

## Kant on Capital Punishment and Suicide

by Attila Ataner, Toronto

In the *Metaphysics of Morals* Kant clearly, and indeed ardently, upholds the state's right to impose the death penalty in accordance with the law of retribution (*ius talionis*). The "principle of equality" as between crime and punishment demands that those who wrongfully kill another should be put to death, for, in having inflicted such an evil upon another, the murderer has effectively killed himself.<sup>1</sup> Kant is quite emphatic on this point: those who have committed murder "must die". Here, he argues, "there is no substitute that will satisfy justice", for there "is no similarity between life, however wretched it may be, and death, hence no likeness between the crime and the retribution unless death is judicially carried out upon the wrongdoer [...]".<sup>2</sup> The *ius talionis* is, for Kant, the basic principle and measure in accordance with which criminal justice functions. Since the *ius talionis* entails a strict equality between crime and punishment, Kant's insistence that only the death penalty serves as the appropriate response to murder (or to any other equally egregious crime) is fairly straightforward.

<sup>1</sup> See MS, AA 06: 332; 473. The English translations consulted are listed below. Citations are to the volume and page of the German text followed immediately by reference to the corresponding page numbers in the English translations.

*Groundwork of the Metaphysics of Morals*. In: *Practical Philosophy*. Ed. and transl. by M. J. Gregor. Cambridge 1996.

*Critique of Practical Reason*. In: *Practical Philosophy*. Ed. and transl. by M. J. Gregor. Cambridge 1996.

*Critique of Pure Reason*. In: *Critique of Pure Reason*. Transl. by P. Guyer and A. Wood. Cambridge 1998.

*The Metaphysics of Morals*. In: *Practical Philosophy*. Ed. and transl. by M. J. Gregor. Cambridge 1996. *On the Common Saying: That May be Correct in Theory, but It Is of No Use in Practice*. In: *Practical Philosophy*. Ed. and transl. by M. J. Gregor. Cambridge 1996.

*Lectures on Ethics*. In: *Lectures on Ethics*. Ed. and transl. by P. Heath. Cambridge 1997.

*On a Supposed Right to Lie From Philanthropy*. In: *Practical Philosophy*. Ed. and transl. by M. J. Gregor. Cambridge 1996

*Perpetual Peace*. In: *Perpetual Peace and Other Essays*. Transl. by T. Humphrey. Indianapolis 1983.

<sup>2</sup> MS, AA 06: 333; 474. "Es giebt hier kein Surrogat zur Befriedigung der Gerechtigkeit. Es ist keine Gleichartigkeit zwischen einem noch so kummervollen Leben und dem Tode, also auch keine Gleichheit des Verbrechens und der Wiedervergeltung, als durch den am Thäter gerichtlich vollzogenen, doch von aller Mißhandlung, welche die Menschheit in der leidenden Person zum Scheusal machen könnte, befreieten Tod."

What remains unclear however is whether judicial killing, even in accordance with the law of retribution, could ever be rightful in the Kantian framework. I would pose the following question in this respect: could practical reason, in its external aspect as the general or universal will, sanction without contradiction the killing, i.e., deliberate execution, of a human being under any circumstances? Or, put differently, could a legislator, bound by the principle of right and the idea of the original contract, frame a law that mandates the execution of certain criminals?<sup>3</sup> I would answer these questions in the negative. Put simply, practical reason, as self-legislating will (*Wille*), cannot posit its own annihilation without contradiction.

This principle is found in Kant's arguments regarding the irrationality of suicide. Practical reason, as it governs the subject internally, cannot sanction the annihilation of the conditions of its own instantiation, namely its embodiment in the physical person. But there is no immediate application of this principle to the death penalty, and Kant obviously did not see an underlying connection between the proscription on suicide and the illegitimacy of the death penalty. Indeed, he would likely deny that there is such a connection to be found within his framework. For, while the proscription on suicide is a question of ethics, and pertains to the (perfect) duties the individual owes to himself, the legitimacy of capital punishment is a question of right, and pertains to the duties of forbearance individuals owe each other by virtue of their co-existence in a system of mutual (external) constraint, i.e., civil society. Thus, Kant does not hesitate to say that public justice – or practical reason as it governs external relations among citizens – may dictate the annihilation of those who have wilfully committed a capital offence. The proscription on suicide would, at first glance, seem to have no effect on this conclusion.

Nevertheless, I believe there is a case to be made against capital punishment that draws broadly on the Kantian proscription against suicide. However, if there is a case to be made from within the Kantian framework at all – i.e., if Kant's own arguments could somehow be used against his endorsement of capital punishment – it must contend with the Kantian division between the ethical and political (or juridi-

<sup>3</sup> These questions presuppose that there is a conceptual interconnection between “practical reason”, the “universal will”, the “principle of right” and the idea of the “original contract”. This relationship has been described by Weinrib (1987, 490) as follows: “Practical reason is the determination of purposive activity by the causality of concepts; similarly, the principle of right, that one person's action must be capable of coexisting with another's freedom, is the determination by the concept of right of the relationships governed by that concept. Both practical reason and the principle of right abstract the form of free choice from whatever content it happens to have, and make this form determine the operation of the free will. The principle of right is therefore the external aspect of practical reason, or practical reason as it pertains to interaction among free wills. Under its external aspect, practical reason or *Wille* becomes the general or universal will (*der allgemeine Wille*). [...] Just as practical reason holds free choice to the requirements of the rational nature of free choice, so the general will, as it functions in accordance with the principle of right, holds the external and practical relationship among those with free choice to the conceptual requirements of that relationship.”

cal) domains, where the former pertains to the inner motives of the individual while the latter pertains strictly to the external relations between members of civil society. For, it could be that the conditions of political autonomy permit the destruction of the (culpable) subject even while ethical autonomy does not. In other words, there may, within the Kantian framework, be no contradiction in the proposition that a citizen, in having committed a heinous crime, effectively commits a kind of suicide by setting himself up for state-imposed capital punishment – for, as Kant puts it, in murdering another, you really “kill yourself”<sup>4</sup> – even if it would be ethically impermissible for such a person to kill himself by way of self-punishment. As Kant insists in his response to Beccaria,<sup>5</sup> to be discussed further below, there is no inherent difficulty with the notion that citizens, as co-legislators of the penal law, may quite legitimately “will” their own punishment *as criminals*; or, that a co-legislator of the penal law may enact and endorse a form of punishment that could, in the event that he becomes a criminal, lead to his own death.

The justification of capital punishment, in light of Kant’s strict division between the ethical and political domains, proceeds along the following lines. The power to punish and coerce rests exclusively within the domain of public justice (a court), and is deployed by way of ensuring that no individual exercises his outer freedom in a manner that is inconsistent with the freedom of another or, more generally, by way of securing a condition of “right”.<sup>6</sup> The death penalty, as a form of punishment, is supposed to be a rightful coercive response to an unrightful hindrance of the freedom of another, i.e., to a violation of right. As an act of public authority, it is supposed to be carried out in the name of right. It would seem, then, that the coercive mechanism by means of which the freedom of all is to be secured may be rightfully deployed to annihilate, absolutely, the freedom of a wrongdoer – who, as a *person*, as a being possessing rational will, is otherwise endowed with an innate right to freedom and, accordingly, to life. It would seem, in other words, that external (or juridical) lawgiving – which, unlike internal lawgiving, constrains the subject via an incentive

<sup>4</sup> MS, AA 06: 332; 473. “[...] tödtest du dich selbst”. Of course, Kant does not here suggest that in murdering another you *will* your own death, but rather that you will an act that is punishable by death; the juridical significance of your culpable act is that, through it, you forfeit your life.

<sup>5</sup> See MS, AA 06: 335; 475.

<sup>6</sup> Right, generally, pertains to the “external and indeed practical relation of one person to another, insofar as their actions, as facts, can have (direct or indirect) influence on each other”. (MS, AA 06: 230; 387. “Der Begriff des Rechts, sofern er sich auf eine ihm correspondirende Verbindlichkeit bezieht, (d. i. der moralische Begriff desselben) betrifft *erstlich* nur das äußere und zwar praktische Verhältnis einer Person gegen eine andere, sofern ihre Handlungen als Facta aufeinander (unmittelbar oder mittelbar) Einfluß haben können.”) The rule of law, and the use of force and coercion in the name of justice, is a matter of securing “right” as the “sum of the conditions under which the choice of one can be united with the choice of another in accordance with a universal law of freedom”. Punishment, then, may be imposed when an individual has exercised their freedom, or made a choice, in a manner that cannot be reconciled with a like freedom on the part of all, i.e., has violated the domain of rightful freedom of another.

drawn from the subject's aversions, i.e., it threatens to deploy force or inflict pain as a consequence of wrongdoing<sup>7</sup> – reserves the power to destroy its subject should he deviate from its core proscriptions. Arguably, external lawgiving may exercise this power despite the fact that, as a person, the subject is also a repository of rational will and thus a source of legislation as such – for, the rational will legislates directly for itself (i.e., it is the source of self-binding duties), and also indirectly, via assent in the original contract, for others (i.e., it is also, ultimately, the source of duties binding others as a matter of right). It is important to note, then, that Kant's endorsement of capital punishment must also effectively legitimate the annihilation, in the right circumstances, of the subject *as rational will* – since the capital offender, as a person, is also a rational and autonomous being, at the very least *in potentia*.

#### *Remarks on the Kantian conception of autonomy*

In order to see in what way Kant's position is problematic, even on his own terms, we may begin by highlighting certain features of his conception of autonomy and rationality. First, we may note that there is an underlying parallel between internal and external lawgiving: in either case, the addressee of the law is at the same time its author.<sup>8</sup> For, according to Kant, we are autonomous beings precisely because the source of the laws that determine our exercise of freedom, or free choice (*Willkür*), ultimately resides within ourselves, i.e., within the will (*Wille*) as practical reason itself.<sup>9</sup> This is to say that there is an essential unity between the morally and politically autonomous subjects: ultimately, it is our capacity for freedom as such, whether internal or external, that places us before the practical law, for the “practical” is “everything that is possible through freedom”.<sup>10</sup> Autonomy, as such, is a matter of exercising our capacity for freedom in accordance with laws issued by reason, which is universally present in all persons.

<sup>7</sup> See MS, AA 06: 219; 383.

<sup>8</sup> In the *Groundwork of the Metaphysics of Morals*, Kant writes: “[...] the will is not merely subject to the law but subject to it in such a way that it must be regarded also as self-legislative and only for this reason as being subject to the law (of which it can regard itself as the author)”. (GMS, AA 04: 431; 81. “Der Wille wird also nicht lediglich dem Gesetze unterworfen, sondern so unterworfen, daß er auch als *selbstgesetzgebend* und eben um deswillen allererst dem Gesetze (davon er selbst sich als Urheber betrachten kann) unterworfen angesehen werden muß.”) And, in the *Critique of Practical Reason*, he echoes this in saying: “[we] are indeed lawgiving members of a kingdom of morals possible through freedom [...] but we are at the same time subjects in it, not sovereigns”. (KpV, AA 05: 82; 206. “Wir sind zwar gesetzgebende Glieder eines durch Freiheit möglichen [...] aber doch zugleich Unterthanen, nicht das Oberhaupt desselben [...].”)

<sup>9</sup> See MS, AA 06: 213; 375. And note: “Since reason is required for the derivation of actions from laws, *the will is nothing other than practical reason*.” (GMS, AA 04: 412; 66; emphasis added. “Da zur Ableitung der Handlungen von Gesetzen *Vernunft* erfordert wird, so ist der Wille nichts anders als praktische Vernunft.”)

<sup>10</sup> KrV: B 828. “Praktisch ist alles, was durch Freiheit möglich ist.”

But practical reason, as our capacity to formulate universally valid principles of conduct and to set unconditionally binding ends (for moral agents), is subject to certain inexorable *formal* constraints internal to itself. At a minimum, practical reason cannot formulate principles of conduct that are self-contradictory, internally (or logically) inconsistent, or incoherent from the standpoint of the will's (*Wille*) own essential attribute as *the* source of universally valid and unconditionally binding moral principles.<sup>11</sup> (This is also evident in light of the universalizability requirement of the Categorical Imperative, which in Kant's first formulation constrains the moral agent to act only in accordance with those maxims that he can at the same time will as a universal law. A maxim will fail the test of validity if, once hypothesized as a universally applicable law, it is shown to contradict its own purported end; and, consequently, any course of action that draws on such a maxim will be morally impermissible.) This is to say, simply, that there are certain things that a rational agent cannot will, *qua* rational agent. Though an agent may freely choose (*Willkür*) to act in accordance with a maxim that contradicts or is inconsistent with the requirements of practical reason, he is something less than fully *autonomous* if he does so.<sup>12</sup>

<sup>11</sup> For Kant, the rational will itself – as opposed to anything independent of, prior or external to it – is the source of its own unconditionally binding ends, i.e., it is *the* source of any possible duty, precisely because it is autonomous and as such stands outside and beyond the blind necessity of nature. The rational will is unique, as it is not part of the heteronomous order of the (external) world. As Kant explains in a statement that precedes the one cited immediately above: “Everything in nature works in accordance with laws. Only a rational being has the capacity to act *in accordance with the representation* of laws, that is, in accordance with principles, or has a *will*.” (GMS, AA 04: 412; 66. “Ein jedes Ding der Natur wirkt nach Gesetzen. Nur ein vernünftiges Wesen hat das Vermögen, *nach der Vorstellung* der Gesetze, d. i. nach Principien, zu handeln, oder einen *Willen*.”) The moral agent is autonomous if and only if his will can bind itself to its own moral decree, which is the basic meaning of “autonomy” in the Kantian idiom. Thus, for Kant, we exhibit our incomparably higher worth as rational beings only if our actions conform to a moral law self-given in our rational will. Moreover, our capacity for autonomy constitutes our *dignity* as human beings: “For, nothing can have a worth other than that which the law determines for it. But the lawgiving itself, which determines all worth, must for that very reason have a dignity, that is, an unconditional, incomparable worth; and the word *respect* alone provides a becoming expression for the estimate of it that a rational being must give. *Autonomy* is therefore the ground of the dignity of human nature and of every rational nature.” (GMS, AA 04: 436; 85. “Denn es hat nichts einen Werth als den, welchen ihm das Gesetz bestimmt. Die Gesetzgebung selbst aber, die allen Werth bestimmt, muß eben darum eine Würde, d. i. unbedingten, unvergleichbaren Werth, haben, für welchen das Wort *Achtung* allein den geziemenden Ausdruck der Schätzung abgiebt, die ein vernünftiges Wesen über sie anzustellen hat. *Autonomie* ist also der Grund der Würde der menschlichen und jeder vernünftigen Natur.”)

<sup>12</sup> This raises a controversial issue that I cannot pursue here but will be of some relevance to the question, to be discussed below, regarding the status of a criminal and his act, i.e., whether it is autonomous or not. The issue is summarized by Schneewind (1992, 330): “[Kant's view of freedom and autonomy in the *Groundwork*] seems to suppose that we are free just when we are acting rationally. But then, if we act irrationally, we are not free. Immoral action is, however, irrational. So it seems to follow that we are responsible only when acting as the moral law requires, and not responsible when we do something wicked. Kant might have had a reply to this objection, but if so he did not give it.”

Thus, for Kant, the irrationality of a practical principle, in the sense suggested here, entails its practical impermissibility. And this applies to the ethical and the juridical standpoints equally, because the operation of reason as a practical faculty is what makes human autonomy as such possible.<sup>13</sup>

Given this position, an argument against capital punishment, broadly based on a Kantian conception of autonomy, would consist in this: as it is irrational (or self-contradictory) and hence impermissible for me to frame a suicidal maxim from an ethical standpoint, it is also irrational and hence impermissible for me to frame a capital penal law as a co-legislator in the social contract from a political standpoint. This is because practical reason, in both its internal and external aspects, cannot sanction without contradiction an (ethical) maxim or a (juridical) law the end of which is human death and the annihilation of the rational will that it necessarily entails (for obvious reasons). Now, let us proceed by disregarding, for the moment, Kant's own affirmation of capital punishment, and try to develop a contrary position in more detail.

#### *Why suicide is irrational according to Kant*

We have noted that, for Kant, the rational will is itself *the* source of moral legislation to which it is subject; and that, accordingly, the capacity for autonomous agency grounds our dignity and unconditional worth as human beings.<sup>14</sup> Because physical embodiment is an essential pre-condition of the rational will (or of the

Kant's response consists, roughly, in drawing a distinction between will (*Wille*) and choice (*Willkür*) as a kind of executive function of will that may or may not follow its dictates, and saying that the free exercise of choice (*Willkür*) is sufficient to ground responsibility or to impute an action to an agent, so that freedom and (moral) autonomy are effectively distinguished. One may act freely and thus be held responsible for, say, a criminal act, even though one was not acting (morally) autonomously. See Beck (1993), Hudson (1991) and Timmons (1994).

<sup>13</sup> It has been noted that the will (*Wille*) is *the* locus of practical reason as the ground determining choice to action (MS, AA 06: 213; 374–375). Thus, the inexorable constraints inherent to practical reason must apply equally in its internal legislative function binding the individual's choice (*Willkür*) and in its external aspect as the general will positing the (positive) laws and policies appropriate to civil society.

<sup>14</sup> See note 11 above and accompanying text. Hill (1999, 417–418) shows concisely that human dignity is connected with our capacity to legislate, which is relevant in our discussion of the Kantian conditions of political autonomy below: "Every rational person has dignity as a legislator of the moral law. [...] Humanity, or rational nature in all persons, is taken to have a special status: it has dignity, an unconditional and incomparable worth, above all price, and without equivalent. [...] [According to Kant] what gives humanity this special status is 'the idea of the will of every rational agent as a will giving universal law'. [GMS, AA 04: 431–432; 81–82] All rational agents are pictured as together law-makers and subjects in an ideal analogue of political community, the kingdom of ends, where they legislate not from private interest or commitment to prior authorities but with an impartial regard for the humanity of each co-legislator."

*noumenal* self) being present in the world,<sup>15</sup> death entails its annihilation; death negates an essential condition under which any exercise of autonomy is possible (at least in this life).<sup>16</sup> The question concerning suicide is therefore this: may the rational will posit its own annihilation as a practical end; or, is self-destruction a conceivable autonomous determination? For Kant, the answer is “no” because the proposition that the rational will as the very source of universally binding normative principles could as such authorize its own destruction involves a contradiction. Put differently: the will as legislator cannot authorize its own destruction as subject because its capacity for freedom and autonomy (i.e., its very constitution), and thus its continued existence, is what makes it possible for the will to be a source of legislation for itself in the first place. Suicide involves the destruction of the self-legislating, *noumenally* free subject whose presence in the world, as instantiated in any given person, is the pre-condition of any possible morality or lawgiving as such (and who is thus the sole entity possessed of unconditional worth).<sup>17</sup>

Kant expresses this point in various ways. In the *Lectures on Ethics*, he suggests that the condition of the possibility of any exercise of moral freedom is the immutability of the rational autonomous will: “[...] freedom can exist only though an immutable condition, which cannot be changed under any circumstances. This condition is that I do not employ my freedom against myself for my own destruction, and that I do not let it be limited by anything external”.<sup>18</sup> He states, further:

[Freedom] has to be restricted, not, though, by other properties and faculties, but by itself. Its supreme rule is: In all self-regarding actions, so to behave that any use of powers is compatible with the greatest use of them. For example, if I have drunk too much today, I am incapable of

<sup>15</sup> In the *Lectures on Ethics*, Kant states that “the body is the total condition of life, so that we have no other concept of our existence save that mediated by our body, and since the use of our freedom is possible only through the body, we see that the body constitutes a part of our self. [...] [Suicide] is contrary to the supreme self-regarding duty, for the condition of all other duties is thereby abolished. It transcends all limits on the use of free choice, *for the latter is only possible insofar as the subject exists*”. (VE, AA 27: 369–370; 144–145, emphasis added. “Nun ist aber der Körper die gänzliche Bedingung des Lebens, so daß wir keinen andern Begriff von unserm Leben haben, als vermittelt unsers Körpers, und da der Gebrauch unserer Freiheit nur durch den Körper möglich ist, so sehn wir, daß der Körper einen Theil unsrer selbst ausmacht. [...] Dieses ist der obersten Pflicht / gegen sich selbst zu wider, denn dadurch wird die Bedingung aller übrigen Pflichten aufgehoben. Dies geht über alle Schranken des Gebrauchs der freyen Willkühr, denn der Gebrauch der freyen Willkühr ist nur dadurch möglich, daß das Subject ist.”)

<sup>16</sup> I disregard, here, Kant’s arguments on the immortality of the soul.

<sup>17</sup> As Kant states, “lawgiving itself, which determines all worth, must for that reason have a dignity, that is an unconditional, incomparable worth”. (Cited at note 11 above.) Suicide negates lawgiving itself, i.e., the autonomous will that is subject to no law other than that which it gives to itself.

<sup>18</sup> VE, AA 27: 374; 148. “[...] die Freyheit nicht durch eine unwandelbare Bedingung bestehn kann, die sich unter keinen Umständen ändern kann. / Diese Bedingung ist, daß ich meine Freyheit nicht wider mich selbst zu meiner destruction gebrauche, sondern daß ich meine Freyheit durch nichts äußeres einschränken lasse”.



making use of my freedom and my powers; or if I do away with myself, I likewise deprive myself of the ability to use them. So this conflicts with the greatest use of freedom, that it abolishes itself, and all use of it, as the highest *principium* of life. Only under certain conditions can freedom be consistent with itself; otherwise it comes into collision with itself. [...] The conditions under which alone the greatest use of freedom is possible, and under which it can be self-consistent, are the essential ends of mankind. With these, freedom must agree. The *principium* of all duties is thus the conformity of the use of freedom with the essential ends of mankind.<sup>19</sup>

This passage pre-figures both the (formal) universalizability and consistency requirements of the Categorical Imperative and the (more substantial) requirements of the Principle of Humanity as articulated in the *Groundwork*.<sup>20</sup> Simply put, the argument is that because we are endowed with autonomy, which constitutes our unconditional worth, it is equally impermissible (contradictory) for us to act in a manner that negates our own autonomy as that of others.<sup>21</sup>

Kant's more definitive treatment of suicide appears in the *Metaphysics of Morals*, where he states:

A human being cannot renounce his personality as long as he is a subject of duty, hence as long as he lives; and it is a contradiction that he should be authorized to withdraw from all obligation, that is, freely to act as if no authorization were needed for this action. To annihilate the subject of morality in one's own person is *to root out the existence of morality itself from the world, as far as one can*, even though morality is an end in itself. Consequently, disposing of

<sup>19</sup> VE, AA 27: 346; 126–127. “[Die Freyheit] muß also restringirt werden, aber nicht durch andre Eigenschaften und Vermögen, sondern durch sich selbst. Ihre oberste Regel ist: In allen Handlungen / in Ansehung seiner selbst so zu verfahren, daß aller Gebrauch der Kräfte mit dem größten Gebrauch derselben möglich ist, z. E. habe ich heute zu viel getrunken, so bin ich ohnmächtig, mich meiner Freyheit, meiner Kräfte zu bedienen, oder bringe ich mich selbst um, so nehme ich mir gleichfalls das Vermögen des Gebrauchs derselben. Es streitet dieses also mit dem größten Gebrauch der Freyheit, daß sie als das höchste principium des Lebens sich selbst und allen ihren Gebrauch aufhebe. Unter gewissen Bedingungen kann nur die Freyheit mit sich selbst übereinstimmen, sonst collidirt sie mit sich selbst. [...] Die Bedingungen unter denen nur allein der größte Gebrauch der Freyheit möglich ist und unter welchen sie mit sich selbst übereinstimmen kann, sind die wesentlichen Zwecke der Menschheit. Mit diesen muß die Freyheit übereinstimmen. Das principium aller Pflichten ist also die Uebereinstimmung des Gebrauchs der Freyheit mit den wesentlichen Zwecken der Menschheit.”

In sum: “What constitutes suicide is the intention to destroy oneself. [...] Suicide evokes revulsion with horror, because everything in nature seeks to preserve itself: a damaged tree, a living body, an animal; in man, then, is freedom, which is the highest degree of life, and constitutes the worth of it, to become now a *principium* for self-destruction?” (VE, AA 27: 371–372; 146. “Die Intention sich selbst zu destruiren macht den Selbstmord aus. [...] Der Selbstmord hat einen Abscheu mit Grausen, denn jede Natur sucht sich selbst zu erhalten. Ein verletzter Baum, ein lebendiger Körper, ein Thier; und nun soll bey Menschen die Freyheit, die der höchste Grad des Lebens ist, und den Werth des Lebens ausmacht, ein principium sein, sich selbst zu zerstören?”)

<sup>20</sup> See GMS, AA 04: 428–429; 78–80.

<sup>21</sup> In effect, the argument states that because autonomy is unique in the world and thus worthy of respect as such (as indicated in note 11 above), an autonomous being must endeavour to preserve, sustain and maximize its presence therein.



oneself as a mere means to some discretionary end is debasing humanity in one's own person (*homo noumenon*), to which the human being (*homo phaenomenon*) was nevertheless entrusted for preservation.<sup>22</sup> (MS, AA 06: 422–423; 547; emphasis added)

Thus, the claim that no authorization of suicide is possible because we are, as possessors of humanity in our person, always already subjects of duty reiterates the argument from (formal) contradiction, which seems to be the common denominator through Kant's pronouncements on suicide. For him, the justification of a maxim or end that involves the negation of its own justificatory ground is (formally) inconceivable; autonomy cannot be exercised to achieve some (discretionary) end that effectively cancels out autonomy itself.

Now, in addition to this, the most significant point to be ascertained here, for our purposes, is found in the phrase "to root out the existence of morality itself from the world, as far as one can". Here, Kant gives full expression to the sheer gravity of suicide as a moral offence: to will one's own death is a terrible thing because it entails renouncing autonomy in the most extreme way. It involves rooting out *morality itself* from the world. The suicide does not just kill himself; rather, his death as a particular person, also brings about an annihilation of morality as such. Because humanity is something we all share, *noumenal* freedom being universally present in all persons, the implications of suicide transcend the particular act. To *will* the annihilation of morality, i.e., of the will itself, is perhaps the most profound contradiction of all.<sup>23</sup>

On a more superficial reading, Kant's prohibition on suicide appears to be based on his view that a rational will cannot give priority to suffering over life itself and that the alleviation of misery is a merely discretionary end that cannot justify self-killing as a means. However, underlying this is a far more important and much stronger point, namely that a rational will cannot contemplate death as a practical end in itself. This is not the same as the argument that a rational will cannot approve misery as a reason to end one's life, but it is discernible and necessary within that argument nonetheless. Otherwise, we could not make sense of Kant's insistence that to "annihilate the subject of morality in one's own person is to root out the existence of morality itself from the world, as far as one can, even though morality

<sup>22</sup> "Der Persönlichkeit kann der Mensch sich nicht entäußern, so lange von Pflichten die Rede ist, folglich so lange er lebt, und es ist ein Widerspruch die Befugniß zu haben sich aller Verbindlichkeit zu entziehen, d. i. frei so zu handeln, als ob es zu dieser Handlung gar keiner Befugniß bedürfte. Das Subject der Sittlichkeit in seiner eigenen Person zernichten, ist eben so viel, als die Sittlichkeit selbst ihrer Existenz nach, so viel an ihm ist, aus der Welt vertilgen, welche doch Zweck an sich selbst ist; mithin über sich als bloßes Mittel zu ihm beliebigen Zweck zu disponiren, heißt die Menschheit in seiner Person (*homo noumenon*) abwürdigen, der doch der Mensch (*homo phaenomenon*) zur Erhaltung anvertrauet war."

<sup>23</sup> It is thus not surprising that Kant is deeply concerned with the problem of suicide and that he discusses it, throughout his works, both by way of articulating his conception of duties to oneself and as a central example in the universalizability and consistency requirements of his ethics. Suicide is perhaps the clearest instance of an inconsistent or contradictory end or maxim, just as the making of false promises is a good example of a course of action that is impermissible because its maxim cannot be universalized, or willed for everyone.

is an end in itself". Suicide constitutes a contradiction in the will in the most radical sense, because it entails the will's contemplation of its own annihilation. So it is not just that the will cannot posit death as a means to a discretionary end, but rather that it cannot posit death as such (in the absence of an overriding imperative, as discussed below).

*Suicide as self-punishment?*

There is some question as to whether Kant's prohibition on suicide is absolute, and at least one prominent commentator argues that it is not.<sup>24</sup> Kant states that "disposing of oneself *as a mere means to some discretionary end* is debasing humanity in one's own person", which suggest that where the end (i.e., killing oneself) is not merely discretionary it could, conceivably, be morally permissible. As Rawls argues, the "casuistical questions Kant lists in this section [in MS, AA 06: 423–424; 548] imply that such a title can be given by conflicting grounds of obligation ([MS, AA 06: 224; 378–379]); for these may be at times stronger than the ground not to take our life".<sup>25</sup> Now, Kant remains silent on the question as to how exactly the duty against suicide could be overridden in situations where there are conflicting grounds of obligation, and he does not actually resolve the casuistical questions he invokes. However, we may nevertheless accept Rawls' argument on this point.

Because I hope to show that the death penalty, as a form of wilful negation of a will, should be impermissible in the Kantian framework, Rawls's interpretation poses a difficulty. It would seem that where the will posits the death of the subject *not* as a practical end in itself, but rather as a means toward the fulfillment of some other rational obligation or moral duty, this death may be permissible. For, the duty of self-preservation may be overridden where the end posited is not merely discretionary but obligatory. In other words, we may have to concede that even though the (direct) annihilation of the autonomous, rational will itself is an absolutely inconceivable end, its (indirect) annihilation is nevertheless conceivable, if the end posited is not death as such, but rather the fulfillment of some other ethical requirement. Kant may have thought that the unconditional dictates of duty must be upheld, even if, in some instances, the existence of morality itself is rooted-out from the world as a consequence. An obvious example is Kant's strict prohibition on lying, the consequence of which is that we must tell the truth even if doing so results in death.<sup>26</sup>

<sup>24</sup> Rawls (2000), 193.

<sup>25</sup> *Ibid.* For Kant's discussion of conflicting duties and grounds of obligation, see the passage referred to by Rawls (namely MS, AA 06: 224; 378–379).

<sup>26</sup> See VRML, AA 08: 427; 612. The classic scenario contemplated in connection with Kant's views in that essay is this: your friend's mortal enemy show up at your door asking if your friend is in the house, which is actually the case; knowing this individual would kill your friend at first sight, do you lie about your friend's whereabouts? Kant would answer in the negative.

(Additionally, I certainly do not mean to suggest that the inherent irrationality of positing death as an end in itself, which informs Kant's prohibition on suicide, is supposed to render irrational a killing in the course of self-defence, or altruistic self-sacrifice.) Rather, the question, for our purposes, is whether *punishment* is an unconditional dictate or requirement wherein the death of the subject may be legitimately contemplated, i.e., without contradiction.

Before discussing capital punishment and suicide in the context of political autonomy, we should consider punishment from an ethical standpoint. At the risk of adding a further casuistical question to Kant's list, I would ask: is it permissible to commit suicide by way of self-punishment for having committed a heinous crime? In other words, is self-punishment a legitimate end for an individual, such that suicide would be consistent with his nature as a rational, autonomous being? If there is no contradiction in the proposition that the rational will may negate itself in self-punishment, i.e., if self-execution as punishment is conceivable for an autonomous will, then the argument I am developing here falls apart. As it happens, Kant is highly unlikely to endorse self-punishment as a rational end, given this claim: "No one suffers punishment because he has willed *it* but because he has willed a *punishable action*; for it is no punishment if what is done to someone is what he wills, and it is impossible *to will* to be punished."<sup>27</sup> Kant makes this claim by way of arguing that capital punishment is compatible with the autonomy of citizens in civil society which, as I intend to show below, is wrong. Here, I cite the passage only to show that self-annihilation as a means to punishment is not an end that a rational will could consistently posit for itself, and that Kant was unequivocal about this. If it is clearly "impossible to will to be punished",<sup>28</sup> then at a minimum it is doubtful that we could have any kind of duty or unconditional obligation to self-punish and, moreover, it is inconceivable that self-execution as an end could be any kind of obligation overriding the grounds of the prohibition on suicide (or taking precedence over the duty of self-preservation).<sup>29</sup> Self-execution as a means to punishment is simply not within the scope of autonomous agency, i.e., it is not an end that a will could rationally legislate for itself.

<sup>27</sup> MS, AA 06: 335; 476. "Strafe erleidet jemand nicht, weil er *sie*, sondern weil er eine *strafbare Handlung* gewollt hat; denn es ist keine Strafe, wenn einem geschieht, was er will, und es ist unmöglich, gestraft werden zu *wollen*."

<sup>28</sup> "Impossibility" in this context, and in the Kantian idiom generally, does not connote empirical impossibility, as a merely contingent matter, but rather rational inconceivability or, simply, irrationality.

<sup>29</sup> Though I do not think Kant discusses this, he may say that, having committed a capital offence, we have an obligation to turn ourselves over to the law and to accept our punishment, even if we suffer death as a consequence. But if it is impossible, as Kant says, for us to directly will to be punished, this must be because punishment as such cannot be a rational end for an individual autonomous being, let alone a punishment that entails the destruction of the will itself. The upshot of the passage cited above is that punishment is not something a rational will could inflict on itself. This is no doubt related to the position that capital punishment is a prerogative of the state, which is an additional reason as to why it falls outside the proper scope of individual autonomy, i.e., besides the fact that it is inherently irrational from the standpoint of an individual ethics and the duties Kant believes we owe to ourselves.

*Capital punishment in civil society and the conditions of political autonomy*

What remains to be determined is whether an execution of the autonomous subject can be justified as a matter of external lawgiving, i.e., from the standpoint of practical reason in its external aspect. As we have noted above, external lawgiving governs the relations between citizens in civil society in so far as the external actions of each may affect the actions of others, or what Kant calls the reciprocal relations of choice between members of civil society.<sup>30</sup> Here, practical reason determines the duties and obligations we have in light of our (external) relations with others; it sets our juridical duties as limits on our comportment so that it may conform with an equal degree of freedom for all.<sup>31</sup> Thus, Kant's Universal Principle of Right states that any "action is *right* if it can coexist with everyone's freedom in accordance with a universal law, or if on its maxim the freedom of choice of each can coexist with everyone's freedom in accordance with a universal law".<sup>32</sup> However, essentially unlike ethical lawgiving, external lawgiving as bound by the Universal Principle of Right is connected with an authorization to use coercion,<sup>33</sup> which is to say that juridical duties may be physically or externally enforced and that, in general, the physical or external coercion of the juridical subject is compatible with the dictates of practical reason.

<sup>30</sup> MS, AA 06: 229ff; 386ff. See note 6 and accompanying text above.

<sup>31</sup> Although juridical duties based on the principle of right are also ethical duties (MS, AA 06: 219–220; 383–385), the reverse is not true for Kant. This is to say that a virtuous person will necessarily obey the laws of the state, but an obedient citizen is not necessarily a virtuous person. Furthermore, we are bound by the ethical dictates of practical reason regardless of whether or not we happen to find ourselves in a civil condition or in the state of nature, for that is the condition of our being rational autonomous agents as such. But we cannot have any juridical duties where we have not entered into (external) relations with others, i.e., if our exercise of freedom could not, in fact, have an affect on others.

<sup>32</sup> MS, AA 06: 230; 387. "Eine jede Handlung ist *recht*, die oder nach deren Maxime die Freiheit der Willkür eines jeden mit jedermanns Freiheit nach einem allgemeinen Gesetze zusammen bestehen kann." (emphasis added).

<sup>33</sup> As Kant states: "Resistance that counteracts the hindering of an effect promotes this effect and is consistent with it. Now whatever is wrong is a hindrance to freedom in accordance with universal laws. But coercion is a hindrance or resistance to freedom. Therefore, if a certain use of freedom is itself a hindrance to freedom in accordance with universal laws (i.e., wrong), coercion that is opposed to this (as a *hindering of a hindrance to freedom*) is consistent with freedom in accordance with universal laws, that is, it is right. Hence, there is connected with right by the principle of contradiction an authorization to coerce someone who infringes upon it." (MS, AA 06: 231; 388. "Der Widerstand, der dem Hindernisse einer Wirkung entgegengesetzt wird, ist eine Beförderung dieser Wirkung und stimmt mit ihr zusammen. Nun ist alles, was unrecht ist, ein Hinderniß der Freiheit nach allgemeinen Gesetzen: der Zwang aber ist ein Hinderniß oder Widerstand, der der Freiheit geschieht. Folglich: wenn ein gewisser Gebrauch der Freiheit selbst ein Hinderniß der Freiheit nach allgemeinen Gesetzen (d. i. unrecht) ist, so ist der Zwang, der diesem entgegengesetzt wird, als *Verhinderung* eines *Hindernisses der Freiheit* mit der Freiheit nach allgemeinen Gesetzen zusammen stimmend, d. i. recht: mithin ist mit dem Rechte zugleich eine Befugniß, den, der ihm Abbruch thut, zu zwingen, nach dem Satze des Widerspruchs verknüpft.")

Now, the justification of *punishment* is understood to be grounded in the Kantian authorization of coercion. For instance, Hill's claim that the "general authority of the state coercive powers, on which the right to punish is based, is the authority to 'hinder hindrances to freedom'",<sup>34</sup> is typical in this respect.<sup>35</sup> Although Kant does not draw an explicit connection between punishment and coercion, we may nevertheless accept this interpretation.<sup>36</sup> But the relevant question is whether, granting the necessity of punishment-as-coercion for the purposes of a secure social order, its justification covers capital punishment? Put differently, is capital punishment rational from a political standpoint given that it is otherwise permissible, indeed necessary, that all rational wills subject themselves to a condition of public lawful external coercion (Right) by way of securing a maximal sphere of external freedom for all?<sup>37</sup> My answer is that, because it involves the *death* of the subject, capital punishment is an extraordinary form of punishment that does not have a place within the Kantian framework.

We have seen that, according to Kant, suicide is irrational, even as a means to self-punishment, because it involves rooting out the existence of morality itself from the world. An autonomous being cannot rationally posit its own death because doing so entails the annihilation of an essential pre-condition, namely life, of any possible autonomous lawgiving and thus the very ground of any rationally conceivable end. For Kant, practical reason determines the law for the moral as well as the juridical subject because, to say again, it is our capacity for autonomous agency as such that places us before the law, which is innate to the rational will itself. Thus, the inexorable (formal) constraints inherent to practical reason apply equally with respect to the ethical and the juridical domains.<sup>38</sup>

<sup>34</sup> Hill (1999) 429.

<sup>35</sup> See also Byrd (1989).

<sup>36</sup> Incidentally, I am not convinced that there is a clear connection between punishment and coercion in the Kantian framework. It seems to me that, given Kant's conception of a "hindering of a hindrance to freedom" or of a "resistance that *counteracts* the hindering of an effect", the paradigmatic form of coercion would be something like the proprietary right to expel trespassers, or perhaps the injunctive remedies in tort. A plain, straightforward reading of the passage cited at note 33 above would be that it contemplates the (justified) use of coercion as means of counteracting a *concurrent* (wrongful) infringement of freedom. Because punishment is a *post facto* response to wrongdoing, it is not clear how exactly it is supposed to *counteract* any given bit of wrongdoing, or how it is meant to *hinder a hindrance* of freedom. I realize this is a rather narrow and restrictive reading, but Kant could certainly have said more about how punishment is supposed to fit in to the scheme of the Doctrine of Right; as they stand, his retributivist arguments are not well integrated with the rest of the work. For a similar argument, see Gorner (2000). Hill, on the other hand, argues that coercion consists in the *threat* of punitive sanctions meant to maintain social order, which is necessarily accompanied by *post facto* punishment (in particular cases) as the carrying out of that threat. The connection he thus draws between coercion and punishment is plausible, but it is not one that Kant explicitly makes.

<sup>37</sup> See MS, AA 06:312; 455–456.

<sup>38</sup> Indeed, it is fair to say that Kant's conception of Reason unifies the ethical and political-juridical aspects of his practical philosophy, or that his conception of autonomous agency as its own

The rational will is ultimately the author of any law addressed to a human being, whether he is simply an individual moral agent or a citizen of civil society. This must mean that as rational autonomous beings we can no more posit our own destruction as moral subjects than as juridical subjects.

It is true that Kant's principle of right underwrites the use of coercion (as a means of securing justice and an equal freedom for all), so that there is no inherent irrationality in our being co-legislators of coercive juridical laws.<sup>39</sup> Nevertheless, capital punishment is not a rationally permissible form of coercion. As citizens and parties to the original contract we are authorized, by the principle of right,<sup>40</sup> to deploy

source of legislation is the common denominator between the otherwise distinct spheres occupied by duties of virtue and duties of right. As Kersting (1992, 143–144) notes: “Kant's political philosophy finds its architectonic place within the complex structure of his entire practical philosophy in the pure philosophy of right and in the philosophy of history. [...] In that Kant's practical philosophy spells out politics in a rationally legal way, it also grounds political philosophy in the end in the concept of pure practical Reason which, *for all realms of practical philosophy*, has equal weight and equal justificatory importance. For Kant there is a necessary relation of logical dependence between the validity of political philosophy and the universal, objective and categorically demanding principle of pure practical Reason.” (emphasis added). For a more detailed explication of the underlying connection between Kant's political and moral philosophy, see Guyer (1997) and Benson (1987).

<sup>39</sup> According to Kant, the conditions of political autonomy require that any given item of legislation or policy governing the relations of civil society must issue from a (hypothetical or idealized) general will, i.e., all political and legal arrangements must have been posited collectively by (rational and autonomous) citizens themselves. As Kant states: “The legislative authority can belong only to the united will of the people. For since all right is to proceed from it, it *cannot* do anyone wrong by its law. Now, when someone makes arrangements about *another*, it is always possible for him to do the other wrong; but he can never do wrong in what he decides upon with regard to himself (for *volenti non fit iniuria*). Therefore only the concurring united will of all, insofar as each decides the same thing for all and all for each, and so only the general will of the people, can be legislative.” (MS, AA 06: 313–314; 457. “Die gesetzgebende Gewalt kann nur dem vereinigten Willen des Volkes zukommen. Denn da von ihr alles Recht ausgehen soll, so muß sie durch ihr Gesetz schlechterdings niemand unrecht thun *können*. Nun ist es, wenn jemand etwas gegen einen *Anderen* verfügt, immer möglich, daß er ihm dadurch unrecht thue, nie aber in dem, was er über sich selbst beschließt (denn *volenti non fit iniuria*). Also kann nur der übereinstimmende und vereinigte Wille Aller, so fern ein jeder über Alle und Alle über einen jeden ebendasselbe beschließen, mithin nur der allgemein vereinigte Volkswille gesetzgebend sein.”)

Thus, any given penal law must be a (formally and rationally) conceivable expression of the general will.

<sup>40</sup> For Kant the legislator, whether the citizen himself or his representative, is bound equally by the principle of right and the idea of the original contract as dictates of practical reason, and may legislate justly only on those terms: “Now the legislator can indeed err in his appraisal of whether those measures are adopted prudently, but not when he asks himself whether the law also harmonizes with the principle of right; for there he has that idea of the original contract at hand as an infallible standard, and indeed has it *a priori*.” (TP, AA 08: 299; 298. “In dieser Beurtheilung, ob jene Maßregel *klüglich* genommen sei oder nicht, kann nun zwar der Gesetzgeber irren, aber nicht in der, da er sich selbst fragt, ob das Gesetz auch mit dem Rechtsprincip zusammen stimme oder nicht; denn da hat er jene Idee des ursprünglichen Vertrags zum unfehlbaren Richtmaße und zwar *a priori* bei der Hand [...]”).

coercion as a means of restricting the external freedom of others, i.e., we may forcefully bring their external actions into conformity with an equal freedom for all, and to impose punishment if they have transgressed, intentionally and thus *culpably*, the bounds of their rightful sphere of external freedom. This means we may give our assent to penal laws and that the (rightful) use of coercion is compatible with our political autonomy.<sup>41</sup> Because capital punishment entails the annihilation of the underlying source of any possible juridical lawgiving, namely the autonomous will of the citizen, it cannot form a part of any rationally constructed civil constitution.

For Kant the general will is a unity of *rational* wills, rather than an aggregate of particular, empirically conditioned or arbitrary wills. Thus, where penal laws are concerned, “it is pure reason in me (*homo noumenon*)”<sup>42</sup> that legislates with regard to rights. The general will, as a founding legislative organ (a kind of Kantian *Grundnorm*) does not consist of a multiplicity of diverse or distinct wills; rather, it is nothing other than practical “reason in its external aspect”. As Dyke explains:

[...] when Kant talks about the united Will of the people to which the legislative function in a state must be attributed, he is talking about the union of self-legislating faculties, not the union of individual preferences generated by inclination or interest (especially self-interest). [...] Individual Wills cannot be individuated as such, because Will is identical to practical reason, and practical reason, in turn, is identical in every person in whom it is found. Reason is, by its very nature, independent of the person who is exercising his faculty of reason. [...] Thus in uniting their Wills, the individuals of a society find unity in the realm of pure reason; but it is the unity of identity. It immediately follows from this that a union of Wills will not produce legislative results (i.e., laws) different from those produced by one Will. And this means that one man, with an autonomous will, could legitimately legislate for a whole society. Others in the society, if they possessed autonomous will, would see that he had legislated correctly, that is, in accordance with the laws of justice.<sup>43</sup>

If ultimately it is practical reason itself, as the Will, that constitutes the legislative function in a state, then it is not surprising that juridical lawmaking should be subject to the same fundamental (formal) constraints as maxim formulation in the context of moral agency. Thus, the universalizability requirement and “contradiction in the will” test of the Categorical Imperative are paralleled, respectively, by the Universal Principle of Right and the basic principle of non-contradiction in the idea of the original contract as a test of juridical or political legitimacy.<sup>44</sup>

<sup>41</sup> For a discussion of Kant’s reconciliation of coercion, punishment and political autonomy, see Dodson (1997) and Carr (1989).

<sup>42</sup> MS, AA 06: 335; 476. “[...] so ist es in mir die reine rechtlich-gesetzgebende Vernunft (*homo noumenon*) [...]”.

<sup>43</sup> Dyke (1969), 28. Note that Dyke is not saying that the General Will is needed to generate the Principle of Right itself, which would be highly misleading. Rather, the point is that because adherence to the Principle of Right is necessarily integral to truly autonomous willing as such, the perfect identity of Wills in the realm of pure reason means that a single will could legislate justly just as well as a unity of multiple Wills.

<sup>44</sup> This test is found in the following passage: “[The original contract is] *only an idea* of reason, which, however, has its undoubted practical reality, namely to bind every legislator to give his laws in such a way that they *could* have arisen from the united will of a whole



In either instance, the legislator is not permitted to frame a law that contradicts or, all the more, negates the underlying conditions of autonomous willing. Capital penal codes contradict the underlying conditions of autonomous willing precisely because the ultimate end they contemplate is the negation, i.e., annihilation, of those conditions in a particular instance.

To say that capital punishment is permissible is to say that practical reason may frame a *singular* kind of penal law, the ultimate purpose of which is the annihilation of practical reason itself as instantiated in a given person (provided certain conditions are met, namely that this person is a criminal). Death is integral to the death penalty, i.e., it is not merely a contingent by-product of applying punishment. The death penalty is a unique form of punishment that necessarily entails the contemplation of death as an end in itself, which, as we have seen, is impossible for a rational will.<sup>45</sup> If we say that capital punishment is legitimate, then we are saying that the rational will may sanction the rooting-out of the existence of morality itself from the world. In this sense, capital punishment is formally equivalent to suicide.

people and to regard each subject, insofar as he wants to be a citizen, as if he has joined in voting for such a will. For this is the touchstone of any public law's conformity with right. In other words, if a public law is so constituted that a whole people *could not possibly* give its consent to it (as, e.g., that a certain class of *subjects* should have the hereditary privilege of *ruling rank*), it is unjust; but if it is only possible that a people could agree to it, it is a duty to consider the law just, even if the people is at present in such a situation or frame of mind that, if consulted about it, it would probably refuse its consent." (TP, AA 08:297; 296–297. "[Der ursprüngliche Contract] ist eine *bloße Idee* der Vernunft, die aber ihre unbezweifelte (praktische) Realität hat: nämlich jeden Gesetzgeber zu verbinden, daß er seine Gesetze so gebe, als sie aus dem vereinigten Willen eines ganzen Volks haben entspringen *können*, und jeden Unterthan, so fern er Bürger sein will, so anzusehen, als ob er zu einem solchen Willen mit zusammen gestimmt habe. Denn das ist der Probestein der Rechtmäßigkeit eines jeden öffentlichen Gesetzes. Ist nämlich dieses so beschaffen, daß ein ganzes Volk *unmöglich* dazu seine Einstimmung geben *könnte* (wie z.B. daß eine gewisse Klasse von *Unterthanen* erblich den Vorzug des *Herrenstandes* haben sollten), so ist es nicht gerecht; ist es aber *nur möglich*, daß ein Volk dazu zusammen stimme, so ist es Pflicht, das Gesetz für gerecht zu halten: gesetzt auch, daß das Volk jetzt in einer solchen Lage, oder Stimmung seiner Denkart wäre, daß es, wenn es darum befragt würde, wahrscheinlicherweise seine Beistimmung verweigern würde.")

<sup>45</sup> At a minimum, the very nature of the death penalty is such that it renders our purpose in applying it ambiguous and indeterminate – is the ultimate goal punishment or death? The two are actually inseparable. As I argue below, *punishment* proper, i.e., as an ultimate end in itself, presupposes the continued existence of the person being punished. But when it comes to capital punishment, our purpose is not merely to punish him – it is to *kill* him. Another way of putting this is that punishment proper admits of substitution; if the point is to hold the criminal responsible for his crime, in accordance with a measure of its gravity, the punishment may take a variety of forms so long as it reflects this measure, e.g., prolonged imprisonment or brief but harsh corporal punishment. (Indeed, some pre-modern legal systems offered the criminal this kind of choice.) But capital punishment admits of no alternative, for the execution just *is* the punishment. Elucidating the absolute uniqueness of capital punishment involves drawing this crucial conceptual distinction between our purpose in applying a form of punishment and the application itself, i.e., the actual *means* of realizing the punishment. With capital punishment, this distinction collapses.

The citizen, as ideal legislator, is supposed to be able to consent to a law under which the very capacity for rational willing that makes it possible for him to grant his consent in the first place may (again, in certain circumstances) be annihilated. Assuming Dyke's interpretation is correct, namely that in the Kantian framework one man with an autonomous will could in principle legitimately legislate for all, Kant's argument as to the irrationality of suicide (and self-punishment) would preclude such a man from subjecting himself to a form of punishment that entails his own death. That is, in this sense, capital punishment just is a form of suicide, and Kant ought to have thereby excluded it as a punitive practice in a rationally constituted republic.<sup>46</sup>

*What is the status of an individual subsequent to having  
committed a crime? Kant and Beccaria*

As it happens, Kant does have a response to this line of argument that requires careful consideration. There is a critical distinction between suicide and capital punishment since the execution of a citizen is contingent upon the commission of a crime, which does not apply in the case of suicide. We have accepted that punitive provisions are indispensable in civil society, so that punishment as such is not some merely arbitrary end. (Nor do I intend to challenge Kant's retributivism in general or his commitment to the strict principle of equality between crime and punishment.) Given this, we must ask whether there is something about the nature of a criminal act itself that modifies the status of the individual who has committed it, such that his execution would be compatible with the dictates of the general legislative will. Put differently: is the citizen who commits a capital offence thereby somehow dissociated from his inherent humanity and deprived of those qualities that previously made him fit to be a co-legislator for all via the general will? Does he cease being a self-legislating faculty, such that his execution, as a *criminal*, would not disclose a contradiction in the general legislative will? Kant seems to think so, but the arguments he offers are deeply problematic.

The arguments I am referring to appear in the context of Kant's comments on Beccaria's *Dei delitti e delle pene* (1764). There, Beccaria develops a number of arguments against the cruel penal codes of the time, in particular against the death penalty, and relies on Rousseau's claim in the *Social Contract* that "all laws must be

<sup>46</sup> All this can be summed up as follows. We may ask: who is the executioner, and what validates or grounds his actions? It is the *noumenal* self as state authority – the general will as expressed in state action. But how can the general will as the unity of identical *noumenal* selves sanction the execution of one of its organs or instantiations? Is not the *noumenal* self – at least as a potentiality for rational autonomous agency – universal and universally present in all persons? So how could the universal *noumenal* self will the annihilation of its own instantiation or embodiment in a particular person? Is not the death sentence, therefore, an irrational form of suicide?

regarded as if they proceed from the unanimous will of the people”, a notion that Kant has, as we have seen, also adopted. Beccaria’s basic argument against capital punishment is very similar to the one I have developed here although, of course, it does not appeal to the Kantian conception of autonomy itself or to the notion of a contradiction in the will. Still, it is essential that I address Kant’s harshly worded rebuttal at this point:

[T]he Marchese Beccaria, moved by overly compassionate feelings of an affected humanity (*compassibilitas*), has put forward his assertion that any capital punishment is wrongful because it could not be contained in the original civil contract; for if it were, everyone in a people would have to have consented to lose his life in case he murdered someone else (in the people), whereas it is impossible for anyone to consent to this because no one can dispose of his own life. This is all sophistry and juristic trickery.<sup>47</sup>

#### *The first part of Kant’s response to Beccaria*

Kant’s rebuttal in fact consists of two distinct arguments, though he runs them together. The first is this:

No one suffers punishment because he has willed *it* but because he has willed a *punishable action*; for it is no punishment if what is done to someone is what he wills, and it is impossible *to will* to be punished. – Saying that I will to be punished if I murder someone is saying nothing more than that I subject myself together with everyone else to the laws, which will naturally also be penal laws if there are any criminals among the people. As a legislator in dictating the *penal law*, I cannot possibly be the same person who, as a subject, is punished in accordance with the law; for as one who is punished, namely as a criminal, I cannot possibly have a voice in legislation (the legislator is holy). Consequently, when I draw up a penal law against myself as a criminal, it is pure reason in me (*homo noumenon*), legislating with regard to rights, which subjects me, as someone capable of crime and so as another person (*homo phaenomenon*), to the penal law, together with all others in a civil union.<sup>48</sup>

<sup>47</sup> MS, AA 06: 334–335; 475–476. “Hiegegen hat nun der Marchese *Beccaria* aus theilnehmender Empfindelheit einer affectirten Humanität (*compassibilitas*) seine Behauptung der *Unrechtmäßigkeit* aller Todesstrafe aufgestellt: weil sie im ursprünglichen bürgerlichen Vertrage nicht enthalten sein könnte; denn da hätte jeder im Volk einwilligen müssen, sein Leben zu verlieren, wenn er etwa einen Anderen (im Volk) ermordete; diese Einwilligung aber sei unmöglich, weil Niemand über sein Leben disponiren könne. Alles Sophisterei und Rechtsverdrehung.”

<sup>48</sup> MS, AA 06: 335; 476. “Strafe erleidet jemand nicht, weil er *sie*, sondern weil er eine *strafbare Handlung* gewollt hat; denn es ist keine Strafe, wenn einem geschieht, was er will, und es ist unmöglich, gestraft werden zu *wollen*. – Sagen: ich will gestraft werden, wenn ich jemand ermorde, heißt nichts mehr als: ich unterwerfe mich sammt allen Übrigen den Gesetzen, welche natürlicherweise, wenn es Verbrecher im Volk giebt, auch Strafgesetze sein werden. Ich als Mitgesetzgeber, der das *Strafgesetz* dictirt, kann unmöglich dieselbe Person sein, die als Unterthan nach dem Gesetz bestraft wird; denn als ein solcher, nämlich als Verbrecher, kann ich unmöglich eine Stimme in der Gesetzgebung haben (der Gesetzgeber ist heilig). Wenn ich also ein Strafgesetz gegen mich als einen Verbrecher abfasse, so ist es in mir die reine rechtlich-gesetzgebende Vernunft (*homo noumenon*), die mich als einen des Verbrechens Fähigen, folglich als eine andere Person (*homo phaenomenon*) sammt allen übrigen in einem Bürgerverein dem Strafgesetze unterwirft.”

Now, the quick and easy response to this is to challenge Kant's distinction between *homo noumenon* as pure rational legislator and *homo phaenomenon* as criminal. If, as the former, I subject myself to a penal law that (given the right circumstances) results in my death as the latter, I am no less dead in the final analysis, i.e., as a whole person. If physical embodiment, i.e., continued living, is an essential precondition of moral agency – if *homo noumenon* can exist in this world only in so far as it is also physically instantiated as *homo phaenomenon* – then the death of either necessarily entails the death of the other.<sup>49</sup> So Kant is not entitled to suggest, as I take him to be, that the author of the capital penal law (*homo noumenon*) is distinguished from the addressee of the law (*homo phaenomenon*) in such a way that the execution of the latter, i.e., his “punishment in accordance with the law”, has no significant effect on the former. Though I may die as a criminal whose will has no place in the original contract, I must also die as a self-legislating faculty that, as such, identifies with the general will (even if, as a criminal, I am severed from the original contract). Kant simply cannot escape the fact that execution extinguishes an autonomous will.<sup>50</sup>

*Does the criminal cease being rational by virtue of his crime?*

However there is more going on in this passage that requires comment. As I see it, Kant wants to say that because the criminal, as such, could not possibly have a voice in legislation, the penal law under which he is executed cannot be said to have issued from his will. The distinction he draws between the criminal and the co-legislator is supposed to mean that neither the criminal nor the legislator is actually willing his own death. The co-legislator merely dictates the penal law under which the criminal is subsequently executed, and because these are supposed to be distinct entities, the annihilation of the latter leaves the former untouched. But the relevant question is not whether the criminal has forfeited his status as a co-legislator in the

<sup>49</sup> See note 15 above and accompanying text.

<sup>50</sup> As we shall see further below, Kant wants to say that execution is acceptable because, even if the rational will cannot contemplate its own death as a possible object of choice, the will must nevertheless subject itself to penal laws and thereby make itself vulnerable to punishment if it commits a wrong, which of course is not the same as choosing its own death. The point, however, is that it is the very *contemplation* of death as an end itself that is irrational. If the will puts itself under a *capital* penal law, it must also contemplate its own death as a possible object of *indirect* choice (for this “choice”, i.e., the execution itself, would of course be attributed to the authority of the state). The ultimate basis of a capital penal code is traced to the consent of the autonomous will itself, but such consent is impossible if a necessary consequence of the will granting such consent is its own execution. The important distinguishing feature of capital punishment, namely that unlike suicide death is accomplished through the mediation of state agency, does not rescue Kant. For, if the suicide cannot rationally approve a self-executory maxim, the citizen likewise cannot posit a capital penal code that legitimates state execution (of an autonomous will).

original contract, but rather whether there is any sense in which the criminal can be said to have continued possession of a rational autonomous will and, thus, whether the legislator, as definitively endowed with an autonomous will, can rationally assent to a penal law the application of which entails the annihilation of an identical autonomous will (albeit it must be one that, as we shall see, is diminished through a forfeiture of “civil personality”). Kant places the emphasis on the status of the criminal whereas it should be on the legislator.

At this point we encounter some rather difficult terrain in Kant’s philosophy of penal law. Above, I mention the idea that a criminal, having committed a crime, forfeits his status as a co-legislator and is thereby severed from the original contract. This just refers to Kant’s definition of a (public) crime as a “transgression of public law that makes someone who commits it *unfit to be a citizen*”.<sup>51</sup> The crucial issue is whether this means that a criminal also thereby ceases to be a rational being, or becomes dispossessed of an autonomous will. Jean-Christophe Merle argues that “according to Kant the criminal is no longer to be dealt with as a rational being” because he has denied his own rational essence and, indeed, ceases to *be* a rational being as soon as he commits a crime.<sup>52</sup> To support this claim, Merle cites Kant’s sole exception to his otherwise unequivocal stance against slavery, namely when it comes to those who have “forfeited [their] personality by a crime”<sup>53</sup>. We have seen that Kant also clearly identifies the criminal with *homo phaenomenon*. If Merle’s position is correct, then my argument falls apart, because it would mean that the death of the criminal (as *homo phaenomenon*) really is inconsequential from a juridical-legislative standpoint.

There are however countervailing considerations. For one, it is utterly out of the question that Kant’s identification of the criminal with *homo phaenomenon* should mean that the criminal lacked *noumenal* freedom during the commission of the crime, since that would mean that the act could not be imputed to him, i.e., that he was not responsible. The suggestion that a criminal, in committing a crime, acts under conditions of heteronomy, i.e., that his choice is determined (and not merely affected) by laws of nature, contradicts a number of fundamental tenets of Kant’s philosophy.<sup>54</sup> Simply put, Kant cannot be saying that the criminal act is involuntary or that the criminal exhibits *no* capacity for rational agency at the time he commits it.<sup>55</sup>

<sup>51</sup> MS, AA 06: 331; 472; emphasis added. “Diejenige Übertretung des öffentlichen Gesetzes, die den, welcher sie begeht, *unfähig macht, Staatsbürger zu sein* [...]”.

<sup>52</sup> Merle (2000), 327–328.

<sup>53</sup> MS, AA 06: 283; 431. “[...] der sich durch ein Verbrechen seiner Persönlichkeit verlustig gemacht hat”.

<sup>54</sup> See KpV, AA 05: 89–106; 211–225.

<sup>55</sup> Indeed, in the *Groundwork*, Kant makes it clear that freedom of the will is an essential presupposition for a rational being as an underlying condition of *all his voluntary actions*: “[T]he rightful claim to freedom of will made even by common human reason is based on the consciousness and the granted presupposition of the independence of reason from merely subjectively determining causes, all of which together constitute what belong only to

So the relevant question is whether the criminal ceases to be a rational being as a *consequence* of having committed a criminal act. And here, there is very little in Kant's writing that explains how this could possibly transpire. In the portions of the *Metaphysics of Morals* wherein Kant discusses the minimum conditions of rationality and freedom in moral wrongdoing, there is nothing to suggest that human beings may cease to *be* rational in having committed a crime.<sup>56</sup> Kant defines freedom of choice, in its negative concept, as independence from being determined by sensible

feeling and hence come under the general name of sensibility. The human being, who this way regards himself as an intelligence, thereby puts himself in a different order of things and in relation to determining grounds of an altogether different kind when he thinks of himself as an intelligence endowed with a will, and consequently with causality, than when he perceives himself as a phenomenon in the world of sense (as he also really is) and subjects his causality to external determination in accordance with laws of nature. [...] On the presupposition of the freedom of the will of an intelligence, however, its *autonomy*, as the formal condition under which alone it can be determined, is a necessary consequence. Moreover, to presuppose this freedom of the will is [...] not only quite *possible* [...], it is also practically *necessary* – that is, necessary in idea, without any further condition – for a rational being who is conscious of his causality through reason and so of a will (which is distinct from desires) to put it under all his voluntary actions as their condition.” (GMS, AA 04: 457–461; 103–107. “Der Rechtsanspruch aber selbst der gemeinen Menschenvernunft auf Freiheit des Willens gründet sich auf das Bewußtsein und die zugestandene Voraussetzung der Unabhängigkeit der Vernunft von bloß subjectiv-bestimmenden Ursachen, die insgesamt das ausmachen, was bloß zur Empfindung, mithin unter die allgemeine Benennung der Sinnlichkeit gehört. Der Mensch, der sich auf solche Weise als Intelligenz betrachtet, setzt sich dadurch in eine andere Ordnung der Dinge und in ein Verhältniß zu bestimmenden Gründen von ganz anderer Art, wenn er sich als Intelligenz mit einem Willen, folglich mit Causalität, begabt denkt, als wenn er sich wie ein Phänomen in der Sinnenwelt (welches er wirklich auch ist) wahrnimmt und seine Causalität äußerer Bestimmung nach Naturgesetzen unterwirft. [...] Unter Voraussetzung der Freiheit des Willens einer Intelligenz aber ist die *Autonomie* desselben, als die formale Bedingung, unter der er allein bestimmt werden kann, eine nothwendige Folge. Diese Freiheit des Willens vorauszusetzen, ist auch nicht allein (ohne in Widerspruch mit dem Princip der Naturnothwendigkeit in der Verknüpfung der Erscheinungen der Sinnenwelt zu gerathen) ganz wohl *möglich* (wie die speculative Philosophie zeigen kann), sondern auch sie praktisch, d.i. in der Idee, allen seinen willkürlichen Handlungen als Bedingung unterzulegen, ist einem vernünftigen Wesen, das sich seiner Causalität durch Vernunft, mithin eines Willens (der von Begierden unterschieden ist) bewußt ist, ohne weitere Bedingung *nothwendig*.”)

But if freedom of the will, as a rational presupposition, is the condition of all voluntary actions, must not the converse hold as well, namely that all voluntary actions are conditional upon the capacity of the rational agent, as such, to be conscious of his causality through reason, i.e., to presuppose his own freedom of the will? So, if we are committed to saying that criminal actions must be voluntary, then we have to admit that the criminal exhibits (some) capacity for rational agency. Of course, there is the controversy I raised in note 12 above, namely that in the *Groundwork* Kant seems to assume that only the actions of a truly rational autonomous agent are actually voluntary, but the solution to this problem – which involves the distinction between *Wille* and *Willkür* as I note above – does not involve denying that voluntary actions presuppose *some* capacity for rational agency, even if it does involve watering-down, minimizing or restricting the Kantian conception of rational agency to the domain of choice (i.e., to *Willkür* as an integral but fallible aspect of *Wille*).

<sup>56</sup> See MS, AA 06: 211–221; 370–376. For a more thorough discussion, see Hill (1992).

impulses, and states that human choice (*Willkür*) can “indeed be *affected* but not *determined* by impulses, and is therefore still of itself (apart from an acquired proficiency of reason) not pure but can still be determined to actions by pure will”<sup>57</sup>. Thus, despite its impurity, human choice retains the *potential* for an acquired proficiency of reason, i.e., to be determined by pure will (by the dictates of reason as generated in *Wille*). Kant offers no account of the conditions under which this *potential* to be determined by pure will, which is apparently inherent to human choice as such, could itself be extinguished through some wrongful choice. (Of course, the successful suicide can be said to have extinguished this potential in himself, but that, arguably, is the sole, and all too obvious, exception.) Furthermore, there is nothing here to suggest that the commission of a criminal act at a given point in time, which could only have transpired if the criminal’s choice, as it were, conceded to some (vicious) impulse, could place the wrongdoer beyond all future redemption by irrevocably and completely extinguishing the capacity for rational agency that is otherwise present within him. A total failure to abide by reason on a given occasion is just not equivalent to the absolute loss of reason.

Now, although it is difficult to make sense of the suggestion that the criminal ceases to *be* rational as soon as he commits a crime, Kant did have some rather callous things to say about the suicide, which could, I presume, pertain to the criminal, too. In the *Lectures on Ethics*, Kant claims that “suicide evokes horror, in that man thereby puts himself below the beasts. We regard a suicide as a carcass [...]”<sup>58</sup> He goes on to say that one who attempts suicide loses his claim to human worth, that “he who takes himself for [a beast], who fails to respect humanity, who turns himself into a thing, becomes an object of free choice for everyone; anyone, thereafter, may do as he pleases with him; he can be treated by others as animal or a thing; he can be dealt with like a horse or a dog, for he is no longer a man; he has turned himself into a thing, and so cannot demand that others should respect the humanity in him”.<sup>59</sup> But having previously said that the horror of suicide consists in one’s making his freedom, as the acme of life, a principle for his own destruction,<sup>60</sup> Kant does

<sup>57</sup> MS, AA 06: 213; 375. “Die menschliche Willkür ist dagegen eine solche, welche durch Antriebe zwar *afficirt*, aber nicht *bestimmt* wird, und ist also für sich (ohne erworbene Fertigkeit der Vernunft) nicht rein, kann aber doch zu Handlungen aus reinem Willen bestimmt werden.”

<sup>58</sup> VE, AA 27: 372; 146. “Es erweckt also der Selbstmord ein Grausen, indem der Mensch sich dadurch unter das Vieh setzt. Wir sehen einen Selbstmörder als ein Aas an [...]”.

<sup>59</sup> VE, AA 27: 373; 147. “Die Thiere werden hier auch als Sachen angesehen; der Mensch aber ist keine Sache; disponirt er demohngeachtet über sein Leben, so versetzt er sich also in den Werth des Viehes. Wer sich aber als so etwas nimmt, der die Menschheit / nicht respectirt, der sich zur Sache macht, der wird ein Object der freyen Willkühr für jedermann; mit dem kann hernach ein jeder machen was er will; [...] man kann sich an ihm exerciren so wie an einem Pferde oder Hunde; denn er ist kein Mensch mehr; er hat sich selbst zur Sache gemacht, demnach kann er nicht fordern, daß andre seine Menschheit in ihm respectiren sollen, da er sie selbst schon weggeworfen hat.”

<sup>60</sup> See note 19 above.



not explain why or how it could be permissible for us to apply the suicide's own principle in our comportment with respect to him. If the suicide's use of freedom to destroy his life is horrific, is our use of freedom to destroy it, in the spectacle of using him as some dog, any less so?

Then, Kant appears to retract his statements just as soon as he makes them, for he states: "Humanity, however, is worthy of respect, and even though somebody may be a bad man, the humanity in his person is entitled to respect."<sup>61</sup> And in the *Critique of Practical Reason*, Kant reaffirms this in saying that a "human being is indeed unholy enough but the *humanity* in his person must be holy to him"<sup>62</sup>. The suggestion that there is a part of us, namely our humanity, that demands our respect despite a vicious (or "unholy") disposition is surely at odds with the claim that the suicide, or any other unsavoury character, may be treated like a dog. Thus, it is difficult not to dismiss Kant's pronouncements regarding the notion that the suicide, or criminal, has altogether *discarded* his humanity, as mere misstatements.

Moreover, Kant's philosophy of punishment as articulated in the *Metaphysics of Morals* explicitly precludes the state from disposing of the criminal as a mere thing. Retributive punishment is imposed only on the assumption that the criminal acted voluntarily, and precisely for that reason – indeed, this is an essential element of any retributivist theory of punishment.<sup>63</sup> For, retributivism holds that the extent to which the criminal act may be imputed to the criminal, i.e., the extent of his responsibility, determines, limits, and *justifies*, the extent to which he suffers punishment. Nothing can be done to the criminal by way of punishment apart from that which he may be said to deserve *as a being responsible for his actions*. Indeed, Kant argues specifically that punishment must be limited in accordance with the principle of balanced retribution because the "innate personality" of the criminal protects him from disproportionate or excessive suffering:

*Punishment by a court (poena fornesis)* [...] can never be inflicted merely as a means to promote some other good for the criminal himself or for civil society. It must always be inflicted upon him only *because he has committed a crime*. For a human being can never be treated merely as a means to the purposes of another or be put among the objects of rights to things: his innate personality protects him from this, even though he may be condemned to lose his civil personality.<sup>64</sup>

<sup>61</sup> VE, AA 27: 373; 147. "Die Menschheit ist aber achtungswerth, und wenn auch der Mensch ein schlechter Mensch ist, so ist doch die Menschheit in seiner Person achtungswerth."

<sup>62</sup> KpV, AA 05: 87; 210. "Der Mensch ist zwar unheilig genug, aber die *Menschheit* in seiner Person muß ihm heilig sein."

<sup>63</sup> See Scheid (1983).

<sup>64</sup> MS, AA 06: 331; 473. "*Richterliche Strafe (poena forensis)*, die von der *natürlichen (poena naturalis)*, dadurch das Laster sich selbst bestraft und auf welche der Gesetzgeber gar nicht Rücksicht nimmt, verschieden, kann niemals bloß als Mittel ein anderes Gute zu befördern für den Verbrecher selbst, oder für die bürgerliche Gesellschaft, sondern muß jederzeit nur darum wider ihn verhängt werden, *weil er verbrochen* hat; denn der Mensch kann nie bloß als Mittel zu den Absichten eines Anderen gehandhabt und unter die Gegenstände des Sachenrechts gemengt werden, wovider ihn seine angeborne Persönlichkeit schützt, ob er gleich die bürgerliche einzubüßen gar wohl verurtheilt werden kann."

This passage clearly discloses (a) that Kant does, after all, identify the criminal as a *human being*, not a mere thing, even subsequent to his having committed the crime, (b) that the criminal, given the demands of Kant's retributivism, must have acted voluntarily, to have been a "free cause" of the act, which is to say that even a criminal act exhibits some (perhaps minimal) degree of *noumenal* freedom sufficient to ground the wrongdoer's responsibility, and, (c) that what the criminal loses is his "civil personality", not his "innate personality".

The fact that the criminal continues, according to Kant, to be endowed with innate personality is sufficient to render capital punishment irrational from a juridical perspective, i.e., from the standpoint of the co-legislator in the original contract (or of the general will). Now, I take Kant to be saying that the forfeiture of his civil personality deprives the criminal of a voice in legislation, i.e., that criminals having been severed from the original contract, their *external* freedom need not figure in the operation of practical reason in its ordering of external relations among citizens.

I would explain this in less abstract terms as follows. **Because the essential purpose of civil society is to secure an equal sphere of external freedom for all citizens, the loss of *civil* personality entails the loss of one's external freedom, or a forfeiture of the protection that civil society is otherwise meant to guarantee and secure for all citizens.** This is just to say that the criminal no longer has a right to be an authoritative participant in decisions pertaining to his own *external* freedom, so that he could, for instance, be chained or locked up for the entirety of his existence. And although this severely diminishes his autonomy – or, more precisely, restricts his capacity to *exercise* it in the world around him – it cannot mean that the criminal irrevocably ceases to *be* rational. A necessary consequence of terminating the criminal's external freedom through execution, is the rooting-out his inner freedom, which I think Kant associates with "innate personality".<sup>65</sup> What makes capital punishment irrational from the standpoint of a citizen-participant in the original contract (i.e., a self-legislating facility worthy of identifying with the general will) is that it purports to extinguish the inner autonomous life of the criminal – the source of morality that always already resides within him, and makes him *capable* of virtue and rational self-perfection *in potentia*, if not in actuality – which the two actually share as *human beings*. Kant's effort to sever the co-legislator of the penal law and the criminal on whom it is applied must, for this reason, fail, and so capital punishment is not a provision that a politically autonomous citizen can rationally consent to.

<sup>65</sup> A full exposition of Kant's distinction between the "innate" and "civil" personalities would require a lengthy consideration of the further distinctions he makes between "moral personality", "psychological personality" and "personality" proper. Thankfully, doing this is unnecessary, because for our purposes what matters most is that the loss of "civil personality" just does not mean that the criminal ceases to *be* rational, or to possess inner freedom. For a valiant though ultimately convoluted effort to disentangle Kant's complex web of distinctions between the various types of "personality", see Merle (2000), 328–331.

*The second part of Kant's response to Beccaria*

In the second part of his response to Beccaria, Kant again stresses the perspective of the criminal. This is coupled with a distinct suggestion that punishment is not within the legislative scope of the general will in the first place, so that Beccaria is supposedly wrong to say that our assent to the original contract entails our granting the state implicit permission to be punished in case we commit a crime. Kant states:

In other words, it is not the people (each individual in [a civil union]) that dictates capital punishment but rather the court (public justice), and so another than the criminal; and the social contract contains no promise to let oneself be punished and so to dispose of oneself and one's life. For if the authorization to punish had to be based on the offender's *promise*, on his *willing* to let himself be punished, it would also have to be left to him to find himself punishable and the criminal would be his own judge. – The chief point of error in this sophistry consists in its confusing the criminal's own judgment (which must necessarily be ascribed to his *reason*) that he has to forfeit his life with a resolve on the part of his *will* to take his own life, and so in representing as united in one and the same person the judgement upon a right and the realization of that right.<sup>66</sup>

There are three points to be made with respect to this passage.

First, we must note Kant's persistent emphasis on the question concerning the appropriate scope of the *offender's* juridical powers. He makes the all too obvious point that the offender's promise cannot authorize punishment, or that the criminal cannot be his own judge. But this is deeply problematic, because the real question is *not* what the citizen can legitimately will once he has become a criminal, but rather what he can rationally assent to *prior* to this, i.e., prior to having forfeited his civil personality (which, I readily concede, is a legitimate consequence of criminal wrongdoing).<sup>67</sup> Indeed, it is difficult to see how Kant expects to defeat Beccaria, who, as I see it, *is* asking the right question, namely: may the *citizen* give his rational consent to a law in accordance with which he could be put to death? The fact that it is not the criminal, *as such*, who makes this determination, "and so another", is beside the point. All in all, Kant does not squarely address Beccaria's question.

<sup>66</sup> MS, AA 06: 335; 476. "Mit andern Worten: nicht das Volk (jeder einzelne in demselben), sondern das Gericht (die öffentliche Gerechtigkeit), mithin ein anderer als der Verbrecher dictirt die Todesstrafe, und im Socialcontract ist gar nicht das Versprechen enthalten, sich strafen zu lassen und so über sich selbst und sein Leben zu disponiren. Denn wenn der Befugniß zu strafen ein *Versprechen* des Missethätters zum Grunde liegen müßte, sich strafen lassen zu *wollen*, so müßte es diesem auch überlassen werden, sich straffällig zu finden, und der Verbrecher würde sein eigener Richter sein. – Der Hauptpunkt des Irrthums (πρωτοψευδος) dieses Sophisms besteht *darin*: daß man das eigene Urtheil des Verbrechers (das man seiner *Vernunft* nothwendig zutrauen muß), des Lebens verlustig werden zu müssen, für einen Beschluß des *Willens* ansieht, es sich selbst zu nehmen, und so sich die Rechtsvollziehung mit der Rechtsbeurtheilung in einer und derselben Person vereinigt vorstellt."

<sup>67</sup> Recall that an essential aspect of rationality is the capacity to be a legislator of the law, and constitutes our dignity as human beings (see note 14 above). But I would argue that the forfeiture of "civil personality" entails a loss of the entitlement for *external* lawgiving (within the civil union), and cannot mean a loss of the *capacity* for internal lawgiving, which for Kant is always already integral to the human mind.

Second, Kant's claim that the "social contract contains no promise to let oneself be punished" is also deeply problematic. For, prior to this he stated that, in a civil union, "I subject myself together with everyone else to the laws, which will naturally also be penal laws". Moreover, we saw earlier that right is inherently connected with an authorization to use coercion, and that this is the basis of our ability to grant rational assent to penal laws, which, in threatening sanctions in case of non-compliance, are coercive in nature.<sup>68</sup> And the notion of a "promise" to let oneself be punished must be qualified in the following terms: it is the idea that if the punitive laws that we collectively place ourselves under ultimately issue from our will, we cannot reasonably deny that these laws may, one day, be applied to ourselves. For how could we subject ourselves to penal laws without contemplating, at least implicitly, the possibility that we may suffer the sanctions they carry, i.e., in case we become criminals?<sup>69</sup> Kant further complicates matters by arguing that it is not the people who dictate punishment, but a court (public justice). But the authority of any court, along with the laws it is meant to apply, ultimately resides in the general will as the ordering principle of any rational, republican constitution. The punitive determination of a court cannot be severed from its justificatory ground, which is the people's assent in the original contract. So, once again, Beccaria's question remains unanswered.

Third, regarding the concluding point in Kant's passage, namely that the criminal's rational judgement must be distinguished from a resolve on the part of his will: the question raised by Beccaria is, precisely, whether there can be any judgement to forfeit one's life on the part of reason in the first place (given that life is an essential precondition of the possibility any judgement). Kant may be right to distinguish the criminal's rational faculty of reason from a definitive resolve on the part of his will (which, I believe refers to the executive function of *Willkür*, and not *Wille* proper),<sup>70</sup> and his claim actually bolsters my contention that even the criminal, as also a human being, retains his capacity for reason. This distinction, however, has no effect on the issue at hand.

<sup>68</sup> See Hill (1999), 428 ff.

<sup>69</sup> The notion of a "promise" in this context is misleading, and Kant capitalizes on this. There is, strictly speaking, no *promise* to let oneself be punished. I owe the following observation to Arthur Ripstein. Kant's insistence that there is no promise to let oneself be punished is based on the idea that we put ourselves under coercive laws and thus bind ourselves coercively (as we bind ourselves by posting collateral for a loan) and not that we undertake to accept the punishment if it comes. (We do not promise to give up the collateral if we default on the loan; we set things up so that the lender can seize it if we default. Promising is not part of the account, even though the transaction is voluntary.) Irrespective of the reference to "promising", my point is simply that there are certain inexorable limits on the *kinds* of coercive laws under which we may bind ourselves. (We cannot, for instance, post our own body as collateral for a loan, or set things up so that the lender can enslave us upon default.) Giving our assent to capital penal laws is not a rational way of binding ourselves coercively. For, as I argue further below, capital punishment is not *merely* a form of coercion in the first place. Execution is not a legitimate instrument within a system of coercion, under which we might otherwise rationally bind ourselves for the purposes of security.

<sup>70</sup> Beck points out that, in the Kant's usage, *Wille* (will) often stands for *Willkür* (human choice), though never the converse. See Beck (1993), 39.

*Capital punishment does not have place in the scheme  
contemplated by the Rechtslehre*

There is a further argument against capital punishment implicit in what I have said above that needs to be drawn out, for it is potentially decisive on its own. As it happens, Kant's own strict division between the ethical and political-judicial domains invalidates his support for the death penalty. Kant delimits the scope of Right in the following terms:<sup>71</sup>

The concept of right, insofar as it is related to an obligation corresponding to it (i.e., the moral concept of right), has to do, *first*, only with the external and indeed practical relation of one person to another, insofar as their actions, as deeds, can have (direct or indirect) influence on each other. [...] [I]n this reciprocal relation of choice no account at all is taken of the *matter* of choice, that is, of the end each has in mind with the object he wants [...] All that is in question is the *form* in the relation of choice [...]<sup>72</sup>.

As we have seen, right is connected with an authorization to use coercion. But if right pertains strictly to the external relation of one person to another, the concept of coercion must be likewise delimited. In other words, coercion must involve restricting the *external* freedom a person, and *only* this. This goes along with a well-accepted interpretation of Kant's *Rechtslehre*, namely that it cannot be the aim of law and politics to force citizens to be virtuous<sup>73</sup> and, indeed, as is often pointed

<sup>71</sup> For a thorough discussion, see Pogge (1997), 165, who argues that: “[I]n associating *Recht* with a *complete* set of conditions [i.e., *Right* as “the whole of the conditions under which the choice of one can coexist with the choice of the other according to a universal law of freedom” (MS 6: 230; 387)], Kant is also suggesting the exclusion of any redundant conditions. The set must only be *inclusive* so that conditions it contains are jointly sufficient, but also be *exclusive* so that each condition it contains is individually necessary, for the maintenance of mutually secure domains of external freedom. *Recht* excludes any conditions that make any contribution to such mutual security [...]”.

<sup>72</sup> MS, AA 06: 230; 387. “Der Begriff des Rechts, sofern er sich auf eine ihm correspondirende Verbindlichkeit bezieht, (d.i. der moralische Begriff desselben) betrifft *erstlich* nur das äußere und zwar praktische Verhältniß einer Person gegen eine andere, sofern ihre Handlungen als Facta aufeinander (unmittelbar oder mittelbar) Einfluß haben können. Aber zweitens bedeutet er nicht das Verhältniß der Willkür auf den *Wunsch* (folglich auch auf das bloße Bedürfniß) des Anderen, wie etwa in den Handlungen der Wohlthätigkeit oder Hartherzigkeit, sondern lediglich auf die Willkür des Anderen. *Drittens*, in diesem wechselseitigen Verhältniß der Willkür kommt auch gar nicht die *Materie* der Willkür, d.i. der Zweck, den ein jeder mit dem Object, was er will, zur Absicht hat, in Betrachtung, z.B. es wird nicht gefragt, ob jemand bei der Waare, die er zu seinem eigenen Handel von mir kauft, auch seinen Vortheil finden möge, oder nicht, sondern nur nach der *Form* im Verhältniß der beiderseitigen Willkür, sofern sie bloß als frei betrachtet wird, und ob durch die Handlung eines von beiden sich mit der Freiheit des andern nach einem allgemeinen Gesetze zusammen vereinigen lasse.”

<sup>73</sup> See Gregor (1963), 26–38. Also, it must be noted that Kant excluded inner motives from the appropriate and practicable scope of judicial assessments of culpability, as Hill (1999), 429 points out: “Kant makes clear that judicial punishment must be for (intentional) ‘external acts’ as they can be assessed in a public court of law. The law cannot assess ‘inner’ moral worth of offenders because that would require knowing more about the agent’s motives

out, even a race of self-serving devils could, in principle, be brought under a condition of right according to Kant.<sup>74</sup>

But *punishment* as a form of coercion must be confined the external manipulation of the wrongdoer, since the governance of external relations between citizens is its justificatory point. Capital punishment exceeds the limits implicit in the Kantian conception of (external) coercion because it reaches the inner life of the individual. It is a truly unique form of punishment in this respect, as, arguably, no other conceivable form of punishment could negate the inner freedom of the subject while it leaves him alive. (Even the Pavlovian mental-conditioning contemplated in Burgess' *A Clockwork Orange* involved manipulating the offender's incentive structure through the introduction of novel impulses and aversions, and thus counts as external coercion in the Kantian sense.) This means, then, that the Kantian authorization of coercion cannot encompass the elimination of the subject of coercion. The very idea of coercion, as a hindrance of a hindrance to freedom, presupposes the continued existence of the subject being coerced, i.e., of a subject *capable* of being externally coerced. Or, simply, capital punishment is just not a form of coercion. Because it involves the elimination of the subject, it cannot be merely a rightful (countervailing) hindrance. (To say that is rather like saying brain-death is merely a hindrance to thinking.) An act that annihilates the external freedom of a person, or his capacity for any kind of freedom at all, simply cannot be placed in the same category as acts that aim to confine, control or otherwise manipulate the external freedom of a person, i.e., acts that we can identify as being *coercive* in a strict and straightforward sense.

There is another way of putting this. The target of coercion, as a means of ordering external relations, is actually the faculty of choice (*Willkür*). As we have seen, the distinguishing feature of external lawgiving is that it constrains the subject via an incentive drawn from the subject's aversions.<sup>75</sup> And what coercion entails, in part, is a threat to impose sanctions, to deploy force or inflict pain. That is, coercion is a means of providing the subject, the addressee of the law, with an incentive to act in accordance with right. For, in so far as he is also susceptible to inclination, the human subject can be expected to have a natural aversion to being punished and will avoid sanctions by complying with the law, i.e., by bringing his external comportment in line with its demands. But the idea of coercion therefore also presupposes the Kantian conception of human choice (*Willkür*), which, as we have seen, can be "*affected*, but not determined by impulses" according to Kant. For coercion is, precisely, an attempt to affect our choice for the purposes of

and 'will' than we can determine with confidence. The justifying purpose of a practice of punishment, then, cannot be to make wrongdoers suffer according to their intrinsic moral deserts. Kant's justification, in fact, lies elsewhere. The general authority for state coercive powers, on which the right to punish is based, is the authority to 'hinder hindrances of freedom'."

<sup>74</sup> See ZeF, AA 08: 366.

<sup>75</sup> See MS, AA 06: 219; 383.

right.<sup>76</sup> Coercive threats would, obviously, be useless if they could not generate an impulse that affects our choice. Thus, to say again, the appropriate target of coercion is *Willkür* – and, in light of Kant’s delimitation of right, it is not *Wille*. Because it also targets *Wille*, capital punishment violates this delimitation.

Now, we have seen that Kant’s view of freedom involves the distinction between *Wille* and *Willkür* – and that it is the latter that accounts for a free, imputable act, which is freely made even where it is made irrationally, i.e., against the dictates of *Wille* as practical reason.<sup>77</sup> So, as long as a criminal act may be said to be a wrongful exercise of *Willkür*, the criminal is culpable and thus liable for punishment. But *Wille* and *Willkür* are two aspects of human freedom as such,<sup>78</sup> and it would make no sense to say that a person may exhibit *Willkür* without being also possessed of *Wille*. *Willkür* is supposed to be merely a kind of executive function of *Wille* (as the legislative function of pure reason), but given human weakness does not always comply with its dictates. Nevertheless, despite his wrongdoing, the criminal, as a person, is possessed of *Wille* (a rational will), and in so far as he is a *person* even after having committed a crime, he continues to be a repository of practical reason, at the very least *in potentia*. So how is it possible that a wrongful exercise of *Willkür* may legitimately expose the wrongdoer, as a person, to a death that also extinguishes *Wille*? How could practical reason sanction this under any conceivable circumstances? Put differently: how could that aspect of humanity that constitutes our utmost dignity ever be justifiably extinguished, even by way of doing justice in the form of the *ius talionis* – i.e., even by way of punishing that aspect of humanity that constitutes the greatest possible evil? In a manner of speaking, the *Wille* cannot be made to pay for the sins of *Willkür*, and that is just what capital punishment purports to do.

### Conclusion

We saw at the outset that Kant insists on a strict equivalence between crime and punishment, and that, for him, there can be no such equivalence between death and imprisonment, no matter how wretched. Thus, per the *ius talionis*, murder must be met with death. Does the argument developed here undermine, then, the notion of the *ius talionis*? Here, we have to consider that there can never be a strict, literal equivalence between crime and punishment in any case, for the simple reason that one is an irrational exercise of freedom on the part of an individual and the other is supposed to be a rational application of force on the part of a state agency. (It is thus

<sup>76</sup> That is, coercion presupposes weakness of *Willkür*, that it may be affected in such a way as to induce compliance with right. The threat of punishment provides the would-be criminal with an external incentive to avoid crime.

<sup>77</sup> See authors cited in note 12 above.

<sup>78</sup> According to Benson (1987), 571: “*Willkür* and *Wille* are, then, necessarily interconnected and distinct aspects of free will. This concept of free will is, according to Kant, basic to both parts of the metaphysics of morals [i.e., the Doctrines of Right and Virtue].”



not surprising that retributivists often say that the requisite proportionality between crime and punishment is a question of *form*.) Crime and punishment cannot be placed in the same category as uses of human freedom in general (as forms of human agency), and so there may be certain inexorable limitations on the latter that preclude the strictness of equivalence that Kant demands. And why should the act of the executioner resemble, even in “spirit” if not “to the letter”, the act of the murderer? Isn’t this disturbing: that the state should take its cue from the murderer himself? How could the actions of the state as punisher *reflect* the actions of the criminal in their character? Doesn’t a residue of “evil” attach to the actions of the executioner, in so far as it is supposed to reflect the magnitude of evil committed by the criminal? Finally, it is difficult to imagine that Kant would have been so inflexible on this point as to deny that an exception to his philosophy of punishment could be made, particularly where it is shown that a given punitive practice is inherently irrational.<sup>79</sup> If the overarching purpose of right is to provide the requisite security for a condition of maximal freedom, wherein human autonomy may ultimately flourish,<sup>80</sup> it is difficult to see how a punitive practice that involves an absolute negation of freedom could be part of a condition of right in this sense.

#### References

- Beck, Lewis White, 1993. Kant’s Two Conceptions of the Will in Their Political Context. In: *Kant and Political Philosophy*. Ed. by R. Beiner. New Haven: Yale University Press.
- Benson, Peter, 1987. External Freedom According to Kant. In: *Columbia Law Review* 87, 559–579.
- Byrd, Sharon B., 1989. Kant’s Theory of Punishment: Deterrence in its Threat, Retribution in its Execution. In: *Law and Philosophy* 8, 151–200.
- Carr, Craig L., 1989. Kant’s Theory of Political Authority. In: *History of Political Thought* 10.4, 719–731.
- Dodson, Kevin E., 1997. Autonomy and Authority in Kant’s *Rechtslehre*. In: *Political Theory* 25.1, 93–111.
- Dyke, C., 1969. Collective Decision Making in Rousseau, Kant, Hegel, and Mill. In: *Ethics* 80.1, 21–37.
- Gorner, Paul, 2000. The Place of Punishment in Kant’s *Rechtslehre*. In: *Kantian Review* 4, 121–130.

<sup>79</sup> Kant ought to have allowed an exception to his demand for a strict, literal equivalence between crime and punishment in the case of murder, as he does for rape, pederasty and bestiality (MS, AA 06: 363; 498), because it is impossible to will a death, i.e., to reproduce at a rational level what the murderer has chosen irrationally

<sup>80</sup> See Guyer (1997), who argues that the *ultimate* purpose of justice is to secure conditions under which morality may be realized (even though justice is not directly concerned with enforcing Kantian morality).

- Gregor, Mary, 1963. *Laws of Freedom*. Oxford: Oxford University Press.
- Guyer, Paul, 1997. Justice and Morality: Comments on Allen Wood. In: *Southern Journal of Philosophy* 36 (Supp.), 21–28.
- Hill, Thomas E., 1992. Kant's Theory of Practical Reason. In: *Dignity and Practical Reason in Kant's Moral Theory*. Ithaca: Cornell University Press.
- , 1999. Kant on Wrongdoing, Desert and Punishment. In: *Law and Philosophy* 18, 407–441.
- Hudson, Hud, 1991. *Wille, Willkür*, and the Imputability of Immoral Actions. In: *Kant-Studien* 82.2, 179–196.
- Kersting, Wolfgang, 1992. Kant's Concept of the State. In: *Essays on Kant's Political Philosophy*. Ed. by H. Williams. Cardiff: University of Wales Press.
- Merle, Jean-Christophe, 2000. A Kantian Critique of Kant's Theory of Punishment. In: *Law and Philosophy* 19, 311–338.
- Pogge, Thomas W., 1997. Is Kant's *Rechtslehre* Comprehensive? In: *The Southern Journal of Philosophy* 36 (Supp.), 161–187.
- Rawls, John., 2000. *Lectures on the History of Moral Philosophy*. Cambridge, Mass.: Harvard University Press.
- Scheid, Don E., 1983. Kant's Retributivism. In: *Ethics* 93, 262–282.
- Schneewind, J. B., 1992. Autonomy, Obligation and Virtue: An Overview of Kant's Moral Philosophy. In: *The Cambridge Companion to Kant*. Ed. by P. Guyer. Cambridge: Cambridge University Press.
- Timmons, Mark, 1994. Evil and Imputation in Kant's Ethics. In: *Jahrbuch für Recht und Ethik* 2, 113–141.
- Weinrib, Ernest, 1987. Law as a Kantian Idea of Reason. In: *Columbia Law Review* 87, 472–508.