

The Concept of *Right*, and Objectivity: Some Preliminary Considerations

A very rough draft for a future paper, that will be able to come forward as a dissertation chapter

Alptekin Sanli

I

1. A salient feature of Kant's methodological reflections on his work during the critical period, is his reliance on legal and political metaphors. He glosses the notion of 'critique' as the establishment of a "true court of justice" for the claims of pure reason, and declares that the purpose of a 'critique of reason' is to usher in an era of "perpetual peace" in philosophy.¹ The quarrels of metaphysicians are described, in his work, in the imagery of a Hobbesian state of nature. He says that the "archetype, in the ideal of *the philosopher*", is that of a "legislator [Gesetzgeber] of human reason".² Accordingly, he transforms the traditional question as to the 'objective reality' of concepts, from one that concerns the representational *in esse*—whether as shapings of the soul, or as modifications of *res cogitans*, or even as the parade of atomic 'simple ideas' on a mental stage—of forms, which also inform things of the world, into one that concerns the *legitimacy of claims to knowledge* in which they are used. The question which launches his 'transcendental deduction of the pure concepts of the understanding' is a *quid juris?* question, borrowed openly from juristic terminology on property rights. In his sole public confrontation with the school of Leibnizian-Wolffian metaphysics, the 'Eberhard controversy', he refutes the charge that his 'categories' and 'pure intuitions' are simply innate ideas that had already received better treatment in Leibniz's hands, saying that

¹ A751/B780

² A838–839/B867–868

they are rather to be understood as “original acquisitions”, another term borrowed from the vocabulary of property rights.³ And it almost goes without saying, that his entire practical thought appears to be grounded in the notion that we are subject to no law or principle of action that we do not ‘legislate for ourselves’.

2. I shall argue that the ‘legalisms’ which permeate Kant’s critical work are more than stylistic devices, or side-illustrations. A very specific understanding of practical objectivity—that realized in the *rightful condition*, as an institution of *equal freedoms*—comes to condition his understanding of objectivity *in general*. This observation, if on the right track, has two important corollaries:

- i) Kant appears to have imported a criterion of *instituted* ‘absolute generality and strict universality’, viz., *exceptionlessness*, into a theory that claims to formulate the ‘absolutely general and strictly universal’ principles that are *constitutive* of the objects of experience.
- ii) He also appears to have made a specific use of practical reason—*knowing how to behave* with respect to ‘impersonally’, or ‘not merely subjectively’, or ‘externally’ legislated norms of conduct; and *letting* one’s claim making to be limited by these—to be an *essential* moment of the exercise of theoretical reason.

I think that the importation described in (i) is, bluntly speaking, a dead end. On systematic grounds, we can show that instituted ‘exceptionlessness’ cannot become the model for a theory of constitutive principles. On historical grounds, we can show that the ‘importation’ takes place via Kant’s radicalization of Newton’s groundbreaking use of apodictic principles in physics, into a ‘transcendental’ register: despite Kant’s

³ Ak VIII.221

protestations, the resulting idealism cannot free itself from the appearance of a merely subjective, 'restrictive idealism', because Newton's masterful sense of the distinction between instituted and constitutive principles gets lost in the radicalization process.

3. On the other hand, the construal of theoretical reason adumbrated under (ii) is exceptionally fruitful, and may indeed offer a better starting point for a 'critical idealism' than the 'transcendentalizing' approach. I think, that a very robust outline for a 'critical idealism' of this sort has been worked out, in the 20th Century, by Wilfrid Sellars. Kant's own, very restricted, development of this notion, on the other hand, takes place mostly under the rubric of his *Transcendental Dialectic*. The significance of this, I think, is twofold: Such a conception of 'disciplined' reason represents the antithesis of the dogmatic conception—thus, we ought to acknowledge, that far from playing an ancillary role with respect to the project outlined in (i), it represents just the 'reformed reason', in relation to which the analytical labors of (i) play a propaedeutic role. On the other hand, Kant's treatment remains restricted, because the focus of the *Transcendental Dialectic* is the doctrine of 'Transcendental Illusion': this is just one dimension of the claim-making life of reason, albeit a dimension that is very privileged from Kant's point of view. An important consequence of starting with such a focus, however, is that it renders nearly invisible the essential involvement, in the exercise of theoretical reason, of 'externally'-oriented, *legal and political* practical reason, in favor of a thesis as to the primacy of *moral* practical reason over the theoretical, *in matters that have to do with the 'supersensible'*. The latter thesis is motivated by an entirely different set of concerns, than those that have to do with the practical dimension in 'justifying and being able to justify' one's claims to knowledge.

4. In what follows, I shall briefly discuss the basic problem as to the conflation of instituted, and constitutive objectivity. The rest of my argument, however, is going to be

a defense of the primacy of legal-political practical reason, over the moral-practical, in Kant's thought. This defense is necessary for several reasons. Kant gave a systematic treatment of the concept of *right* only at the very end of his career; his *Doctrine of Right* was published, as the first part of his *Metaphysics of Morals*, only in 1797. Thus, trying to interpret the first *Kritik* in the light of concepts borrowed from the *Doctrine of Right* might appear to be a case of 'putting the cart before the horse'. Furthermore, it is generally surmised that his *Universal Principle of Right* is, in some sense, a 'derivation' from the *Categorical Imperative*—and it certainly seems that way. This, is the 1788 formulation of the *Categorical Imperative*:

So act, that the maxim of your will could always hold at the same time as a principle in a giving of universal law.⁴

And this, the *Universal Principle of Right*, from 1797:

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.⁵

But, if that is so, then any attempt to understand how reason can be practical, or issue 'ought's, or how practical reason can be *essentially involved* in the use of theoretical reason, should start with the Categorical Imperative, and not the concept of *right*. Onora O'Neill, for example, who is among the very few commentators who pay attention to the political metaphors of KrV's *Doctrine of Method*, argues that although Kant's political thought is not a mere side-project to his moral philosophy, it is "the Categorical Imperative [that is] the fundamental principle of all reasoning and acting".⁶

⁴ Ak V.30

⁵ Ak VI.230

⁶ O'Neill 1989

5. *Sed contra*, I shall try to show that although by 1764 Kant had already formulated the distinction between hypothetical and categorical imperatives, and declared moral precepts to be exclusively of the latter sort, he was able to justify their unconditional ‘bindingness’ [Verbindlichkeit], hence ‘objective necessity’, only after he became acquainted with Rousseau’s notion of the *volonté générale*. And, as we shall discuss below in more detail, Kant’s ‘critical’ understanding of the the Categorical Imperative was ‘derived’ from a version of the Universal Principle of Right—insofar as the latter was inherited from Rousseau. In order to do this, Kant ‘internalized’ Rousseau’s ‘externally’ legislated system of equal freedoms, converting the public character of his *volonté générale* into a matter of the moral agent’s necessarily presupposed ‘citizenship’ in a supersensible, *world-independent* community of intelligences.

II

6. Let us return to the theme of institution and constitution. Now, ‘exceptionlessness’, is a necessary trait of the rules that govern our instituted practices; these are practices, within which the only standards of correctness are those of fulfilling properly the roles set by the rules of the practice. For people, these rules concern performance; for inanimate things, measuring up to the requirements that make ‘the game playable’. Participants are expected to fit the practice, and not the other way around: They exert no independent ‘normative constraint’ on the rules. This is not to say, of course, that an ‘instituted practice’ is exempt from criticism altogether. Inside a game of basketball, for instance, a player cannot “do things right, and get things wrong”, but the game as an institution certainly gets ‘something’ right—it is exceptionally good at providing a certain sort of entertainment.

7. As we shall discuss in more detail below, Kant found, in Rousseau's work, the idea of an institution that is its own end. Rousseau had offered a solution, to a problem that had found its most powerful formulation in Hobbes; we may call this the problem of 'civic objectivity'. While Hobbes's solution was to interpret the 'civic game' to be a human institution that finds its justification in providing 'self-preservation'; Rousseau developed the notion that the 'civic game' cannot simply be the means to an arbitrary end: insofar as it is understood as an institution of equal freedoms, it is its own end. Furthermore, it is the kind of game which we cannot help but play, despite the tendency of most of us, to cheat all the time: its 'exceptionless'-instituted rules command unconditionally, insofar as we are the owners of our selves, and of our actions. For Kant, this was a wonderful example of a normative sphere that can be justified a priori: An instituted practice—the 'rightful condition'—that can justify itself to itself without having to rely on an authority outside of itself, while realizing the conditions that are constitutive of human reason as such.

8. In Kant's usage in the *Critique*, the concept of "constitution" is contrasted with that of "regulation"; e.g. the mathematical categories are *constitutive* of the objects of intuition; the dynamical categories, by contrast, are *regulative* of intuition, but *constitutive* of experience. The "Regulative Ideas of Reason" are *regulative* with respect to judgments (whether they might be 'constitutive' of anything at all, is a matter of controversy). I shall work for the moment, however, with John Haugeland's use of the term: "Constitutive Principles", in contrast to rules of institution, are rules, which we formulate to one another, that 'things themselves' obey. The formulation of rules to one another is, indeed, an 'instituted' practice; yet, once we begin to play *our* game, of 'getting right' the rules of the *independent* game which things themselves must be 'playing', we have to deal with a more complicated set of criteria of correctness. It must

be our postulated rules, that are answerable to 'how things are', and 'why they 'behave' the way they do'. However, things do not speak to us, to tell us their reasons; therefore, lest it arbitrarily take 'the way we happen to take things to be' to reflect 'how things are in themselves', our instituted practice of making claims to one another must develop the resources to *let things resist* our ways of trying to make sense of them, make it *possible* for us to 'do things right', but 'get things wrong'. Neither does this suffice all by itself, since the instituted practice of letting oneself be wrong must be held in check by another instituted practice, that of doing all this with the intention of eventually getting things right. It would appear, then, that for the 'game of truth-telling', the *only* exceptionless rule is to have a commitment to truth, which is very similar in this regard, to the highest principle of the 'civic game', that of commitment to equal freedoms. *By contrast, it must be possible to challenge 'constitutive principles', even as we do our best to make sure they survive challenges. Declaring our 'constitutive principles' to be exceptionless appear, in Sellars's apt phrase, to be "a myth, a snare, and a delusion"*⁷.

9. The legal conception of 'absolute generality and strict universality' comes, indeed, to condition Kant's adoption of Newton as his guide for a revolution in metaphysics. Throughout his entire career, Kant regarded Newton as having set the 'special metaphysics of nature' on the 'secure path of a science'. Newton's groundbreaking approach to constitutive apodictic principles in physics had not only succeeded, *for the first time* in the course of the scientific revolution, in applying mathematics to the study of the infinite complexities of empirical nature, but also introduced a mode of causal explanation that made no recourse to the inner nature of things. In bringing a decisive end to the dispute between Tychonians and Copernicans, and explaining Keplerian

⁷ Is there a synthetic a priori?

regularities without reference to a mysterious cosmic magnetism or the inner nature of matter as extension, Newton had, in a sense, effected a ‘treaty of perpetual peace’ in physics, and paved the way for an unprecedented flourishing of empirical knowledge in the same act. In Kant’s eyes, the same approach could revolutionize the ‘general metaphysics of nature’, which, when specified with respect to a very peculiar sort of empirical concept, the ‘movable in space’, would subsume, with the necessary adjustments, the metaphysical core of Newton’s work. My claim, however, is that translating, or rather *radicalizing*, Newton’s conception of constitutive apodictic principles into a ‘transcendental idealism’, leads precisely to the familiar puzzles concerning the ‘unknowableness of things-in-themselves’, and the exact sense in which the categories are more than merely subjective forms of consciousness. Newton’s constitutive principles, while ‘apodictic’, *were not* intended to be *exceptionlessness*: rather, empirical exceptions to their predictions were welcomed, as opportunities for *testing* their claims. Eventually, a growing body of ‘exceptions’—the problem concerning the ‘perihelion of Mercury’—came to play a decisive role in the replacement, bluntly speaking, of the Newtonian framework with the Einsteinian.

III

10. How, then, does the concept of *Recht* come to play such a decisive role in Kant’s understanding of objectivity? Is it not somehow ‘derivative’ from his moral philosophy? I have claimed above, that Kant’s moral philosophy is the result of his ‘internalizing’ Rousseau’s ‘externally’ legislated system of equal freedoms, and converting the public character of his *volonté générale* into a matter of ‘citizenship’ in a supersensible, *world-independent* community of intelligences. I can now begin to flesh out this claim, while looking for clues in the service of our question of ‘objectivity’.

11. Kant is commonly interpreted to be an ‘internalist’, in his moral philosophy. We might begin with one such interpretation, and see where it leads us. Stephen Darwall refers frequently to Kant in his 1995 study, “The British Moralists and the Internal ‘Ought’, 1640-1750”, characterizing his moral thought to be the continuation of a project which emerged within the tradition of “British internalism”. Thinkers who were part of this project, Darwall explains, held that before attempting to “vindicate moral obligation”⁸, one had to “understand the nature of normativity itself”⁹—what makes *any* “ought” valid, binding, or obligating¹⁰ (terms which Darwall appears to use interchangeably); or, what makes reason practical *at all*—and, do this in the guidance of the following core commitments:

that the existence of motive, perhaps of a certain kind or under certain circumstances, is (at least part of) what it is for a normative proposition to be true[; that obligation] consists in something internal to the moral agent in some suitable sense; [that] it can be realized only in motives available to a deliberating agent, from a practical point of view. [that] a valid or binding ‘ought’ is ultimately an internal ‘ought’.¹¹

These commitments were groundbreaking, according to Darwall, in that they shifted the sources of “the ought”, “obligation”, or “normativity” from the world (traditionally, the divinely ordained inherent rationality of the world, its ‘purposiveness’), into motives that had to be “realized” in an agent’s deliberations, hence, “internally”. This, Darwall says, was an anticipation of Kant’s distinction between “the two viewpoints: theoretical and practical, or observer’s and agent’s”¹². He adds that against this background, it became possible to develop two further notions, of which the

⁸ Darwall 1995, p.206

⁹ *ibid.*, p.11

¹⁰ *ibid.*, p.2

¹¹ *ibid.*, p.11

¹² *ibid.*, p.13

combination would give us Kant, at least in outline. The first was Locke's 'autonomist accountability view', which established being "free", or "capable of a Law", or "self-determining"¹³ (terms used interchangeably) to be necessary conditions for being accountable under obligations in general. The second notion came with Shaftesbury's formulation of the nexus which holds the "deliberations of an agent", "self-determining agency", and his or her "internal motivation" together, in a "normative theory the will":

[A]ccording to a normative theory of the will, the requisite motivation comes from a motivational state that is both distinctive of will and different from the desire for any good, namely, the agent's recognizing a consideration or demand as authoritative.¹⁴

And thus, Darwall thinks, with such a thesis as to self-motivating rationality in place, the path to a "vindication of moral obligation" was opened—this is the path that Kant would follow:

Whereas the accountability view asserts the *morality entails free rational will* half of Kant's reciprocity thesis, the idea that guidance by a normative conception is intrinsic to free rational will might be deployed in an attempt to vindicate the *free rational will entails morality* half. For the latter project to work, it would, of course, have to be argued that in being guided by a conception of law, free agents are committed to being guided, more specifically, by moral laws. This would be Butler's project and, of course, Kant's.¹⁵

12. Darwall's use of "of course" is puzzling, however. Indeed, the idea that the 'sources of the ought/normativity/obligation' be relocated 'in the agent' sounds Kantian; but, for the connection between Kant and "British Internalism" to be as self-evident as Darwall suggests, Kant would have had to subscribe to certain principles which all the British internalists appear to have deemed axiomatic: i) that the 'deliberating agent'

¹³ *ibid.*, p.17

¹⁴ *ibid.*, p.206

¹⁵ *ibid.*, p.206

needs, over and above his or her 'deliberations as to what to do', *something extra, a motive*, to begin to exercise his or her causal efficacy in the world; ii) that reason can be practical—respond to, or issue any kind of 'ought'—only insofar as it can 'generate a motive' for the will; and, iii) that the question as to the bindingness, or validity of a practical precept, is a question of its capacity to motivate an agent. At least on the surface, however, it is difficult to find such commitments in Kant's practical work. To find such commitments, perhaps we need to look deeper into the texts, and do this in the guidance of a Kant scholar: Darwall relies heavily on Christine Korsgaard's interpretation of what makes reason practical for Kant; thus we should turn to her next.

13. One of Korsgaard's accounts as to what makes reason practical for Kant, is developed via an interpretation of the following passage in GMS. It is worth quoting the passage in full, as we shall have to revisit it, in the sections to follow:

[After having introduced the notion of an 'imperative', and identified three species of it: the *technisch*, *pragmatisch*, and the *moralisch*, Kant says:] Now the question arises: how are all these imperatives possible? This question does not inquire how the performance of the action that the imperative commands can be thought, but only how the necessitation of the will, which the imperative expresses in the problem, can be thought. How an imperative of skill is possible requires no special discussion. Whoever wills the end also wills (insofar as reason has decisive influence on his actions) the indispensably necessary means to it that are within his power. This proposition is, as regards the volition, analytic; for in the volition of an object as my effect, my causality as acting cause, that is the use of means, is already thought, and the imperative extracts the concept of actions necessary to this end merely from the concept of a volition of this end (synthetic propositions no doubt belong to determining the means themselves to a purpose intended, but they do not have to do with the ground for actualizing the act of will but for actualizing the object).¹⁶

¹⁶ Ak IV.417

In this passage, Korsgaard finds the *Aufhebung* (so to speak) of Hume's theory of action. In "The Normativity of Instrumental Reason", where the passage is interpreted, she takes it for granted, that Hume had got the basic structure of volition right (that we *desire* an end, *believe* that certain means will help us attain it; and thus, get *motivated* to take up those means) but consistently failed to make the links between the terms work, because he had inadequate conceptions of "the instrumental principle". The best version that he *could have* come up with—"if you *have a reason* to pursue an end, then *you have a reason* to take the means to that end"—although on the right track, leaves unexplained the connection between having reasons (beliefs) and getting motivated.¹⁷ Kant, Korsgaard holds, fixes this problem: He first replaces the ill-defined notion of 'desiring' an end, with the notion of *committing oneself* to an end, or *willing* it, properly speaking. Then, he says, since the proposition "whoever wills the end, wills the means" is analytic, we can simply *derive the necessitation of the will* (the motivation) as to the pursuit of our means, of which, we already have a synthetic proposition (the belief). Hence, the formulation "the imperative extracts the concept of actions necessary to this end merely from the concept of a volition of this end". In Korsgaard's own terms:

[W]illing an end just is *committing* yourself to realizing the end. Willing an end, in other words, is an essentially first-personal and normative act. To will an end is to give oneself a law, hence, to govern oneself. That law is not the instrumental principle; it is some law of the form: realize this end. That of course is equivalent to "Take the means to this end." So willing an end is equivalent to committing yourself, first personally, to taking the means to that end. In willing an end, just as Kant says, your causality—the use of means—is already thought. What is constitutive of willing the end is not the outward act

¹⁷ Korsgaard 2008, pp.32–46

of actually taking the means but rather the inward, volitional act of prescribing the end along with the means it requires to yourself.¹⁸

In this vindication of the practicality of reason, Korsgaard thinks, we also find a vindication of morality—because in showing ‘giving laws to ourselves’, or ‘being able to act only under the idea of law’ to be *constitutive of selfhood*, Kant has found the grounds for the bindingness of specifically ‘moral’ laws. In acting purely out of respect for our end-setting (by lawgiving) capacities in general, we act morally: this, must be the meaning of the *Principle of Humanity* of GMS III. Each of us is an “end in itself”, since each of us can rise above his or her inclinations, and be himself or herself, to the extent that he or she can set his or her own ends, by self-motivating via self-legislation. This capacity commands absolute respect, and is the foundation of morality.

14. This interpretation does seem to capture the spirit of GMS III, and help us get a grip on the difficult notion that a ‘will’ is ‘free’, only to the extent that it is a “universally lawgiving will”, and that each will of this sort, is an inestimably worthy “end in itself”. However, Korsgaard too starts with the presupposition as to the continuity of Kant’s moral thought with the British ‘internalist’ project: one cannot cite her explanation of the *Principle of Humanity*, as a justification of it—that would be begging the question. If we can refute her account of the *Principle*, however, then we can refute the presuppositions that are the conditions for its assertion. I think that such a refutation is possible, and that it reveals important clues as to Kant’s conception of reason.

15. Korsgaard claims that we are ‘ends in ourselves’, because we are ‘endsetters in general’; and we are ‘endsetters in general’, because we are ‘lawgivers, in general’ (we rise above our inclinations etc., by lawgiving). Granted, Korsgaard does not equate ‘the

¹⁸ *ibid.*, pp.57-58

law', with the hypothetical imperative: she says that the former contains a bare 'realize this end'. However, it simply contradicts Kant's notion of practical *lawgiving*, to suppose that a law can be issued for the realization of any *particular* end. It is *Allgemeingültigkeit/universal validity* which characterizes a practical law—and, that means neither 'necessarily presupposed in any act of volition', nor 'essential to what makes us human', but rather *unconditional validity*. To anticipate some of what we shall discuss below, such 'conditionless', or 'particular-endless' volition is precisely the way Rousseau's *volonté générale* exercises its will—in Rousseau's words, "the general will changes in nature when it has a particular object, and it cannot, being general, pronounce judgment on a particular man or fact"¹⁹. First, we ought to note that the 'necessitation' contained in a practical precept for the realization of any particular end, is *by definition* a *conditioned* necessitation, and for this reason alone, unsuitable for the formulation of a *law*; Kant explains in KpV:

Reason, from which alone can arise any rule that is to contain necessity, does indeed put necessity even into this precept (for otherwise it would not be an imperative), though it is only a subjectively conditioned necessity and cannot be presupposed in the same degree in all subjects. But it is requisite to reason's lawgiving that it should need to presuppose only *itself*, because a rule is objectively and universally valid only when it holds without the contingent, subjective conditions that distinguish one rational being from another.²⁰

Then, we ought to ask, whether the 'conditioned necessitation' in question is a matter of the reason-guided generation of a 'motive'. In order to answer this question, we need to look at Theorem I of KpV:

¹⁹ Contrat 57

²⁰ Ak V.20–21

All practical principles that presuppose an object (matter) of the faculty of desire as the determining ground of the will are, without exception, empirical [*insgesamt empirisch*] and can furnish no practical laws.

Does this mean, that if the “determining ground of the will” is “altogether empirical”, then there is no ‘real willing’, no self-motivation, hence no agency involved—but merely a ‘being drawn’ in the absence of lawgiving? Kant continues:

By “the matter of the faculty of desire” I understand an object whose reality is desired. Now, when desire for this object precedes the practical rule and is the condition of its becoming a principle, then I say (*first*) that this principle is in that case always empirical. For, the determining ground of choice [*Willkür*] is then the representation of an object and that relation of the representation to the subject by which the faculty of desire is determined to realized the object. Such a relation to the subject, however, is called *pleasure* in the reality of an object. This would therefore have to be presupposed as a condition of the possibility of the determination of choice.

Now, what ‘determines the will’, here, appears to be the representation of a non-real *object* (perhaps, the future state of a real object) which promises the satisfaction of a present need; pleasure ensues, (assuming that we could diagnose our need well, and project a good solution) when alterations in the state of the world come to fit our projected representation. There is no suggestion here, either of lawgiving, or, of ‘desire’, ‘pleasure’, or any other ‘empirical’ factor *working on* an agent. It is not insinuated in any way, that the agent is the owner of his or her actions, only to the extent that he or she can ‘rise above’ such factors: rather, these appear to provide mundane uses of practical reason with external criteria of satisfaction, or conditions to be fulfilled, or tasks to be done. Conspicuous by its absence, is the need for a motive, or any kind of intermediary device between the will-determining representation of a non-real object, and the “realization” [*Wirklichmachung*]. Kant ends the section thus:

Now (*second*) a principle that is based only on the subjective condition of receptivity to a pleasure or displeasure (which can always be cognized only empirically and cannot be valid in the same way for all rational beings) can indeed serve as his *maxim* for the subject who possesses this receptivity but not as a *law* even for him (because it is lacking in objective necessity, which must be cognized a priori); such a principle can, accordingly, never furnish a practical law.²¹

It does not seem likely, that for Kant, the question as to the *objective necessity* of an action is a question as to the possibility of practical ‘principles’, or ‘ought’s *in general*. Perhaps, for Kant, ‘volition’ is simply a matter of what Sellars has called “language-exit transitions”. Perhaps, we need rather to start with the question, why the British internalists required the mediacy of ‘motives’ in order to make reason practical.

IV

16. The British internalists needed ‘motives’ to make reason practical, because they were working within an intellectual milieu, in which reason had been *sanitized* with respect to ‘the ought’ in general. Their commitment to the notion that unaided reason could not be the source of any substantive knowledge, had them conclude that a true proposition concerning matters of fact *could not*, on its own, *in any way* authorize the assertion of another such proposition. Similarly, their commitment to the notion that there were no rationally discoverable ‘goods in themselves’, had them conclude that between the rational representation of *any* ‘good’ and its pursuit, an empirically attestable moving factor had to intervene. I think, however, that while Kant certainly did share their commitments, he repudiated their conclusions in the theoretical, and the practical domain alike. How, then, did the British ‘sanitation of reason with respect to the ought’ take place? Why was Kant’s ‘reformed reason’ radically different from theirs?

²¹ Ak V.38–40

17. The British empiricists found a way of suppressing the *independent* claim-making proclivities of reason, by divesting it of the capacity to contribute to the 'content' of experience altogether. Roughly speaking, reason was relegated to the role of a tool for the arrangement of signs, and the 'signs' themselves—the content of experience proper—had to be supplied directly by the senses. Furthermore, these signs had to be construed atomistically, as 'absolute particulars', since they could neither be internally complex (otherwise, 'substantial forms' would make a return), nor related to one another (otherwise, inherently rational necessary connections between things, and first and final causes would be back). Representations of relations could only be justified by reference to the properties of our 'arrangement tool's; and, general representations, or concepts, were nothing but complex particulars of the mind's own making, concocted out of faint copies of the 'true content' of experience.

18. How, then, did this reason *qua* 'arrangement tool' function? We might begin by noting, that the notion of 'arrangement' that seems to be at work here, is very similar to our contemporary notion of 'logical form'. We specify the logical form of a language, by specifying its truth-functional connectives, and the signs that must be put to use in order to express 'atomic' truth claims. On their own, the connectives do not carry any sort of information about the world; rather, each represents a possible configuration of truth values. E.g., using the signs of propositional logic, to say ' $P \rightarrow Q$ ' is not to say, *the truth of P entails the truth of Q*; rather, it is simply to make the kind of claim that is false when P is true and Q is false, and true on any other configuration. Now, if we suppose that our 'early-modern' British empiricist could wield the resources of our first-order logic, how would he respond, if we were to ask him how he makes sense of the claim, "if you heat a piece of metal, then it expands"? He would probably say the following:

Let $M(x)$ stand for x is a metal, $H(x)$ for x 'is heated', and $E(x)$ for 'expands'. We find after repeated observations, that in cases where $M(a, \text{ or } b, \text{ or } c \dots)$ and $H(a, \text{ or } b, \text{ or } c \dots)$ both obtain, $E(a, \text{ or } b, \text{ or } c \dots)$ obtains soon afterwards; so in order to explain this kind of phenomenon, we come up with a formula: although we know that it is 'underdetermined' by experience, we assert that $\forall x((M(x) \wedge H(x)) \rightarrow E(x))$. The formula does all of its explanatory work, by universal instantiation, and the meanings of the truth-functional connectives involved. We also know when to abandon it; if, say, we find that $M(z) \wedge H(z)$ is true, but $E(z)$ is false, we understand that we have made a mistake. Now, compare all of this, with what the rationalist would say: he would search for 'expandability' inside 'heatedness' and speculate as to their connection with 'metalhood'; or even worse, try to argue that 'heatedness followed by expandedness at this precise moment in the unfolding of the history of the universe is part of the infinitely specified individual essence of the individual constant in question'.

19. Now, we can imagine the rationalist's response to be thus: Well, take a look at this—a piece of metal 'p' that shrinks when cooled. Does your formula 'explain' this phenomenon as well? The Empiricist replies: Of course not; all that my formula says, is that if this piece of metal were heated, then, it would have expanded. The rationalist says: But that cannot be what your formula really wants to say, if it is to do all of its 'explanatory' work by universal instantiation and the meanings of the truth-functional connectives involved. In this case, we just 'instantiate', and get the following: $M(p) \wedge H(p)$ is false, and so is $E(p)$; thus, your conditional comes out true. Have you thus 'explained' this phenomenon? If you really want to say, "if this piece of metal were heated, then, it would have expanded", then you would have to resort to the sort of formulation, which you have been persistently avoiding, lest you authorize 'pure reason' to make an independent claim: that the truth of $M(x) \wedge H(x)$ entails the truth of

E(x). But, how would Kant respond? About ‘hypothetical judgments’ in the published logic lectures, he says,

The matter of *hypothetical* judgments consists of two judgments that are connected with one another as ground and consequence. One of these judgments, which contains the ground, is the *antecedent* (*antecedents, prius*), the other, which is related to it as consequence, is the *consequent* (*consequens, posterius*), and the representation of this kind of connection of two judgments to one another for the unity of consciousness is called the *consequentia* [*Konsequenz*], which constitutes the *form* of hypothetical judgments.

In categorical judgments nothing is problematic, rather, everything is assertoric, but in hypotheticals only the *consequentia* is assertoric. In the latter I can thus connect two false judgments with one another, for here it is only a matter of the correctness of the connection – the form of the *consequentia*, on which the logical truth of these judgments rests. There is an essential difference between the two propositions, All bodies are divisible, and, If all bodies are composite, then they are divisible. In the former proposition I maintain the thing directly, in the latter only under a condition expressed problematically.²²

What Kant appears to be saying here, is that the consciousness of the ‘Konsequenz’ comes *before* any kind of commitment to the truth-values of the antecedent and the consequent. And by ‘Konsequenz’, Kant appears to want to say ‘assertability under a condition’. Now, we ought to recall, from the Transcendental Dialectic of KrV, Kant’s dictum that our search for ‘conditions of assertion’ can never come to an end (there simply are no categorical judgments at which we can *stop asking why*): Every assertion-candidate gets its ‘permission’ from another assertion-candidate, and issues a ‘permission’ for still another assertion-candidate. Thus, it appears that in the case of the heated piece of metal, the Kantian position is closer to that of the rationalist—but while the rationalist does not probe any further the ‘entailment’ relation which he postulates,

²² Ak IX.105

Kant understands this in terms of the permissibility of the sort of inference that is sanctioned at the transcendental level by the category of causality, and held in check at the empirical level, by the independent (often intractable, irregular) behavior of the things that we are trying to understand.

20. What, then, do we mean by 'permissibility'? Now, we ought to recall, that one of the ways in which Kant describes his pure concepts of the understanding, *and* his pure intuitions of space and time, is that they are "original acquisition"s, rather than 'innate' representations²³. This dictum has been interpreted to say, that the 'pure' representations in question are *acquired*, because, first, they have to be occasioned by an independent factor (the givenness of sensation), and second, because it takes the detailed analyses of a work such as the *Kritik*, to make these into fully transparent, *explicit*, concepts. They are a priori on the other hand, or "original", because their content is not derived from experience, but is intrinsic to the nature of the human mind.²⁴ Now, we cannot address the question as to whether this is a good interpretation here; but we can remind ourselves, that "original acquisition" is a technical term in Kant's *Doctrine of Right*. Speaking very roughly, an "original acquisition" is an 'original claim': For a person to claim possession of a part of the world, that is not part of his or her body, is for that person to claim i) that it is to remain his or hers, when he or she is not physically present in that part of the world; and ii) that an infringement of this claim constitutes an injury to his or her person. An original claim, then, is the kind of act which opens up the space of *permissions and prohibitions*, and contains at least a provisional notion as to the conditions for their justification. Now, I shall try to figure out the exact sense in which the categories and the pure intuitions of space and time

²³ Ak VIII.223

²⁴ Grüne 2008, Longuenesse

might be 'original claim's in a later chapter, but the point which I would like to stress here, is that for Kant, an ambience of cognitive *norms of conduct* (opened up, in one way or another, by the pure concepts and intuitions) appears to be *constitutive* for concept use in general. Kant calls a concept that has no real content, a "usurpatory" concept—i.e., a concept which tries, as it were, to make an exception of itself within the space of cognitive 'permissions and prohibitions'. The question as to whether a general representation has any 'objective reality' for Kant, is not a question as to whether it contains a mental copy of a thing of the world, but rather, a *quid juris* question. It follows, that for Kant the simplest claim-making *essentially* involves knowing how to behave with respect to 'ought's, or knowing how to partake in a system of externally legislated permissions and prohibitions.

21. Let us, then, return to the above-quoted passage from GMS, about the 'analyticity' involved in the hypothetical imperative. The reason why Kant thinks that the possibility of such an imperative requires no "special discussion", may be that it concerns simply another use of the conceptual 'ought': here the 'norms of conduct' govern not intralinguistic moves, but language-exit transitions. Let us, then, take Kant's claim, that "the imperative extracts the concept of actions necessary to this end merely from the concept of a volition of this end" at face value, and see if we can make sense of this 'derivation' as a conceptual derivation, rather than the derivation of a 'motive'.

22. We ought to begin, I think, with the notion of a maxim. Kant describes a 'maxim' as a "subjective principle", and this is usually glossed as 'general policy of action'. Both 'principle', and 'policy', however, are loaded terms which obscure the crucial distinction between a maxim and an imperative. I suggest, that a maxim is very similar to a universal/assertoric/categorical *judgment*, with the difference that its copula is in the subjunctive, rather than the indicative mood. E.g., 'clothes be clean', 'promises be

kept', 'profit be made'; after one of the examples that Kant gives, 'injuries be revenged' (states of the world where I find myself injured, to become states of the world where I am 'even'), and so on. The 'material' of a maxim is the representation of a non-real object—an 'S as P', rather than an 'S is P'—and its 'Form' consists in its 'to be realizedness'. Borrowing a term of Sellars's, I shall call maxims *oughts-to-be*.

23. For the 'material' of a maxim to "determine the will", Kant says, is for the agent to act in accordance with a hypothetical imperative. Thus on the model just suggested, the transition from an ought-to-be to an ought-to-do would lead to an imperative of this sort:

If *S as P* is to be realized, then, I ought to make *the content of P (the 'marks', of which P is comprised) true of S*.

This is to say the following: Nothing takes on a new quality, merely by my imagining; and neither can I instantaneously cause things to get new qualities. So, I have to interfere in the business of the world one step at a time: I have to come up with a 'to do list', to guide my language-exit transitions. This relation between my 'to do list', and my conception of the end to be realized, is *similar to* the relation between the subject and predicate terms of an analytical judgment: In special cases, I may indeed simply derive a 'to do list' analytically from my conception of the end; but much more importantly, my conception of the end remains the only standard against which not only my 'to do lists' (however obtained), and alterations in the state of things get compared, but my 'doings' get *evaluated* as well—just as, for an analytical judgment, the *only* criterion of truth is that the predicate term not contradict the content of the subject term.

24. One reason why such 'determinations of the will' are not suitable to yield universally valid laws, is that they prescribe that something be done, under a condition. There is, however, another reason why Kant thinks 'Materie-determinations of the will'

are entirely unsuitable to yield *moral* laws. However refined such end-setting and means-pursuing behavior may be, it is, in the last analysis, nothing more than *an effort to blend successfully into the causal economy of the universe, to let oneself be ruled by it, to go willingly under the tutelage of 'things'*. Kant finds this self-negating for practical reason. As we have discussed in the context of Theorem I of KpV, however, this cannot be because he thinks that in trying to realize particular ends, we forgo reason, dissolve our 'selves' altogether, and abandon ourselves to the sway of the inclinations; I suggest, then, that it is rather a matter of remaining in a state of *minority/immaturity [Unmündigkeit]* with respect to the universe. Kant does say on a number of occasions²⁵, that the precepts of *technical*-practical reason, into which precepts of prudence are reducible, are in turn reducible into *theoretical* propositions. The real evidence, however, in support of this interpretation lies in the analysis of Kant's eventual answer to the question as to the *bindingness* of moral precepts.

25. Now, if there are moral laws, then we ought to look for them in 'determinations of one's will' with respect only to the *form* of a maxim, or its sheer *to-be-realizedness*. Thus, 'oughts-to-do' which result from such determinations are—to introduce another clumsy term—*simply-to-be-dones*. An even more clumsy term would be, 'to-be-dones-for-the-sake-of-the-deed-alone'. That there are *simply-do-be-dones*, that we are conscious of their demands, and that we are capable of acting in response to such demands, do not even come into question for Kant; GMS begins with the very notion, arguing that the simplest human understanding is capable of acting out of *duty*. KpV makes the same point using more technical terms:

²⁵ Reference from KU, GMS

The most common understanding can distinguish without instruction what form in a maxim makes it fit for a giving of universal law and what does not.²⁶

The philosopher's task consists in accounting for the *synthetic* connection between having a will, and being subject to unconditional demands; viz., in developing a reflective understanding of what justifies them. Thus, his or her question is a *quid juris* question; and it has to be answered *a priori*.

V

26. It is helpful, I think, to take a look at the development of Kant's thinking on this question. Kant was committed, long before the 'critical period'—i.e., long before autonomy and freedom had become the central concepts of his practical thought—to the principles, that moral precepts make *unconditional* demands, and that the moral philosopher's task consists in accounting for their *Verbindlichkeit*. These commitments are expressed unequivocally in his 1764 Preisschrift, the *Untersuchung über die Deutlichkeit der Grundsätze der natürlichen Theologie und der Moral*. The relevant passage is worth quoting at length:

§2. *The fundamental principles of morality in their present state are not capable of all the certainty necessary to produce conviction.* In order to make this claim clear I shall merely show how little even the fundamental concept of obligation is yet known, and how far practical philosophy must still be from furnishing the distinctness and the certainty of the fundamental concepts and the fundamental principles which are necessary for certainty in these matters. The formula by means of which every obligation is expressed is this: one ought to do this or that and abstain from doing the other. Now, every ought expresses a necessity of the action and is capable of two meanings. To be specific: either I ought to do something (as a means) if I want something else (as an end), or I ought immediately to do something else (as an end) and make it actual. The former may be

²⁶ Ak V.27

called the necessity of the means (*necessitas problematica*); and the latter the necessity of the ends (*necessitas legalis*). The first kind of necessity does not indicate any obligation at all. It merely specifies a prescription as the solution to the problem concerning the means I must employ if I am to attain a certain end. If one person tells another what actions he must perform or what actions he must abstain from performing if he wishes to advance his happiness, he might perhaps be able, I suppose, to subsume all the teachings of morality under his prescription. They are not, however, obligations any longer except in the sense, say, in which it would be my obligation to draw two intersecting arcs if I wanted to bisect a straight line into two equal parts. In other words, they would not be obligations at all; they would simply be recommendations to adopt a suitable procedure, if one wished to attain a given end. Now since no other necessity attaches to the employment of means than that which belongs to the end, all the actions which are prescribed by morality under the condition of certain ends are contingent. They cannot be called obligations as long as they are not subordinated to an end which is necessary in itself. Take the following examples: I ought to advance the total greatest perfections or: I ought to act in accordance with the will of God. To whichever of these two principles the whole practical philosophy is to be subordinated, the principle chosen must, if it is to be a rule and ground of obligation, command the action as being immediately necessary and not conditional upon some end. And here we find that such an immediate supreme rule of all obligation must be absolutely indemonstrable. For it is impossible, by contemplating a thing or a concept of any kind whatever, to recognise or infer what one ought to do, if that which is presupposed is not an end, and if the action is a means. But this cannot be the case; if it were, our principle would not be a formula of obligation; it would be a formula of problematic skills.²⁷

Now, we ought to keep in mind that already in the *Nova Dilucidatio* of 1755, Kant had repudiated the Leibnizian theses as to pre-established harmony, and the infinite analyticity of contingent truths²⁸. He had, in short, divested Leibniz's Principle of Perfection—the point at which *all questions why* were supposed to come to an end—of any authority whatsoever: thus, he could not appeal to reason for the discovery of an “in itself necessary end”. Therefore, he first turned to Hutcheson, who had spoken of a

²⁷ Ak II.298–299

²⁸ *Nova Dilucidatio*

special 'moral sense' for the 'disinterested approbation' of deeds; but he clearly did not *fully endorse* Hutcheson's theory either, in that he did not think that a 'pure feeling' could be isolated, that would pick out a 'purely good end, necessary in itself'.²⁹ The Preisschrift says, of Hutcheson, only the following: "Hutcheson and other have, under the name of moral feeling, provided us with a starting point from which to develop some excellent observations"³⁰ And, in his *Observations on the feeling of the beautiful and sublime* which had appeared in the same year—a work that is generally considered the epitome of Kant's 'sentimentalist' period—he says:

Even the vices and moral failings often carry with them some of the traits of the sublime or the beautiful, at least as they appear to our sensory feeling, without having been examined by reason.³¹

In human nature there are never to be found praiseworthy qualities that do not at the same time degenerate through endless gradations into the most extreme imperfection.³²

[True] virtue can only be grafted upon principles, and it will become the more sublime and noble the more general they are. These principles are not speculative rules, but the consciousness of a feeling that lives in every human breast and that extends much further than to the special grounds of sympathy and complaisance. I believe that I can bring all this together if I say that it is the *feeling of the beauty and the dignity of human nature*. The first is a ground of universal affection, the second of universal respect, and if this feeling had the greatest perfection in any human heart then this human being would certainly love and value even himself, but only in so far as he is one among all to whom his widespread and noble feeling extends itself. Only when one subordinates one's own particular inclination to such an enlarged one can our kindly drives be proportionately applied and bring about the noble attitude that is the beauty of virtue.³³

²⁹ See Ak II.299ff.

³⁰ Ak II.300

³¹ Ak II.212

³² Ak II.213

³³ Ak II.217

Here, too, the grounds for ‘true virtue’ are to be found in ‘principles’ that cannot be subjects matter, of theoretical reason: the *Observations* provides only a commentary on the feelings that *accompany* the consciousness of such principles, without reducing the latter into the former. How, then, did Kant get from this point, to the *Principle of Humanity*? One answer, which would vindicate Korsgaard, immediately suggests itself: At this point in his career, Kant was increasingly under the influence of Rousseau. So he must have found his ‘in itself necessary end’, in the *equal* (he would later say: *because, inestimable*) *worth* of all human beings. He must have found, that is, the ‘unconditional good’ that would justify the *bindingness* of moral precepts, in the purity of heart of Rousseau’s *noble savage*, before the alleged advances of civilization corrupt his will, and make it partial to himself. Thus, the *Principle* reminds us of this ideal of purity, to which we must strive, by learning to act only out of ‘universal respect’ for the end-setting/means-pursuing capacities of mankind, etc.

27. A conclusion of this sort, however, rests on several unwarranted assumptions. To begin to unravel these, we need to start, again, from 1755.

28. In the *Nova Dilucidatio*, Kant had rejected pre-established harmony, the infinite analyticity of contingent truths, the ‘windowlessness’ of monads, and the Principle of Perfection; furthermore, he had radically modified the Leibnizian understanding of the Principle of Sufficient Reason—and, he had done all of this work, in the context of a *propaedeutic* to a *reformed metaphysics*, which would provide ‘foundations’ for Newtonian physics. That reformed metaphysics appeared in the shape of a *Physical Monadology*: its chief ambition was to account for the central ‘unexplained explainer’ of Newton’s celestial mechanics, that motions had to be described with respect to ‘Absolute Space’. Kant’s monads, just like those of Leibniz, were simple substances, and themselves non-extended—but, by virtue of their ‘inner activity’, they generated a certain “sphere”

around themselves; thus, they could occupy space (and thus, we could explain the impenetrability of matter). Unlike Leibniz's monads, however, Kant's monads were in relations of reciprocal *gravitational attraction*. What the physicist called 'Absolute Space', then, could be accounted for, in terms of their gravitational, as opposed to 'pre-established' relations with an eye to the realization of the best of all possible worlds.

29. The very notion of a 'physical monadology' is, of course, absurd. One piece of evidence that we have, in support of the idea that the work became a serious source of embarrassment for Kant, is that he dedicates part of the Dynamics chapter of his 1784 *Metaphysical Foundations of Natural Science*—after all those years, *and* the *Kritik*—to prove the impossibility of physical monads. However, for our present discussion, a more relevant piece of evidence comes in Alison Laywine's work³⁴ on Kant's 'pre-critical' career, and especially the *Dreams of a Spirit-Seer*, of 1766. Laywine argues, I think convincingly, that the chief 'spirit-seer' under attack in that work, is the Kant of 1756, who had tried to establish a system of gravitational attraction between intelligences. I think, that this makes it all the more noteworthy that a subsidiary concern of the *Dreams* appears to be to conduct thought-experiments, as to whether there might *at all* be a sense in which one could *legitimately* talk about 'intelligible beings', and their 'intelligible world' as distinct from the sensible world. The first chapter raises the question, and suggests an answer with distinct echoes of the 20th Century—an 'intelligible being' is the *whole person*, considered under an aspect.³⁵ The second chapter, however, brings up even more interesting ideas:

A secret power forces us to direct our will towards the well-being of others or regulate it in accordance with the will of another, although this often happens contrary to our will,

³⁴ Laywine 1992

³⁵ Ak II.324–325

and in strong opposition to our selfish inclination. The focal point at which the lines which indicate the direction of our drives converge, is therefore not merely to be found within us; there are, in addition, other forces which move us and which are to be found in the will of others outside ourselves. This is the source from which the moral impulses take their rise. ... As a result, we recognise that, in ourmost secret motives, we are dependent upon the rule of the general will. It is this rule which confers upon the world of all thinking beings its moral unity and invests it with a systematic constitution, drawn up in accordance with purely spiritual laws. ... Thus it was that Newton called the certain law governing the tendencies inherent in all particles of matter to draw closer to each other the gravitation of matter, not wishing to entangle his mathematical demonstrations in possible vexatious philosophical disputes concerning the cause of those tendencies. ... Are we, then, to suppose that it would not in the same way be possible to represent the phenomenon of the moral impulses in thinking natures, who are reciprocally related to each other, as the effect of a genuinely active force, in virtue of which spirit-natures exercise an influence on each other? ³⁶

Now, this passage occurs in the second chapter of the first part of the *Dreams*, which part is supposed to be a satire of dogmatic metaphysics—but here, Kant appears to be speaking, or at least conducting a thought-experiment, *in his own voice*. We can say this, on two accounts: In the marginal notes to the *Observations*, he says:

Newton was the first to see systematicity and regularity bound together with the greatest simplicity, where, before him, there was to be found disorder, and with it, an awful multiplicity—and since then, comets move in geometrical paths.

Rousseau was the first to discover, under the multiplicity of forms which humanity assumes, its deeply concealed nature, and the hidden law, according to which providence, through his observations, would be justified.

[...] After Newton and Rousseau, god is justified³⁷

³⁶ Ak II.334-335

³⁷ Ak XX.58-59

And, as we shall discover in a moment, although the 1770 *Dissertation* does not mention any 'intelligible beings' in its short discussion of the 'intelligible world', this Rousseauian reshaping of Kant's earlier monadology makes an unmistakable return, *mutatis mutandis*, in the *Canon* chapter of the *Doctrine of Method* of KrV, and from there onwards, contributes significantly to the over-arching argument of the GMS. Before we return to the critical period, however, we still need to understand how Rousseau could have led Kant, from the problematic of the 1764 Preisschrift, to what I have called the 'thought-experiment' of the *Dreams*.

30. We might begin, by taking a cue from the excerpt just quoted: what could it mean, for Rousseau to have 'justified providence'? This, I suggest, is a way of referring to Rousseau's solution to a problem, that had found its most incisive formulation in Hobbes. Now, Hobbes certainly does seem to suggest in Chapter XIII of his *Leviathan*, 'Of the Natural Condition of Mankind, as concerning their Felicity, and Misery', that his 'state of nature' is one of sheer savagery:

where every man is Enemy to every man ... there is no place for Industry ... no Culture of the Earth ... no Navigation ... no Arts; no Letters; no Society; and which is worst of all, continuall feare, and danger of violent death; And the life of man, solitary, poore, nasty, brutish and short.

To this warre of every man against every man, this also is consequent; that nothing can be Unjust. The notions of Right and Wrong, Justice and Injustice have there no place.³⁸

Yet, on the other hand, it is clear that he is not in the least bit interested in providing a 'conjectural history' of the origins of civilization; and neither do the parties involved in his 'state of war' appear to be sub-lingual savages. To the contrary, they clearly regard themselves *and* their rivals as rational agents, who have *incommensurable* conceptions of

³⁸ *Leviathan* 89-90 (Cambridge Edition)

'the good', and what they are entitled to, in order to attain it—from this follows, that a conception of 'the good' that is not held in check by *any* means, leads inevitably to a right to absolutely *everything*. And, all of this seems perfectly reasonable, in that all human beings have "equal hope in attaining"³⁹ their ends, given Hobbes's radical anti-Aristotelian stance. The state of war, then, is generated by the clash of such incommensurable 'rational' conceptions. Now, Hobbes appeals openly to a sort of *pure reason* (not to a conjectural history, or *any* kind of empirical consideration) to figure out *the only possible principle* upon which the war could at all end. Reason, then, comes up with a *theorem*:

That every man, ought to endeavour Peace, as farre as he has hope of obtaining it [*seek Peace, and follow it*]; and when he cannot obtain it, that he may seek, and use, all helps, and advantages of Warre [*by all means we can, to defend our selves*].⁴⁰

These dictates of Reason, men use to call by the name of Lawes, but improperly: for they are but Conclusions, or Theoremes concerning what conduceth to the conservation and defence of themselves. ... [If] we consider the same Theoremes, as delivered in the word of God, that by right commandeth all things; then are they properly called Lawes.⁴¹

And this, of course, is a point of momentous importance in the history of ideas. Hobbes plucks, as it were, the notion of 'natural law' (in the political sense) from the inherently purposive rationality of the world, and turns it into a matter of *human institution*. However, the idea of *social contract* derived from the 'theorem' in question paradoxically negates the very ability to think for oneself, or have one's own ends: the sides to the contract do not agree to mutually limit their claims to the things of the world, and within the confines of such mutual limitation, to let every man pursue what

³⁹ *ibid.*, p.87

⁴⁰ *ibid.*, p.92

⁴¹ *ibid.*, p.111

he thinks is 'good'. Rather they transfer *all* their rights, with the exception of a single inalienable right, that of *self-preservation*, to the sovereign. It appears, then, that for Hobbes, we either play the 'civic game' by its proper rules—those of *obedience*, i.e., deliberate minority: what Kant would call *Unmündigkeit*—and attain the end of *peace*; or, we insist on 'playing by our own book' and face the consequences, i.e. *war*.

31. Now, Rousseau's earlier political thought may be considered an attempt to turn this picture upside down. Instead of starting with some 'purely rational' theorem as to practical objectivity, Rousseau resorts, in his second *Discours* (1755), to a conjectural history—to work out the 'rules' of the 'civic game' from the inside out. Thus the occupants of *his* state of nature are *noble savages*, who do not even have any consciousness of themselves as rational end-setters/means-pursuers. That consciousness is awakened only through the 'unfortunate' advance of culture; and only when it is fully awakened, do human beings find themselves in a *Hobbesian* state of nature. In this early work, however, there is no systematic solution as to a way out of such a state. But, if that is so, then we cannot just suppose the 'purity' of the *noble savage* to be Rousseau's last account of practical objectivity. That account comes seven years later in *Du Contrat social* (1762), of which the project Rousseau describes thus:

To find a form of association that will defend and protect the person and 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.⁴²

Hobbes, of course, had disregarded the crucial last phrase of this formulation—one's 'liberté' had to be given up altogether, in order to put an end to the war of incommensurable understandings of the good. For Rousseau, however, one cannot

⁴² Contrat 39

simply renounce a “*liberté naturelle*”: rather, it transforms itself through the act of contract; and, it does this by becoming conscious of its inalienability. Thus, the criterion for the objective validity of a ‘doing’, a ‘claiming’, and any other mode of behavior of which a rational agent is considered capable, becomes its *compatibility* with the absolutely inalienable right, of every rational agent, to self-determination. It follows from this, that the fundamental criterion for the legitimacy of a social order is that it be a system of *equal freedoms*. In committing themselves to such a criterion of legitimacy, parties to the contract make each other’s ends, ‘their own’ ends; and, this does not mean that they unite their wills into a collective, or that they discover the ‘essential element in all human volition’, and make it guide their actions. Rather, a reflective dimension is added to their rational agency, and it becomes a duty on the part of each participant in the ‘civic game’, to regard himself or herself not simply as a ‘particular will’ setting his or her particular ends and pursuing the means for them (and either doing this well or badly), but at the same time as the organ of an absolutely impartial, *impersonal*, kind of volition which maintains the system of equal freedoms by holding such particular strivings in check. It is this impersonal volition, that is to be called the *volonté générale*.

32. Thus, for Rousseau, the ‘civic game’, understood as a system of *equal freedoms*, can justify itself to itself, without reference to anything outside of itself; neither to Hobbes’s ‘biological’, even *visceral* criterion, nor to Leibniz’s supernatural criterion, viz., the Principle of Perfection. The absolute impartiality—*exceptionlessness*—of the legislation of the general will solves the problem of the 1764 Preisschrift: Unconditional practical demands, insofar as they can be justified, have to be understood as the demands of an absolutely impartial will. Kant is so elated to have found such a solution, that he promotes Rousseau to the rank of ‘the Newton of practical reason’.

33. Now, what does it mean to say, that he applies Rousseau's principle to the inhabitants of an intelligible world? We can now turn to this question.

34. The thought-experiment of the *Dreams* becomes a positive thesis of the *Kritik*. To see how, we ought to start with the *Amphiboly* chapter. The main contention of this chapter of KrV is that both the 'dogmatic rationalist' and the 'empiricist' were incapable of "transcendental reflection" into the two distinct "sources" of our cognition, the "receptivity of impressions", and the "spontaneity of concepts"; as a result, Kant says, they conflated the functions of the one, with the other. The rationalist thought that our cognition consists in the comparison and contrast of *received* conceptual content; and, that in its infinite complexity, this content has to appear to our limited intellects, as 'sense perception'. Accordingly, for the rationalist, judgmental form could have no constitutive role, but only one of 'arranging' given content; and, the forms of our sensibility had to have an even lesser role, 'arranging' nothing more than confused conceptual representations. Such commitments, Kant says, lead *inevitably* to 'monadology':

If ... I abstract from all conditions of intuition, and restrict myself solely to the concept of a thing in general, then I can abstract from every outer relation, and yet there must remain a concept of it, that signifies no relation but merely inner determinations. Now it seems as if it follows from this that in every thing (substance) there is something that is absolutely internal and precedes all outer determinations, first making them possible, thus that this substratum is something that contains no more outer relations in itself, consequently that it is simple (for corporeal things are still always only relations, at least of the parts outside one another); and since we are not acquainted with any absolutely inner determinations except through our inner sense, this substratum would be not only simple, but also (according to the analogy with our inner sense) determined through

representations, i.e. all things would really be monads, or simple beings endowed with representations.⁴³

It goes without saying, that for the purposes of theoretical philosophy, this argument serves as the major premise of a *reductio* of Leibniz's monadology: such simple substances cannot be the objects of *any* sort of theoretical cognition. Kant dedicates one third of his *Transcendental Dialectic* to a refutation of the very idea of a 'science' of such 'entities', or a 'transcendental psychology'; he also repeats, on several occasions, that the cognition we have of our empirical selves cannot be a cognition of the absolutely 'inner' determinations of a 'self in itself'. However, he also seems to be suggesting, that if there were a context, in which we *could* legitimately talk of 'objects' that are independent of the system of outer relations which constitute the objects of *experience*, then, we would *have to* treat those objects as monads. However, there certainly is such a context: we are *required to* think of our 'moral selves' to be independent of the outer relations which constitute the objects of experience. The *Canon* chapter of the first *Kritik's Doctrine of Method* develops the practical, as opposed to the theoretical/problematic, notion of *freedom*, thus:

I assume that there are really pure moral laws, which determine completely *a priori* (without regard to empirical motives, i.e., happiness) the action and omission, i.e., the use of the freedom of a rational being in general, and that these laws command *absolutely* (not merely hypothetically under the presupposition of other empirical ends), and are thus necessary in every respect.⁴⁴

I call the world as it would be if it were in conformity with all moral laws (as it *can* be in accordance with the *freedom* of rational beings and *should* be in accordance with the necessary laws of morality) a *moral world*. This is conceived thus far merely as an intelligible world, since abstraction is made therein from all conditions (ends) and even

⁴³ A283/B339–340

⁴⁴ A807/B835

from all hindrances to morality in it (weakness or impurity of human nature) ... The idea of a moral world thus has objective reality, not as if it pertained to an object of an intelligible intuition (for we cannot even think of such a thing), but as pertaining to the sensible world, although as an object of pure reason in its practical use and a *corpus mysticum* of the rational beings in it, insofar as their free choice under moral laws has thoroughgoing systematic unity in itself as well as with the freedom of everyone else.⁴⁵

Shortly afterwards, Kant confirms that he is offering a ‘rousseauified’, and *practically sanctioned* version of Leibniz’s ‘Realm of Grace’:

Leibniz called the world, insofar as in it one attends only to rational beings and their interconnection in accordance with moral laws under the rule of the highest good, the *realm of grace*, and distinguished it from the *realm of nature*, where, to be sure, rational beings stand under moral laws but cannot expect any successes for their conduct except in accordance with the course of nature in our sensible world. Thus to regard ourselves as in the realm of grace, where every happiness awaits us as long as we do not ourselves limit our share of it through the unworthiness to be happy, is a practically necessary idea of reason.⁴⁶

Thus, in GMS, when Kant is referring to an “end in itself”, or to a “universally legislating will”, he must be referring to a *very specific* sort of *will*. The will that *can only* act under the ‘idea of law’, the will that is free, or ‘truly itself’, only when *all* of its maxims can become universal laws, the will that can only be bound by the laws that it legislates for itself, is the will of an agent considered, *from the practical perspective perfectly literally*, as a citizen of the republic of *monads*, legislating in the capacity of an organ of its *volonté générale*. Kant traces the very source of the moral law, and the grounds for its *bindingness*, to the *volonté générale* of the perfectly just order of a projected, supersensible community:

⁴⁵ A808/B836

⁴⁶ A812/B840

The concept of every rational being as one who must regard himself as giving universal law through all the maxims of his will, so as to appraise himself and his actions from this point of view, leads to a very fruitful concept dependent on it, namely that of a *kingdom of ends*.⁴⁷

Thus, as he rounds up the argument of GMS:

And so categorical imperatives are possible by this: that the idea of freedom makes me a member of an intelligible world and consequently, if I were only this, all my actions *would* always be in conformity with the autonomy of the will; but since at the same time I intuit myself as a member of the world of sense, they *ought* to be in conformity with it; and this *categorical* ought represents a synthetic proposition a priori, since to my will affected by sensible desires there is added the idea of the same will but belonging to the world of the understanding—a will pure and practical of itself, which contains the supreme condition, in accordance with reason, of the former will [because, as Kant explains shortly before this passage: the ‘world of understanding’ contains the ‘ground’ for the ‘world of the senses’, and thus, its laws as well]; this is roughly like the way in which concepts of the understanding, which by themselves signify nothing but lawful form in general, are added to intuitions of the world of sense and thereby make possible synthetic propositions a priori on which all cognition of a nature rests.⁴⁸

And, we should also note that although in the second *Kritik* Kant renounces the idea that the ‘possibility of categorical imperatives’ requires a dedicated deduction, and establishes this possibility solely by reference to a “Faktum der Vernunft”, he *still* accounts for the ‘inseparable’ ties between the consciousness of the moral law and that of one’s freedom of the will using the language of GMS, and of the *Canon*:

[The *Analytik* of KpV, Kant says,] shows that pure reason can be practical—that is, can of itself, independently of anything empirical, determine the will—and it does so by a fact in which pure reason in us proves itself actually practical, namely autonomy in the principle of morality by which reason determines the will to deeds. At the same time it

⁴⁷ Ak IV.433

⁴⁸ Ak IV.454

shows that this fact is inseparably connected with, and indeed identical with, consciousness of freedom of the will, whereby the will of a rational being that, as belonging to the sensible world cognizes itself as, like other efficient causes, necessarily subject to laws of causality, yet in the practical is also conscious of itself on another side, namely as a being in itself, conscious of its existence as determinable in an intelligible order of things—conscious of this not, indeed, by a special intuition of itself but according to certain dynamic laws that can determine its causality in the sensible world; for it has been sufficiently proved elsewhere that freedom, if it is attributed to us, transfers us to an intelligible order of things.⁴⁹

On the basis of these observations, then, I repeat the claim with which I began: Rather than derive the *Universal Principle of Right* from the *Categorical Imperative*, Kant derives the latter from a version of the former (insofar as he had inherited it, from Rousseau). Kant's moral thought, is quite literally a 'supersensible edition' of what later became his *Doctrine of Right*.

VI

35. If this interpretation is true, what of significance does it reveal about Kant's moral thought? We might begin answering this question, by reflecting on the very last passage quoted in the previous section. Why should the idea of 'freedom', if it is attributed to us, place us in an intelligible order of things? Let us rephrase the question in the light of the foregoing analysis: If there are no rational grounds within the Kantian framework, for us either to doubt, or to try to prove to ourselves our ownership to our actions, then the sort of 'freedom', which can be attributed to us *only* to the extent that we can count ourselves as partaking in an *extra-worldly volonté générale*, has to be the sort of 'freedom' that is qualitatively different from *anything* we can experience in the mundane exercises of our will—so, of what sort is this 'freedom'? What are we deprived of, when we are

⁴⁹ Ak V.42

deprived of it? What becomes of us, if this 'freedom' cannot be attributed to us? I think we have already rehearsed the answer in our discussion above: We end up being *Unmündig* with respect to *the universe*; we try to blend into its ways, thus *let* ourselves be ruled by it. We shall return to this point in a moment; but before we do that, we ought to ask several follow-up questions: What does it mean to be part of an extra-worldly, or world-independent general will? That sort of will *exists only in public*. It is a matter of the mutual limitation of agents' power of choice, exercised publicly, with clear 'empirical' demands, 'empirical' criteria of success and failure, and so on. The legislation of that kind of will cannot somehow be 'emulated' by a person on his or her own; trying to reproduce the dynamics of mutual limitation within one's own 'moral psyche' leads to bizarre notions such as countering the 'claims' of our desires and inclinations, with the claims of reason. Besides, the core of the *Prinzip der Menschheit*, the idea that "we are subject to no law or principle of action that we do not 'legislate for ourselves'"⁵⁰, while it makes perfect sense in a Rousseauian *public* setting, turns into a veritable puzzle in Kant's moral monadology. Roughly, the 'public' expansion of this principle would be, that in the exercise of our 'private' wills, we can be legitimately constrained only by those principles that the absolutely impartial 'general will', of which we are part if we are to play the 'civic game' at all, has legislated. Kant's monadic will, on the other hand, if it takes this principle to heart, would not know what to do, about *any* 'principle of action' that it encounters: its only option, it seems, would be to construct possible worlds, or situations, and *keep* asking itself, "what would the general will do?"

⁵⁰ This formulation, from Pippin 2008, p.65

36. What, then, does it mean to 'self-negatingly blend into the causal economy of the world', or be *Unmündig* with respect to the universe? Can we even make sense of such notions? We can imagine how a 'virtue ethicist', who has an unfavorable opinion of Kant, would respond: *There is no reasonable answer to such questions, and all that they do, is to bring out how utterly impoverished Kant's practical thought really is. The only method at its disposal for the evaluation of human actions, is to measure them as performances against the demands of an imperative. The giving and taking of imperatives, however, constitutes merely one aspect of the rich fabric of 'practically rational' human life. To illustrate, let us pick a simple example from an ancient tradition that Kant was very familiar with: Vergil's Aeneas, and his commitment to the ideal of being pious. On the Kantian scheme, it would appear that being pious is either demanded by a universally valid, unconditional command, or, it is some state of the world to be "realized". If it is the latter, then the unpalatable conclusion would follow, that Aeneas does what he does, in the guidance of a 'checklist' of sorts—carry father on your back, do not listen to the Carthaginian queen, etc.—gathering points, as it were, on the way to a piusness certificate. What is even worse, getting that certificate would be, according to Kant, a self-negating effort to 'blend successfully into the causal economy of the world'. But, neither can we say that being pious is an unconditional demand, on all rational beings, insofar as they are capable of commitments in general: There is nothing 'general' or 'unconditional' about aspiring to the ideal of pious. This is a requirement on Aeneas, insofar as he is the 'head of a household'. The problem appears to be that Kant is wholly unable to recognize such 'material final ends', of which the normative import can only be understood in terms of the concrete reality in which they are embedded, yet still be worthy 'in themselves'. Things that truly matter to people leading human lives, are formulated in the language of such 'material final ends'.—this, however, is not the conclusion that I shall draw, because it is not true that Kant is unable to recognize 'material final ends', and it is not true that he is trying blindly to*

impose empty formal criteria on the ‘evaluative’ dimension of human life. The ‘problem’ with Kant’s ‘pure principles’ is not their emptiness, or their contrivedness: rather, it is their ambition to *comprehensiveness*.

37. Both in his practical thought, and as I shall try to show in the rest of my work, his theoretical thought, Kant appears to pick out apodictic principles that make perfect sense in a well-defined domain, and generalize, or *radicalize* them into metaphysical principles. The absolute ‘formality’ of the *Universal Principle of Right*, for instance, works perfectly well in the political domain; and Kant has given a wonderfully concise account as to how, in his *What is Enlightenment?*: The idea that there are *given* ‘material final ends’ is rejected; rather, the specification of such ends becomes a matter of our *negotiating* them, in the guidance of an absolutely impartial principle, among ourselves, and, *with the ‘independent behavior’ of the things of the world*. On the ‘material’ side of things, we design ‘functional’ methods to cope with each other, and with our ever-changing needs in an ever-changing world—we set up *institutions*, and ‘bind’ our *private use of reason* (just as the parties to Hobbes’s contract bind themselves to the sovereign) to their functions, for the sake of the ends we have agreed to pursue. The one ‘formal’ constraint on our institution designs, however, is that they be compatible with the right to self-determination of *all*. Now, it follows from this principle, first of all, that no one may be used as a mere means, a *thing*, in the pursuit of anyone else’s ends—an institution loses its legitimacy altogether, if the slightest leeway is encountered, among the principles of its constitution, as to making persons into things. Just as importantly, however, it follows that if an institution is *failing* to realize the ends for the sake of which it was designed, then the participants have every right to make the institution’s shortcomings *public*, and call for change (as long as the act of calling for change does not infringe on anyone else’s right to self-determination)—that, is the *public use of reason*, or

the 'formal' mode of reflection into how well we are doing, in negotiating our 'ways of coping with the world' with the 'independent behavior of things'. This, is a formal principle that holds a check on how well we are doing 'materially', and do this from within those material practices themselves, rather than from a wholly independent point of view.

38. I have suggested above, that the simplest sort of claim making in Kant involves knowing how to behave with respect to an 'ought'; and, I have added, that this is a rather 'political' ought, one that exists only in an ambiance of externally legislated permissions and prohibitions. In the rest of this work, we shall see how fruitful this model is, in trying to make sense of the concept of truth; and, we shall not have to start from scratch either: in the latter half of the 20th Century, Wilfrid Sellars has developed a theory of truth along precisely these lines. We shall also see, however, that such a functional/'normative' conception, although it is deeply rooted in Kant's work, is incompatible with his *transcendental* idealism.