this way as if it were a law of nature. On the other hand, we are also not attributing a purposive or teleological plan to nature (as some interpretations have proposed).\textsuperscript{50} The laws of nature serve as a \textit{type} for practical laws.\textsuperscript{51} What he means by this is that we should take no more from the world of sense than “the form of lawfulness” (\textit{KpV}, 71–2), that is, the form of order or regularity in the production of an effect.\textsuperscript{52}

However, and this brings us to the third step in the CI-procedure, it is still not clear how a contradiction might arise since it would seem that the maxim “I will break promises if it is advantageous,” is only a particular instantiation of its universalized typified counterpart, “Everyone will break promises if it is advantageous.” The answer is found in the phrase “at the same time,” contained in the first formulation. We are asked to consider if it is possible without contradiction to will simultaneously the intended maxim (together with its necessary conditions and reasonably predictable consequences) and its universalized typified counterpart. Nell offers the following summary of the CI-procedure:

It asks whether we can simultaneously intend to do ‘x’ (assuming that we must intend some set of conditions sufficient for the successful carrying out of our intentions and the normal and predictable results of such execution) and intend everyone else to do ‘x’ (assuming again that we must intend some conditions sufficient for the successful execution of their intentions and the normal and predictable results of such execution). (p. 73)

For example, we are to ask ourselves if we could simultaneously will to break promises if it is advantageous and will a system of nature, of which we are a part, in which everyone breaks promises if it is advantageous. Kant claims this is not possible since it would undermine the conditions necessary for such an intention.

There are a number of difficulties and further considerations that a full discussion of Nell’s reconstruction of the CI-procedure would have to take up, but I must pass over these here.\textsuperscript{53} The above summary is sufficient to enable us to see how Nell interprets the UPJ and how she views its relation to the Moral Law.

One of the stronger claims in her study (which has not yet been mentioned) is that the distinction between the two kinds of tests introduced by Kant in the \textit{Groundwork} (the test for a contradiction-in-conception and the test for a contradiction-in-will) provides a criterion for distinguishing
between duties of justice and duties of virtue. Accordingly, the contrary of maxims that fail the contradiction-in-conception test (that is, maxims that cannot even be conceived as universal laws) define the class of duties of justice, while the contrary of maxims that may survive the contradiction-in-conception test but fail the contradiction-in-will test define the class of duties of virtue. Nell then suggests that the UPJ may be viewed as a variant on the contradiction-in-conception test. That is, “Act externally in such a way that the free use of your will is compatible with the freedom of everyone according to a universal law,” is materially equivalent to (i.e., yields the same results as) the contradiction-in-conception test (which, according to Nell, itself only shows certain external acts to be forbidden).

There are several difficulties that this claim needs to address. I will briefly note two here. First, Kant introduced the two tests not as a way of distinguishing between duties of justice and duties of virtue (which he does not mention in the *Groundwork*), but as a way of distinguishing between perfect and imperfect duties. Moreover, in contrast to the traditional (natural law) distinction between perfect and imperfect duties, which does parallel Kant’s later distinction (in the *MdS*) between duties of justice and virtue (i.e., those that may and those that may not be externally legislated), Kant’s own distinction between imperfect and perfect duties in the *Groundwork* is based on a different criterion, namely, one that distinguishes between duties that allow for an exception in the interest of inclination and those that do not (*GMS*, 422n).

Second, it is significant that in his later ethical writings Kant does not make use of this distinction between the two tests and, in particular, the *Metaphysics of Morals* does not distinguish between duties of justice and duties of virtue by means of this device. Rather, in the *Metaphysics of Morals* Kant introduces a material doctrine of ends to contrast duties of virtue with those of justice (*MdS*, 395).

What is more important for our own investigation, however, is Nell’s further claim that the UPJ can be viewed as a “subsidiary formula” of the categorical imperative and thus that it is a test procedure for determining duties of justice (p. 39; see also p. 45). This interpretation of the principle conflicts with my own claim that Kant’s major concern in the *Rechtslehre* is not to define duties of justice, but to provide an answer to the natural law question about the legitimate use of coercion. It also overlooks the unique status that the UPJ possesses in Kant’s theory. Two considerations may be raised against Nell’s interpretation. First, Kant does not use the UPJ as a test for maxims of action and, as we have seen, it is very difficult to read it in this way. The determination of duties of justice (as well as duties of virtue) is the task of ethics, not of the *Rechtslehre*. Second, such an interpretation of the UPJ makes it very difficult to understand the subsequent uses Kant makes of it, specifically with respect to the justification of property rights and the moral obligation (which
can be coercively enforced) to enter into a civil society or social contract. These are applications of the UPJ that will be considered in the next chapter.

Let me briefly summarize the several points which this section has sought to establish concerning Kant’s theory of justice: First, the Rechtslehre introduces the discussion of rights in connection with a conception of practical reason that cannot be reduced to a narrower notion of rational self-interest. (Rawls will later make essentially the same point in his own distinction between the Reasonable and the Rational.) The Rechtslehre thus presupposes the notion of moral personality or autonomy (positive freedom). To an extent yet to be considered, this may mean that Kant’s theory of justice stands or falls with the success of his justification of the Moral Law as a Fact of Reason and of the related notion of transcendental freedom.

Second, Kant’s theory of justice (and his practical philosophy in general) is constructivist, not teleological. A teleological interpretation of his theory undermines the fundamental principle of autonomy.

Third, one implication of his moral constructivism is that Kant does not view the question of political justice as less important or “merely for the sake of” his theory of morality. This claim was supported via a discussion of Kant’s distinction between the two forms of legislation and the distinction between a principle of judgment and a principle of execution. Laws of justice have the same source of justification as do the laws of ethics, namely the Moral Law (or categorical imperative). The distinction between laws of justice and laws of virtue is a distinction within the principle of execution, that is, the incentive or motivation for compliance, and not within the principle of judgment. The UPJ provides a justification for the use of coercion with respect to a certain class of external actions, namely, duties of justice, but is not itself introduced as a test for determining or specifying those actions.

Finally, Kant’s theory of justice is a theory about the legitimate use of force or coercion. However, unlike some of his contemporaries, notably Fichte, the use of force is not derived or deduced directly from a moral principle. Rather, Kant more or less simply appropriates this problem from existing discussions within the natural law tradition and proposes a solution that invokes the basic categories of his moral philosophy. Thus the UPJ is introduced primarily as a justification for the use of coercion.

III. PROPERTY RIGHTS AND THE SOCIAL CONTRACT

For all contract theorists, including Kant, political legitimacy is based upon the (counterfactual or hypothetical) consent of the governed. The differences among them begin to emerge when we inquire into the motivations and considerations which lead up to the agreement. For Kant, consent to the social
contract is not based upon considerations of rational self-interest or prudence (Hobbes), nor upon a natural right to self-preservation and the guarantee of absolute property rights (Locke), but upon a moral obligation to institutionalize and make conclusive in a social contract property rights that in the state of nature have only a provisional character. Whether this approach ultimately makes the idea of a voluntary agreement superfluous is a question I will address below.\textsuperscript{56} First, I would like to consider the unique status Kant assigns to property rights in his theory of the social contract. This status has for the most part not been observed in the secondary literature, and Kant has even been accused of falling behind the achievements of Locke and Rousseau by grounding property rights on an element of brute force.\textsuperscript{57} However, since Kant himself espoused a labor theory of property rights in his earlier writings, it is important to consider carefully the basis for its later rejection in the Rechtslehre. Further, since Kant’s theory of property rights constitutes an alternative to Locke (as well as to Hobbes and Rousseau), it may be possible to find within the contract tradition an alternative to the view of property rights found in some of its contemporary versions (e.g., Rawls, Nozick, or Buchanan).

In the natural law tradition (as well as in Kant) the right to private property implies an authorization on the part of the owner to prohibit others from its use by coercion.\textsuperscript{58} For Kant, therefore, it is a matter of “strict justice,” and falls under the “universal principle of justice.”\textsuperscript{59} Since, according to this tradition, the earth was originally given by God to all men in common, the central question for the theory of property rights is how such an obligation/authorization could arise. Furthermore, since rights entail a moral authority over others and since one cannot acquire an authority over others without their consent, Grotius and Pufendorf argued that property rights could only come about through an actual contract or agreement in which everyone relinquishes his common right to the use of the land and agrees to a principle for its individual or private distribution. For these theorists, the validity of private property rights presupposes that this unique and irrevocable contract actually took place at some point in the past. As is well known, Locke (and many others) rejected this notion of an original contract: “If such a consent as that was necessary, Man had starved, notwithstanding the Plenty God had given him.”\textsuperscript{60} According to Locke, individuals acquire a direct right to property, apart from the consent of others, simply by “mixing” their labor with it or by otherwise “joining” it to themselves. It is this direct and absolute right that then provides the basis for the social contract. Thus, to a certain degree, Kant’s rejection of Locke’s theory represents a return to the natural law conception that a person can acquire a right to an object—and hence the authority to prohibit others from its use—only on the basis of their consent. However, unlike natural law theorists, Kant does not assume that this consent actually took place in the past. Rather, he introduces the notion of a united agreement as an idea of prac-